Chicago-Kent Law Review

Volume 100
Issue 1 *AI Disrupting Law*

Article 5

8-15-2025

Win-Win: How to Remove Copyright Obstacles to AI Training While Ensuring Author Remuneration (and Why the AI Act Fails to do the Magic)

Martin Senftleben
University of Amsterdam, M.r.f.senftleben@uva.nl

Win-Win: How to Remove Copyright Obstacles to AI Training While Ensuring Author Remuneration (and Why the AI Act Fails to do the Magic)

Cover Page Footnote

The author wishes to thank Carys Craig, Graeme Dinwoodie, Katharina de la Durantaye, Niva Elkin-Koren, Christophe Geiger, Bernt Hugenholtz, Edward Lee, Jessica Litman, and Pam Samuelson for invaluable comments and feedback on earlier versions of this article. Thanks are also due to the editors and staff of the Chicago-Kent Law Review for their thoughtful comments and diligent work on the article manuscript. All views expressed in the article (and all mistakes) are the author's own.

WIN-WIN: HOW TO REMOVE COPYRIGHT OBSTACLES TO AI TRAINING WHILE ENSURING AUTHOR REMUNERATION (AND WHY THE AI ACT FAILS TO DO THE MAGIC)

MARTIN SENFTLEBEN*

INTRODUCTION

With the adoption of the AI Act ("AIA"),[1] the EU has substantially enhanced the rules governing the training of generative AI systems and, more specifically, the interface with copyright protection. The AI Act clarifies that, from an EU perspective, reproductions carried out for AI training purposes have copyright relevance and require the authorization of right holders unless a copyright exception permits the use, such as the general exemption of text and data mining ("TDM") in Article 4 or the more specific rule for scientific TDM in Article 3 of the 2019 Directive on Copyright in the Digital Single Market ("CDSMD").[2] The AI Act also confirms that the rights reservation system of Article 4(3) CDSMD must be observed when TDM falls outside the scope of the scientific TDM exemption and goes beyond mere temporary copying: copyright owners seeking to prevent the use of their works for AI training purposes can reserve their rights by declaring an "opt-out" in an appropriate—machine-readable—manner.

Remarkably, the AI Act seeks to universalize this approach and achieve a "Brussels effect."[3] Regardless of whether the training has taken place in

* The author wishes to thank Carys Craig, Graeme Dinwoodie, Katharina de la Durantaye, Niva Elkin-Koren, Christophe Geiger, Bernt Hugenholtz, Edward Lee, Jessica Litman, and Pam Samuelson for invaluable comments and feedback on earlier versions of this article. Thanks are also due to the editors and staff of the Chicago-Kent Law Review for their thoughtful comments and diligent work on the article manuscript. All views expressed in the article (and all mistakes) are the author's own.

1. Parliament and Council Regulation (EU) 2024/1689, 2024 O.J. 1 (L number not yet assigned) [hereinafter Regulation 2024/1689].

2. Parliament and Council Directive (EU) 2019/790, 2019 O.J. (L 130) 92 [hereinafter Directive 2019/790]. For an overview of relevant copyright exceptions in EU copyright law, see Martin Senftleben, *Study on EU Copyright and Related Rights and Access to and Reuse of Data*, at 27, 36–37 (2022), https://data.europa.eu/doi/10.2777/78973 [https://perma.cc/9P8R-WCAD].

3. *See* A. BRADFORD, THE BRUSSELS EFFECT: HOW THE EUROPEAN UNION RULES THE WORLD (2020). For a discussion of recent EU online platform regulation that may have repercussions in other regions, including the U.S., see Martin Husovec & Jennifer Urban, *Will the DSA have the Brussels Effect?*, VERFASSUNGSBLOG (Feb. 21, 2024), https://verfassungsblog.de/will-the-dsa-have-the-brussels-effect/ [https://perma.cc/4RHK-U488].

the EU or elsewhere, it imposes a market ban on AI systems that have not been trained in accordance with EU requirements, including the obligation to observe opt-outs declared under Article 4(3) CDSMD. To enable right holders to police AI training processes, the AI Act also introduces a new transparency obligation. Developers of generative AI systems must submit sufficiently detailed information on work repertoires that have been used for training purposes.

Before embarking on an analysis of these new rules, Part I sheds light on the policy objective underlying this copyright package in the AI Act: the intention to ensure that authors are properly remunerated for the use of their works in AI training processes. The discussion contrasts this policy goal with the societal interest in AI innovation and unbiased, high-quality AI systems. Part II weighs the potential benefits to authors against the regulatory burdens the AI Act imposes on AI trainers. More specifically, it will ask whether the new rules encourage not only rights clearance at the industry level but also payments to individual authors. It explains the EU strategy to extend opt-outs on EU territory to other regions and discusses the new transparency obligations which the EU legislator deems necessary to enforce copyright in AI training contexts. The analysis leads to the insight that the EU approach seeking to ensure the payment of remuneration upfront, as a prerequisite for lawful AI training, is likely to impede, if not thwart, AI innovation.

Against this background, Part III explores alternative solutions. It shows that, instead of imposing heavy burdens on AI development, lawmakers can use the offer and commercialization of fully trained AI systems as a reference point for remuneration systems. Following this alternative avenue, the remuneration obligation concerns the final stage when generative AI products and services are brought to the market.[4] In Part IV, I conclude that, in contrast to the upfront payment approach underlying the AI Act, this alternative solution refrains from encumbering the AI training process with obligations to observe opt-outs, establish lists of training resources, and pay remuneration. Before following in the footsteps of the European AI Act, policymakers in other regions should evaluate the advantages of this alternative avenue. Arguably, it offers considerable flexibility to reconcile author remuneration interests with the broader societal interest in AI innovation.

4. Martin Senftleben, *Generative AI and Author Remuneration*, 54 IIC – INT'L REV. INTELL. PROP. & COMPETITION L. 1535, 1549–56 (2023).

I. COMPETING POLICY OBJECTIVES

With the evolution of generative AI systems,[5] machine-made productions in the literary and artistic field have reached a level of refinement that allows them to replace human creations.[6] Inevitably, the increasing sophistication of AI systems will disrupt the market for human literary and artistic creations. Generative AI systems provide literary and artistic output much faster and cheaper. They are capable of mimicking human creativity because human works have been used as training material.[7] Analysing existing literary and artistic creations that serve as input data, machine-learning algorithms are able to recognize patterns and similarities. Following this deductive method, a generative AI system learns how to produce novel literary and artistic output by imitating the style of human works.[8] The machine-

5. For an attempt of a legal definition of generative AI, see *Draft Compromise Amendments on the Draft Report Proposal for a Regulation of the European Parliament and of the Council on Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, at 40, COM (2021) 0206 (Sept. 5, 2023). For examples of current generative AI products and services, see *Image Models*, STABILITY.AI, https://stability.ai/stablediffusion [https://perma.cc/M78F-BZ27]; MIDJOURNEY, https://www.midjourney.com/home/ [https://perma.cc/876C-KQXK]; *Dall-E* 2, OpenAI, https://openai.com/dall-e-2 [https://perma.cc/C45G-S3W7]; *Adobe Firefly*, ADOBE, https://www.adobe.com/sensei/generative-ai/firefly.html [https://perma.cc/2KFD-8KV4].

6. *Cf.* Pamela Samuelson, *Fair Use Defenses in Disruptive Technology Cases*, 72 UCLA L. REV. (forthcoming 2025); Christophe Geiger, *When the Robots (Try to) Take Over: Of Artificial Intelligence, Authors, Creativity and Copyright Protection*, *in* KREATION INNOVATION MÄRKTE - CREATION INNOVATION MARKETS: FESTSCHRIFT FÜR RETO M. HILTY 67, 67-68 (Florent Thousvenin et al. eds., 2024); Christophe Geiger & Vincenzo Iaia, *The Forgotten Creator: Towards a Statutory Remuneration Right for Machine Learning of Generative AI*, 52 COMPUT. L. & SEC. REV. 1, 1-2 (2024), https://doi.org/10.1016/j.clsr.2023.105925 [https://perma.cc/4BK2-WRDF]; Giancarlo Frosio, *Should We Ban Generative AI, Incentivise It or Make It a Medium for Inclusive Creativity?*, *in* A RESEARCH AGENDA FOR EU COPYRIGHT LAW (Enrico Bonadio & Caterina Sganga eds., 2025) (forthcoming Mar. 2025); Senftleben, *supra* note 4, at 1535; Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 TEX. L. REV. 743, 766–67 (2021).

7. In this category, a distinction can be drawn between "machine-learning" and "deep-learning" algorithms. *See* Daniel J. Gervais, *The Machine as Author*, 105 IOWA L. REV. 2053, 2055–59 (2020); Jane C. Ginsburg & Luke Ali Budiardjo, *Authors and Machines*, 34 BERKELEY TECH. L.J. 343, 401–02 (2019); Jean-Marc Deltorn, *Disentangling Deep Learning and Copyright*, *in* 5 AMI - TIJDSCHRIFT VOOR AUTEURS-, MEDIA- EN INFORMATIRECHT, 172, 173–74 (2018); Annemarie Bridy, *Coding Creativity: Copyright and the Artificially Intelligent Author*, 5 STAN. TECH. L. REV. 1, 3 (2012); Margaret A. Boden, *Computer Models of Creativity*, 30 AI MAGAZINE 23, 23 (2009). For a practical example of AI-generated imitations of human vocals, see Laura Snapes, *AI Song Featuring Fake Drake and Weeknd Vocals Pulled from Streaming Services*, THE GUARDIAN (Sept. 8, 2024, 05:37 PM), https://www.theguardian.com/music/2023/apr/18/ai-song-featuring-fake-drake-and-weeknd-vocals-pulled-from-streaming-services[https://perma.cc/95NC-TR5D].

8. *See* Jacob Zunkula & Aaron Mok, *ChatGPT May be Coming for our Jobs. Here Are the 10 Roles That AI is Most Likely to Replace*, BUS. INSIDER (Mar. 6, 2024, 11:18 AM), https://www.businessinsider.com/chatgpt-jobs-at-risk-replacement-artificial-intelligence-ai-labor-trends-2023-02 [https://perma.cc/VJN7-YQL8]; João Pedro Quintais & Nick Diakopoulos, *A Primer and FAQ on Copyright Law and Generative AI for News Media*, GENERATIVE AI NEWSROOM (Apr. 26, 2023), https://generative-ai-newsroom.com/a-primer-and-faq-on-copyright-law-and-generative-ai-for-news-media-f1349f514883 [https://perma.cc/7ZQ6-Y48H]; Charlie Beckett, *New Powers, New Responsibilities: A Global Survey of Journalism and Artificial Intelligence*, *in* THE LONDON SCHOOL OF BUSINESS AND

learning algorithm enables the generative AI system to generate literary and artistic content on its own—based on the computational analysis of human works that served as training material.[9] Considering this dependency of the machine on human training material, remuneration claims by authors and rights holders in the creative industry do not come as a surprise: generative AI systems are no true creators. They can only imitate human literary and artistic expression because they had the chance of analysing human creations. Against this background, the argument can be made that human authors should be compensated for the use of their works during AI training which, ultimately, may culminate in the usurpation of the market for literary and artistic productions, reduce the market share for human creations and cause loss of income.[10]

The discussion on author remuneration has brought to light a rich spectrum of policy considerations that can be invoked to support the introduction of payment obligations and remuneration systems (Section A). At the same time, it is clear that generative AI systems have a remarkable potential to contribute to economic growth by enabling new products and services. It can also be said that generative AI enhances freedom of expression by democratizing the process of creating literary and artistic content, thus broadening access to the literary and artistic discourse (Section B). Ideally, policymakers develop a regulatory approach that strikes a proper balance between the interest in copyright protection and author remuneration on the one hand and the broader societal interest in AI innovation and high-quality AI systems on the other (Section C).

POLITICAL SCIENCE 1, 24–25 (2019); Katharine Trendacosta & Cory Doctorow, *AI Art Generators and the Online Image Market*, ELECTRONIC FRONTIER FOUND. BLOG (Apr. 3, 2023), tps://www.eff.org/deeplinks/2023/04/ai-art-generators-and-online-image-market [https://perma.cc/2CHR-V226]; Deltorn, *supra* note 7, at 173-74; Shlomit Yanisky-Ravid, *Generating Rembrandt: Artificial Intelligence, Copyright and Accountability in the 3A Era–The Human-Like Authors Are Already Here–A New Model,* 2017 MICH. ST. L. REV. 659, 662 (2017); Annemarie Bridy, *The Evolution of Authorship: Work Made by Code*, COLUM. J.L. & ARTS 395, 397 (2016); Robert C. Denicola, *Ex Machina: Copyright Protection for Computer-Generated Works*, 69 RUTGERS U. L. REV. 251, 251 (2016); William T. Ralston*, Copyright in Computer-Composed Music: HAL Meets Handel,* 52 J. COPYRIGHT SOC'Y U.S.A. 281, 283 (2005).

9. *See* STUART RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH 693-717 (3rd ed., Pearson Education 2010).

10. As to the potential of generative AI systems to replace human creativity in different sectors, see Complaint, at ¶ 113-22, Authors Guild v. OpenAI, Inc., No. 1:23-cv-8292 (S.D.N.Y. Sept. 19, 2023); Zinkula & Mok, *supra* note 8 (explicitly listing "[m]edia jobs (advertising, content creation, technical writing, journalism)" and "[g]raphic designers" as risk categories); Giancarlo Frosio, *The Artificial Creatives: The Rise of Combinatorial Creativity from Dall-E to GPT-3*, *in* HANDBOOK OF ARTIFICIAL INTELLIGENCE AT WORK: INTERCONNECTIONS AND POLICY IMPLICATIONS 225, 225–27 (Martha Garcia-Murillo et al. eds., 2024); Quintais & Diakopoulos, *supra* note 8; Beckett, *supra* note 8, at 24–25; Trendacosta & Doctorow, *supra* note 8; Deltorn, *supra* note 7, at 173–74; Yanisky-Ravid, *supra* note 8, at 662; Bridy, *supra* note 8, at 397; Denicola, *supra* note 8, at 251; Ralston, *supra* note 8, at 283.

A. Six Arguments for Author Remuneration

The discussion on generative AI and author remuneration has produced various arguments seeking to support the remuneration claims of authors and rights holders in the creative industry. These arguments range from the parasitic use of human literary and artistic works and central societal functions of human literary and artistic productions to broader socio-political objectives and considerations relating to AI improvement.

First, it can be said that authors should be compensated for the parasitic usurpation of the market for human creative labor. The machine is only capable of mimicking human literary and artistic works after it had the opportunity to derive patterns for its own literary and artistic productions from myriad human creations that served as resources for training purposes. From this perspective, it is only fair that human authors, who are providing the source material for AI ingenuity, receive remuneration when AI productions finally kill the demand for the same human creativity that empowered the AI system to become a competitor in the first place.[11] This line of argument plays a central role in the lawsuits, which several U.S. newspapers, including the *New York Times*, have brought against OpenAI and Microsoft.[12] It also features prominently in the lawsuit between Getty Images and Stability AI in the UK.[13]

11. *Authors and Performers Call for Safeguards Around Generative AI in the European AI Act*, INITIATIVE URHEBERRECHT (Apr. 19, 2023), https://urheber.info/diskurs/call-for-safeguards-around-generative-ai [https://perma.cc/XS2Z-N8SU]; *Joint Statement from Authors' and Performers' Organisations on Artificial Intelligence and the AI Act – True Culture Needs Originals: Transparency and Consent are Key to the Ethical Use of AI*, FED'N EUR. SCREENDIRECTORS (Feb. 9, 2023), https://screendirectors.eu/joint-statement-from-authors-and-performers-organisations-on-artificial-intelligence-and-the-ai-act/ [https://perma.cc/GUH9-8S9Q]; *Our Manifesto for AI Companies Regulation in Europe*, 2023, EUR. GUILD FOR A.I. REGUL. (Nov. 04, 2023), https://www.egair.eu/ - manifesto [https://perma.cc/64LL-K3WE].

12. Michael M. Grynbaum & Ryan Mac, *The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work*, N.Y. TIMES (Dec. 27, 2023), https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html [https://perma.cc/BB69-DXWA] (reporting the argument that "[m]illions of articles from The New York Times [sic] were used to train chatbots that now compete with it . . ."). As to similar arguments advanced in more recent lawsuits launched by U.S. newspapers, see Katie Robertson, *8 Daily Newspapers Sue OpenAI and Microsoft Over A.I.*, N.Y. TIMES (Apr. 30, 2024) , https://www.nytimes.com/2024/04/30/business/media/newspapers-sued-microsoft-openai.html [https://perma.cc/AJ2G-E58Y]. For an overview of AI lawsuits in the U.S., see *Copyright: Status*, CHATGPT IS EATING THE WORLD, https://chatgptiseatingtheworld.com/category/status/ [https://perma.cc/H2QD-J9B6].

