ARTICLE

Infringing AI: Liability for AI-Generated Outputs under International, EU, and UK Copyright Law

Eleonora Rosati

Law Department, Stockholm University, Stockholm, Sweden; Queen Mary University of London, CCLS, London, UK; Glion Institute of Higher Education (GIHE), Glion/Bulle, Switzerland; Católica Global School of Law, Lisbon, Portugal
Email: eleonora@elawnora.com

Abstract

The analysis of liability aspects facing Artificial Intelligence ('AI')-generated outputs under copyright and related rights has been overlooked compared to other issues connected to the development and use of AI. This study fills this gap by exploring pertinent questions under international, EU and UK law. Specifically, the study tackles actionable reproduction, allocation of liability, and availability of defences. The analysis ultimately shows that, while it is clear that each case will need to be decided on its own merits, the generative AI output phase raises several profiles of liability under copyright law. If the goal of policymakers and relevant stakeholders is to ensure the balanced and sustainable development of AI, then the issues related to the generation and dissemination of AI outputs need to be given ample attention and a greater role in the debate than what has been the case so far, whether it is in the context of risk assessment and compliance, licensing initiatives, or in contentious scenarios.

Keywords: AI act; artificial intelligence; copyright; liability

I. Executive summary

Empirical research shows that large generative AI models may memorize training data which may include or consist of copyright-protected works and other protected subject-matter, or parts thereof. When prompted appropriately, these models may produce outputs that closely resemble such works and other subject-matter. Under what conditions may the resemblance between such pre-existing works or subject-matter and an AI-generated output, which in technical terms is referred to as "plagiaristic output," be regarded as an actionable reproduction? Who would be prima facie liable for the doing of such acts of reproduction: would it be solely the user of the AI model inputting the prompt resulting in the infringing output or could it be that the developer and provider are also to be deemed liable? If prima facie liability is established, who may benefit from exceptions and limitations ("**exceptions**") under copyright/related rights, and under what conditions?

By also considering the requirement that a fair balance is struck between protection of copyright and related rights, freedom to conduct a business and innovate, and freedom of expression and information alike, this study starts by mapping relevant acts under copyright and related rights in the transition from input/training to output generation. It then considers actionable reproduction, allocation of liability, and potential defences to

prima facie infringement and beneficiaries thereof under EU and UK law. Key findings include:

- **Input/training phase and TDM exceptions**: Exceptions for text and data mining ("**TDM**") under EU and UK laws allow, at certain conditions, the extraction and reproduction for TDM purposes, not subsequent restricted acts, e.g. reproduction and/or communication/making available to the public through output generation. Furthermore, Article 53(1)(c) and recital 106 of Regulation 2024/1689[1] ("**AI Act**") indicate that TDM is not an end in itself, but rather a step in the development and offering of AI models. The EU AI Act requires providers of general-purpose AI models to put in place a policy to comply with EU copyright and related rights rules, including Article 4(3) of Directive 2019/790 ("**DSM Directive**"). Subject to transitional periods, such an obligation also applies to (i) new versions/releases of general-purpose AI models placed on the EU market prior to the entry into force of the AI Act and, by operation of Article 111(3) of the AI Act generally, to (ii) general-purpose AI models put in the market in the EU twelve months before the entry into force of the AI Act.
- **Infringement through output generation**: The test for actionable reproduction differs between copyright and related rights because of the different rationale of protection. The taking of a sufficiently original part of a copyright-protected work is actionable, while the taking of any part of subject-matter protected by related rights or, alternatively, the taking of a part that reflects the investment of the relevant rightholder is actionable.
- **Liability**: Liability for infringing outputs may extend beyond users of AI models to developers and providers of such models, based on factual considerations. This is supported by case law of UK courts and the Court of Justice of the European Union ("**CJEU**"). AI developers and providers could be held liable as secondary infringers, accessories/joint tortfeasors, or even as primary infringers. A finding of primary/ direct liability may be also foreseeable having regard to CJEU case law on internet platform operators. This case law also suggests that the contractual limitation/ exclusion of liability of, e.g. AI model providers in relation to infringing activities performed by users of their services may turn out to be ineffective vis-á-vis rightholders in some instances, with the result that liability could be found to subsist alongside users of such models.
- **Defences**: For an unlicensed act to fall within the scope of application of a given exception, relevant conditions thereunder must be satisfied, including having regard to the three-step test and, insofar as the UK is concerned, the requirement that the dealing at hand is fair. The use and the volume of the use must be justified in light of its purpose and having regard to its effects on the market for the original work or protected subject-matter.

While it is clear that each case will need to be assessed on its own merits and facts and that no sweeping conclusions should be drawn, this study shows that the generative AI output phase raises several issues under copyright and related rights. If AI is to develop in a sustainable manner and in compliance with the aforementioned fair balance principle, then relevant issues related to this phase deserve more careful scrutiny, whether it is in the context of risk assessment and compliance, licensing initiatives, or in contentious scenarios.

---

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024.

II. Introduction

One of the best-known frames of Todd Phillips's 2019 *Joker* film starring Joaquin Phoenix is that of Phoenix's "Joker" inside a lift.[2] Let's imagine a situation in which the user of a generative Artificial Intelligence ("**AI**") model inputted the following prompt: "Create an image of Joaquin Phoenix Joker movie, 2019, screenshot from a movie, movie scene." As discussed in detail in a recent *New York Times* article[3], the output could look like shown in Figure 1.

This situation is neither unique nor unprecedented.[4] If, by using another AI image generator tool, a request was made to provide "a video game plumber in the style of Mario" or "an image of Dua Lipa," the results could be those shown in Figures 2 and 3.

In all the examples above, the AI-generated outputs undeniably *resemble* the appearance of characters "Joker," as played by Phoenix, and Nintendo's "Mario" from the *Super Mario* franchise, as well as the likeness of singer-songwriter Dua Lipa.

Empirical research suggests that generative AI models may sometimes memorize training data and, when appropriately prompted, reproduce such data verbatim or nearly verbatim.[5] While the first two examples above may be also indicative of what has been referred to in literature as a "Snoopy problem,"[6] "Italian plumber problem"[7] or "Pikachu Paradox,"[8] claimants in generative AI lawsuits in both the US and the UK have been submitting evidence, which the reader is referred to in order to appreciate technical aspects of generative AI,[9] of alleged verbatim/semi-verbatim regurgitation of input data by generative AI models beyond issues of character copyright.[10]

Furthermore, the generation of an output that closely resembles an actual person or their personal attributes, e.g. voice, demonstrates how legal questions extend beyond copyright and related rights (as well as trade marks) into areas like personality and image/publicity rights.[11] Issues connected to deepfakes and AI clones have generated several

---

[2] See https://www.youtube.com/watch?v=9fjpdDxZZcE&ab_channel=WarnerBrosPortugal.

[3] SA Thompson, "We asked A.I. to create the Joker. It generated a copyrighted image." (25 January 2024) The New York Times, available at https://www.nytimes.com/interactive/2024/01/25/business/ai-image-generators-openai-microsoft-midjourney-copyright.html.

[4] See A Susaria, "Generative AI could leave users holding the bag for copyright violations" (22 March 2024) The Conversation, available at https://theconversation.com/generative-ai-could-leave-users-holding-the-bag-for-co pyright-violations-225760.

[5] See *inter alia* N Carlini and Others, "Quantifying memorization across neural language models" (6 May 2023), available at https://arxiv.org/abs/2202.07646; and S Biderman and Others, "Emergent and predictable memorization in large language models" (31 May 2023), available at https://arxiv.org/abs/2304.11158. See, critically, A Feder Cooper – J Grimmelmann, "The files are in the computer: On copyright, memorization, and generative AI" (June 2024) ChiKentLRev (forthcoming).

[6] M Sag, "Copyright safety for generative AI?" (2023) 61 Hous L Rev 295, p 330–332.

[7] TB Lee – J Grimmelmann, "Why The New York Times might win its copyright lawsuit against OpenAI" (20 February 2024) Understanding AI, available at https://www.understandingai.org/p/the-ai-community-needs-to-take-copyright.

[8] A Guadamuz, "Snoopy, Mario, Pikachu, and reproduction in generative AI" (18 August 2024) TechnoLlama, available at https://www.technollama.co.uk/snoopy-mario-pikachu-and-reproduction-in-generative-ai.

[9] A continuously updated repository of US generative AI copyright lawsuits is run and maintained by students at George Washington University: see https://blogs.gwu.edu/law-eti/ai-litigation-database/. For a discussion of how AI models work, see D Gervais and Others, 'The heart of the matter: Copyright, AI trainings and LLMs' (2024) J Copyright Soc'y USA (forthcoming), §II.

[10] M Sag, "Copyright safety for generative AI?" (2023) 61 Hous L Rev 295, p 336, explaining that "LLMs will face this problem whenever the training data includes multiple versions of simple copyrightable images tightly associated with labels, regardless of whether those images are copyrightable characters or not."

[11] From a US perspective, see United States Copyright Office, Copyright and Artificial Intelligence – Part 1: Digital Replicas (July 2024), available at https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf, concerning unauthorized digital replicas.

Figure 1. Source: The New York Times.

Figure 2. Source: https://askaichat.app as of 31 May 2024.

headlines over the past few months alone.[12] Therefore, it is unlikely that a statement like "the circumstances in which the output is similar to any given input will be rare"[13] may be regarded as generally accurate.

All this raises a crucial question: Could the similarity seen in the examples above, often referred to as "plagiaristic outputs," be regarded in legal terms as actionable reproduction under copyright and related rights? If so, who would be prima facie liable for the doing of such acts of reproduction: would it be solely the user of the AI model inputting the prompt

---

[12] Examples may be made of AI clones of musicians like Drake, The Weeknd, and Tupac Shakur (see, e.g., J Coscarelli, "An A.I. hit of fake 'Drake' and 'The Weeknd' rattles the music world" (19 April 2024) The New York Times, available at https://www.nytimes.com/2023/04/19/arts/music/ai-drake-the-weeknd-fake.html; B Donahue, "Tupac Shakur's estate threatens to sue Drake over diss track featuring AI-generated Tupac voice" (24 April 2024) Billboard, available at https://www.billboard.com/pro/tupac-shakur-estate-drake-diss-track-ai-ge nerated-voice/) and actors like Scarlett Johansson (see, e.g. E Shanfeld, "Scarlett Johansson takes legal action against AI app that ripped off her likeness in advertisement" (1 November 2023), Variety, available at https://va riety.com/2023/digital/news/scarlett-johansson-legal-action-ai-app-ad-likeness-1235773489/). See also the discussion in A Baris, "AI covers: legal notes on audio mining and voice cloning" (2024) 19(7) JIPLP 571.