13. Getty Images (US) Inc. v. Stability AI [2023] EWHC 3090, [8] (Ch) (Eng.); Samuelson, *supra* note 6, at 64–65.

Second, it has been demonstrated in the cultural sciences that human literary and artistic creations have particular value to society as a whole.[14] Artwork made by flesh and blood authors provide important impulses for social and political changes by modelling experimental practices that open up new horizons for the development of society.[15] Human literary and artistic expression can mirror the shortcomings of present society, unmask defects of existing social and political conditions, and prepare society for the transition to a better community.[16] Arguably, AI-generated productions in the literary and artistic field are incapable of providing comparable impulses for the improvement of societal conditions. An AI system may manage to mimic human creativity and generate comparable literary and artistic output.[17] However, an AI system does not have the capacity to permeate the surface of a human artwork, go beyond its mere form of appearance, and critically assess its message and meaning in the light of current societal conditions. AI systems do not perceive and experience social and political conditions as humans do. They are simply not affected by societal conditions in the same way as humans.[18] Unable to experience and suffer contemporary societal conditions like a human, an AI system will inevitably fail to evoke visions of a new consensus on ethical norms that corresponds with people's current desires.[19]

Third, support for human authors is a good investment in new, innovative directions in literature, art and music. While human authors can initiate avant-garde movements that lead to new forms of expression, AI systems cannot free themselves from the data input fuelling their algorithm. They have difficulty refusing rule obedience, negating historical work templates,

14. Martin Senftleben, *The Copyright/Trademark Interface – How the Expansion of Trademark Protection Is Stifling Cultural Creativity*, *in* KLUWER L. INT'L 54–64 (Information Law Series Volume 44, 2020).

15. PETER OSBORNE, ANYWHERE OR NOT AT ALL – PHILOSOPHY OF CONTEMPORARY ART 208-11 (2013); Barton Beebe, *Bleistein, the Problem of Aesthetic Progress, and the Making of American Copyright Law*, 117 COLUM. L. REV. 319, 346–47 (2017).

16. FREDERICK SCHILLER, AESTHETICAL AND PHILOSOPHICAL ESSAYS, 35–36, 92, 120–21 (Tapio Riikonen & David Widger eds., 2006); THEODOR ADORNO, AESTHETIC THEORY, 9–10, 19, 25–26, 55–56, 127, 199 (Gretel Adorono & Rolf Tiedemann eds., Robert Hullot-Kentor trans., 2002).

17. For a description of the functioning of "creative" AI systems, see Martin Senftleben & Laurens Buijtelaar, *Robot Creativity: An Incentive-Based Neighbouring Rights Approach*, 42 EUR. INTELL. PROP. REV. 797, 802-04 (2020).

18. *Cf.* M.R.F. Senftleben, *Works of Authorship and the Single Equitable Remuneration for AI Substitutes*, *in* DESIGN OF THE INFORMATION LEGISLATION (V. Fischer et al. eds., 2022), [hereinafter *Single Equitable Remuneration*]; Martin Senftleben, A Tax on Machines for the Purpose of Giving a Bounty to the Dethroned Human Author – Towards an AI Levy for the Substitution of Human Literary and Artistic Works, 10–11 (Jun. 2, 2022) (unpublished manuscript) https://ssrn.com/abstract=4123309 [https://perma.cc/3TAX-4QDK] [hereinafter *Tax on Machines*].

19. As to the contribution of artworks to the improvement of societal conditions, see Schiller, *supra* note 16, at 120–21.

and autonomously creating something that falls outside existing aesthetical categories—something that brings chaos in the established order to shed light on tensions and conflicts in society and propose changes.[20] AI avant-garde experiments striving for societal relevance are doomed to fail. AI output cannot transcend the horizon of expectation that has evolved from known societal conditions.[21] To preserve the central societal function of new, unexpected directions in the literary and artistic realm, it is thus advisable to ensure that human creativity survives the dethroning of the human author by generative AI systems. The introduction of a remuneration system that channels money to human art projects makes sense from this perspective. It prevents the loss of avant-garde movements and the loss of important impulses for improving social and political conditions that can follow from the critical impetus of new, surprising directions in the literary and artistic field. Leaving literary and artistic productions to AI systems, society deprives itself of human impulses for future creativity and weakens its ability to evaluate and renew itself. With the introduction of an AI remuneration system, society can halt this trend.[22]

Fourth, there is a broader socio-political dimension. Inevitably, the replacement of the human author and the disruption of the market for literary and artistic productions require adequate countermeasures and investment. Authors who lose their jobs will need financial support. Literary and artistic projects focusing on human contributions can provide new job opportunities. Investment in training activities can enable authors to change course and obtain new skills and credentials. In this situation, introducing a remuneration system that provides money for new projects and training is an important and desirable step.

Compared with the preceding second and third arguments, this broader socio-political rationale has a more universal field of application. AI productions may win prizes and enter literary journals, concert halls, museums, and galleries.[23] Nonetheless, it seems unlikely (at least at this point in time) that AI output will replace human creations in the fine arts segment altogether. Creators of avant-garde artworks with the above-described potential to provide impetus for social and political changes may be exposed to substitution effects to a lesser extent than authors of literary and artistic everyday

20. As to this characteristic of human artworks, see ADORNO, *supra* note 16, at 27-28, 32-34, 41, 197–98, 337–38.

21. As to this requirement for societal relevance of literary and artistic productions, see OSBORNE, *supra* note 15, at 203–11.

22. *Single Equitable Remuneration*, *supra* note 18, at 122–24; *Tax on Machines*, *supra* note 18, at 10–11.

23. Senftleben & Buijtelaar, *supra* note 17, at 797–798.

products and works of applied art. The impact of AI will likely be felt much more strongly in areas such as news articles, illustrations and decorations, background music for bars and restaurants, and so on.[24]

In the latter segments, the mirror-of-society rationale may have less power of persuasion. Considering the substantially higher risk of substitution, however, the general socio-political objective to soften replacement effects gains more importance. Admittedly, general tax money could enable humans in affected creative sectors to adapt to the challenges of generative AI systems. Compared to a tax-based model, however, the copyright framework offers crucial advantages. With collecting societies and their remuneration and repartitioning schemes, the copyright system offers a well-established infrastructure for appropriately distributing collected money.[25] Moreover, a copyright-based solution seems much more stable than a general tax measure that could be undone in the next financial crisis or when the tax system is reformed.

Fifth, it can be added that human literary and artistic practice has societal value in and of itself. Relying on insights from pragmatist aesthetics,[26] Barton Beebe has argued that it is particularly important to the everyday individual to be involved in aesthetic practice and aesthetic play.[27] The active assimilation, appropriation, and creative recombination of aesthetic expression in the aesthetic play has intrinsic value. It constitutes a source of pleasure, moral and political cultivation, imaginative freedom, and self-actualization.[28] To the extent to which the aesthetic play is left to machines, humans in society lose opportunities for experiencing well-being, moral and political cultivation, imaginative freedom, and self-actualization. When the machine displaces the human author from the literary and artistic field, it also deprives society of role models for human aesthetic engagement.

Admittedly, generative AI systems provide tools for human users to experiment with different styles and motifs for art production. The act of developing and entering a prompt for an AI system, however, must not be confused with aesthetic play. The act of creation, which is the central element of aesthetic engagement, is not carried out by the human user. Instead, it becomes the task of the AI system. This has worrisome societal

24. *See* Zinkula & Mok, *supra* note 8.

25. A. Dietz, *A Modern Concept for the Right of the Community of Authors (Domaine Public Payant)*, *in* 24 COPYRIGHT BULL. 13, 15–16 (Evgueni Guerassimove & Nicole Paudras eds., 1990).

26. Beebe, *supra* note 15, at 346–347, 373–374, 384–385. *Cf.* JOHN DEWEY, ART AS EXPERIENCE 4–10 (1934); RICHARD SHUSTERMAN, PRAGMATIST AESTHETICS – LIVING BEAUTY, RETHINKING ART 143, 208 (2d ed., 2000).

27. Beebe, *supra* note 15, at 347.

28. *Id.* at 346–47.

repercussions. Once literary and artistic production is primarily seen as the domain of the machine, people may no longer have any reason to develop an aesthetic practice and play with different forms of expression themselves. The active assimilation, appropriation, and creative recombination of literary and artistic works becomes the machine's area of expertise. As a result, the potential of this practice to promote imaginative freedom and contribute to the cultivation and self-actualization of the individual in modern society is lost.[29]

From this perspective, it is not decisive that generative AI systems can imitate human literary and artistic works. This is only the final result of human creativity. The decisive factor in the equation, however, is the creative process: the aesthetic play. Giving instructions and pressing the button is not enough. The crucial element is the creative remix and reuse of literary and artistic sources of inspiration.[30] AI systems mimicking human works degrade the remix and reuse of literary and artistic source material to an automated process that can be left to machines. By establishing a remuneration system that provides human creators with financial means to survive in the field of aesthetic engagement, society can give the important signal that aesthetic practice is and remains an important human activity with particular value. Enabling human authors to stay in the literary and artistic field, this regulatory measure ensures that the role model of the human creator does not sink into oblivion and can inspire others to embark on aesthetic practice.

Sixth, the promotion of human literary and artistic productions is good for the AI industry itself. It is an important and wise investment in the continuous improvement of generative AI systems. By financially supporting the continuous flow of new human creations, the AI industry can ensure that a rich spectrum of fresh human training material for generative AI systems is constantly available. A continuously enriched reservoir of human source material appears as an important complement to known literary and artistic expressions of the past. Based on the analysis of historical human source materials, a generative AI system may be capable of producing endless recombinations of expressions that we have already seen. Adding the output of other AI systems to the training material, an AI system may also manage to recombine the recombinations of other machines. Ultimately, however, the generative AI process remains in a permanent loop. If the source repertoire for AI training is not constantly refreshed and enriched, the AI output

29. As to the political dimension of this educational effect of art, see *id.* at 336–37 (describing the belief of "early-republic Americans" that the progress of the fine arts promises to promote the overall progress of civic virtue and good government).

30. *Id.* at 390–91.

can hardly be expected to go beyond monochrome variations of known forms and styles. Fresh human literary and artistic productions, thus, have particular value for the AI industry itself. To break out of the spiral of endless repetition of "the same old thing" it makes sense to invest in human creativity. From this point of view, the payment of remuneration to support and foster human literary and artistic projects constitutes a legitimate policy goal that is in the AI industry's own interest.

B. Counterarguments Emphasizing Benefits for Society

As the discussion in the preceding section has shown, there are several good reasons for the introduction of legal mechanisms that ensure a fair remuneration for authors. However, it must not be overlooked that generative AI systems also offer substantial economic and socio-cultural benefits. When tracing the conceptual contours of appropriate regulatory solutions, these countervailing values must be taken into account to arrive at a balanced approach.

First, the implementation of generative AI systems offers remarkable potential for improving products and services while contributing to economic growth in various sectors, including the media sector and the creative industries more broadly.[31] Like other disruptive technologies that impacted copyright in the past,[32] generative AI is not only a threat to the routines of incumbent firms but also a technological tool that offers opportunities to develop new products and services.[33] The enhanced functionality that generative AI systems make available to human authors in creative industry sectors can reduce production costs and broaden the spectrum of literary and artistic content that can be brought to the market. As *Zarya of the Dawn* has shown, an author can employ generative AI to add dimensions to human production. While Kris Kashtanova wrote the story underlying the comic book, Midjourney's AI engine provided the illustrations.[34]

Second, generative AI paves the way for the further democratization of content production. It broadens access to the literary and artistic discourse.

31. Lemley & Casey, *supra* note 6, at 744–45.

32. For an overview of more detailed discussion, see generally Samuelson, *supra* note 6.

33. Not surprisingly, law and policymakers around the globe seek to devise a legal framework that is attractive to AI high-tech industries. *Cf.* Martin Senftleben et al., *Ensuring the Visibility and Accessibility of European Creative Content on the World Market: The Need for Copyright Data Improvement in the Light of New Technologies and the Opportunity Arising from Article 17 of the CDSM Directive*, 13 J. INTELL. PROP. INFO. TECH. & ELEC. COM. L. 67, 72–73 (2022).

34. *Zarya of the Dawn*, WIKIPEDIA, https://en.wikipedia.org/wiki/Zarya_of_the_Dawn [https://perma.cc/GF47-8KGU]; *see* Geiger & Iaia, *supra* note 6, at 5–6 (pointing out that generative AI systems can serve as tools for human creativity and expression).

With the opportunity to upload photos, films, music, and text to user-generated content platforms, formerly passive users have already become active contributors to content portals, wikis, online marketplaces, discussion and news fora, social networking sites, virtual worlds, and academic paper repositories.[35] Generative AI appears as a further step in this democratization process. It enables users to improve and refine their contributions. Victoria Kraetzig describes this further step in the democratization of the literary and artistic discourse as follows:

> Generative artificial intelligence (AI) serves the human behind the machine: the artist who prompts artificially creative art. It has been able to create new content from pre-existing content for years. A relatively recent development, however, is that the average consumer can operate it: at least since OpenAI's AI system ChatGPT has been on the market, laypeople can let the AI create by entering text modules (so-called prompts). The AI-generated content appears as if by magic a few minutes after the prompts have been entered. The models unleash creativity on an unprecedented scale. In one fell swoop, anyone can make art, anyone can be an artist. AI will change the economics of creativity. Endless creative content can be produced at no cost. A significant proportion of this content does not enjoy copyright protection: AI art expands the digital commons. Contemporary art in the public domain – that has not happened since the time of Goethe. More art means more culture. For us as a society, this can only be good.[36]

35. An overview of user generated content provided can be found in Organisation for Economic Co-operation and Development [OECD], *Participative Web: User-Created Content*, OECD Doc. DSTI/ICCP/IE(2006)7/FINAL (Apr. 12, 2007). As to the debate on user-generated content and the need for the copyright debate in this area, see Martin Senftleben, *User-Generated Content – Towards a New Use Privilege in EU Copyright Law*, *in* RESEARCH HANDBOOK ON INTELLECTUAL PROPERTY AND DIGITAL TECHNOLOGIES 136, 136–39 (Tanya Aplin ed., 2020). *See* Jean Paul Triaille et al., *Study on the Application of Directive 2001/29/EC on Copyright and Related Rights in the Information Society*, at 455–457 (2013), https://pure.unamur.be/ws/portalfiles/portal/54977637/7363.pdf [https://perma.cc/JTM2-RN5W]; Edward Lee, *Warming Up to User-Generated Content*, 2008 U. ILL. L. REV. 1459, 1506–13 (2008); Steven D. Jamar, *Crafting Copyright Law to Encourage and Protect User-Generated Content in the Internet Social Networking Context*, 19 WIDENER L.J. 843, 846–48 (2010); Natali Helberger et al., *Legal Aspects of UCC*, (Nov. 14, 2009), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1499333 [https://perma.cc/BTR4-494F]; Mary W. S. Wong, *Transformative User-Generated Content in Copyright Law: Infringing Derivative Works or Fair Use?,* 11 VAND. J. ENT. & TECH. L. 1075, 1075–76 (2009); Branwen Buckley, *SueTube: Web 2.0 and Copyright Infringement*, 31 COLUM. J.L. & ARTS 235, 235 (2008); Tom W. Bell, *The Specter of Copyism v. Blockheaded Authors: How User-Generated Content Affects Copyright Policy*, 10 VAND. J. ENT. & TECH. L. 841, 841-42 (2008); Steven Hechter, *User-Generated Content and the Future of Copyright: Part One – Investiture of Ownership*, 10 VAND. J. ENT. & TECH. L. 864, 864 (2008); Greg Lastowka, *User-Generated Content and Virtual Worlds*, 10 VAND. J. ENT. & TECH. L. 893, 893 (2008).