[13] MA Lemley, "How generative AI turns copyright upside down" (2024) 25 Colum Sci & tech L Rev, 21, p 23.
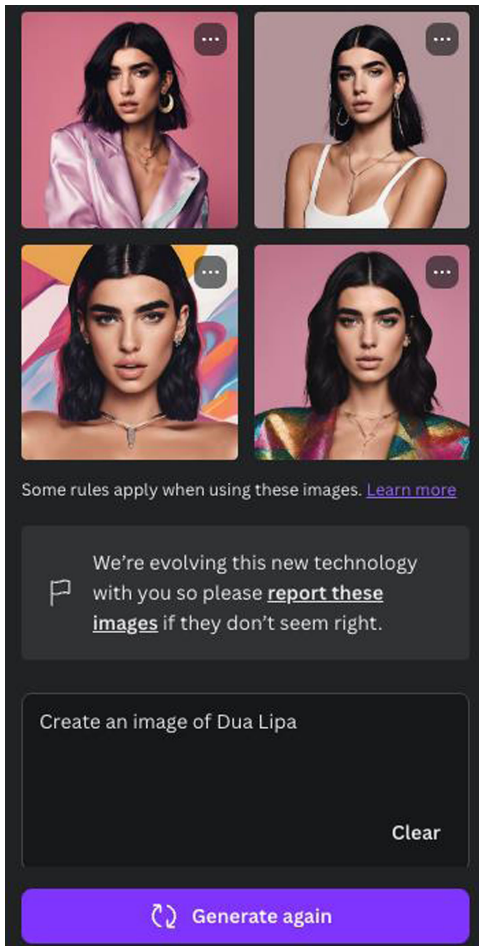
Figure 3. Source: canva.com as of 31 August 2024.

resulting in the prima facie infringing output or could it be that the developer and provider are also to be deemed liable? If prima facie liability is established, who would be the beneficiary of exception, and under what conditions?

So far, the discussions surrounding AI and copyright/related rights have mainly focused on the training/input phase and issues such as the lawfulness of unlicensed TDM. Insofar as the output phase is concerned, the discourse has mostly revolved around protectability considerations. In turn, the analysis of liability aspects connected to AI-generated outputs has remained largely underdeveloped.[14]

This study intends to fill this gap and explore pertinent aspects arising from AI-generated outputs that reproduce third-party protected content. The analysis is limited to copyright and related rights, although it is evident – as seen above – that relevant legal issues exceed them. Furthermore, the present work takes an international, EU and UK perspective. While other jurisdictions could have also been considered – notably the US,

---

[14] While literature is also emerging, in judicial practice a notable exception is the recent decision of the Guangzhou Internet Court (China) in the *Ultraman* case: see https://m.21jingji.com/article/20240226/herald/133a 6c2f9c0b045899e4dea10c5778eb.html. Also acknowledging the relevance and yet under-researched character of this area, see JP Quintais, "Generative AI, copyright and the AI Act" (1 August 2024), available https://ssrn.com/ abstract=4912701, §4.1.

Singapore and Japan given their relevance to the AI development discourse – the analysis does not specifically encompass those legal systems given the substantial differences with EU/UK copyright laws, notably having regard to defences to copyright infringement. Insofar as the US is concerned, there are nevertheless references to fair use case law in what follows, as appropriate. As regards jurisdictions like Singapore and Japan, with regard to TDM-specific exceptions relevant to the input/training phase, the reader is referred to earlier work of the same author of the present contribution.[15]

The preambles to the WIPO Internet Treaties emphasize the need to fairly balance, on the one hand, the rights of authors, performers and producers of phonograms and, on the other hand, the larger public interest, particularly education, research and access to information. Similarly, the fair balance principle is referred in multiple EU copyright instruments[16] and the Court of Justice of the European Union ("**CJEU**") has employed it as a material standard of interpretation in its case law.[17] Building on this foundational requirement in copyright law, this study is structured as follows:

- **Part III** considers copyright and related rights issues in AI development processes. Section 1 maps relevant acts in the transition from unlicensed TDM/AI training to output generation. It also critically discusses and rejects the argument advanced by some commentators that no expression would be copied in the input/training phase and, thus, no reproduction aspects would come into consideration in either that or the output generation phase. Section 2 analyzes Article 53(1)(c) of the recently adopted AI Act, read in combination with recital 106 in the preamble thereof. The provision mandates that general-purpose AI models trained outside of the EU must comply with EU law, including Article 4(3) of the DSM Directive, if made available for use there. Issues of temporal and territorial application of the provision are also tackled in this section.
- Turning to the test for actionable reproduction, **Part IV** shows that a distinction must be drawn between the authorial right of reproduction (Section 1), which is premised on originality, and the right of reproduction of other rightholders (Section 2), which is instead rooted within an investment logic. While the latter has traditionally been viewed as having no threshold conditions, the former requires a qualitative assessment of what has been taken. This is also consistent with pre-*Infopaq* case law of UK courts,[18] as well as more recent judgments. In any event, it is clear that no reproduction is actionable without proof of derivation.
- **Part V** discusses allocation of liability for infringing outputs. While any such determinations will depend on specific circumstances and contractual relationships, existing case law on platform liability in the EU and UK suggests that liability for the use of AI models may not always solely rest with the user but could instead extend to developers and providers of such models (Section 1). In turn, terms of service of AI model providers that seek to remove any liability potentially arising from the use of their models might turn out to be ineffective towards third-party rightholders or users of their services (Section 2).

---

[15] E Rosati, "No step-free copyright exceptions: the role of the three-step in defining permitted uses of protected content (including TDM for AI-training purposes)" (2024) 46(5) EIPR 262, p. 270–71.

[16] See inter alia recitals 31 of Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, OJ L 167, 22.6.2001, p. 10–19 ("**InfoSoc Directive**") and 6 of Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC, OJ L 130, 17.5.2019, p. 92–125 ("**DSM Directive**").

[17] See the discussion in E Rosati, *Copyright and the Court of Justice of the European Union* (Oxford University Press 2023), 2nd edn, p 77–85.

[18] *Infopaq*, C-5/08, EU:C:2009:465.

- Defences to *prima facie* infringement are subsequently reviewed in **Part VI**. CJEU case law indicates that exceptions must be construed so as to ensure a fair balance of rights and interests and compliance with the three-step test. Furthermore, exceptions would be likely available to users of AI models, rather than providers thereof (Section 1). In any event, a justification for the use made of a third-party work/protected subject-matter without authorization is required, as is the consideration of the amount taken and the effects on the market for the original work/protected subject-matter. This is exemplified by reference to quotation and pastiche (Section 2).

On a final, preliminary note, a couple of observations are warranted given the terminology adopted in the AI Act. First, the more general term "user" is used instead of "deployer," given the definition of the latter in the AI Act (recital 13 and Article 3(4)), notably the circumstance that the notion of "deployer" there excludes users in the course of a personal and non-professional activity. Those are instead considered in the present contribution. Second, while the notion of "developer" encompasses that of "provider" in the AI Act (Article 3(3)), the two roles are not conflated here, as it is foreseen that it might not always be the case that providers of AI models have also developed them and vice versa.[19]

III. From AI training to output generation

When the "data" consists of protected works and/or subject-matter, or parts thereof, training of generative AI models is premised on the doing of acts restricted by copyright and related rights. TDM activities, defined as "any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations" (Article 2 No 2 of the DSM Directive), are an essential step in any AI development project. Considering the legislative history of Articles 3 and 4 of the DSM Directive and the rapid developments in the field of generative AI since the adoption of that piece of legislation, it has been discussed whether these provisions were intended to cover activities related to the development of generative AI.[20] The AI Act now appears to have settled this issue for good by linking the DSM Directive's TDM exceptions to the development of general purpose AI models, including generative AI models.[21] Importantly, the AI Act recognizes the relevance of TDM to AI training, but in no way does it indicate that TDM is synonymous with AI training or

---

[19] On the circumstance that the AI Act flattens the notion of "developer" onto that of "provider," see critically A Peukert, "Copyright in the Artificial Intelligence Act – A primer" (2024) 73(6) GRUR Int 497, p 501.

[20] On the history of Article 4 of the DSM Directive, see further E Rosati, *Copyright in the Digital Single Market. Article-by-Article Commentary to the Provisions of Directive 2019/790* (Oxford University Press 2021), p 63–68. See also: C Villani and Others, *Donner un Sens à l'Intelligence Artificielle: Pour une Stratégie Nationale et Européenne* (8 September 2017 – 8 March 2018), available at https://www.vie-publique.fr/files/rapport/pdf/184000159.pdf, p 35–36, and S Dusollier, 'The 2019 Directive on Copyright in the Digital Single Market: Some progress, a few bad choices, and an overall failed ambition' (2020) 57 CMLRev 979, p 984, further showing how at that time the link between provisions allowing unlicensed TDM and AI development was already clearly understood. Recently, see GB Abbamonte, "The application of the copyright TDM exceptions and transparency requirements in the AI Act to the training of generative AI" (2024) 46(7) EIPR 479, p 479, and H Hamann, 'Artificial Intelligence and the law of machine-readability: A review of human-to-machine communication protocols and their (in)compatibility with Article 4(3) of the Copyright DSM Directive' (2024) 15(2) JIPITEC 102, p 105-106. Nevertheless, according to some social media reports, the Acting Head of the Copyright Unit at the European Commission, Emmanuelle Du Chalard, recently spoke publicly of how the TDM exceptions in the DSM Directive were drafted with AI in mind, but not generative AI: see https://www.linkedin.com/feed/update/urn:li:activity:7202919364067479552/.

[21] In substantially the same sense, see also M Senftleben, "AI Act and author remuneration – A model for other regions?" (24 February 2024), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4740268, p 7–9, and A Peukert, "Copyright in the Artificial Intelligence Act – A primer" (2024) 73(6) GRUR Int 497, p 503.

that everything in-between TDM and AI training is covered by Articles 3 or 4 of the DSM Directive.

The above said, in the context of generative AI, a distinction needs to be drawn between the input/training and the output phase, with latter entailing the generation of content – whether it is text, audio, image, or video – in response to instructions (prompts) given by users of resulting AI models. What follows is a review of relevant acts under copyright and related rights with regard to these two key phases. After that, a discussion is provided of Article 53(1)(c) of the AI Act, which requires compliance with Article 4 of the DSM Directive – including the need to identify and comply with the reservations of rights by rightholders under paragraph 3 – for providers looking to place general-purpose AI models on the EU market, "regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of those general-purpose AI models take place."

1. Mapping of restricted acts under copyright and related rights

As discussed in greater detail elsewhere,[22] during the TDM activities the main acts under UK and EU law are those of extraction (e.g. if the data is incorporated in a database) and reproduction. Insofar as the latter is concerned, the claim that copyright would not restrict non-expressive uses of protected content appears erroneous.[23] Several considerations support this conclusion, with two in particular worth highlighting: the broad construction of the right of reproduction and the fact that specific exceptions to the right of reproduction to allow (at certain conditions) TDM have been adopted in multiple jurisdictions over the past several years.

Starting with the scope of the right of reproduction (whether of authors or other rightholders), this must receive a broad construction and interpretation, aligning with relevant statutes. At the international level, Article 9(1) of the Berne Convention mandates the protection of reproduction "in any manner or form." The Agreed Statements to the WIPO Internet Treaties further clarify that the right "fully appl[ies] in the digital environment."[24] Article 2 of the InfoSoc Directive, which was adopted to implement the WIPO Internet Treaties (recitals 15 and 61) into the EU legal order, states that the right encompasses "the direct or indirect, temporary or permanent reproduction by any means and in any form, in whole or in part" of works and other protected subject-matter.

Consistent with the requirement of a high level of protection (recital 9–11), CJEU case law also indicates that the concept of "reproduction" must be construed broadly. This is because (i) recital 21 links a broad understanding of "reproduction" to legal certainty, and (ii) the already mentioned wording of Article 2 uses expressions such as "direct or indirect," "temporary or permanent," "by any means" and "in any form."[25] While, as it is discussed in Part IV, a distinction appears warranted as regards the test of actionable reproduction under copyright vis-à-vis related rights, in both cases the right of reproduction is engaged regardless of the end goal of the act in question.

Turning to TDM-specific exceptions adopted in some jurisdictions, if no copyright-relevant act was undertaken, then the provisions introduced by legislatures over the past

---

[22] E Rosati, *The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Technical Aspects*, available at http://www.europarl.europa.eu/RegData/etudes/BRIE/2018/604942/IPOL_BRI(2018)604942_EN.pdf.

[23] See, amongst others and from a US perspective, M Sag, "The new landscape for text mining and machine learning" (2019) 66 J COPYRIGHT Soc'y USA 291; MW Carroll, "Copyright and the progress of science: Why text and data mining is lawful" (2019) 59 UC Davis Law Rev 893; MA Lemley – B Casey, "Fair learning" (2021) 99 Tex L Rev 743.

[24] WIPO Copyright Treaty, Agreed Statements concerning Article 1(4); WIPO Performances and Phonograms Treaty, Agreed statement concerning Articles 7, 11 and 16.

[25] Recently, *Austro-Mechana*, C-433/20, EU:C:2022:217, paras 16–18.

several years – ranging from Japan (Article 30-4 of the Copyright Act) to Singapore (Section 244 of the Copyright Act 2021), from the UK (section 29A of the Copyright, Designs and Patents Act ("**CDPA**")) to other EU Member States prior to the adoption of the DSM Directive[26] and the DSM Directive itself – would have been unnecessary, if not altogether misleading. The fact that legislatures have deemed such exceptions necessary indicates that TDM activities entail the doing of copyright-relevant acts.[27]

The same conclusion applies in jurisdictions, like the US, where no specific exceptions for TDM exist. Even the decision of the 2nd Circuit in *Google Books*, which is often referenced in support of the view that unlicensed TDM could generally fall within the scope of application of the fair use doctrine under Section 107 of the US Copyright Act,[28] in considering the possibility to mine the Google Books corpus as favouring a finding of fair use, does not dispute that all this was premised on the prior scanning – that is: reproduction – by Google of third-party books and, thus, a prima facie infringing activity.[29] Recent case law, notably the US Supreme Court ruling in *Andy Warhol Foundation,*[30] confirms that the analysis of whether the unlicensed use of protected content is fair use also substantially depends on whether there is a market for licensing content for training data. In light of emerging licensing practices and models, it is thus doubtful that fair use shall be broadly applicable to AI training on protected content.[31]

A further important point to note is that the scope of the TDM exceptions (including in the EU and the UK) is *limited* to *specified* restricted acts, as shown in Table 1. The subsequent undertaking of, e.g. acts of reproduction and communication/making available to the public, including but not necessarily limited to output generation, is thus not covered by either section 29A CDPA or Articles 3 and 4 of the DSM Directive. To exemplify: if an AI model lawfully trained on the repertoire of a certain Italian singer-songwriter under the Italian transposition of Article 4 of the DSM Directive (and the lack of appropriate reservation on the side of relevant rightholders[32]) allowed a user of that model to reproduce fully or partly a song of that artist, such an act of reproduction would not fall within the scope of (the Italian transposition of) Article 4 of the DSM Directive.

2. Compliance with EU law regardless of the jurisdiction in which the training is performed

Before considering actionable reproduction in the context of AI-generated outputs, it is worth examining the requirements under Article 53(1)(c) of the AI Act.[33] The provision

---

[26] See further E Rosati, "No step-free copyright exceptions: the role of the three-step in defining permitted uses of protected content (including TDM for AI-training purposes)" (2024) 46(5) EIPR 262, p. 269–70.

[27] See critically M Senftleben, "Generative AI and author remuneration" (2023) 54(10) IIC 1535, 1542–43, and literature referenced there.

[28] See, e.g. M Sag, "Copyright safety for generative AI? (2023) 61 Hous L Rev 295, p 303–307, and literature referenced there.

[29] *Authors Guild v. Google, Inc*., No. 13-4829-cv (2d Cir. 16 October 2015). The same is true of *Authors Guild v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014), also often relied in support of the general lawfulness of unlicensed TDM under US law.

[30] *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508.

[31] In this sense, JC Ginsburg, "Fair use in the US redux: Reformed or still deformed?" (March 2024 Online) Sing JLS 1, available at https://law.nus.edu.sg/sjls/wp-content/uploads/sites/14/2024/05/firstviewmarch24-JaneGinsburg.pdf, p 22–23.

[32] Recent research suggests that rights reservations are becoming increasingly frequent for online content, though there are also substantial differences in the extent of restrictions: see the discussion in S Longpre, "Consent in crisis: The rapid decline of the AI data commons" (24 July 2024), available at https://arxiv.org/abs/2407.14933.

[33] The AI Act also imposes an obligation to publicly disclose a sufficiently detailed summary of the content used for training the general-purpose AI model: see recital 107 and Article 53(1)(d).

Table 1 TDM E&Ls under UK and EU law

Jurisdiction	Provision	Restrictions on beneficiaries?	Acts covered	Does the provision extend to output generation?	Purpose-specific	Express limitation to non-commercial use	Express requirement of lawful access	Express possibility of reservation (opt-out)	Reference to the three-step test
UK	Sec. 29A CDPA	No	Reproduction	No	Yes (research)	Yes	Yes	No	No
EU	Art. 3 DSM Directive	Yes (research organizations and cultural heritage institutions)	Extraction and reproduction	No	Yes (scientific research)	No	Yes	No	Yes (Art. 7 DSM Directive)
EU	Art. 4 DSM Directive	No	Extraction and reproduction	No	No	No	Yes	Yes	Yes (Art. 7 DSM Directive)

mandates compliance with inter alia Articles 3 and 4 of the DSM Directive, irrespective of where the acts of extraction and reproduction for TDM purposes have taken place, if the resulting AI model is made available in the EU. Recital 106 further clarifies that this is necessary in order to create a level playing field among AI providers. The AI Act also imposes an obligation to publicly disclose a sufficiently detailed summary of the content used for training the general-purpose AI model (recital 107 and Article 53(1)(d)). Two aspects of Article 53(1)(c) warrant closer examination: its temporal application and territorial reach.

Starting with the temporal application of the provision: Article 2(8) of the AI Act provides that the Regulation will inter alia not apply to development activities "regarding AI systems or models prior to their being placed on the market or put into service." Thus, for systems or models trained lawfully outside of the EU, there is no requirement to comply with Article 4 of the DSM Directive until (if ever) the resulting systems or models are placed on the EU market. A question then arises whether the obligations under Article 53 of the AI Act, including the one at (c), apply to AI systems and models (lawfully) trained outside the EU and made available in the EU *before* the entry into force of the AI Act. Although some commentators have ruled out that this could be the case based on the general EU principle of non-retroactivity,[34] it appears more appropriate to qualify the answer.

The principle of non-retroactivity is expression of legal certainty and the protection of legitimate expectations. That said, as also recently recalled by the Grand Chamber of the CJEU, "according to *settled case-law*, a new rule of law applies from the entry into force of the act introducing it, and, while it does not apply to legal situations that have arisen and become final under the old law, it does apply to their *future effects*, and to new legal situations."[35] It follows that it would not be a breach of the principle of non-retroactivity to extend the application of rules, including those under Article 53(1) of the AI Act, to future effects of situations (such as the placing of an AI model on the EU market) that arose before the entry into force of the AI Act. To exemplify: if an AI model was updated/upgraded following its "first" entry on the EU market prior to the AI Act becoming enforceable, then the AI Act could apply to the effects resulting from such subsequent updates/upgrades. That is so because those situations could not be considered as having become definitive before that date. Such a conclusion is also in line with the Regulation itself. Articles 111(3) and 113(b) indicate that the obligation on general-purpose AI models to keep and disclose sufficiently detailed summaries of the content used for training will apply from 2 August 2025 for new models and from 2 August 2027 for models placed on the market or put into service before 2 August 2024.