36. "Generative Künstliche Intelligenz (KI) dient dem Menschen hinter der Maschine: dem Künstler, der die künstlich kreative Kunst promptet. Schon seit Jahren kann sie aus Vorbestehendem neue Inhalte erstellen. Eine recht junge Entwicklung ist jedoch, dass der Otto-Normalverbraucher sie bedienen kann: Spätestens seitdem OpenAI's KI-System ChatGPT auf dem Markt ist, können Laien die KI durch die Eingabe von Textbausteinen (sog. Prompts) gestalten lassen. Wie von Geisterhand erscheint wenige Minute nach Eingabe der Prompts der von KI generierte Inhalt. Die Modelle setzen Kreativität ungehörigen Ausmaßes frei. Mit einem Schlag kann jeder Kunst machen, jeder Künstler sein. KI wird die Ökonomie der Kreativität verändern. Ohne Kosten können endlos kreative Inhalte produziert werden. Ein

This democratization argument adds an important nuance to the criticism in the preceding section. As explained, generative AI systems are likely to corrode human aesthetic practice. The act of developing and entering a prompt must not be confused with aesthetic play. The act of creation, the central element of aesthetic engagement, is not carried out by the human user. It becomes the task of the AI system (fifth argument above). As already conceded in the preceding section, however, the fact remains that generative AI systems provide tools for human users to experiment with different styles and motifs for art production. Even users without any literary or artistic skills can produce results that appear as valid contributions to the literary and artistic discourse. Hence, generative AI gives everybody access to the societal subsystem of literary and artistic production.[37] Kraetzig's statement focuses on this enabling function. While AI will predominantly produce endless combinations and recombinations of known ideas, concepts and styles, these variations of known cultural expressions can serve as entrance tickets to the literary and artistic discourse. Next to economic arguments, it is thus important to consider the potential of generative AI systems to offer users access to the field of literary and artistic production and enrich the digital public domain.[38]

erheblicher Teil von ihnen genießt keinen urheberrechtlichen Schutz: KI-Kunst vergrößert die digitale Allmende. Gemeinfreie zeitgenössische Kunst – das hat es seit der Goethe-Zeit nicht mehr gegeben. Mehr Kunst bedeutet mehr Kultur. Für uns als Gesellschaft kann das nur gut sein." (English translation by DeepL.com (free version) with minor corrections by the author). Viktoria Kraetzig, *KI-Kunst als schöpferische Zerstörung [AI Art as Creative Destruction]*, 2024 NEUE JURISTISCHE WOCHENSCHRIFT 697, 697; *see* Samuelson, *supra* note 6 (referring to "planning gardens, workouts, and meals, writing wedding speeches and emails, sorting through archives of photographs, organizing research for one's thesis, fixing software bugs, and building Spotify playlists" as examples of socially beneficial uses of generative AI systems).

37. As to the different fields and subsystems of society and the role of the field of literary and artistic production, see Senftleben, *supra* note 14, at 54–64; NIKLAS LUHMANN, SOCIAL SYSTEMS 34–70 (John Bednarz, Jr. & Dirk Baecker trans., 1995); PIERRE BOURDIEU & LOIC J.D. WACQUANT, AN INVITATION TO REFLEXIVE SOCIOLOGY 94 (1992); PIERRE BOURDIEU, THE RULES OF ART: GENESIS AND STRUCTURE OF THE LITERARY FIELD 214–218, 289 (Susan Emanuel trans., 1996).

38. Geiger, *supra* note 6, at 80 (pointing out that AI systems authors "might very well cohabitate in the future and support each other."). As to the question which human contribution is necessary to obtain copyright protection for interactions with, and modifications of, AI-generated output, see Case C-05/08, Infopaq Int'l A/S v. Danske Dagblades Forening, 2009 E.C.R. I-6624, I-6643; Case C-145/10, Painer v. Standard Verlags GmbH, ECLI:EU:C:2011:798, ¶ 89; P. Bernt Hugenholtz & Jose P. Quintais, *Copyright and Artificial Creation: Does EU Copyright Law Protect AI-Assisted Output?*, 52 INT'L REV. INTELL. PROP. & COMPETITION L. 1190, 1212-13 (2021). *But see* Dan L. Burk, *Thirty-Six Views of Copyright Authorship, by Jackson Pollock*, 58 HOUS. L. REV. 263, 270-321 (2020); Ginsburg & Budiardjo, *supra* note 7, at 395–96; Mary-Christine Janssens & Frank Gotzen, *Kunstmatige Kunst. Bedenkingen bij de toepassing van het auteursrecht op Artificiële Intelligentie [Artificial Art: Cautions on Applying Copyright to Artificial Intelligence]*, AUTEURS EN MEDIA 323, 325–27 (2020), https://lirias.kuleuven.be/2933893&lang=en [https://perma.cc/RTW3-V63T]; Russ Pearlman, *Recognizing Artificial Intelligence (AI) as Authors and Investors under U.S. Intellectual Property Law*, 24 RICH. J.L. & TECH. i, 4 (2018).

Third, AI can enhance information about social and political conditions. In the discussion on text and data mining, for instance, it has been emphasized that superhuman possibilities of data analysis may give journalists unprecedented powers of discovery.[39] While generative AI functions may play an increasingly important and perhaps threatening role in routine journalism labor, AI tools also give journalists "new powers of discovery, creation and connection."[40] Hence, the discussion about the role of AI systems must not be confined to the risk of spreading misinformation and disinformation, and amplifying false, harmful and extreme messages.[41] AI can also support the watchdog function, which quality journalism fulfils in democratic societies, and lead to important discoveries and news stories.[42]

C. Need for a Proper Balance

The discussion of the pros and cons of generative AI systems in the preceding sections sheds light on the need for a proper balance. Policymakers must reconcile divergent economic objectives and industry interests: a favorable climate for AI development is desirable, and a proper incentive and reward for investment in human literary and artistic productions as well.

39. See the examples given by Charlie Beckett, *New Powers, New Responsibilities: A Global Survey of Journalism and Artificial Intelligence*, *in* THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE 1, 1, 26 (2019). *See also* Sean M. Fiil-Flynn et al., *Legal Reform to Enhance Global Text and Data Mining Research: Outdated Copyright Laws Around the World Hinder Research*, 378 SCI. 951, 951 (2022); Christophe Geiger et al., *Text and Data Mining: Articles 3 and 4 of the Directive 2019/790/EU*, CTR. INT'L INTELL. PROP. STUD. 5, 31 (Jan. 21, 2019), https://ssrn.com/abstract=3470653 [https://perma.cc/7C4E-8NZB]. *But see* Thomas Margoni & Martin Kretschmer, *A Deeper Look Into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology*, 71 GEWERBLICHER RECHTSSCHUTZ UND URHEBERRECHT – INTERNATIONAL [GRUR INT'L] 685, 694 (2022).

40. Beckett, *supra* note 39, at 1.

41. *See* Bill Tomlinson et al., *Turning Fake Data Into Fake News: The AI Training Set as a Trojan Horse of Misinformation*, 60 SAN DIEGO L. REV. 641, 649-57 (2023); Erik Derner & Kristina Batistič, *Beyond the Safeguards: Exploring the Security Risks of ChatGPT*, ARXIV (May 13, 2023), https://arxiv.org/abs/2305.08005 [https://perma.cc/A2VR-7HYQ]; Robert Post, *The Internet, Democracy and Misinformation*, *in* DISINFORMATION, MISINFORMATION AND DEMOCRACY (A. Koltay et al. eds.) (forthcoming 2025) (manuscript at 10–11); Spencer Overton, *Overcoming Racial Harms to Democracy from Artificial Intelligence*, IOWA L. REV. (forthcoming 2025); Jon M. Garon, *An AI's Picture Paints a Thousand Lies: Designating Responsibility for Visual Libel*, 3 J. FREE SPEECH L. 425, 428–29, 435–42 (2023); Rory Gillis et al., *Trust and Trustworthiness in Artificial Intelligence*, *in* HANDBOOK ON ARTIFICIAL INTELLIGENCE AND PUBLIC POLICY 1–2 (Regine Paul et al. eds., 2024); Anirban Mukherjee, *Safeguarding Marketing Research: The Generation, Identification, and Mitigation of AI-Fabricated Disinformation*, ARXIV 17–20 (Mar. 17, 2024), https://arxiv.org/abs/2403.14706 [https://perma.cc/2LJA-RMSD].

42. For a discussion of the role and importance of "public interest journalism," see Jan van Cuilenburg & Denis McQuail, *Media Policy Paradigm Shifts: Towards a New Communications Policy Paradigm*, 18 EUR. J. COMMC'N 181, 182 (2003); DENIS MCQUAIL, MEDIA PERFORMANCE: MASS COMMUNICATION AND THE PUBLIC INTEREST 25 (1992); VIRGINIA HELD, THE PUBLIC INTEREST AND INDIVIDUAL INTERESTS 18–54 (1970).

In addition, the impact on freedom of expression and information must not be overlooked. Potential corrosive effects on human creativity must be weighed against broader access to the literary and artistic discourse and the enrichment of the digital public domain.[43] Finally, the general societal interest in high-quality AI systems must be factored into the equation. Inevitably, limited access to human training resources restricts the ability of AI trainers to develop models capable of producing fair, unbiased results: AI output that reflects all cultures, traditions, and values expressed in human literary and artistic works. AI training based on mainstream works will lead to mainstream AI output that marginalizes niche repertoires and opinions. AI training based on a specific segment of literary and artistic production will lead to AI output focusing on this specific segment and neglecting other expressions.[44]

In the thicket of societal challenges and opportunities, the question arises which regulatory tools lawmakers have available to achieve appropriate results. Considering the whole process from AI development to AI exploitation,[45] two central starting points for regulatory interventions can be

43. As to the reconciliation of copyright protection with freedom of expression and information, including the right to research in text and data mining cases, see generally Senftleben, *supra* note 2, at 12–15; Christophe Geiger & Bernd J. Jütte, *The Right to Research as Guarantor for Sustainability, Innovation and Justice in EU Copyright Law*, *in* INTELLECTUAL PROPERTY RIGHTS IN THE POST PANDEMIC WORLD: AN INTEGRATED FRAMEWORK OF SUSTAINABILITY, INNOVATION AND GLOBAL JUSTICE 138 (Taina Pihlajarinne et al. eds., 2023); Christophe Geiger & Bernd J. Jütte, *Conceptualizing a "Right to Research" and Its Implications for Copyright Law: An International and European Perspective*, 38 AM. U. INT'L L. REV. 1 (2023); Daniel Jongsma, *Creating EU Copyright Law: Striking a Fair Balance*, HELSINKI: HANKEN SCHOOL OF ECONOMICS (June 9, 2020), https://harisportal.hanken.fi/en/publications/creating-eu-copyright-law-striking-a-fair-balance [https://perma.cc/PG4F-5ZNY]; Jonathan Griffiths & Luke McDonagh, *Fundamental Rights and European IP Law: The Case of Art 17(2) of the EU Charter*, *in* CONSTRUCTING EUROPEAN INTELLECTUAL PROPERTY: ACHIEVEMENTS AND NEW PERSPECTIVES 75 (Christophe Geiger ed., 2013); Christophe Geiger & Elena Izyumenko, *Copyright on the Human Rights' Trial: Redefining the Boundaries of Exclusivity Through Freedom of Expression*, 45 INT'L REV. INTELL. PROP. & COMPETITION L. 316, 316 (2014); Jonathan Griffiths, *European Union Copyright Law and the Charter of Fundamental Rights Advocate General Szpunar's Opinions in (C-469/17) Funke Medien, (C-476/17) Pelham GmbH and (C-516/17) Spiegel Online*, 20 ERA F. 35, 46–49 (2019); Thomas Dreier, *Balancing Proprietary and Public Domain Interests: Inside or Outside of Proprietary Rights?*, *in* EXPANDING THE BOUNDARIES OF INTELLECTUAL PROPERTY: INNOVATION POLICY FOR THE KNOWLEDGE SOCIETY 295 (Rochelle Cooper Dreyfuss et al. eds., 2001); Christophe Geiger, *"Constitutionalising" Intellectual Property Law? The Influence of Fundamental Rights on Intellectual Property in the European Union*, 37 INT'L REV. INTELL. PROP. & COMPETITION L. 371 (2006); P. Bernt Hugenholtz, *Copyright and Freedom of Expression in Europe*, *in* THE COMMODIFICATION OF INFORMATION 239 (Niva Elkin-Koren & Neil W. Netanel eds., 2002); Yochai Benkler, *Free as the Air to Common Use: First Amendment Constraints on Enclosure of the Public Domain*, 74 N.Y.U. L. REV. 354, 355 (1999); Neil W. Netanel, *Copyright and a Democratic Civil Society*, 106 YALE L.J. 283, 347–48 (1996).

44. *Cf.* Carys Craig, *The AI-Copyright Challenge: Tech-Neutrality, Authorship, and the Public Interest*, OSGOODE DIGITAL COMMONS (2022), https://digitalcommons.osgoode.yorku.ca/all_papers/360 [https://perma.cc/6DFC-7VYK]; Geiger & Iaia, *supra* note 6, at 5; Lemley & Casey, *supra* note 6, at 770.

45. For a description of the different steps in this process, see generally Kacper Szkalej & Martin Senftleben, *Generative AI and Creative Commons Licences: The Application of Share Alike Obligations*

identified. On the one hand, the input dimension can be brought into focus: the use of human literary and artistic creations for AI training purposes. On the other hand, the final result leading to the generation of AI output—the offer of generative AI products and services in the marketplace—may serve as a reference point. Quite clearly, the complexity of the outlined interests and policy goals implies that the ultimate answer can hardly be black or white:

- *categorically prohibiting use in AI training*: the categorical prohibition of the use of copyrighted material for AI training is unlikely to offer an appropriate basis for the development of fair, unbiased AI models. A legal obligation to license each and every content item before inclusion in an AI training dataset can easily lead to a focus on mainstream repertories that may become available—without excessive transaction costs[46]—after negotiations with several large rights holders. It will lead to biased AI output that fails to reflect the full spectrum of literary and artistic traditions and expressions;
- *categorically permitting use in AI training*: placing AI training outside the realm of copyright altogether and offering unbridled freedom to use copyrighted works for AI training without remuneration can hardly be expected to satisfy the societal interest in human literature and art. This solution refuses to provide support for human literary and artistic projects. It leads to a loss of impulse for the improvement of social and political conditions that human works can provide. Moreover, this approach fails to address the interests and concerns of authors facing displacement effects and seeking remuneration for the use of their works in AI training processes.

Hence, appropriate strategies for adding shades of grey are particularly important. Between the two poles, categorical prohibition and unbridled freedom, there may be legal mechanisms that allow lawmakers to square the circle. These legal mechanisms can ensure fair remuneration for authors without stifling AI innovation and compromising the quality of AI systems. To identify these legal mechanisms, it is necessary to embark on a closer inspection of input-based (focus on AI training) and output-based (focus on the final offer of AI products and services) legal solutions capable of

to Trained Models, Curated Datasets and AI Output (June 20, 2024), https://ssrn.com/abstract=4872366 [https://perma.cc/2FGU-QFLQ].

46. Geiger & Iaia, *supra* note 6, at 5–6; Craig, *supra* note 44; Lemley & Casey, *supra* note 6, at 770–771. *Cf.* Reto M. Hilty & Heiko Richter, *Position Statement of the Max Planck Institute for Innovation and Competition on the Proposed Modernisation of European Copyright Rules*, MAX PLANCK INST. FOR INNOVATION & COMPETITION (Jan. 20, 2017), https://ssrn.com/abstract=2900110 [https://perma.cc/5ETS-LTAU].

generating a money stream that supports authors without imposing too heavy a burden on AI developers. With its focus on AI training, the European AI Act allows the analysis of an approach seeking to ensure the payment of remuneration upfront: at the AI training stage where human works are used as input for machine learning. In the following part II, this approach will be scrutinized before turning to output-oriented alternatives in part III.

II. EX ANTE APPROACH: REMUNERATION AS PREREQUISITE FOR AI DEVELOPMENT

Remuneration arguments focusing on AI training (input dimension) emphasize the fact that myriad human works have been used without prior authorization to train generative AI systems capable of replacing human literary and artistic productions. This line of argument had a deep impact on the legislative process leading to the AI Act. In Europe, artists and rights holders expressed serious concerns over the parasitic use of human works for AI training purposes.[47] In its *Call for Safeguards Around Generative AI*, the Authors' Rights Initiative—over forty associations and trade unions representing authors, performers, and copyright holders in various creative industry segments—stressed that:

> [t]he output of AI systems depends on the input they are trained with; this includes texts, images, videos and other material from authors, performers and other copyright holders: Our entire digital repertoire serves training purposes, often without consent, without remuneration and not always for legitimate uses. The unauthorised usage of protected training material, its non-transparent processing, and the foreseeable substitution of the sources by the output of generative AI raise fundamental questions of accountability, liability and remuneration, which need to be addressed before irreversible harm occurs.[48]

Similarly, the European Guild for Artificial Intelligence Regulation ("EGAIR") adopted the *EGAIR Manifesto* pointing out that:

> the products sold by AI companies are the result of operations on datasets, which contain all sorts of data, including millions of copyrighted images, private pictures and other sensitive material. These files were collected by indiscriminately scraping the internet without the consent of the owners and people portrayed in them and are currently being used by AI companies for profit.[49]

The *Joint Statement from Authors' and Performers' Organisations on Artificial Intelligence and the AI Act* warned that:

47. As to similar expressions of concern in the U.S., see Samuelson, *supra* note 6.
48. INITIATIVE URHEBERRECHT, *supra* note 11.
49. EUR. GUILD FOR A.I. REGUL., *supra* note 11.