Turning to the territorial reach of the provision: Article 53(1)(c) of the AI Act implies that it is irrelevant *where* the acts of extraction and reproduction at the basis of the TDM processes described above take place and whether such acts are to be deemed lawful under the laws of those jurisdictions. *If* the resulting, general-purpose AI model is made available in the EU, all the copyright-relevant steps leading to such a situation must comply with EU law. Conversely, if no AI system or model is placed on the EU market, relevant TDM processes undertaken outside of the EU shall not be affected by EU rules. While such an approach raises concerns from the perspective of general principles, including sovereignty and international comity, as well as potentially the principle of territoriality under copyright law,[36] it is not unprecedented, including having inter alia regard to international and EU copyright law.

---

[34] A Peukert, "Copyright in the Artificial Intelligence Act – A primer" (2024) 73(6) GRUR Int 497, p 505.

[35] *VD*, C-339/20 and *SR*, C-397/20, EU:C:2022:703, para 62 (emphasis added). See also the discussion and case law cited in Opinion of Advocate General Kokott in *VYSOČINA WIND a.s.*, C-181/20, EU:C:2021:619, paras 42–46.

[36] Specifically regarding territoriality of copyright, see JP Quintais, "Generative AI, copyright and the AI Act" (1 August 2024), available https://ssrn.com/abstract=4912701 §4.4. See critically, M Rättzén, "Where to train your

At the international level, for example, Article 2 of the Phonograms Convention requires contracting parties to give rightholders the right to prohibit the importation of copies of recordings made "without the consent" of rightholders. The obligation applies even if the act of copying would not have required rightholders consent in the first place in the country of exportation. Turning to the Berne Convention, it is commonly[37] understood that Article 5(2) therein adopts a lex loci protectionis (law of the place of protection) approach to the localization of copyright-relevant acts.[38] At the EU level, the determination of the law applicable to copyright infringements is made through the lex generalis found in Regulation 864/2007[39] (**"Rome II"**). Like the Berne Convention, Article 8(1) Rome II adopts a lex loci protectionis approach to determine the applicable law in cross-border infringements.[40]

Over time, courts around the world – including in the EU and the UK – have considered themselves competent to rule, under local laws, in disputes concerning situations consisting of multiple infringing acts committed in different jurisdictions, that is: in decentralized situations. As discussed in greater detail elsewhere,[41] various criteria have been adopted to this end, including accessibility (the court seized is competent if it is located in a country from which the allegedly infringing content is accessible[42]) and targeting (courts located in the territory at which the allegedly infringing content is targeted have jurisdiction). The latter in particular has gained traction as a criterion to determine applicable law and jurisdiction in several legal systems, including in the UK and at the level of the CJEU.

Insofar as the UK is concerned, case law indicates that, where a certain act originates outside the UK territory but its effects are felt in the UK, then such act shall be deemed to be targeted at the UK territory, with the result that UK law is applicable.[43] In a copyright context, this approach was approved of in *TuneIn,* where the Court of Appeal of England and Wales upheld the decision at first instance that the operators an online platform would be infringing UK law when giving access to UK customers to radio stations that had not been licensed for the UK.[44] It was also confirmed and discussed from a broader intellectual property perspective in the recent judgment of the UK Supreme Court in *Lifestyle Equities.*[45]

AI: Transnational AI development and copyright extraterritoriality" (2024) COLUM SCI & TECH L REV (forthcoming), §IV.

[37] Though not uncontroversially so: see ME Ancel and Others, "International Law Association's Guidelines on intellectual property and private international law ('Kyoto Guidelines'): Applicable law" (2021) 12(1) JIPITEC 44, p 45.

[38] See the discussion in S Ricketson – JC Ginsburg, *International Copyright and Neighbouring Rights. The Berne Convention and Beyond* (OUP:2022), 3rd edn, §§20.02-20.03.

[39] Regulation (EC) No 864/2007 of the European Parliament and of the Council of 11 July 2007 on the law applicable to non-contractual obligations, OJ L 199, 31 July 2007, p 40–49.

[40] Under EU law, the localization of the infringing activity may *also* serve to establish judicial competence through the rule found in Article 7(2) of Regulation 1215/2012. It is considered "settled case-law" (see, recently, *Gtflix Tv,* C-251/20, EU:C:2021:1036, para 27) that the place where the harmful event occurred or may occur can be either the place where the damage occurred or the place of the event giving rise to it, so that the defendant may be sued – at the option of the plaintiff – in the courts for either of those places.

[41] E Rosati, *The Localization of IP Infringements in the Online Environment: from Web 2.0 to Web 3.0 and the Metaverse* (30 November 2023) WIPO/ACE/16/10, available at https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=622296, paras 36–54.

[42] This was for example that approach recommended in Opinion of Advocate General Szpunar in *Grand Production,* C-423/21, EU:C:2022:818, para 25.

[43] See, with regard to UK law, *EMI Records Ltd & Ors v British Sky Broadcasting Ltd & Ors* [2013] EWHC 379 (Ch) (28 February 2013), para 38.

[44] *Tunein Inc v Warner Music UK Ltd & Anor* [2021] EWCA Civ 441 (26 March 2021).

[45] *Lifestyle Equities CV v Amazon UK Services* [2024] UKSC 8.

Turning to the CJEU, in the copyright field, that court has expressly adopted a targeting approach for the sui generis (database) right[46] and the right of distribution.[47] The CJEU has also indicated that EU copyright rules apply to infringements whose effects are ultimately *felt* in the EU, irrespective of where the users' acts of reproduction and the servers are located.[48] Along the same lines, the CJEU has indicated that the focus should be on the *intended final result* of the various acts, thus adopting a functional approach. For example, in *Mircom*, C-597/19, the particular technique used in peer to peer ("**P2P**") file sharing activities, known as "seeding," was considered. This consists of the downloading of a file through a P2P network and the simultaneous provision for uploading of parts ("pieces" or "seeds") of it, which can be very fragmentary when compared to the whole. As a result, the pieces that are downloaded and re-uploaded by users are not of the copyright-protected works or other protected subject-matter themselves; instead, they are pieces of the files containing such works or subject-matter. Endorsing the analysis contained in Advocate General ("**AG**") Szpunar's Opinion and reasoning by analogy with how the World Wide Web functions,[49] the CJEU considered that the fact that the pieces are unusable per se is irrelevant. What is made available through the seeding technique – *irrespective of where* the seeds are downloaded and re-uploaded and, thus, where the various acts of reproduction and communication/making available to the public take place – is the file containing the protected content, with the result that users have ultimately access to the complete work or subject-matter. The technical means or process used to undertake an act of communication are irrelevant. If the users have consented to the automatic re-uploading of the pieces that they download, they act in full knowledge of the consequences of their actions and, as a result, perform acts restricted by Article 3 of the InfoSoc Directive. It is not required that users manually initiate the re-uploading: what matters is that they have consented to the use of the relevant software after receiving information regarding its characteristics.[50]

In sum: consistent case law of the CJEU indicates that consideration of the ultimate objective sought is necessary to characterize a given situation from a legal standpoint.[51] It follows from all this that, if the acts of extraction and reproduction undertaken during TDM processes are *functional* – as they are – to the training of AI models, which, in turn, are made available for use in the EU, the application of EU law also to those acts appears justified having regard to the broader process of which they are part.

IV. Reproduction issues through output generation

Following the analysis above, the focus shall now shift to the output generation phase: "[i]f an AI output is identical or similar to parts of an original copyrighted work included in the training data, copyright issues may arise regarding whether the original work's copyright

---

[46] *Football Dataco and Others*, C-173/11, EU:C:2012:642.

[47] *Donner*, C-5/11, EU:C:2012:370.

[48] See, e.g. *Ziggo*, C-610/15, EU:C:2017:456, and *YouTube*, C-682/18 and C-683/18, EU:C:2021:503. See further the discussion in Part V.

[49] Opinion of Advocate General Szpunar in *Mircom*, C-597/19, EU:C:2020:1063, paras 48–49.

[50] *Mircom*, C-597/19, EU:C:2021:492, para 47, referring by analogy to *VCAST*, C-265/16, EU:C:2017:913, para 42 and the case law cited there.

[51] This logic is also clear having regard to, e.g. Opinion of Advocate General Szpunar in *Gema*, C-135/23, EU:C:2024:151, paras 39–41 (substantially endorsed in the resulting judgment: *Gema*, C-135/23, EU:C:2024:526) warning that "overly formalistic reasoning could lead to an outcome that, although coherent according to the logic adopted, is unfair or just substantively incorrect." On the functional approach taken by the CJEU, see further E Rosati, *Copyright and the Court of Justice of the European Union* (OUP:2023), 2nd edn, p 165–66.

was infringed, and if so, who is responsible."[52] To address these issues the test for actionable reproduction must be therefore determined. If an AI model generated an output like those discussed in Part II – that is outputs similar (if not, highly similar) to the content used during the input/training phase – would such an output be prima facie infringing under EU or UK law? And would it matter if the training of the AI model did lawfully occur under the UK or EU TDM exceptions? While the answer to the first question might be in the affirmative, that to the second would be in the negative because, as stated, TDM exceptions only encompass the input/training phase (whether fully or only in part).

1. The right of reproduction under copyright

In accordance with international law, under EU and UK law, the right of reproduction encompasses a work in its entirety or part thereof. It applies to instances of literal and non-literal copying and, therefore, also to situations that, formally, might fall under the right of adaptation.[53] That said, copyright does not protect against independent creation. While this aspect has been indicated as having the potential to be problematic in the context of generative AI outputs,[54] recent studies suggest that methods can be devised to identify if specific content, such as text sequences, were included in certain training data.[55] Moreover, the EU AI Act transparency obligation is intended to mitigate against the evidentiary problems related to establishing access to and/or derivation from pre-existing works through the transparency obligation under Article 53(1)(d).