> AI technologies increasingly use authors' and performers' works and creations to "feed" and train their applications without their consent or knowledge, in breach of authors' and performers' rights granted under international, EU or national laws. In this era of rapidly advancing AI technologies, whose principle consists solely of copying and mixing, we must highlight the urgent need to protect the works and performances of professional authors and performers from misappropriation. Not only to preserve their livelihoods, but also to inform citizens about the use of original works by AI applications.[50]

The final text of the AI Act shows that these initiatives informed the parliamentary debate and the trilogue phase in which the European Commission, the Council, and the European Parliament established the definite version of the AI Act. Recital 105 AIA addresses "[g]eneral-purpose AI models, in particular large generative AI models, capable of generating text, images, and other content."[51] Recognizing potential corrosive effects on human creativity, it points out that these models "present unique innovation opportunities but also challenges to artists, authors, and other creators and the way their creative content is created, distributed, used and consumed."[52] The Recital also emphasizes that the development and training of generative AI models "require access to vast amounts of text, images, videos, and other data. Text and data mining techniques may be used extensively in this context for the retrieval and analysis of such content, which may be protected by copyright and related rights."[53]

After this problem statement, Recital 105 confirms that the use of literary and artistic works for AI training purposes has copyright relevance and involves TDM that requires the authorization of rights holders: "[a]ny use of copyright protected content requires the authorisation of the rightsholder concerned unless relevant copyright exceptions and limitations apply."[54] As requested by authors, performers and creative industries, the EU legislator, thus, clarified that authors and industry rights holders can exercise control over the use of human works during AI training processes on the basis of copyright protection—unless a copyright exception applies.

Prior to the adoption of the AI Act, the TDM discussion in EU copyright law already culminated in the introduction of rules that could be understood to confirm the copyright relevance of AI development processes, such as the use of protected works during the training of generative AI systems. The specific TDM provisions in Articles 3 and 4 CDSMD set forth two specific

50. FED'N EUR. SCREENDIRECTORS, *supra* note 11.
51. Regulation 2024/1689, ¶ 105.
52. *Id.*
53. *Id.*
54. *Id.*

exceptions to copyright, related rights, and database protection that play an important role in the context of TDM projects that require the extraction of data from protected literary and artistic works. Arguably, it would not have been necessary to adopt specific copyright exceptions if TDM had no copyright relevance.

Addressing scientific research directly, Article 3(1) CDSMD sets forth an obligation for Member States to permit "reproductions and extractions made by research organisations and cultural heritage institutions in order to carry out, for the purposes of scientific research, text and data mining of works or other subject matter to which they have lawful access."[55]

Article 2(2) CDSMD adds a harmonized TDM definition covering "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."[56]

In addition to the exemption of scientific TDM in Article 3 CDSMD, Article 4(1) CDSMD contains a general exemption of TDM that is not limited to scientific research. Under this additional provision, anyone, including commercial AI system developers and trainers, may make copies of works for the purposes of TDM and retain them as long as necessary for the AI training process.[57] With regard to this broader category of TDM outside the scope of the scientific research rule in Article 3 CDSMD, Article 4(3) CDSMD adds an important nuance by stipulating that rights holders can reserve their rights. The provision contains the following opt-out mechanism:

> The exception or limitation provided for in paragraph 1 shall apply on condition that the use of works and other subject matter referred to in that paragraph has not been expressly reserved by their right holders in an appropriate manner, such as machine-readable means in the case of content made publicly available online.[58]

As indicated, it can be argued that the introduction of specific copyright exceptions for TDM in the CDSM Directive already established the copyright relevance of TDM and related AI training processes prior to the adoption of the AI Act. As the CDSM Directive dates back to 2019, however, it could also be emphasized that the EU legislator did not have in mind the use of copyrighted material as mere data input for generative AI training

55. Directive 2019/790, art. 3(1), 113.

56. *Id.* art. 2(2), 112.

57. *Id.* art. 4(1)–(2), 113–14. As to the relevance of Article 4 CDSMD to generative AI systems, see João Pedro Quintais, *Generative AI, Copyright and the AI Act*, KLUWER COPYRIGHT BLOG (May 9, 2023), https://copyrightblog.kluweriplaw.com/2023/05/09/generative-ai-copyright-and-the-ai-act/ [https://perma.cc/G7ZQ-ZLFW].

58. Directive 2019/790, art. 4(3), 114.

purposes.[59] In the TDM debate, it has been underlined around the globe that TDM copies have a specific nature. They fall outside the concept of reproduction in the traditional sense of making copies for the purpose of consulting and enjoying a work.[60] From a U.S. perspective, Michael Carroll has pointed out that in the context of TDM, "copies are made only for computational research and the durable outputs of any text and data mining analysis would be factual data and would not contain enough of the original expression in the analyzed articles to be copies that count."[61]

Explaining the outright exemption of TDM activities in Article 30-4(ii) of the Japanese Copyright Act, Tatsuhiro Ueno has pointed out that:

> if an exploitation of a work is aimed at neither enjoying it nor causing another person to enjoy it (e.g. text-and-data mining, reverse engineering), there is no need to guarantee the opportunity of an author or copyright holder to receive compensation and thus copyright does not need to cover such exploitation. In other words, exploitation of this kind does not prejudice the copyright holder's interests protected by a copyright law.[62]

Criticizing the regulation of TDM in the EU, Rosanna Ducato and Alain Strowel described the following alternative approach:

> when acts of reproduction are carried out for the purpose of search and TDM, the work, although it might be reproduced in part, is not used as a work: the work only serves as a tool or data for deriving other relevant information. The expressive features of the work are not used, and there is no public to enjoy the work, as the work is only an input in a process for searching a corpus and identifying occurrences and possible trends or patterns.[63]

In fact, the distinction between use of "works as works" and use "as data" is not entirely new in the European copyright debate. In 2011, Mauricio Borghi and Stavroula Karapapa already developed the concept of "de-intellectualized use"[64] against the background of mass digitization projects, such

59. *See* Geiger, *supra* note 6, at 79–80. For a discussion of this argument with regard to the right of reproduction in international copyright law, see Martin Senftleben, *Compliance of National TDM Rules with International Copyright Law: An Overrated Nonissue?*, 53 INT'L REV. INTELL. PROP. COMPETITION L. 1477, 1493–502 (2022).

60. Senftleben, *supra* note 59, at 1495–502; Craig, *supra* note 44.

61. Michael W. Carroll, *Copyright and the Progress of Science: Why Text and Data Mining Is Lawful*, 53 UC DAVIS L. REV. 893, 954 (2019). *See also* Lemley & Casey, *supra* note 6, at 772–73, 779–80.

62. Tatsuhiro Ueno, *The Flexible Copyright Exception for 'Non-Enjoyment' Purposes – Recent Amendment in Japan and Its Implication*, 70 GEWERBLICHER RECHTSSCHUTZ UND URHEBERRECHT – INTERNATIONAL [GRUR INT'L] 145, 150–51 (2021).

63. Rossana Ducato & Alain Strowel, *Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out*, 43 EUR. INTELL. PROP. REV. 322 (2021), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3829858 [https://perma.cc/P5Z4-MQ7C] (manuscript at 26-27).

64. Maurizio Borghi & Stavroula Karapapa, *Non-display Uses of Copyright Works: Google Books and Beyond*, 1 QUEEN MARY J. INTELL. PROP. 21, 45 (2011).

as the Google Book Search. As Borghi and Karapapa point out, mass digitization turns protected content into mere data with the result that "the expression of the idea embodied in the work is not primarily used to communicate the 'speech' of the author to the public but rather to form the basis of machine-workable algorithms."[65]

In the light of these comments, it could be argued that, at least as long as no protected traces of literary and artistic works find their way into the trained model,[66] the use of copyrighted works during AI training fell outside the scope of copyright from the outset because it did not constitute use of the author's individual expression for communication purposes. After the clarification in Recital 105 AIA, however, the power of persuasion of this argument vanishes at least in the EU. Without distinguishing between use of "works as works" and use of "works as data," the AI Act confirms that EU copyright law brings all forms of TDM under the umbrella of the right of reproduction and, thus, requires the invocation of a copyright exception, such as the scientific research rule in Article 3 CDSMD, the broader exemption in Article 4 CDSMD, or the long-standing temporary copying rule in Article 5(1) of the 2001 Information Society Directive ("ISD").[67] In the case of commercial AI training falling under Article 4(1) CDSMD, this configuration of the right of reproduction also means that EU copyright law brings TDM activities within the reach of rights holders seeking to receive a remuneration for the use of their works.[68] Referring to the opt-out mechanism in Article 4(3) CDSMD, the AI Act confirms the intention to give rights holders the opportunity to exercise control over the use of their works for AI training

65. *Id.* at 44–45.

66. *See* Matthew Sag, *Copyright Safety for Generative AI*, 61 HOUS. L. REV. 295, 321–37 (2023); *cf.* Andres Guadamuz, *A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs*, 73 GEWERBLICHER RECHTSSCHUTZ UND URHEBERRECHT – INTERNATIONAL [GRUR INT'L] 111, 121–22 (2024); Szkalej & Senftleben, *supra* note 45; *contra* Nicholas Carlini et al., *Quantifying Memorization Across Neural Language Models*, ARXIV (Mar. 6, 2023), https://arxiv.org/abs/2202.07646 [https://perma.cc/JV6U-BKM9]; Nicholas Carlini et al., *Extracting Training Data from Diffusion Models*, ARXIV (Jan. 30, 2023), https://arxiv.org/abs/2301.13188 [https://perma.cc/9ARA-6LNG]; Ivo Emanuilov & Thomas Margoni, Forget Me Not: Memorisation in Generative Sequence Models Trained on Open Source Licensed Code (Mar. 2024) (unpublished manuscript) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4720990 [https://perma.cc/G78C-44TB].

67. Parliament and Council Directive 2001/29/EC, art. 5(1), 2001 O.J. (L 167) 1, 16. For an overview of the exceptions in EU copyright law that can be invoked with regard to TDM, see Senftleben, *supra* note 2, at 27–28 (temporary copying), 36–37 (scientific TDM).

68. Paul Keller, *Protecting Creatives or Impeding Progress? Machine Learning and the EU Copyright Framework*, KLUWER COPYRIGHT BLOG (Feb. 20, 2023), https://copyrightblog.kluweriplaw.com/2023/02/20/protecting-creatives-or-impeding-progress-machine-learning-and-the-eu-copyright-framework/ [https://perma.cc/UN2G-YPHD]; *Policy Paper #15 on Using Copyrighted Works for Teaching the Machine*, COMMUNIA (Apr. 26, 2023), https://communia-association.org/policy-paper/policy-paper-15-on-using-copyrighted-works-for-teaching-the-machine/ [https://perma.cc/93FF-NDJY].

purposes in Article 4 CDSMD scenarios: "[w]here the rights to opt out has been expressly reserved in an appropriate manner, providers of general-purpose AI models need to obtain an authorisation from right holders if they want to carry out text and data mining over such works."[69]

In accordance with Article 4(3) CDSMD, rights holders can thus prohibit TDM via a machine-readable rights reservation. This means that AI trainers must take into account robots.txt files, but also the terms and conditions of a website or online service, in order to assess whether TDM is permitted with regard to a particular work.[70] In principle, rights holders can thus rely on technical safeguards to prevent the use of human creations for AI training purposes.

As in other cases where copyright holders can refuse permission for a given form of use, this veto right can pave the way for remuneration.[71] It is conceivable that the rights reservation option in Article 4(3) CDSMD leads to the evolution of machine-readable rights reservation protocols that express different rights holder standpoints. One standpoint could be robots.txt, which signals an outright exclusion of any use of the literary and artistic work at issue for AI training purposes. Using this version of robots.txt, right holders can express their preference for an outright prohibition and prevent TDM of copyrighted material altogether. An alternative standpoint, however, could be robots.txt, which prohibits use for AI training purposes only if the AI trainer behind the crawler is reluctant to pay remuneration. Using this alternative version, rights holders can thus express their willingness to permit the use against the payment of remuneration. In other words, in an ideal world, the rights reservation option in Article 4(3) CDSMD serves as a catalyst to arrive at generally agreed, machine-readable remuneration protocols that trigger an automated process for the payment of remuneration. Unfortunately, it may be quite difficult to achieve this ideal result in the real world.

First, the rights clearance infrastructure in the EU is highly fragmented.[72] Even if standardized rights reservation protocols, capable of

69. Regulation 2024/1689, ¶ 105; Geiger & Iaia, *supra* note 6, at 7.

70. Directive 2019/790, ¶ 18 (clarifying that "[i]n the case of content that has been made publicly available online, it should only be considered appropriate to reserve those rights by the use of machine-readable means, including metadata and terms and conditions of a website or a service"). *Cf.* Bernt Hugenholtz, *Artikelen 3 en 4 DSM-richtlijn: Tekst- en Datamining [Articles 3 and 4 DSM Directive: Text and Datamining]*, 5 AMI - TIJDSCHRIFT VOOR AUTEURS-, MEDIA- EN INFORMATIERECHT 167, 170 (2019) (NL).

71. *See* Keller, *supra* note 68; *Policy Paper #15 on Using Copyrighted Works for Teaching the Machine*, *supra* note 68.

72. Senftleben et al., *supra* note 33, at 70; *cf.* Martin Senftleben et al., *How the European Union Outsources the Task of Human Rights Protection to Platforms and Users: The Case of User-Generated Content Monetization*, 38 BERKELEY TECH. L.J. 933, 946–47 (2023).

expressing remuneration wishes and modalities become available, it is unclear whether copyright holders and collecting societies will ever manage to create efficient, pan-European rights clearance solutions that offer reliable and well-functioning payment interfaces. As long as the automated, machine-based identification of rights holders and the automated processing of payments remains complicated or unreliable, the rights reservation option in Article 4(3) CDSMD is unlikely to pave the way for a remuneration system that has success in practice. TDM requires the availability of myriad literary and artistic works. The moment AI trainers are obliged to check rights ownership, observe specific payment conditions, and obtain permission at the level of individual works or databases, the burden of rights clearance will inevitably put an end to the whole remuneration endeavor.[73]

Second, the described need for standardized, machine-readable remuneration protocols under Article 4(3) CDSMD indicates that, if satisfactory rights clearance solutions become available at all, these solutions will most probably be the result of industry collaboration: the creative industry agrees with the high-tech industry on conditional rights reservation protocols that make use of protected material possible the moment the desired remuneration has been paid.[74] The reference to "large private or public databases or data archives" in Recital 107 AIA also demonstrates that fears about right clearance and remuneration payments at the industry level between the AI industry and content majors in the creative industry are not unfounded. If collecting societies with repartitioning schemes ensuring direct payments to individual artists are not at the negotiation table,[75] the results of the new remuneration infrastructure may disappoint individual creators. Additional income from TDM may fill the pockets of large companies in the creative industries that own impressive repertoires of literary and artistic works.[76]

73. Craig, *supra* note 44; Lemley & Casey, *supra* note 6 at 770–71; *see* Hilty & Richter, *supra* note 46, at 1; *Cf.* Geiger & Iaia, *supra* note 6 at 5–6.

74. *See* Samuelson, *supra* note 6, at 81.

75. Bundesgerichtshof [BGH] [Federal Court of Justice] July 11, 2002, I ZR 255/00 1, 14–15 (Ger.); Guido Westkamp, *The "Three-Step Test" and Copyright Limitations in Europe: European Copyright Law Between Approximation and National Decision Making*, 56 J. COPYRIGHT SOC'Y U.S.A. 1, 55–59 (2008); João Pedro Quintais, *Copyright in the Age of Online Access – Alternative Compensation Systems in EU Law*, *in* KLUWER LAW INTERNATIONAL, at 335–36, 340–41, 347–49, 356–57 (P. Bernt Hugenholtz ed., 2017); *Opinion on Reprobel*, EUR. COPYRIGHT SOC'Y (Sept. 5, 2015), https://european-copyrightsociety.org/opinion-on-reprobel/ [https://perma.cc/WYG9-H3X4]; Christophe Geiger, *Promoting Creativity through Copyright Limitations: Reflections on the Concept of Exclusivity in Copyright Law*, 12 VAND. J. ENT. & TECH. L. 515, 532–33 (2010); R.M. Hilty, *Verbotsrecht vs. Vergütungsanspruch: Suche Nach Konsequenzen der Tripolaren Interessenlage im Urheberrecht*, *in* PERSPEKTIVEN DES GEISTIGEN EIGENTUMS – FESTSCHRIFT FÜR GERHARD SCHRICKER ZUM 70, 325, 325-353 (A. Ohly et al. eds., 2005).

76. *See* Trendacosta & Doctorow, *supra* note 8 (predicting "the perverse effect of limiting this technology development to the very largest companies, who can assemble a data set by compelling their

Individual creators whose works form part of these repertoires, however, will not necessarily receive higher honoraria or an appropriate share of the TDM income.[77]

Third, it is foreseeable that, with or without generally agreed rights reservation protocols, trust issues will remain. Who can guarantee that AI trainers observe rights reservations that are made in accordance with Article 4(3) CDSMD? And who can convincingly prove that a given work has been part of the AI training dataset when the final AI-generated output only reflects general style elements and bears no direct resemblance to a specific pre-existing work? Even if machine-readable remuneration protocols evolve from industry negotiations, it will remain difficult to control compliance with remuneration requirements and ensure remuneration payments that are accurate in the sense of capturing all works that have been used for AI training purposes.[78] The AI Act seeks to solve this dilemma by imposing a transparency obligation on AI trainers:

> [i]n order to increase transparency on the data that is used in the pre-training and training of general purpose AI models, including text and data protected by copyright law, it is adequate that providers of such models draw up and make publicly available a sufficiently detailed summary of the content used for training the general-purpose AI model.[79]

More concretely, the summary must offer sufficient detail to allow copyright enforcement. It is intended "to facilitate parties with legitimate interests, including copyright holders, to exercise and enforce their rights."[80] The AI Act gives the example of "listing the main data collections or sets that went into training the model, such as large private or public databases or data archives."[81] It also mentions the option of "providing a narrative explanation

workers to assign the 'training right' as a condition of employment or content creation"). However, see also Nicola Lucchi, *ChatGPT: A Case Study on Copyright Challenges for Generative Artificial Intelligence Systems*, EUR. J. RISK REGUL. 1, 17–18 (2023), and Mauritz Kop, *The Right to Process Data for Machine Learning Purposes in the EU*, 34 HARV. J.L. & TECH. 1, 7 (2021), who point out that data sharing agreements or data altruism initiatives may soften the predominance of large industry players.

77. As to the weak position of authors in the relationship with exploiters of their works and the limited success of copyright contract rules seeking to strengthen the position of authors, see Stef J. van Gompel et al., *Evaluatie Wet Auteurscontractenrecht: Eindrapport [Evaluation Copyright Contract Law]*, AMSTERDAM: IViR 1, 36, 44, 63–64, 96–97 (2020); Séverine Dusollier, *EU Contractual Protection of Creators: Blind Spots and Shortcomings*, 41 COLUM. J.L. & ARTS 435, 447–48, 454–55 (2018); Martin Senftleben, *More Money for Creators and More Support for Copyright in Society – Fair Remuneration Rights in Germany and the Netherlands*, 41 COLUM. J.L. & ARTS 413, 429 (2018).

78. Geiger & Iaia, *supra* note 6, at 5; Quintais, *supra* note 57.

79. Regulation 2024/1689, ¶ 107; *see also* Regulation 2024/1689, art. 53(1)(c)–(d).

80. *Id.* ¶ 107.

81. *Id.*

about other data sources used."[82] To harmonize the modes of reporting, the newly established AI Office will provide a template for summaries. The introduction of these measures confirms the existence of transparency and trust issues. For the AI industry, the submission of detailed summaries of copyrighted training resources is a substantial administrative burden. Moreover, these summaries represent considerable risk factors. Conceding the use of certain work repertoires in reports submitted in the EU, AI companies provide ammunition for copyright infringement claims in other regions.[83] Without doubt, the creative industry will study the EU summaries carefully and assess the chances of obtaining damages and licence fees not only in the EU but also elsewhere.

Fourth, it must not be overlooked that the main international competitors of the EU have chosen TDM approaches that markedly depart from the focus on licensing in the EU. The U.S., Canada, Singapore, South Korea, Japan, Israel, and Taiwan have opted for broader, more flexible copyright limitations.[84] Arguably, this regulatory approach enhances the innovation potential of high-tech companies in these countries in comparison with their EU counterparts. In the U.S., TDM has routinely been considered to be transformative fair use that is permissible without the prior authorization of the right holder and which does not generate claims for remuneration.[85] Japan has implemented in its copyright legislation a broad TDM exception in 2009.[86] Both countries offer particularly interesting examples. They belong to different copyright traditions. Both have successful and important creative and cultural industries. With the specific legal mechanisms available in their respective copyright systems—the fair use doctrine in the U.S. and a specific rule in Japan—they seek to support innovation in the AI sector.[87]

82. *Id.*

83. For an overview of lawsuits in the U.S., see *Copyright: Status*, *supra* note 12.

84. Senftleben et al., *supra* note 33, at 72–73.

85. Samuelson, *supra* note 6, at 74–76; Lemley & Casey, *supra* note 6, at 760–79; Matthew Sag, *The New Legal Landscape for Text Mining and Machine Learning*, 66 J. COPYRIGHT SOC'Y U.S.A. 291, 314–34 (2019); Pamela Samuelson, *Text and Data Mining of In-Copyright Works: Is It Legal?*, 64 COMMC'NS ACM 20, 20–22 (2021); Matthew Sag, *Copyright and Copy-Reliant Technology*, 103 NW. U. L. REV. 1607, 1613 (2009).

86. The Japanese Copyright Act envisages an exception for TDM that is not limited to non-commercial or to research only purposes. Lucie Guibault & Thomas Margoni, *Legal Aspects of Open Access to Publicly Funded Research*, *in* ENQUIRIES INTO INTELLECTUAL PROPERTY'S ECONOMIC IMPACT 373, 396 (2015) (reporting and discussing Article 47-septies Japanese Copyright Act); *see also* Marco Caspers & Lucie Guibault, *Deliverable 3.3 Baseline Report of Policies and Barriers of TDM in Europe*, FUTURETDM (May 13, 2016), https://dare.uva.nl/search?identifier=2e26ec5c-257c-4cfe-96eb-e44355dd2847 [https://perma.cc/3JQX-G7TB]; Ueno, *supra* note 62, at 148.

87. Senftleben et al., supra note 33, at 72–73.

Considering this global scenario, it is clear that impractical, complicated remuneration systems may disadvantage EU-based high-tech industries in comparison with their peers in other legal systems.[88] The remuneration architecture built on the rights reservation option in Article 4(3) CDSMD and the accompanying rules in the AI Act can easily lead to an unfortunate lose-lose scenario: no remuneration for authors and no access to copyrighted resources for AI trainers in the EU.[89] By virtue of Article 4(3) CDSMD, copyright holders may have success in reserving their rights and preventing the use of their works for AI training purposes. However, a successful rights reservation need not lead to licensing agreements and extra income. Instead, the high-tech industry may decide to move AI training activities to other regions that offer a more favorable training environment. As a result, the remuneration claim fails: the EU right holder does not receive money; the AI trainer looks for training resources elsewhere.

To counter this risk, the AI Act seeks to bypass the principle of territoriality and universalize the obligation to ensure compliance with opt-outs in the EU, regardless of where on the planet the AI system has been trained:

> Providers that place general purpose AI models on the EU market should ensure compliance with the relevant obligations in this Regulation. . . . [and] should put in place a policy to respect Union law on copyright and related rights, in particular to identify and respect the reservations of rights expressed by right holders pursuant to Article 4(3) of [CDSM] Directive.[90]

With regard to this feature of the new legislation, the AI Act itself makes no secret of the fact that a "Brussels effect"[91] is intended:

> Any provider placing a general-purpose AI model on the EU market should comply with this obligation, regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of those AI models take place. This is necessary to ensure a level playing field among providers of general-purpose AI models where no provider should be able to gain a competitive advantage in the EU market by applying lower copyright standards . . . .[92]

88. For a critique of the approach taken in the EU, see Christophe Geiger, *The Missing Goal-Scorers in the Artificial Intelligence Team: Of Big Data, the Right to Research and the Failed Text-and-Data Mining Limitations in the CSDM Directive*, 66 PIJIP/TLS RSCH. PAPER SERIES 5–7 (2021), https://digitalcommons.wcl.american.edu/research/66/ [https://perma.cc/H2B7-22R2]); Christian Handke et al., *Is Europe Falling Behind in Data Mining? Copyright's Impact on Data Mining in Academic Research*, *in* PROCEEDINGS OF THE 19TH INTERNATIONAL CONFERENCE ON ELECTRONIC PUBLISHING 120, 128–29 (Birgit Schmidt & Milena Dobreva eds., 2015).

89. *See* Geiger & Iaia, *supra* note 6, at 5–6; Craig, *supra* note 44; Lemley & Casey, *supra* note 6, at 770–71.

90. Regulation 2024/1689, ¶ 106; Regulation 2024/1689, art. 53(1)(c).

91. *See* BRADFORD, *supra* note 3; Husovec & Urban, *supra* note 3.

92. Regulation 2024/1689, ¶ 106.

This additional facet of the EU approach intensifies concerns about an unattractive, perhaps even deterrent, regulatory framework.[93] With the universalization of right holder opt-outs beyond the borders of the EU, the AI Act is likely to cause unwelcome surprises for AI companies finally seeking to offer their goods and services in the EU after a start in another region.

For example, a Japanese start-up company that has trained a highly successful text-to-manga model in full compliance with Japanese law may find it surprising to learn that it cannot enter the EU market unless it first creates an alternative model that observes all opt-outs in the EU. The need to ensure compliance with EU policy when entering the EU market raises the delicate question of "unlearning." Is it possible to remove EU threads woven into the fabric of an AI model that has been trained outside the EU? And what about opt-outs that are declared after the AI development phase? Is it legitimate to keep offering a generative AI system that has been trained prior to the opt-out?[94] And, if so, can the training of new versions of the AI system still be based on the model trained prior to the opt-out? Or is the AI company obliged to start the whole training process from scratch and observe all opt-outs that are valid at that point in time?

In sum, the AI Act package may lead to a situation where the spectrum of generative AI systems on the EU market is smaller than in other regions and where the generative AI systems that become available in the EU are less sophisticated and powerful than in other regions. In an endeavor to provide rights holders with a watertight remuneration claim at the input level, the EU has established a system that must appear highly complex and burdensome from the perspective of AI companies.

Perhaps even more importantly, the various obstacles posed by EU law give rise to a final, fifth concern: the complexity of legal requirements resulting from the amalgamation of opt-outs under Article 4(3) CDSMD and AI Act rules may lead to a situation where AI trainers shy away from the use of literary and artistic resources stemming from the EU. The described legal obligations and restrictions may render EU repertoires unattractive, if not simply unavailable. Hence, the EU approach implies the risk of marginalizing EU literary and artistic input and EU cultural heritage in generative AI

93. For a critical analysis of the practical implications of the EU opt-out systems, see Peter Mezei, *A Saviour or a Dead End? Reservation of Rights in the Age of Generative AI*, EUR. INTELL. PROP. Rev. 461 (2024) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4695119 [https://perma.cc/PD3D-GNW6]. Also, see the critique of copyright obstacles thwarting AI development by Craig, *supra* note 44.

94. Regulation 2024/1689, art. 111(3) (stating that "[p]roviders of general purpose AI models that have been placed on the market before [date of entry into force of the AI Act] shall take the necessary steps in order to comply with the obligations laid down in this Regulation by" two years after the date of entry into force of the AI Act).

models. In consequence, EU culture will be less visible in AI output. The more that AI-mediated communication becomes the rule and the *lingua franca* of coming generations of "genAI natives," the more problematic the unavailability of EU training resources will become. With its restrictive approach to AI development, the EU substantially reduces the impact of its literary and artistic thinking and the messages and values conveyed by European literary and artistic productions on the evolving new mode of AI-supported human communication. The same can be said about the non-EU repertoire that is covered by an opt-out declared by a European or foreign rights holder in accordance with Article 4(3) CDSMD. The rights reservation option impoverishes the spectrum of AI training resources—and the spectrum of expressions that become part of the *lingua franca* of future generations of generative AI users in the EU.

In a less pessimistic scenario, the complexity of EU regulations does not completely extinguish the appetite for EU training resources. However, the transaction costs evolving from licensing and reporting obligations can easily lead to a focus on big repertoire holders and "high-level deals" between large AI companies and content majors in the creative industry.[95] Owners of niche repertoires in the creative industry may have difficulty convincing AI trainers that the drop in training resources they provide will make a difference in the sea of literary and artistic input required for AI development. In the end, the EU approach may thus promote targeted licensing initiatives covering predominantly mainstream repertoires of large rights holders. AI training based on mainstream repertoire, however, will inevitably lead to AI models generating mainstream output. With regard to EU knowledge resources, generative AI systems may thus be biased.[96] Instead of reflecting the full spectrum of literary and artistic traditions and expressions in EU Member States, they only offer content based on mainstream productions—content that neglects smaller repertoires and cultural traditions.

95. For an example of an existing licensing success at large company and big repertoire level, see the agreement concluded between Universal Music and Google/YouTube, as described by Anna Nicolaou & Madhumita Murgia, *Google and Universal Music Negotiate Deal over AI "Deepfakes,"* FIN. TIMES (August 8, 2023), https://www.ft.com/content/6f022306-2f83-4da7-8066-51386e8fe63b [https://perma.cc/K93G-6SV6].

96. *See* Geiger & Iaia, *supra* note 6, at 18–19; Craig, *supra* note 44; Lemley & Casey, *supra* note 6, at 770.

III. EX POST APPROACH: REMUNERATION AS A COMPONENT OF AI EXPLOITATION

Considering these drawbacks of the EU system requiring rights clearance and remuneration as a prerequisite for AI development, it is important to explore alternative approaches. As indicated, a remuneration mechanism in favor of human authors need not focus on and impose heavy administrative burdens on the AI training phase. Alternatively, the final offer of AI services and products on the market can serve as a reference point for a legal obligation to pay remuneration (output dimension).

More specifically, it seems possible to establish a lumpsum remuneration system that channels a certain share of revenue that accrues from the supply and use of generative AI systems in the literary and artistic field to human authors.[97] Instead of imposing payment obligations and administrative burdens on AI developers during the training phase (see the discussion in the preceding section), output-based remuneration systems may give AI trainers far-reaching freedom. Without exposure to any payment or administrative obligation, lawmakers can permit the use of the full spectrum of available literary and artistic resources for AI training purposes.[98] As a result, AI trainers can develop the best and most powerful AI models. Relying on diverse literary and artistic sources for training purposes, they can also ensure that these models can produce fair, unbiased output: content that reflects all cultures, traditions, and values expressed in human artistic and literary works. Once these fully developed AI systems are brought to the market, however, the question of fair remuneration for authors arises. As compensation for the unbridled freedom to use human creations for training purposes, an output-based system requires that authors be remunerated when the supply and use of AI systems in the marketplace generate income.

A. Towards a Lumpsum Remuneration System

Following this alternative approach, providers of generative AI systems would be obligated to pay equitable remuneration for the production of literary and artistic content which has the potential to serve as a substitute for

97. As to theoretical groundwork for this approach, see generally Geiger, *supra* note 75, at 532–33; Christophe Geiger, *Freedom of Artistic Creativity and Copyright Law: A Compatible Combination?*, 8 U.C. IRVINE L. REV. 413, 448–58 (2018); Christophe Geiger, *Elaborating a Human Rights Friendly Copyright Framework for Generative AI*, 55 IIC – INT'L REV. INTELL. PROP. & COMPETITION L. 1129 (2024). For concrete lumpsum remuneration proposals in the AI debate, see Senftleben, *supra* note 4, at 1549–56; Frosio, *supra* note 6, at 19–21; Geiger & Iaia, *supra* note 6, at 5–6; Geiger, *supra* note 6, at 79–80; Lucchi, *supra* note 76, at 18–19.