In *Infopaq*, C-5/08 the CJEU introduced an EU-wide test of actionable reproduction for authorial works[56]: there is reproduction in part when the part of a work that has been copied is original in the sense that it is its author's own intellectual creation.[57] This broad understanding – which extends to the copying of short extracts of a work,[58] as long as the choice, sequence and combination of elements is sufficiently original[59] – is in line with the objective of the InfoSoc Directive to introduce a high level of protection of authors.[60] Therefore, the prima facie infringement test requires to determine, first, if the claimant's work or part thereof is protected and, secondly, if the defendant took a protected work or part thereof. The appropriateness of such an approach has also been upheld by members of the UK judiciary writing extra-judicially,[61] as well as by UK courts.

---

[52] Korean Ministry of Culture, Sports and Tourism and Korea Copyright Commission, *A Guide on Generative AI and Copyright* (16 January 2024), available at https://www.copyright.or.kr/information-materials/publication/research-report/view.do?brdctsno=52591#, p 11.

[53] See further E Rosati, *Copyright and the Court of Justice of the European Union* (OUP:2023), 2nd edn, p 156–57, and literature cited there.

[54] M Lemley, "How generative AI turns copyright upside down" (2024) Colum Sci & Tech L Rev, p 34–40; JB Nordemann, "EU law: Generative AI, copyright infringements and liability – My guess for a hot topic in 2024" (23 January 2024), Kluwer Copyright Blog, available at ~https://copyrightblog.kluweriplaw.com/2024/01/23/eu-law-generative-ai-copyright-infringements-and-liability-my-guess-for-a-hot-topic-in-2024/#:~:text=AI%20output%20can%20be%20deemed,can%20be%20recognised%20in%20it; A Guadamuz, "A scanner darkly: Copyright liability and exceptions in Artificial Intelligence inputs and outputs" (2024) 73(2) GRUR Int 111, p. 123–24.

[55] See, e.g. P Maini and Others, "LLM dataset inference: Did you train on my dataset?" (10 June 2024), available at https://arxiv.org/pdf/2406.06443.

[56] In its subsequent ruling in *Football Association Premier League*, C-403/08 and C-429/08, EU:C:2011:631, para 154, the CJEU confirmed that the concept of "reproduction" must be given an autonomous and uniform interpretation throughout the EU.

[57] *Infopaq*, C-5/08, EU:C:2009:465, para 39.

[58] Ibid, paras 47 and 49–50.

[59] Ibid, para 45.

[60] Ibid, paras 40–43. See also *Nintendo and Others*, C-355/12, EU:C:2014:25, paras 21–22.

[61] R Arnold, "Paintings from photographs: a copyright conundrum" (2019) 50(7) IIC 860, 875.

To establish prima facie infringement under section 16 CDPA, a claimant needs to demonstrate that the defendant has done any of the restricted acts in relation to the work as a whole or any substantial part of it. Even before the CJEU judgment in *Infopaq*, C-5/08, the notion of "substantial taking" was not intended in a quantitative sense.[62] Put succinctly: "the quality relevant for the purposes of substantiality is the literary originality of that which has been copied."[63] As of today, the UK test remains informed by CJEU case law: when considering whether a substantial part has been reproduced, it is necessary to focus on (a) what has been reproduced and (b) whether that part expresses the author's own intellectual creation.[64] In *Sheeran*, Zacaroli J confirmed that "[t]o amount to an infringement [. . .] the copying must be of either the original work or a "substantial part" of it [. . .] This is a qualitative, not quantitative, question. The test is whether the part in question contains elements which are the expression of the intellectual creation of the author of the work [. . .] The essential consideration is to ask whether a defendant has taken that which conferred originality on the claimant's copyright work (or a substantial part of it)."[65] In *Pasternak*, Johnson J also approved of this approach and regarded the cited passage from *Sheeran* as summarizing the test of infringement in a useful and succinct fashion.[66]

2. The right of reproduction under related rights

The notion of reproduction "in part" as adopted by the CJEU in *Infopaq*, C-5/08 in relation to works in Article 2(a) of the InfoSoc Directive does not extend to the other subject-matter listed in that provision. That is because the right of reproduction holders of related rights (performers in relation to fixations of their performances; phonogram producers in relation to their phonograms; producers of the first fixations of films in respect of the original and copies of their films; broadcasting organizations in respect of fixations of their broadcasts) protects "not intellectual creation but financial investment."[67]

The traditional and broadly accepted view is that related rights are not subject to any threshold condition.[68] In turn, any reproduction of protected subject-matter would be actionable.[69] That said, in *Pelham I*, C-476/17, the CJEU adopted an admittedly odd test to determine actionable reproduction of the phonogram producer's related right vis-à-vis

---

[62] See, e.g. *Designer Guild Limited v Russell Williams (Textiles) Limited (Trading As Washington Dc)* [2000] UKHL 58; [2001] 1 All ER 700; [2000] 1 WLR 2416 (23 November, 2000)(Lord Millett); *Ladbroke v William Hill* [1964] 1 WLR 273, (Lord Reid), para 276; *Macmillan & Co Ltd v Cooper* (1923) 93 LJPC 113, (Atkinson J), para 121; *Allen v Bloomsbury Publishing Plc & Anor* [2010] EWHC 2560 (Ch) (14 October 2010), para 85.

[63] *Newspaper Licensing Agency Limited v Marks and Spencer Plc* [2001] UKHL 38 (12 July 2001), (Lord Hoffmann), para 19.

[64] *SAS Institute Inc v World Programming Ltd* [2010] EWHC 1829 (Ch) (23 July 2010), para 43.

[65] *Sheeran & Ors v Chokri & Ors* [2022] EWHC 827 (Ch) (06 April 2022), para 21.

[66] *Pasternak v Prescott* [2022] EWHC 2695 (Ch) (25 October 2022), para 107. The same was reiterated more recently in *Fay Evans v John Lewis Plc and DBB UK Ltd* [2023] EWHC 766 (IPEC), para 40.

[67] Opinion of Advocate General Szpunar in *Pelham I*, C-476/17, EU:C:2018:1002, para 28.

[68] See (albeit critically) PB Hugenholtz, "Neighbouring rights are obsolete" (2019) 50(8) IIC 1006, p. 1008-1010, and C Sganga, "The many metamorphoses of related rights in EU copyright law: Unintended consequences or inevitable developments?" (2021) 70(9) GRUR Int 821, p 827.

[69] Such an approach appears consistent with the traditional view under UK, even though the CDPA does not recognize related rights as such. Section 1(1)(a) CDPA provides that literary, dramatic, musical or artistic works are protected if they are "original." There is no mention of originality for the other subject-matter protected by the CDPA as "copyright" – that is: sound recordings, films or broadcasts, and the typographical arrangement of published editions. The explanation given for this choice is that, while literary, dramatic, musical or artistic works may be classified as contents (works that are protected regardless of the signals by which they are carried), sound recordings, films or broadcasts, and the typographical arrangement of published editions are signals (what is protected is the signal itself, as distinct from the content it carries): see R Arnold, "Content copyrights and signal copyrights: the case for a rational scheme of protection" (2011) 1(3) QMJIP 272, 276.

third-party freedom of artistic expression. The Court ruled that any reproduction of a phonogram is actionable insofar as it is *recognizable to one's ear* (the ear of whom, however, is unclear[70]) and the part reproduced *reflects the investment* made by the rightholder. This finding would be consistent with the objective of guaranteeing a high level of protection and safeguarding the specific objective of the exclusive right of the phonogram producer.[71]

Despite the *specific* context of *Pelham I*, C-476/17, some commentators have suggested that the test of recognizability could be generally applicable as a limitation to the scope of the right of reproduction, at least for rightholders other than authors.[72] This suggestion is flawed and should be rejected, including by the CJEU when it decides the referral in *Mio*, C-580/23, concerning the authorial right of reproduction.[73] There are two key reasons for this conclusion.

The first is the inherent difficulties and shortcomings of a recognizability-based approach, including the uncertainty regarding the benchmark to use to this end: should the part be recognizable to, e.g. an average person, a specialized person/expert (e.g. a musicologist), or should it be machine-detectable? Depending on the choice made, the scope of protection would vary (very) significantly. The second is because, as seen, the main – though not necessarily only (think of performers' rights[74]) – rationale of protection of related rights is one of investment. Even if one were to reject the view that related rights are not subject to any threshold condition, the need to safeguard investment as the benchmark adopted to determine if the reproduction at hand is actionable would be preferable to a broader applicability of the approach adopted by the CJEU in *Pelham I*, C-476/17. This way, there would be reproduction "in part" within the meaning of Article 2(b)-(e) of the InfoSoc Directive when what is reproduced without the permission of the relevant reproduction rightholder interferes with the opportunity, which the rightholder has, of realising meaningful returns on their investment.[75] This approach could be employed in non-InfoSoc situations too[76] and would be also aligned with the qualified infringement test adopted by the CJEU in *'CV-Online Latvia'*, C-762/19 in relation to the EU sui generis database right.[77]

---

[70] On this aspect of the judgment, see also: J Parish, "Sampling and copyright – did the CJEU make the right noises?" (2020) 79(1) CLJ 32, p 33; M Senftleben, "Flexibility grave – partial reproduction focus and closed system fetishism in CJEU, *Pelham*" (2020) 51(6) IIC 751, p 757; BJ Jütte – JP Quintais, "The *Pelham* Chronicles: sampling, copyright and fundamental rights" (2021) 16(3) JIPLP 213, p 216; G Westkamp, "Two constitutional cultures, technological enforcement and user creativity: the impending collapse of the EU copyright regime?" (2022) 53(1) IIC 62, p 73–74.

[71] *Pelham I*, C-476/17, EU:C:2019:624, paras 29–30.