98. As to the advantages of this approach, see Craig, *supra* note 44.

human creations.[99] Surveying the canon of international rules in the field of copyright and neighboring rights, it becomes apparent that a lumpsum remuneration approach is not entirely alien to the protection system. Article 15(1) of the WIPO Performances and Phonograms Treaty ("WPPT") reads as follows: "[p]erformers and producers of phonograms shall enjoy the right to a single equitable remuneration for the direct or indirect use of phonograms published for commercial purposes for broadcasting or for any communication to the public."[100]

Hence, the international copyright community has already recognized that there might be specific circumstances requiring a switch from a right to prohibit use to a mere remuneration claim. As Article 15(1) WPPT is a norm of international law, it seems safe to assume that domestic counterparts of this right to receive equitable remuneration can be found in countries and regions around the globe.[101] Taking the EU exponent of Article 15(1) WPPT as an example, it can be demonstrated how an equitable remuneration rule relating to AI output could be modeled on the equitable remuneration rule in the area of phonograms. In the EU, Article 8(2) of the Rental, Lending and Related Rights Directive ("RLRRD")[102] implements the international norm into harmonized copyright and neighboring rights law:

> Member States shall provide a right in order to ensure that a single equitable remuneration is paid by the user, if a phonogram published for commercial purposes, or a reproduction of such phonogram, is used for broadcasting by wireless means or for any communication to the public, and to ensure that this remuneration is shared between the relevant performers and phonogram producers.[103]

Using this formulation as a blueprint, a new remuneration rule in the area of generative AI systems could take the following shape:

> Member States shall provide a right in order to ensure that a single equitable remuneration is paid by the provider of a generative AI system, if a

99. Senftleben, *supra* note 4, at 1549–56. For more information on remuneration programs refer to Lucchi, *supra* note 76, at 18–19.

100. World Intellectual Property Organization Performances and Phonograms Treaty, *opened for signature* Dec. 20, 1996, T.I.A.S. No. 02-306.1, 2186 U.N.T.S. 38543.

101. As to the distinction between rights "of a preventive nature" and rights "of a compensatory nature," see Case C-135/10, Societa Consortile Fonografici v. Marco Del Corso, ECLI:EU:C:2012:140, ¶ 75 (Mar. 15, 2012); Case C-351/12, OSA – Ochranný svaz autorský pro práva k dílům hudebním o.s. v. Léčebné lázně Mariánské Lázně a.s., ECLI:EU:C:2014:110, ¶ 35 (Feb. 27, 2014); Case C-117/15, Reha Training Gesellschaft für Sport- und Unfallrehabilitation mbH v. Gesellschaft für musikalische Aufführungs- und mechanische Vervielfältigungsrechte eV (GEMA), ECLI:EU:C:2016:379, ¶¶ 33–34 (May 31, 2016). *See also* Martin Senftleben, *Flexibility Grave – Partial Reproduction Focus and Closed System Fetishism in CJEU*, Pelham, 51 IIC – INT'L REV. INTELL. PROP. & COMPETITION L. 751, 762 (2020); D.J.G. Visser, Openbaar maken met Ketchup [*Disclosure with Ketchup*], TIJDSCHRIFT VOOR AUTEURS-, MEDIA- EN INFORMATIERECHT 41, 46 (2013) (Neth.).

102. Council Directive 92/100, art. 8, 1992 O.J. (L 346) 61 (EC).

103. *Id.*

> literary and artistic output generated by the system, has the potential to serve as a substitute for a work made by a human author, and to ensure that this remuneration is paid to social and cultural funds of collective management organizations[104] for the purpose of fostering and supporting human literary and artistic work.[105]

Admittedly, each country considering the introduction of an output-based remuneration regime and the use of the equitable remuneration rule relating to phonograms as a template will have to refine further and clarify the lumpsum remuneration rule governing AI output before it can be adopted to support human authors. Potential definition hurdles, however, seem surmountable.

As to the question of which output quality is necessary to assume a substitution risk, for instance, it must be considered that, in line with the proposal developed here, the lawmaking process would aim at establishing a *lumpsum* remuneration system. Therefore, a general, abstract assessment of whether an AI system is capable of substituting human literary and artistic productions is sufficient to confirm that it has a disruptive effect and requires the payment of remuneration. For instance, a relevant substitution effect could be assumed whenever an AI system is capable of generating content that resembles human literary and artistic productions. A fictitious originality test could be applied in this context: if the AI output had been made by a human author and not by a machine, would the AI-generated content fulfill copyright law's originality test?

The general conceptual contours of the proposed lumpsum remuneration approach could be as follows: the system would serve the overarching purpose of creating a new revenue stream to support the work of flesh-and-blood authors. Revenue accruing from remuneration payments for the use of generative AI systems would then be channelled to collecting societies that use the money to improve the living and working conditions of human authors. In addition, the mandatory, inescapable obligation to pay equitable remuneration is intended to make the use of AI-generated content more expensive. AI system providers can no longer offer generative AI tools for free unless they are willing to pay the remuneration out of their own pocket. Hence, introducing a remuneration obligation also reduces the attractiveness

104. As to the room for social and cultural funds of collecting societies in EU copyright law see, Case C-521/11, Amazon.com Int'l Sales, Inc. v. Austro-Mechana Gesellschaft zur Wahrnehmung mechanisch-musikalischer Urheberrechte Gesellschaft mbH, ECLI:EU:C:2013:515, ¶¶ 49–52 (July 11, 2013). *See also* Martin Senftleben, *Copyright, Creators and Society's Need for Autonomous Art – The Blessing and Curse of Monetary Incentives*, *in* WHAT IF WE COULD REIMAGINE COPYRIGHT? 25, 64–68 (Rebecca Giblin & Kimberlee Weatherall eds., 2017).

105. For an earlier discussion of this proposal, see Senftleben, *Single Equitable Remuneration*, *supra* note 18, at 111–13; Senftleben, *Tax on Machines*, *supra* note 18, at 1–3.

of less expensive, automated AI content production. Theoretically, AI remuneration could even be set at a level that counterbalances lower production costs and enhances the chances of human authors to compete with generative AI systems. On its merits, the proposed system, thus, seeks to *transform AI content revenue into human content revenue*.[106]

More concrete guidelines for the use of the collected revenue can be derived from the six author remuneration objectives described in section I.A above.[107] Following the argument that the remuneration system offers compensation for the parasitic use of human works to enable AI systems to kill the market for human creativity (first argument), collected money could be used broadly to support human literary and artistic productions. For instance, it is conceivable to distribute revenue in accordance with a general repartitioning scheme that is based on the use of certain work repertoires or work genres for AI training purposes or aligned with the number of references to repertoires or genres in prompts entered by AI system users.[108] Insights into user prompts can provide important vectors for calibrating the repartitioning scheme. Data showing that certain work categories or genres are featured prominently in user prompts could offer a basis for increasing the revenue share for human authors whose creative labor concerns these categories and genres. The adoption of a general repartitioning scheme can also make sense from the perspective of the AI industry's own interest in the continuous evolution of fresh human creations that can become training material for the further improvement and diversification of AI output (the sixth argument).

An approach seeking to support literature and art projects may follow from the insight that AI-generated content may lead to a loss of human works and avant-garde movements that, as a mirror of social and political conditions, can provide new directions for future creativity and impulses for improving society (the second and third arguments). To the extent to which AI output can replace this type of avant-garde human creativity with particular societal value,[109] the establishment of cultural funds seeking to promote human works in the high arts sector seems warranted. Support for this

106. *See* Senftleben, *Single Equitable Remuneration*, *supra* note 18, at 113–14; Senftleben, *Tax on Machine*, *supra* note 18 , at 2–3.

107. Senftleben, *supra* note 4, at 1549–56. *See also* Lucchi, *supra* note 76, at 18–19.

108. However, see Geiger & Iaia, *supra* note 6, at 8 who warn of potential imbalances that might be caused in the case of highly popular works and artists.

109. See F. Hoffmann, *Zehn Thesen zu Künstlicher Intelligenz (KI) und Urheberrecht* [Ten Theses on Artificial Intelligence (AI) and Copyright], WETTBEWERB IN RECHT UND PRAXIS, 11, 17–18 (2024); Remy Chavannes, *De Bescherming van Deep Learning-Systemen door het Intellectuele Eigendomsrecht* [The Protection of Deep Learning Systems by Intellectual Property], 5 AMI - TIJDSCHRIFT VOOR AUTEURS-, MEDIA- EN INFORMATIERECHT 179, 182 (2018) (Neth.), who are confident that human fine art and avant-garde productions will survive the generative AI revolution.

approach, which involves investing in projects and activities in the field of literature and art, can also follow from the objective to stimulate human aesthetic engagement and ensure that human role models remain visible in society to inspire everyday human literary and artistic practice (the fifth argument). Considering the overarching goal to avoid the impression that the remix and reuse of literature and art is a task for the machine, the use of AI revenue for projects that further human creativity and allow participants to experience the positive effects of aesthetic play makes sense. Finally, the general socio-political goal of supporting human authors who lose their jobs due to competing AI content (the fourth argument) can justify investment in literature and art projects that offer new job opportunities for freelancers and small production companies.

The six rationales developed above, thus, offer a basis for different measures ranging from the establishment of a general repartitioning scheme to investment in social and cultural funds that support literature and art projects. In *Amazon International Sales v. Austro-Mechana Gesellschaft*, a case about the payment and repartitioning of private copying levies in Austria, the Court of Justice of the European Union ("CJEU") confirmed that EU law offers considerable flexibility with regard to the use of collected funds for social and cultural purposes.[110] One of the prejudicial questions asked by the Austrian Supreme Court was whether a collecting society lost its right to the payment of fair compensation if, in relation to half of the funds received, the collecting society was required by law not to pay the levy income to the persons entitled to compensation but to distribute it to social and cultural institutions.[111]

In answering this question, the CJEU held the view that EU law did not set forth an obligation to pay all the lumpsum remuneration collected via a levy system directly to rights owners in cash. By contrast, a Member State was free to provide that part of the lumpsum remuneration be distributed in the form of indirect compensation through social and cultural institutions set up for the benefit of authors and performing artists.[112] The fact that the remuneration had to be regarded as recompense for a specific harm did not constitute an obstacle to establishing such an indirect payment mechanism through the intermediary of social and cultural institutions.[113] The Court also stated that a system of indirect distribution of collected funds was conducive

110. Case C-521/11, Amazon.com International Sales, Inc. v. Austro-Mechana Gesellschaft zur Wahrnehmung mechanisch-musikalischer Urheberrechte Gesellschaft mbH, ECLI:EU:C:2013:515.

111. *Id.* ¶ 15

112. *Id.* ¶ 49.

113. *Id.* ¶ 50.

to ensuring that European cultural creativity and production received the necessary resources. It also safeguarded the independence and dignity of authors and performers.[114] The Court made it a condition, however, that the social and cultural establishments involved actually benefit those entitled to the lumpsum remuneration (human authors in the case of the AI levy system proposed here). Moreover, it was necessary that the detailed arrangements for the operation of social and cultural institutions were not discriminatory. Benefits had to be granted to those persons entitled to remuneration, and the system had to be open to nationals and foreigners alike.[115]

Arguably, this decision makes it possible to invest a substantial share of AI remuneration in social and cultural funds that, instead of paying collected money to individual rights holders, support human authors and human creativity by financing cultural events and contributing to literary and artistic productions.[116] This insight is important in light of doubts about the beneficial effects of collective licensing solutions that have been expressed in the AI remuneration debate. For instance, Pam Samuelson has asserted that a collective licensing regime would be "very difficult to administer given the staggeringly large number of copyright owners and of works and of types of works used as AI training data literally in the billions."[117] She also fears that "[a] significant proportion of revenues collected from generative AI companies would be needed to cover administrative costs,"[118] "amounts paid to individual copyright owners would likely be very modest, and would be unlikely to provide significant financial support to authors and artists"[119] and that "[a] proliferation of collective license regimes [in countries and regions around the globe], each of which would aim to collect significant sums to support authors and artists, might make development of generative AI too expensive to be feasible,"[120] in particular for small and medium-sized companies.[121]

While these points are particularly important, the criticism overlooks that, from the outset, a lumpsum remuneration regime that is administered collectively does not aim at meticulously determining the most accurate remuneration amount for each and every individual work used for AI training. As explained above, the system would seek to distribute money at the more

114. *Id.* ¶ 52.
115. *Id.* ¶¶ 53–54.
116. *See* Samuelson, *supra* note 6, at 81.
117. *Id.* at 79; *cf.* Craig, *supra* note 44, at 26–28.
118. Samuelson, *supra* note 6, at 79.
119. *Id.*
120. *Id.* at 80.
121. *Id.*

aggregated level of work categories and genres. Collecting societies administering these work categories and genres can pay AI revenue to their members together with other collected money they are distributing without too many additional administrative costs.

Moreover, the CJEU's *Amazon* decision shows that a substantial share of collected AI revenue can be invested in social and cultural funds that invest in impactful literature and art projects.[122] The CJEU recognized that it was very difficult, if not impossible, to calculate the individual damage that an author suffered because of private copying. Considering this difficulty, the Court underlined that Member States enjoyed "wide discretion" in determining the form, the detailed arrangements, and the possible level of lumpsum remuneration.[123] The difficulty in calculating the damage and the required compensation at the level of individual authors was the very reason for the Court to hold that Member States were free to establish a system of indirect remuneration via social and cultural institutions.[124] Hence, the Court itself did not insist on a system that distributes collected money meticulously on the basis of the individual harm suffered by an author because such a detailed calculation of individual damage was hardly possible. A parallel between this aspect of the *Amazon* case and output-based AI remuneration can easily be drawn. As in the case of private copying, it is hardly possible to calculate the exact damage that the offer and use of generative AI inflicts on an individual author or the specific value of that author's works for AI training.[125]

Regardless of this imprecision, a lumpsum remuneration approach has important advantages. In comparison to tax-based support for literature and art, cultural funds of collecting societies offer authors and performers the opportunity to establish rules for the use of collected money themselves and decide autonomously within the creative sector on measures to support human creativity. The censorship risk arising from dependence on general tax

122. Case C-521/11, Amazon.com International Sales, Inc. v. Austro-Mechana Gesellschaft zur Wahrnehmung mechanisch-musikalischer Urheberrechte Gesellschaft mbH, ECLI:EU:C:2013:515, ¶ 50.

123. *Id.* ¶¶ 20, 40.

124. *Id.* ¶ 49.

125. *See* Klaus Goldhammer et al., *AI and Music – Market Development of AI in the Music Sector and Impact on Music Authors and Creators in Germany and France*, GOLDMEDIA 1, 63 (2024), https://www.gema.de/en/news/ai-study [https://perma.cc/998V-RBCN] (report commissioned by SAVEN and GEMA) (explaining that AI training "results in a calculated AI model in which the scraping data is not available as copies. During processing, the parameters (weights) derived from the scraping database are used instead. New content is generated based on the parameters and patterns learnt. The exact way in which the parameters are categorised in the model cannot currently be clearly described from a technical perspective."). *See generally* Geiger & Iaia, *supra* note 6.

money and government decisions that may favor certain types of literature and art, and discriminate against others can be avoided.

Considering the described options for configuring and implementing lumpsum remuneration systems, it is inaccurate to assume that collective licensing solutions would sacrifice innovation in AI for meaningless micropayments to individual rights holders and enormous administrative costs of collecting societies. Even if collective licensing approaches metastasize across countries and regions, they will not thwart AI development. As explained, the output-based approach proposed here would avoid burdens on AI training and give AI trainers far-reaching freedom to use human works for AI development. The obligation to pay remuneration only arises the moment fully trained AI systems are brought to the market. At that stage, payment obligations in different countries and regions do not appear overly burdensome. They are a corollary of income prospects in all territories involved. If an AI company successfully launches its products and services—and makes money—in several countries, it can be expected to contribute part of its revenue to the remuneration of authors in all these countries.

B. Foundation in Copyright Law

As the proposed lumpsum remuneration system focuses on AI output, it raises a delicate legal-doctrinal question as to the copyright basis for the remuneration claim. Content produced by a generative AI system need not display protected traces of individual human expression.[126] Compared to the AI training (input) perspective, the situation is different. During the AI training phase, protected human works are used as learning resources for the AI system. Hence, there is a direct link between the machine-learning process and the use of protected human literary and artistic works. Qualifying copies made for AI training purposes as relevant reproductions,[127] the lawmaker can create a legal basis for a remuneration claim in copyright law. With regard to AI output, however, the copyright basis for equitable remuneration is less clear. Instead of reproducing protected individual expression ("free and creative choices") by a human author,[128] AI output may merely reflect unprotected ideas, concepts and styles.[129]

126. Lemley & Casey, *supra* note 6, at 772–76.

127. For a more detailed discussion of this question, see Senftleben, *supra* note 59, at 1495–1502.

128. Case C-5/08, Infopaq Int'l A/S v. Danske Dagblades Forening, ECLI:EU:C:2009:465, ¶ 45 (July 16, 2009); Case C-145/10, Eva-Maria Painer v. Standard Verlags GmbH, ECLI:EU:C:2011:798, ¶ 89 (Dec. 1, 2011).

129. Agreement on Trade-Related Aspects of Intellectual Property Rights, art. 9(2), Apr. 15 1994, 1869 U.N.T.S. 299; World Intellectual Property Organization Copyright Treaty, *opened for signature* Dec. 20, 1996, T.I.A.S. No. 02-306.1, 2186 U.N.T.S. 38542. *See also* HR 29 maart 2013, NJ 2013, 504

The absence of protected human expression in AI output, however, does not pose an insuperable obstacle. In fact, a copyright concept that, by analogy, can be invoked as a legal-doctrinal basis for the introduction of a lump-sum remuneration system focusing on AI output has already been developed in the last century. In the discussion on the so-called *domaine public payant*, Adolf Dietz explained in a 1990 landmark article that, in addition to traditional exploitation and remuneration rights of individual authors, it is consistent and advisable to recognize in copyright law a new right to which a different right holder, the "community of authors,"[130] is entitled as a collective. Dietz pointed out that this step could be regarded as a corollary of a modern understanding of copyright law "as part of a more comprehensive concept of culture law."[131] Once this broader role and responsibility of copyright is taken as a starting point, the law is no longer condemned to accept "harmful discrepancies"[132] between substantial profits made by exploiters of public domain works on the one hand, and precarious working and living conditions of current authors on the other.[133] Instead, copyright can be employed as a legal tool to introduce a remuneration right for the community of living and creating authors as a means of redress:

> What we finally propose is simply to introduce another right owner, namely the community of living and creating authors, among several kinds of right owners already existing in copyright law. This community of authors should have the direct right to participate in the income from exploitation of works of dead authors after the individual term of copyright protection has expired.[134]

As this statement indicates, Dietz developed his concept of a new right for the community of authors with a focus on the exploitation of works in the public domain. He placed his proposal in the context of the discussion on the *domaine public payant* that had gained momentum after the Second World War.[135] From his perspective, soaring prices and income from the exploitation of public domain works in the field of literature, music, and art

m.nt. P.B. Hugenholtz (Broeren/Duijsens), ECLI:NL:HR:2013:BY8661, ¶ 3.5 (Neth.); Senftleben, *supra* note 14, at 27–28.

130. Dietz, *supra* note 25, at 15.

131. *Id.* at 13.

132. *Id.*

133. *Id.*

134. *Id.* at 14.

135. As to the historical origin and development of the *domaine public payant*, see Walter Dillenz, *Überlegungen zum Domaine Public Payant* [Reflections on the Domaine Public Payant], 1983 GEWERBLICHER RECHTSSCHUTZ UND URHEBERRECHT – INTERNATIONAL [GRUR INT'L] 920, 920–22 (Ger.).

should, "at least partly and proportionally, also serve the living and creating generation of authors."[136]

Evidently, the introduction of a new collective right to participate in revenue accruing from the exploitation of public domain works raises the question of how this new right of the community of living and creating authors might be exercised in practice. Dietz solves this problem by relying on the well-established system of collective rights management in Europe.

> [T]here must be a natural or legal person or body ready to interfere and, in particular, to control the market and claim the participation right, if necessary in a lawsuit. In addition, this body must be able to distribute the incoming money according to statutory purposes and rules, preferably under government supervision. We should not forget, however, that these kind of bodies already exist, and have done so for decades, in the form of collecting societies.[137]

Before turning to parallels between this remuneration concept and the AI remuneration system discussed here, it is noteworthy that in the second half of the last century, the proposal of a *domaine public payant* did not remain a mere theoretical option. In Germany, it formed part of the official government proposal for new copyright legislation discussed in 1965.[138] Although the German legislators finally refrained from introducing a new remuneration right for the community of authors in the 1965 Copyright Act,[139] the fact that the *domaine public payant* was included in the government proposal shows that the concept and the underlying objective to improve the working and living conditions of authors had broad support in Germany.[140] An international UNESCO/WIPO survey conducted in 1982 also brought to light several starting points for implementing the *domaine public payant* in copyright law.[141]

In more recent debates on recalibrating copyright, Rebecca Giblin confirmed the concept's continued relevance and importance. In a critical assessment of the term of copyright protection, she qualified the *domaine public payant* as a useful reference point for her proposal to draw a clearer

136. Dietz, *supra* note 25, at 14.

137. *Id.* at 15.

138. Adolf Dietz, Die sozialen Bestrebungen der Schriftsteller und Künstler und das Urheberrecht, [The Social Aspirations of Writers and Artists and Copyright], 11 GEWERBLICHER RECHTSSCHUTZ UND URHEBERRECHT [GRUR] 12, 14-15 (Ger.).

139. Gesetz über Urhebeerrecht und verandte Schutzrechte [Urheberrechtsgesetz] [UrhG] [German Copyright Act] 9, 1965 BGBI at 1273 (Ger.), *last amended by* Art. 25, June 23, 2021, BGBI at 1858, https://www.gesetze-im-internet.de/englisch_urhg/ [https://perma.cc/4LLP-D82M].

140. *Cf.* Dietz, *supra* note 25, at 14-15.

141. Committee of Non-Governmental Experts on the 'Domaine Public Payant,': *Analysis of the Replies to the Survey of Existing Provisions for the Application of the System of 'Domaine Public Payant' in National Legislation*, U.N. Doc. UNESCO/WIPO/DPP/CE/I/2 (Mar. 10, 1982), https://unesdoc.unesco.org/ark:/48223/pf0000048044.

distinction between incentive and reward goals and introduce an opt-in "creator-right" that would give authors access to remuneration systems in return for the registration of their works after an initial term of protection.[142]

The parallels between the *domaine public payant* and the proposed output-based remuneration system in the area of generative AI are striking. Both concepts concern creations that fall outside the scope of the exploitation rights of individual authors: literary and artistic works that never or no longer enjoy copyright protection in the case of the *domaine public payant*; general ideas, concepts, and styles in the case of AI output that does not reproduce individual expression of a human author. At the same time, it is clear that both concepts concern literary and artistic subject matter: public-domain works and public-domain ideas, concepts, and styles. With regard to AI output, it can even be added that pre-existing human creations have been a *conditio sine qua non* for the literary and artistic productions at issue. Without human training material, the machine could not have generated the content. The same cannot be said about public domain masterpieces made by authors from the past. Current authors can hardly assert that these masterpieces depended on their creative input.

The precursor of the *domaine public payant*, thus, shows that potential legal-doctrinal concerns need not thwart introducing a remuneration system focusing on AI output. Even if AI output merely reflects unprotected ideas, concepts, and styles, it is still possible and consistent to incorporate a lump-sum remuneration right in copyright law, which would be a new right that is subject to mandatory collective rights management. As a new right holder, the community of living authors[143] should be entitled to benefit from payments made under this new system. In line with this approach, the collective remuneration right should be administered and enforced by collecting societies that distribute collected money through repartitioning schemes and social and cultural funds.

Alternatively, it is possible to forge a link with the input dimension, the use of copyrighted works for AI training purposes, and focus on the use of human training material as an indispensable precondition for AI output that resembles human literary and artistic productions. As already pointed out above, generative AI systems are only capable of mimicking human creativity because human works have been used as training material at some

142. Rebecca Giblin, *Reimaging Copyright's Duration*, *in* WHAT IF WE COULD REIMAGINE COPYRIGHT? 177, 200–03, 207–08 (Rebecca Giblin & Kimberlee Weatherall eds., 2017).

143. Dietz, *supra* note 25, at 15.

stage.[144] Even in the case of AI systems trained on synthetic, machine-made literary and artistic material, the system's capability to mimic human creativity can only be explained by the fact that human training resources played a role somewhere in the whole chain of training processes leading to the generative AI system producing output that resembles a human work.

Considering this connection between input and output, it can be argued that remuneration for literary and artistic AI output must be paid because, directly or indirectly, this output is the result of the use of human works for AI training. Once again: without the use of copyrighted training material at some stage in the chain of training processes leading to a generative AI system, the literary and artistic output would not be possible. As input and output are thus two sides of the same coin, the payment of remuneration at the output level simply constitutes a deliberate choice of the legislator. Instead of placing heavy administrative and financial burdens on AI trainers, the lawmaker can leave the training process (input dimension) unencumbered and take measures to compensate authors when final AI products and services are offered in the marketplace and produce literary and artistic content (the output dimension).

This detachment of the act triggering the payment obligation from the act that provides the legal basis for the compensation claim is not unusual in the area of lumpsum remuneration systems. In the context of private copying, for instance, the CJEU has explicitly recognized that EU Member States are free to impose an obligation to pay compensation for reproductions made by private users on manufacturers and importers of relevant copying equipment, devices, and media. Even though the act with copyright relevance, the private copying, will only occur after the equipment, devices, and media have reached end consumers, the payment obligation can be imposed on manufacturers and importers:

> given the practical difficulties in identifying private users and obliging them to compensate rightholders for the harm caused to them, and bearing in mind the fact that the harm which may arise from each private use, considered separately, may be minimal and therefore does not give rise to an obligation for payment . . . , it is open to the Member States to establish a "private copying levy" for the purposes of financing fair compensation chargeable not to the private persons concerned, but to those who have the

144. In this category, a distinction can be drawn between "machine-learning" and "deep-learning" algorithms. Gervais, *supra* note 7, at 2055–59; Ginsburg & Budiardjo, *supra* note 7, at 401–02; Deltorn, *supra* note 7, at 173–74; Boden, *supra* note 7, at 23. For a practical example of AI-generated imitations of human vocals, see Laura Snapes, *AI Song Featuring Fake Drake and Weekend Vocals Pulled From Streaming Services*, THE GUARDIAN (Apr. 18, 2023, 5:37 PM), https://www.theguardian.com/music/2023/apr/18/ai-song-featuring-fake-drake-and-weeknd-vocals-pulled-from-streaming-services [https://perma.cc/69CD-WCSY]. *Contra* Bridy, *supra* note 7, at 3.

> digital reproduction equipment, devices and media and who, on that basis, in law or in fact, make that equipment available to private users or who provide copying services for them.[145]

In the light of this existing configuration of levy systems in the area of private copying,[146] it does not seem unusual (and perhaps even less unusual than a legal-doctrinal solution based on the *domaine public payant*) to simply delay the remuneration payment and take the production of literary and artistic AI output as a reference point for compensating human authors for the use of their works during AI training. This alternative legal-doctrinal approach forges a link with proposals to introduce a statutory license and a remuneration regime at the AI training stage.[147]

Considering the practical problems arising from the opt-out mechanism in Article 4(3) CDSMD and the transparency and training requirements in the AI Act, Christophe Geiger and Vincenzo Iaia have recommended the adoption of statutory licenses and remuneration rules for AI training.[148] They propose this regulatory approach as an alternative to the overly burdensome rights reservation system in the EU with all the shortcomings and difficulties described above. Focusing on input for AI development, Geiger and Iaia argue that the switch to a right to fair remuneration "allows maximization of the copyright content exploitable for [machine learning] purposes while taking into account the interests of the authors to be remunerated for the commercial use of their intellectual efforts . . . ."[149] Based on an analysis of the need to reconcile AI innovation with authors' remuneration interests in the light of fundamental rights, they conclude that "[t]he introduction of a remunerated copyright limitation for Generative AI commercial purposes represents a compelling solution to meet the endangered remuneration right of creators without disproportionately sacrificing the interest of AI developers to offer increasingly high-performing services able to foster human creativity."[150]

Similarly, Giancarlo Frosio has stated that the adoption of a lumpsum remuneration approach for AI training offers the opportunity to reconcile the societal interest in rich training resources for high-quality AI systems with

145. Case C-467/08, Padawan SL v. Sociedad General de Autores y Editores de España (SGAE), ELCI:EU:C:2010:620, ¶ 46 (Oct. 21, 2010).

146. As to parallels with lumpsum remuneration systems for private copying in the debate on generative AI systems and author remuneration, see Geiger & Iaia, *supra* note 6, at 6.

147. Frosio, *supra* note 6, at 19–21; Geiger & Iaia, *supra* note 6, at 11; Geiger, *supra* note 6, at 79–80; Geiger, *supra* note 97, manuscript at 29–33; Kop, *supra* note 76, at 7.

148. Geiger & Iaia, *supra* note 6, at 11.

149. *Id.* at 14.

150. *Id.* at 16.

the legitimate interest of authors to receive fair remuneration for the use of their works:

> [t]o address this tension, while simultaneously (1) addressing market substitution of human creations by AI-generated creativity, (2) considering the mentioned inherently combinatorial nature of such creativity, (3) lowering transaction costs of traditional copyright exclusivity models, and (4) potentially spurring further innovation by encouraging the creation of new original works that could be used for AI training, a levy system should be implemented. This system would allocate revenue from AI productions to human literary and artistic endeavors, with collective management of revenues that might proceed from Generative AI platform facilitating such AI creativity.[151]

Frosio's analysis leaves no doubt that he sees AI development involving the use of human works as input for AI training purposes as the central reference point for the payment of remuneration. He points out that the levy "would pertain specifically to the use of content as input for training foundation models, rather than serving as compensation for potential output infringements."[152] In his view, only using protected works as input for training foundational models can give rise to liability of AI trainers and an obligation to pay fair remuneration.[153]

Arguing for the payment of fair remuneration at a later stage when AI products and services are finally brought to the market, the output-based proposal developed here seems irreconcilable with this approach. However, once output-based remuneration is seen as delayed compensation for the use of human work during the AI training phase, this contradiction vanishes. Rightly understood, output-based remuneration approaches have the potential to integrate all lumpsum remuneration proposals that have been made in the AI debate. Output-based remuneration approaches simply delay the payment of remuneration until AI products and services are finally offered in the marketplace. The remuneration claim, however, can be traced back to the insight that AI output mimicking human creativity is only possible because human creations have been used as training material for generative AI systems.

This basic insight provides common ground for all lumpsum remuneration proposals that seek to offer fair compensation for the use of human works for AI training. The difference between the described approaches focusing on AI development (the proposals made by Geiger, Iaia, and Frosio) and the approach focusing on AI exploitation (the proposal made here)

151. Frosio, *supra* note 6, at 19.
152. *Id.*
153. *Id.*

focuses on the point in time when the remuneration payment is due. While the proposals focusing on the development phase require the payment of fair remuneration as a precondition for the use of human works for AI training, the output-based approach advocated here offers maximum access to human training resources without administrative and payment obligations. It only requires the payment of fair remuneration when a fully trained AI system is finally exploited on the market.

In sum, there are thus two legal-doctrinal avenues that can lead to the introduction of a lumpsum remuneration system focusing on AI output. On the one hand, the concept of *domaine public payant* offers a basis for establishing a collective right of the community of living authors, a new right that is subject to mandatory collective rights management, to receive remuneration for AI output reflecting literary and artistic ideas, concepts, and styles. On the other hand, the focus can be on the use of human training material as an indispensable precondition for AI output that resembles human literary and artistic productions. Following the example of levy systems in the area of private copying, it is possible to uncouple the act triggering the payment obligation from the act that provides the legal basis for the compensation claim. Hence, the legislator is free to delay the remuneration payment and take the production of literary and artistic AI output as a reference point for compensating human authors for the use of their works during AI training.

C. Legal and Practical Advantages

An output-oriented remuneration system offers essential practical advantages compared to the above-described input-based remuneration architecture, which the EU legislator has erected based on the rights reservation option in Article 4(3) CDSMD and fortified with the additional rules in the AI Act.

First, an output-oriented remuneration system can be applied uniformly to all providers of generative AI systems. In contrast to a remuneration obligation focusing on the input dimension and AI training activities, an output-oriented approach avoids the risk of disadvantaging domestic high-tech industries and does not encourage them to look for a more favorable training environment elsewhere. All providers of generative AI systems are equally exposed to the payment obligation the moment they offer their products and services in the marketplace.