[72] K Grisse, "On the significance of (un)recognisability for the reproduction right in European copyright law" (2022) 44(2) EIPR 78.

[73] In its referral, the Svea Hovrätt, Patent- och marknadsöverdomstolen (Sweden) asked the CJEU how the assessment of similarity must be carried out under Article 2(a) of the InfoSoc Directive, notably whether recognizability of the claimant's work in the allegedly infringing subject-matter is relevant.

[74] To this effect, see recital 5 of Directive 2006/115/EC of the European Parliament and of the Council of 12 December 2006 on rental right and lending right and on certain rights related to copyright in the field of intellectual property (codified version), OJ L 376, 27.12.2006, p 28–35.

[75] *Pelham I*, C-476/17, EU:C:2019:624, para 38.

[76] See further E Rosati, *Copyright in the Digital Single Market. Article-by-Article Commentary to the Provisions of Directive 2019/790* (Oxford University Press:2021), p 266–67.

[77] SIA *'CV-Online Latvia'*, C-762/19, EU:C:2021:434, para 47. There the CJEU held that not any extraction of re-utilisation without consent of a database is potentially relevant under the sui generis database right. What is required, instead, is that the extraction and/or re-utilisation at issue adversely affect(s) the investment made in the obtaining, verification or presentation of the content of a database. The test is thus whether a risk has been posed to the possibility of redeeming that investment through the normal operation of the database at issue.

V. Liability for infringing AI-generated outputs

If AI-generated outputs that reproduce third-party copyright works and/or protected subject-matter may be infringing, allocation of resulting liability needs to be considered next. Specifically: how is the traditional divide between primary/direct and secondary/indirect/accessory liability to be construed in the context of generative AI? If the terms of service of AI model providers exclude any liability resulting from uses thereof, are – and, if so, under what circumstances – such terms enforceable against users of the models and/or effective in relation to third-party rightholders whose rights may have been infringed?

Both aspects will be tackled in what follows. Insofar as the first issue is concerned, drawing on case law of the CJEU and UK courts, it will be shown how not only users of AI models, but also developers and providers thereof might be considered liable for the doing of acts restricted by copyright and/or related rights through output generation. As a result, turning to the second issue, at least in certain circumstances, limitations of liability by AI developers and providers vis-à-vis users could be deemed unenforceable against the latter, and ineffective vis-à-vis third parties whose rights have been infringed.

1. Liability for the provision of AI models that could be used to infringe third-party rights

It is clear by now that AI-generated outputs may raise questions of actionable reproduction under copyright and related rights. The user inputting the prompt resulting in prima facie infringing output (and subsequently using that output) might be regarded as the one directly undertaking restricted acts.[78] Nevertheless, any resulting liability could also encompass parties other than users. Insofar as AI model developers are concerned, as seen above in Part III, any potentially applicable TDM exception under EU and UK law would only cover the extraction and reproduction during the input/training phase, not other acts. As for providers of AI models, their liability could potentially be established not only on a secondary/indirect/accessory basis, but also on a primary/direct basis, consistent with case law in the UK and EU.

Under UK law, an AI model provider could be held secondarily liable for, e.g. possessing and/or dealing with infringing copy (created, for example, by the developer of the AI model at hand) in the course of business (section 23 CDPA), or for supplying articles *specifically* designed or adapted for making copies of a copyright work (section 24(1) CDPA). As in all cases of secondary infringement, actual or constructive knowledge would need in any event to be proved. Liability could be also established where the provider of an AI model was in principle eligible for the safe harbours under the Electronic Commerce Regulations 2002 and yet played an "active role" as per relevant case law.[79] The same would be true under EU law and the regime now found in the Digital Services Act.[80]

---

[78] In substantially the same sense, see also Korean Ministry of Culture, Sports and Tourism and Korea Copyright Commission, *A Guide on Generative AI and Copyright* (16 January 2024), available at =https://www.copyright.or.kr/information-materials/publication/research-report/view.do?brdctsno=52591#, p 31–33.

[79] *Google France*, C-236/08 to C-238/08, EU:C:2008:389, para 114; *L'Oréal*, C-324/09, EU:C:2011:474, para 123. In the same sense, see also *SNB-REACT*, C-521/17, EU:C:2018:639, para 52; *YouTube*, C-682/18 and C-683/18, EU:C:2021:503, para 105.

[80] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act), OJ L 277, 27.10.2022, p. 1-102. On the CJEU construction of safe harbour immunities in CJEU case law, see further E Rosati, *Copyright and the Court of Justice of the European Union* (OUP:2023), 2ⁿᵈ edn, p. 266–73. For a discussion of the interplay between copyright-specific legislation and the obligations under the Digital Services Act in the context of AI providers (irrespective of safe harbour availability), see A Peukert, "Copyright and the meta-regulation of intermediary services and artificial intelligence" (13 June 2024) Kluwer Copyright Blog, available at https://copyrightblog.klu

Besides secondary liability, AI model providers could, in principle, also be held liable as accessories or even primary infringers. Accessory liability could arise through authorization (section 16(2) CDPA). Another possibility might be joint tortfeasance by means of procuring the commission of the tort, as it has been considered to be the case in a number of situations concerning providers of P2P software/services,[81] or through a common design between multiple subjects pursuant to which the tort was committed. Recently, the UK Supreme Court considered these as distinct bases for imposing accessory liability and suggested a narrower scope of application of liability based on the construction of the subjective element needed in this regard. To this end, it would be necessary that a person knew the essential facts which made the act done wrongful, even if the tort is one of strict liability. This applies whether the claim is based on procuring the infringement of a right or on participation in a common design.[82] Even from this perspective, it is arguable that – at least in certain situations – the provision of a prompt requesting the generation of an output that is substantially identical or highly similar to a pre-existing work (as in the examples given in the opening to Part II) could be considered such as to confer on the provider of the AI model, which is used to infringe, knowledge of the features of the tort committed by the user. Specifically in a copyright context, accessory and primary liability were reviewed and found to subsist in the already mentioned judgment of the Court of Appeal of England and Wales in *TuneIn*.[83]

The primary/direct liability of internet platforms for communicating/making available to the public infringing content uploaded and shared by users of their services is also well-established at the EU level, having regard to both legislation (notably but not exclusively Article 17 of the DSM Directive) and case law, including in relation to Article 3 of the InfoSoc Directive (which the UK transposed through section 20 CDPA). In *YouTube*, C-682/18 and C-683/18, the CJEU upheld the finding in the 2017 judgment in *Ziggo*, C-610/15 that a platform operator may be directly liable for users' infringing acts. To this end, however, several factual elements must be considered to determine the subsistence or not of the indispensable/essential role played by the platform operator and the deliberate nature of its intervention. Among other things, the CJEU gave weight to the consideration whether a platform would implement appropriate technological measures, as they would be expected from a reasonably diligent operator, to counter – credibly and effectively – the undertaking of infringing acts.[84]

Translating all this to generative AI, and also considering the approach taken in the proposed AI Liability Directive,[85] to mitigate the risk of liability a diligent AI model provider would need to implement effective measures to prevent users from using certain expressions, names, words and phrases when inputting prompts. Obviously, if

---

weriplaw.com/2024/06/13/copyright-and-the-meta-regulation-of-intermediary-services-and-artificial-intelligence/.

[81] *Twentieth Century Fox Film Corp v Newzbin Ltd* [2010] EWHC 608 (Ch); *Dramatico Entertainment Ltd & Ors v British Sky Broadcasting Ltd & Ors* [2012] EWHC 268 (Ch); *EMI Records Ltd v British Sky Broadcasting Ltd* [2013] EWHC 379 (Ch); *Twentieth Century Fox Film Corp v. Sky UK Ltd* [2015] EWHC 1082; *1967 Ltd & Ors v British Broadcasting & Ors* [2014] EWHC 3444 (Ch).

[82] A person who causes another person to do a wrongful act will only be jointly liable as an accessory for the wrong done if they have knowledge of the essential facts which make the act done wrongful (*Lifestyle Equities CV v Ahmed* [2024] UKSC 17, paras 107–108). This is so whether the claim is based on procuring the infringement of a trade mark (or other wrong) or on participation in a common design (ibid, paras 111–37).

[83] *Tunein Inc v Warner Music UK Ltd & Anor* [2021] EWCA Civ 441 (26 March 2021), paras 165–68.

[84] *YouTube*, C-682/18 and C-683/18, EU:C:2021:503, para 84.

[85] Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), COM/2022/496 final. For a general discussion of AI liability from a broader risk regulation perspective, see B Botero Arcila, "AI liability in Europe: How does it complement risk regulation and deal with the problem of human oversight?" (2024) 54 CSLR 106012, and A Racki Marinković, 'Liability for AI-related IP infringements in the European Union' (2024) 19(10) JIPLP 784, p 788–789.

primary liability was established on the side of a platform operator or – for what concerns here – AI model provider under the InfoSoc Directive or EU trade mark instruments,[86] no immunity (e.g. hosting safe harbour) would be available. A somewhat similar approach would be required where the AI provider fell within the scope of application of Article 17 of the DSM Directive.[87] In such a situation, and where no authorization was provided by relevant rightholders, the conditions of Article 17(4) would need to be cumulatively satisfied for an AI provider to remove the risk of its own direct liability, including by implementing measures that "distinguish adequately between unlawful content and lawful content."[88]

2. Liability aspects via-á-vis contractual considerations

Having considered the above, another issue that arises is the enforceability (and effectiveness) of terms of service used by AI model providers to exclude their liability for infringements committed by users.[89] While the answer will obviously depend on the specific circumstances at hand, in *YouTube*, C-682/18 and C-683/18, the CJEU referred to the terms of service of a platform operator, and found that – even if the terms require (i) users to respect third-party rights and (ii) that the user holds all the necessary rights, agreements, consents and licences for the content that they upload – the existence of such terms alone could not be enough to exclude the operator's own liability.