Second, an output-based remuneration system raises less trust and transparency issues. As already pointed out, the remuneration system can be based on the payment of a global lumpsum. For instance, the remuneration could consist of a certain percentage of revenue which AI companies derive

from advertising, subscription fees or other payments they receive from users.[154] In the case of profit-oriented providers of generative AI systems, the remuneration may also consist of a certain percentage of the annual turnover. It also seems possible to explore the possibility of aligning the payment with the number of AI-generated literary and artistic products or the number of prompts entered by users.

Third, the involvement of collecting societies in the area of remuneration for AI output can ensure that the original creators of human works can benefit directly from the extra income accruing from payments. The repartitioning schemes of collecting societies consider not only industry rights holders, but also—and in particular—individual authors.[155] In international law, remuneration rules in the area of neighboring rights reflect this approach. With regard to the single equitable remuneration that is due for the broadcasting and other communication to the public of published phonograms, Article 15(2) WPPT, for instance, explicitly points out that "[c]ontracting Parties may enact national legislation that, in the absence of an agreement between the performer and the producer of a phonogram, sets the terms according to which performers and producers of phonograms shall share the single equitable remuneration."[156] In contrast to industry collaboration evolving from the rights reservation option in Article 4(3) of the CDSMD, the proposed lumpsum remuneration approach at the output stage need not give rise to concerns that collected money will hardly ever reach individual creators.[157]

Fourth, a lumpsum remuneration approach does not require the management of use permissions at the level of individual works. As explained above, the reservation of copyright on the basis of Article 4(3) CDSMD will only lead to the payment of remuneration if a machine-readable rights reservation is combined with the offer of a TDM permission against the payment of remuneration. To achieve this goal, however, it is necessary to establish a well-functioning rights clearance infrastructure that is capable of interacting

154. *See* Lucchi, *supra* note 76, at 18–19.

155. *See* Case C-521/11, Amazon.com International Sales Inc. v. Austro-Mechana Gesellschaft mbH, ECLI:EU:C:2013:145, ¶¶ 49–52 (Mar. 7, 2013).

156. World Intellectual Property Organization Performances and Phonograms Treaty art. 15, *opened for signature* Dec. 20, 1996, T.I.A.S. No. 02-306.1, 2186 U.N.T.S. 38543.

157. In the context of repartitioning schemes of collecting societies, the individual creator has a relatively strong position. As to national case law explicitly stating that a remuneration right leads to an improvement of the income situation of the individual creator, see Bundesgerichtshof [BGH], Federal Court of Justice, July 11, 2002, I ZR 255/00 1, 14–15 (Ger.). For a discussion of the individual creator's entitlement to income from the payment of equitable remuneration, see the literature references, Westkamp *supra* note 75, at 55–59; Quintais *supra* note 75, at 335–36, 340–41, 347–49, 356–57; Geiger *supra* note 75, at 532–33; Hilty *supra* note 75, 325–353.

with content crawlers that are used for AI training purposes. An output-oriented lumpsum remuneration approach, by contrast, need not pose comparable practical and administrative obstacles.

Instead of imposing payment obligations and administrative burdens during the AI development phase, lawmakers can exempt the AI training process from the control of copyright holders. The law can permit the use of the full spectrum of available literary and artistic resources for AI training purposes and allow AI trainers to develop the best and most powerful AI models. Only at a later stage, when these fully developed models are exploited in the marketplace, the payment obligation arises. As compensation for the unbridled freedom to use human creations for training purposes, the output-based lumpsum system proposed here requires the payment of fair remuneration when the supply and use of AI systems in the market finally generates income.

The reduction of financial and administrative burdens leads to a fifth advantage: with an output-based remuneration system, it is no longer necessary to introduce opt-out mechanisms and burdensome upfront transparency and licensing obligations. Lawmakers adopting an output-based remuneration approach need not copy the rights reservation system of Article 4(3) CDSMD and the accompanying AI Act rules. They can avoid these highly complex AI training rules that may deter AI trainers and render literary and artistic resources unavailable or unattractive in the light of practical and administrative hurdles. In this way, an output-based remuneration approach bans the risk of marginalizing a country's literary and artistic repertoire in AI training datasets and AI output. By making the full spectrum of literary and artistic works available for training purposes, it removes access obstacles, contributes to the reduction of biases in AI models that may result from the unavailability of work repertoires, and, as a corollary, broadens the spectrum of messages and values which AI systems can reflect when producing literary and artistic output. By removing access obstacles at the training stage, an output-based remuneration approach thus contributes to better, unbiased AI models.

D. Tool for Ensuring Compliance With Three-Step Test

As a final remark, it is important to point out that an output-based approach—ensuring author remuneration when AI products and services are finally offered in the marketplace—also complies with the so-called three-step test.[158] In EU copyright law, Article 5(5) ISD stipulates that copyright

158. *Cf.* Geiger & Iaia, *supra* note 6, at 7.

"exceptions and limitations shall only be applied in certain special cases[,] which do not conflict with a normal exploitation of the work or other subject-matter and do not unreasonably prejudice the legitimate interests of the right holder."[159]

With regard to the TDM exceptions in EU copyright law, Article 7(2) CDSMD confirms that this three-step test of Article 5(5) must be taken into account.[160] The decision to rely on the test as an additional assessment tool reflects the international obligations of the EU and its Member States. The three-step test of Article 5(5) ISD is an offspring of Article 9(2) of the Berne Convention for the Protection of Literary and Artistic Works ("BC"); Article 13 of the Agreement on Trade-Related Aspects of Intellectual Property Rights ("TRIPS"); Articles 10(1) and (2) of the WIPO Copyright Treaty ("WCT"); and Articles 16(1) and (2) of the WIPO Performances and Phonograms Treaty ("WPPT").[161] Its three assessment criteria are in line with the substantive criteria set forth in these international provisions.

This international dimension of the provision yields important insights with regard to the role of equitable remuneration in ensuring compliance with the three test criteria. When the first three-step test in international copyright law, the test of Article 9(2) BC, was adopted at the 1967 Stockholm Conference for the Revision of the Berne Convention, the report on the work of Main Committee I (the Committee dealing with the three-step test) provided the following example of the functioning of the test:

> A practical example might be photocopying for various purposes. If it consists of producing a very large number of copies, it may not be permitted, as it conflicts with a normal exploitation of the work. If it implies a rather large number of copies for use in industrial undertakings, it may not unreasonably prejudice the legitimate interests of the author, provided that, according to national legislation, an equitable remuneration is paid. If a

159. Directive 2001/29/EC, art. 5(5), 2001 O.J. (L. 167) 10, 16.

160. For a more detailed discussion of the interplay between the TDM provisions and the three-step test, see Martin Senftleben, *EU Copyright 20 Years After the InfoSoc Directive – Flexibility Needed More Than Ever, in* REFORMING INTELLECTUAL PROPERTY, 185, 204–05 (G. Ghidini & V. Falce eds., 2022).

161. Amendments Adopted on the Proposal for a Directive on Copyright in the Digital Market, at 8, COM (2016) 0593 final (Sept. 12, 2018). With regard to the evolution of this "family" of copyright three-step tests in international copyright law, see generally M.R.F. Senftleben, *Copyright, Limitations and the Three-Step Test – An Analysis of the Three-Step Test*, *in* INTERNATIONAL AND EC COPYRIGHT LAW 43 (2004); Daniel Gervais, *Fair Use, Fair Dealing, Fair Principles: Efforts to Conceptualize Exceptions and Limitations to Copyright*, 57 J. COPYRIGHT SOC'Y 499, 510–11 (2010); Annette Kur, *Of Oceans, Islands, and Inland Water – How Much Room for Exceptions and Limitations Under the Three-Step Test?*, 8 RICH. J. GLOB. L. & BUS. 287, 307–08 (2009); SAM RICKETSON & JANE C. GINSBURG, INTERNATIONAL COPYRIGHT AND NEIGHBOURING RIGHTS: THE BERNE CONVENTION AND BEYOND 759–63 (3rd ed. 2022); Christopher Geiger et al., *The Three-Step Test Revisited: How to Use the Test's Flexibility in National Copyright Law*, 29 AM. U. INT'L. L. REV. 581, 583–91 (2014); Joachim Bornkamm, *Der Dreistufentest als urheberrechtliche Schrankenbestimmung – Karriere eines Begriffs*, *in* FESTSCHRIFT FÜR WILLI ERDMANN ZUM 65. GEBURTSTAG 29 (Hans-Jürgen Ahrens et al. eds., 2002).

> small number of copies is made, photocopying may be permitted without payment, particularly for individual or scientific use.[162]

Hence, the payment of equitable remuneration was regarded as a factor capable of tipping the scales in favor of a finding of compliance; it reduces the unreasonable prejudice arising from the exemption of "a rather large number of copies for use in industrial undertakings" to a permissible reasonable level. A parallel between this example in the drafting history of the three-step test and the role of author remuneration in the debate on generative AI systems can easily be drawn: providing for the payment of equitable remuneration when generative AI products and services are brought to the market, the legislator softens displacement effects and avoids unreasonable prejudices that are exponents of the training of generative AI systems with human works.

Once fully trained AI systems are brought to the marketplace, the law can ensure that AI system providers pay equitable remuneration to compensate authors for displacement effects. They can pass these costs on to users or use a part of their own revenue. The legislator can thus employ output-oriented remuneration systems as legal tools to reduce the prejudice caused by the unbridled use of human material for AI training to a level that can be deemed reasonable and permissible in the three-step test analysis. In light of the example given at the 1967 Stockholm Conference, it can be concluded that the three-step test offers flexibility for giving broad access to literary and artistic works in the AI development phase. Providing for the payment of equitable remuneration in the AI exploitation phase, the lawmaker avoids an unreasonable prejudice to legitimate rights holder interests that may be caused by fully trained AI systems capable of mimicking and replacing human creations. This solution follows in the footsteps of the existing regulation of private copying in the EU; Article 5(2)(b) ISD permits the mass phenomenon of unauthorized digital private copying "on condition that the right holders receive fair compensation."[163]

Admittedly, the three-step test sets forth additional assessment criteria: the tests of "certain special cases" (first criterion) and "conflict with a normal exploitation of the work" (second criterion) precede the assessment of prejudices to legitimate right holder interests (third criterion).[164] The question of an unreasonable prejudice to legitimate interests and the option of using equitable remuneration as a tool to reduce the prejudice to a permissible,

162. WIPO Intellectual Property Conference, *Report on the Work of Main Committee I*, Records of the Intellectual Property Conference of Stockholm, Vol. II, 1145–1146 (1971).

163. Directive 2001/29/EC, art. 5(2)(b), 2001 O.J. (L. 167) 10, 16.

164. *Id.*

reasonable level only arises once the analysis reaches the final, third test criterion. Considering the risk of AI-generated content replacing human creations and exposing human authors to substitution effects, one may wonder whether the use of human works for AI training violates other test criteria. Why not assuming a "conflict with a normal exploitation of the work"?[165]

The answer to this question lies in the fact that, apart from outlier cases where the user prompt triggers a result that copies individual protected expression of a specific human work, AI output only reproduces unprotected ideas, concepts, and styles. Hence, a conflict with the normal exploitation "of the work"[166] is unlikely. Instead, the harm caused is more diffuse. Unless protected individual creative choices have been copied, rights holders cannot demonstrate that the market for a specific work has been usurped. In outlier cases where this is different and an AI system enters into direct competition by reproducing substantial parts of a human work, rights holders can take individual action and enforce their rights to prevent these AI imitations from eroding the market for the affected human work.[167]

In general, however, no conflict with the normal exploitation "of the work"[168] can be found in cases where AI output only reflects general ideas, concepts, and styles. In the context of the three-step test, rights holders, thus, must resort to the final criterion of the three-step test and argue that, even though the use of their works during AI training only leads to style imitations in AI output, the use for AI development unreasonably prejudices their legitimate interests. Considering the replacement risks that have been discussed above, it can be said that authors have a legitimate interest in not being exposed to displacement effects arising from the use of their works to build a machine capable of killing the market for the type of works that enabled the machine to become a competitor in the first place. Further legitimate interest arguments may be deduced from the six arguments for author remuneration developed in section I.A. As explained, however, the payment of equitable remuneration when fully trained AI systems are brought to the market

165. *Id.*

166. Directive 2001/29/EC, art. 5(5), 2001 O.J. (L. 167) 10, 16; Berne Convention for the Protection of Literary and Artistic Works art. 9(2), *opened for signature* Sept. 9, 1886, 828 U.N.T.S. 221; Agreement on Trade-Related Aspects of Intellectual Property Rights, art. 13, Apr. 15 1994, 1869 U.N.T.S. 299; WIPO Copyright Treaty art. 10, *opened for signature* Dec. 20, 1996, T.I.A.S. No. 02-306.1, 2186 U.N.T.S. 38542.

167. *Cf.* Szkalej & Senftleben, *supra* note 45.

168. Directive 2001/29/EC, art. 5(5), 2001 O.J. (L. 167) 10, 16; Berne Convention for the Protection of Literary and Artistic Works art. 9(2), *opened for signature* Sept. 9, 1886, 828 U.N.T.S. 221; Agreement on Trade-Related Aspects of Intellectual Property Rights, art. 13, Apr. 15 1994, 1869 U.N.T.S. 299; WIPO Copyright Treaty art. 10, *opened for signature* Dec. 20, 1996, T.I.A.S. No. 02-306.1, 2186 U.N.T.S. 38542.

constitutes a factor capable of tipping the scales in favor of a finding of compliance with the three-step test: providing for the payment of equitable remuneration, the lawmaker reduces potential unreasonable prejudices to a permissible, reasonable level. As no conflict with the normal exploitation "of the work" can be found, this lumpsum remuneration approach is sufficient to dispel concerns about incompliance with the three-step test.

CONCLUSION

By implementing output-based remuneration systems, lawmakers can establish a legal framework that supports the development of unbiased, high-quality AI models while, at the same time, ensuring that authors receive a fair remuneration for the use of literary and artistic works for AI training purposes—a fair remuneration that softens displacement effects in the market for literary and artistic creations where human authors face shrinking market share and loss of income. Instead of imposing payment obligations and administrative burdens on AI developers during the AI training phase, output-based remuneration systems offer the chance to give AI trainers far-reaching freedom. Without exposing AI developers to heavy administrative and financial burdens, lawmakers can permit the use of the full spectrum of human literary and artistic resources. Once fully developed AI systems are brought to the market, however, providers of these systems are obliged to compensate authors for the unbridled freedom to use human creations during the AI training phase and displacement effects caused by AI systems capable of mimicking human literary and artistic works.

As the analysis has shown, the input-based remuneration approach in the EU—with rights reservations and complex transparency rules blocking access to AI training resources—is likely to reduce the attractiveness of the EU as a region for AI development. Considering the legal and practical difficulties resulting from the EU approach, lawmakers in other regions should refrain from following the EU model. As an alternative, they should explore output-based remuneration mechanisms.[169]

In contrast to the burdensome EU system that requires the payment of remuneration for access to human works as AI training resources, an output-based approach does not weaken the position of the domestic high-tech sector because AI developers are free to use human creations as training material. Once fully developed AI systems are offered in the marketplace, all

169. As to the chances of U.S. policy makers considering this alternative approach, see Samuelson, *supra* note 6, at 81, who predicts that, "[a]s challenging as collective license design issues may be, the political reality is that legislation to adopt such a regime is unlikely, at least in the U.S."

providers of AI systems capable of producing literary and artistic output are subject to the same payment obligation and remuneration scheme regardless of whether they are local or foreign companies. The advantages of this alternative approach are evident. Offering broad freedom to use human creations for AI training, an output-based approach is conducive to AI development.

In practice, collective rights management can play a central role in output-oriented lumpsum remuneration systems. On the basis of jointly established repartitioning schemes or statutory rules for revenue sharing,[170] collecting societies can distribute AI revenue appropriately between individual creators and industry rights holders. While input-based rights clearance in the EU will likely benefit industry rights holders primarily, an output-based approach can ensure that a substantial share of collected money reaches individual authors directly. Depending on the legislation and statutes governing a country's collecting societies, it may also be possible to use a part of the remuneration to finance social and cultural funds that support human literary and artistic projects and activities.

170. As to the option of enacting legislation to set the terms for remuneration sharing, *see* World Intellectual Property Organization Performances and Phonograms Treaty art. 15(2), *opened for signature* Dec. 20, 1996, T.I.A.S. No. 02-306.1, 2186 U.N.T.S. 38543.