This is because the terms of service are merely one of the several criteria that need to be taken into account and balanced in order to determine if the platform operator in question is liable for the copyright-restricted acts. With specific regard to Cyando's cyberlocker Uploaded, the CJEU noted that its general terms and conditions prohibited users from infringing third-party copyrights.[90] Yet, the Court suggested that the liability of the platform should be assessed considering all the relevant aspects of the platform's functions and activities, thus materially broadening the assessment beyond the wording of the platform's terms of service. Consequently, the referring court (German Federal Court of Justice) applied the guidance received from the CJEU and concluded that liability for unauthorized acts of communication to the public would exist under certain conditions for the operator of Uploaded.[91]

The above thus suggests that even if the terms of service of an AI model provider seek to exclude the provider from liability for users' infringing acts, such terms alone may not rule out the liability of the provider. Therefore, rightholders could potentially enforce their rights against both users and providers of AI models.

VI. Availability of exceptions

Having reviewed the elements of prima facie liability for outputs incorporating third-party protected content, this final part will consider the availability of exceptions. To start,

---

[86] *Louboutin*, C-148/21 and C-184/21, EU:C:2022:1016.

[87] On the application of the regime found in Article 17 of the DSM Directive to AI providers, see Z Krokida, "Large language models and EU intermediary copyright liability: quo vadis?" (2024) 46(6) EIPR 361, p. 371.

[88] *Poland*, C-401/19, EU:C:2022:297, para 86.

[89] A review of different providers' terms of service is available in P Wymeersch, "Terms of use on the commercialisation of AI-produced images and copyright protection" (2024) 46(6) EIPR 374, p. 380–81. A discussion of indemnifications offered to users by some AI providers is instead undertaken in D Gervais, "Licensing of copyright-protected material for AI makes business sense" (4 December 2023), available at https://www.copyright.com/wp-content/uploads/2023/12/Gervais-Licensing-of-Copyright-Protected-Material-for-AI-Makes-Business-Sense.pdf.

[90] *YouTube*, C-682/18 and C-683/18, EU:C:2021:503, para. 44.

[91] BGH, I ZR 135/18 *uploaded III*.

besides recalling that under EU law it is not possible to exceed the catalogue of available exceptions under national law (e.g. by invoking directly EU Charter of Fundamental Rights[92]), it is worth noting that all exceptions need to be construed in light of the three-step test. Furthermore, in the context of AI-generated outputs produced at the request of users, the beneficiaries of exceptions will likely be them rather than AI model developers or providers.

The analysis that follows will consider also the applicability of the quotation exception under Article 5(3)(d) of the InfoSoc Directive and the exception for pastiche under Article 5(3)(k) of the InfoSoc Directive[93], as also implemented into UK law. The reason for this focus is that the former has been claimed to be applicable without the need to consider the purpose of the use of the protected content at issue,[94] while the latter has been invoked in some ongoing contentious proceedings concerning AI-generated outputs. It is furthermore at the centre of the pending CJEU referral in *Pelham II*, C-590/23.

1. Compliance with the three-step test and beneficiaries of exceptions

The requirement that a fair balance is struck between protection of copyright and related rights and rights and interests of third parties is a fundamental tenet of the copyright system. The CJEU frequently refers to this requirement, including when construing exceptions. This, in turn, has inter alia led to the characterization of the three-step test in Article 5(5) of the InfoSoc Directive as an instrument that also contributes to the fair balance between exclusive rights and exceptions.[95] As it is discussed in greater detail elsewhere,[96] Article 5(5) of the InfoSoc Directive mandates that the exceptions provided for in paragraphs 1 to 4 therein shall only apply 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 rightholder.

According to Article 5 and to the exclusion of Article 5(1), EU Member States enjoy the discretion to decide whether to incorporate an exception from the list into their own laws. Such a freedom is nevertheless limited in two fundamental ways: first, it cannot be used to compromise the objectives of that legislation and the functioning of the EU internal market; second, a national exception must also comply with Article 5(5).[97] From a UK perspective it is noted that, while the CDPA does not refer to the three-step test, case law indicates that the requirements thereunder have been traditionally considered covered by the notion of "fair dealing."[98] That said, more recent case law indicates that UK defences also need to be construed in light of the three-step test.[99]

A further, key consideration is that the availability of exceptions is premised on the *lawful* access to the copyright-protected work or other subject-matter: if the source from

---

[92] Charter of Fundamental Rights of the European Union, OJ C 326, 26.10.2012, p 391–407. See *Spiegel Online*, C-516/17, EU:C:2019:625, paras 40–49 and *Funke Medien*, C-469/17, EU:C:2019:623, paras 55–64.

[93] These are the same exceptions as found under Article 17(7) DSM Directive.

[94] See (with specific reference to Article 10(1) of the Berne Convention) T Aplin – L Bently, *Global Mandatory Fair Use* (Cambridge University Press:2020), p 69–71.

[95] *Pelham I*, C-476/17, EU:C:2019:624, para 62; *Spiegel Online*, C-516/17, EU:C:2019:625, para 46; *Funke Medien*, C-469/17, EU:C:2019:623, para 61.

[96] E Rosati, "No step-free copyright exceptions: the role of the three-step in defining permitted uses of protected content (including TDM for AI-training purposes)" (2024) 46(5) EIPR 262, p. 263–68.

[97] Opinion 3/15 of the Court (Grand Chamber), EU:C:2017:114, paras 123–25.

[98] *England And Wales Cricket Board Ltd & Anor v Tixdaq Ltd & Anor* [2016] EWHC 575 (Ch) (18 March 2016), paras 88–89. On the notion of "fair dealing," see generally L Bently and Others, *Intellectual Property Law* (OUP:2022), 6th edn, p. 245–49.

[99] *Shazam Productions Ltd v Only Fools The Dining Experience Ltd & Ors (Rev1)* [2022] EWHC 1379 (IPEC) (08 June 2022), para 189. On the interplay between fair dealing and the three-step test, see further E Rosati, *Copyright and the Court of Justice of the European Union* (Oxford University Press:2023), 2nd edn, p. 340–41.

which the user has derived a certain work or subject-matter is unlawful, no exception will be available to them. This conclusion is compliant with the requirements under the three-step test: holding otherwise would encourage the circulation of unlicensed works and inevitably reduce the volume of sales or of other lawful transactions relating to the protected works. This would be contrary to exceptions not being in conflict with a normal exploitation of the work and not unreasonably prejudicing the legitimate interests of rightholders.[100] Let's take the exception for quotation: under both Article 5(3)(d) of the InfoSoc Directive and the UK corresponding provision[101] (discussed below at §2), this requires that the copyright-protected work or other subject-matter being quoted "has already been lawfully made available to the public." If a user generated an output – whether text, visual, or music – through an AI model and such an output reproduced third-party protected content used unlawfully during the model's input/training phase, no quotation exception could be arguably invoked by that user, given that the work or subject-matter at issue would have been derived from an unlawful source.

It follows that all exceptions require fulfilling relevant requirements thereunder, as well as the three-step test. To exemplify: if a music historian generated music through an AI model "in the style of Oasis" (as in the Britpop "lads") in order to review and discuss the latter's technique and features in the context of their PhD dissertation, then a defence, e.g. quotation, might be potentially available to them if the AI-generated output reproduced protectable elements from Oasis's repertoire. Conversely, the owner of a night club who generated "a song in the style of Oasis" to exploit commercially in their establishment without relevant licences would most likely be ineligible for the application of any exception if the AI-generated output reproduced protected elements of Oasis's repertoire.

Assuming that an exception is potentially applicable, a further preliminary point to clarify is who is entitled to benefit from the exception: would it be solely the user of the AI model who generates the output, or would it be also the developer and/or provider of the model? In light of relevant case law and recent legislative developments, the answer is that, in principle, it is the former and not the latter who may benefit from the exception. It is the user generating a certain output through an AI model who enters into a dialogue with an earlier work or other protected subject-matter, not the developer or provider of the model. Similarly, having regard to the construction of the right of communication to the public under Article 3 of the InfoSoc Directive in the context of internet platforms, it is the users who upload content incorporating third-party copyright works and other protected subject-matter that could invoke the application of an exception, not the operator of the platform where such content is being uploaded and shared. In the same sense, Article 17 of the DSM Directive provides for mandatory exceptions and makes it clear that the beneficiaries thereof are not online content-sharing service providers ("**OCSSPs**") but rather users of their services. Along similar lines, it could be also recalled that the action launched by the Republic of Poland against Article 17 of the DSM Directive was based on the freedom of expression and information of *users*, not OCCSPs.[102]

2. "Quotation" and "pastiche" under EU and UK law

Exceptions under Article 5(3)(d) and (k) of the InfoSoc Directive and Article 17(7) of the DSM Directive are rooted within freedom of expression and information.[103] Insofar as the quotation exception is concerned, its applicability is subject to fulfilment of the following

---

[100] *ACI Adam and Others*, C-435/12, EU:C:2014:254, paras 39–40; *VCAST*, C-265/16, EU:C:2017:913, paras 51–52.

[101] Regarding the requirement of 'lawful' access also under section 30(1ZA) CDPA, see L Bently and Others, *Intellectual Property Law* (OUP:2022), 6th edn, p. 263–64.

[102] *Poland*, C-401/19, EU:C:2022:297, para 39.

[103] For Article 5(3)(d), see *Painer*, C-145/10, EU:C:2011:798, paras 134–35.

requirements: first, the quotation must relate to a work or other subject-matter which has already been lawfully made available to the public; second, unless this turns out to be impossible, the source (including the author's name) must be indicated; third, the use at hand must be in accordance with fair practice and to the extent required by the specific purpose. In *Pelham I*, C-476/17, the CJEU considered that the concept of "quotation" must be understood with reference to its usual meaning in everyday language and be justified by the purpose "of illustrating an assertion, of defending an opinion or of allowing an intellectual comparison between that work and the assertions of that user." In summary, there must be "the intention of entering into 'dialogue' with that work."[104]

The approach adopted in *Pelham I*, C-476/17 may appear more permissive than that in *Spiegel Online*, C-516/17, in which the Grand Chamber held that the user must "establish a direct and close link between the quoted work and [their] own reflections, thereby allowing for an intellectual comparison to be made with the work of another." This would be so because Article 5(3)(d) of the InfoSoc Directive allows quotations insofar as they enable criticism or review, with all this suggesting that "use of the quoted work must be secondary in relation to the assertions of that user, since the quotation of a protected work cannot [. . .] be so extensive as to conflict with a normal exploitation of the work or another subject matter or prejudices unreasonably the legitimate interests of the rightholder."[105] Despite their seeming different nuances, the judgments can nevertheless be reconciled: they both suggest that a permissible quotation is one that is justified by its *purpose.*

The same understanding of quotation is also found under UK law: the provision in section 30(1ZA) CDPA was derived from Article 5(3)(d) of the InfoSoc Directive and framed within "fair dealing."[106] By referring to "use of a quotation from that work," section 30(1ZA) implies that the defence be only available to limited, partial reproductions of a work.[107] Recently, the High Court of England and Wales considered – and ruled out – the applicability of section 30(1ZA) in *Pasternak* due to lack of acknowledgment of the work being quoted by the defendant.[108] Given that the conditions under section 30(1ZA) – like those under Article 5(3)(d) – are cumulative, the quotation defence could not find application.[109]

Turning to pastiche, it can first be noted that at the time of writing the notion has not yet been defined at the EU level by the CJEU. Nevertheless, the UK Intellectual Property Enterprise Court ("**IPEC**") recently considered and rejected the application of the corresponding UK defence (section 30A CDPA) in *Shazam*, in a case concerning copyright protection of a fictional character and reproduction thereof through an interactive dining show. Having established that copyright would subsist in the work at hand, the analysis turned to whether section 30A CDPA could be invoked successfully given that the defendants had argued that their show would not be infringing because it qualified as parody or, alternatively, pastiche. Insofar as the latter is concerned, relevant guidance on section 30A released by the UK Intellectual Property Office explains that, further to the introduction of this defence in 2014, an artist would be able use small fragments from a

---

[104] *Pelham I*, C-476/17, EU:C:2019:624, para 71.

[105] *Spiegel Online*, C-516/17, EU:C:2019:625, para 79. In this sense, A Hui – F Dohl, 'Collateral damage: reuse in the arts and the new role of quotation provisions in countries with free use provisions after the ECJ's *Pelham*, *Funke Medien* and *Spiegel Online* judgments' (2021) 52(7) IIC 852, 873–74.

[106] On the notion of "fair dealing," see generally L Bently and Others, *Intellectual Property Law* (Oxford University Press:2022), 6th edn, p 245–49.

[107] A Cameron, "Copyright exceptions for the digital age: new rights of private copying, parody and quotation" (2014) 9(12) JIPLP 1002, 1007.

[108] *Pasternak v Prescott* [2022] EWHC 2695 (Ch) (25 October 2022), para 476.

[109] Ibid, para 487.

range of films to compose a larger pastiche artwork.[110] This is consistent with the framing of the defence within fair dealing.

The IPEC noted that, in his Opinion in *Pelham I*, C-476/17, AG Szpunar held that: "[a]s for the concept of pastiche, it consists in the imitation of the style of a work or an author without necessarily taking any elements of that work."[111] Agreeing with scholarly literature[112] and thus substantially discarding the position of AG Cruz Villalón in *Deckmyn*, C-201/13[113], Deputy High Court Judge John Kimbell KC considered that the everyday meaning of pastiche indicates that it is distinct from, and operates outside of, the genres of parody and caricature. Pastiche entails the imitation of the style of pre-existing works and the use or assemblage of pre-existing works in new works. It also needs to be *noticeably different* from the original work.[114]

The approach taken by the IPEC appears to be overall correct. In any event, the difference between "parody," "caricature" and "pastiche" refers to a different *field* of application of these concepts, not a broader scope of "pastiche" vis-à-vis "parody" or "caricature"[115] or even that "caricature" and "pastiche" are not also autonomous concepts of EU law.[116] "Parody," "caricature" and "pastiche" are ways through which one's own freedom of expression may be exercised by entering into a "dialogue" with an earlier work or protected subject-matter and/or the ideas conveyed therein. Considering the requirements of the three-step test, a national exception for pastiche still requires a *justification* for both the *use* of someone else's work or other protected subject-matter and the *amount* thereof that has been reproduced, thus also entailing a consideration of the *effects* of the use on the market for the original work or protected subject-matter.

VII. Conclusion

Generative AI has been booming over the past several months, yet the legal uncertainties surrounding the lawfulness of model training and resulting applications, including output generation, has the potential to impact very substantially on the development, investment,[117] and use of resulting models. As the study has explained, the liability landscape is likely to become increasingly complex, risky, and touch upon multiple stages of the "life" of an AI model.

---

[110] HM Government, *Technical Review of Draft Legislation on Copyright Exceptions: Government Response*, available at https://data.parliament.uk/DepositedPapers/Files/DEP2014-0529/Copyright_Technical_Review_Government_Response_as_published.pdf.

[111] Opinion of Advocate General Szpunar in *Pelham I*, C-476/17, EU:C:2018:1002, fn 31.

[112] The IPEC referred extensively to E Hudson, "The pastiche exception in copyright law: a case of mashed-up drafting?" (2017) 2017/4 IPQ 346. See also S Jacques, *The Parody Exception in Copyright Law* (Oxford Universsity Press:2019), p 11.

[113] The AG had proposed not to distinguish between the various concepts referred to in Article 5(3)(k) of the InfoSoc Directive: Opinion of Advocate General Cruz Villalón in *Deckmyn*, C-201/13, EU:C:2014:458, para 46.

[114] *Shazam Productions Ltd v Only Fools the Dining Experience Ltd & Ors (Rev1)* [2022] EWHC 1379 (IPEC) (08 June 2022), paras 186–88.

[115] Contra, e.g., M Senftleben, "User-generated content – Towards a new use privilege in EU copyright law," in T Aplin (ed), *Research Handbook on Intellectual Property and Digital Technologies* (Edward Elgar:2020), p 158, suggesting that the exception for pastiche could be the basis of an EU-wide exception for user-generated content.

[116] In a similar sense, see P Mezei and Others, "Oops, I sampled again … the meaning of 'pastiche' as an autonomous concept under EU copyright law" (2024) IIC (advance access), §3.3.

[117] From a financial perspective, commentators have analyzed the size of the economic investment made on AI so far and its overall appropriateness: Goldman Sachs Exchanges, *Unlocking the AI M&A Supercycle* (24 October 2023), available at https://open.spotify.com/episode/19IpU9k4YdFpgBkjHA5FFq?si=POVIYSyURtCgO-gNBpfy2Q&nd=1; Goldman Sachs Exchanges, *A Skeptical Look at AI Investment* (16 July 2024), available at https://open.spotify.com/episode/3I3UDjgnEFUEw52klrvn5H?si=udY6-dSySZm_P3MY_RsOuA&nd=1&dlsi=d24cfecb9b1248d4.

By focusing on liability issues surrounding AI-generated outputs, the analysis has mapped relevant acts under copyright and related rights, from an international and EU/UK perspective. It has then turned to actionable reproduction, allocation of liability, and potential defences to prima facie infringement, along with their beneficiaries. The main findings may be summarized as follows.

Starting with the input/training phase and TDM exceptions under EU and UK laws, these cover acts of extraction and reproduction for TDM purposes (where relevant conditions have been met), not subsequent restricted acts, including having regard to the input/training phase itself. Article 53(1)(c) of the AI Act, read alongside recital 106 in the preamble thereof, indicates that TDM activities are not – at least insofar as AI development is concerned – an end in themselves but are rather a step in and functional to the training of AI models and the subsequent development and offering of the models. The obligations under that article, including that to respect EU copyright rules and the one of transparency, also apply (subject to transitional provisions) to new releases of generative AI models placed on the EU market prior to the entry into force of the AI Act, as well as to all models placed on the market twelve months before the entry into force of the Act.

Turning to actionable reproduction, the test differs between copyright and related rights because of the different rationale of protection. The copying of a sufficiently original part of a copyright-protected work is actionable, whereas (at least) any copying of a part that reflects the investment of the relevant rightholder is covered by the reproduction rights of related rightholders. Despite the approach in *Pelham I*, C-476/17, recognizability cannot and should not be intended as a general benchmark under either regime.

Prima facie liability for infringing outputs may extend beyond users of generative AI models to their developers and providers, depending on the assessment of the relevant facts and circumstances. This is supported by case law of UK courts and the CJEU. Under UK law, AI developers and providers could be liable as secondary infringers, accessories/joint tortfeasors, or primary infringers. A finding of primary/direct liability is also foreseeable having regard to case law of the CJEU in relation to internet platform operators. Furthermore, contractual limitations or exclusions of liability for providers may be unenforceable.

Finally, for an unlicensed act to be covered by an exception, relevant requirements must be fulfilled, including having regard to the three-step test and/or, insofar as the UK is concerned, the notion of fair dealing. The unauthorized taking must be overall justified in light of the requirement that a "fair balance" is struck between different rights and interests: the purpose and amount of the taking, as well as effects on the market for the original work or protected subject-matter, need to be considered too.

In sum: while it is clear that each case will need to be decided on its own merits and that, therefore one should refrain from making sweeping conclusions, the discussion undertaken here shows how the generative AI output phase raises several questions of liability under copyright law. If the goal of policymakers and relevant stakeholders is to ensure the balanced and sustainable development of AI, including in the context of the seemingly revived proposal for an AI Liability Directive, then the issues related to the generation and dissemination of AI outputs need to be given ample attention and a greater role in the debate than what has been the case so far, whether it is in the context of risk assessment and compliance, licensing initiatives,[118] or in contentious scenarios.

---

[118] Voluntary collective licences for AI are currently being developed: see, e.g. https://cla.co.uk/ai-and-copyright/ and https://www.copyright.com/media-press-releases/ccc-pioneers-collective-licensing-solution-for-content-usage-in-internal-ai-systems/.

Cite this article: E Rosati (2025). "Infringing AI: Liability for AI-Generated Outputs under International, EU, and UK Copyright Law". *European Journal of Risk Regulation* **16**, 603–627. https://doi.org/10.1017/err.2024.72