

Comment of the European Copyright Society on the request for preliminary ruling in Case C-250/25 (Like Company)

February 2026

The European Copyright Society (ECS) was founded in 2012 with the aim of creating a platform for critical and independent scholarly thinking on European Copyright Law and policy. Its members are scholars and academics from various countries of Europe, seeking to articulate and promote their views of the overall public interest on all topics in the field of authors rights, neighbouring rights and related matters. The ECS is neither funded nor instructed by any particular stakeholders. Its Opinions represent the independent views of a majority of ECS members.

Executive summary

The reference in *Like Company v Google (Case C-250/25)* is seen as a potential landmark case, giving the EU's highest court the opportunity to define the scope and conditions of permitted artificial intelligence (AI) training and develop an infringement test for AI outputs. The European Copyright Society (ECS) urges the Court of Justice (sitting as a Grand Chamber) to exercise caution.

While the reference stems from a plausible complaint by a press publisher against the provider of an AI powered chatbot reproducing and communicating its editorial content, the implications of this problematic reference could be far-reaching.

- (1) The reference is factually murky with respect to the technology and services at stake, conflating concepts of 'chatbot', 'large language model', and 'search engine'.
- (2) The reference fails to identify consistently the subject matter at stake, which is the press publishers' right under Article 15 of the Copyright in the Digital Single Market Directive (2019/790/EU, hereinafter CDSMD), not authorial works.

Specifically, the reference conflates questions relating to the training phase (Questions 2 and 3) with the legal characterisation of the use of press publications by an LLM-based chatbot (Question 4 but also Question 1, referring to the right of communication to the public and the right of reproduction under Directive 2001/29/EC, hereinafter InfoSoc Directive).

If the reference is found admissible, it is suggested that the Court of Justice should address jointly Questions 4 and 1, which relate to the legal characterisation of the use of press publications in the display. Here it is important to correctly understand next-token prediction in large language models, augmented retrieval technology (where the use of data does not generally form part of the learning process) as well as ‘online use’, defining the scope of the press publishers’ right under Article 15 of the CDSMD.

In the Opinion of the ECS, the ambiguous characterisation of a fast-moving technology may result in the failure to realise the societal benefits of AI as a potential general-purpose technology. There are risks that a rash decision will push Europe towards a licensing economy in which AI systems are offered as a service by (non-European) multinationals, without solving issues of equity such as creator consent and distribution of revenues.

OPINION

1. Introduction: the implications of the case for generative AI

The reference by the Budapest Környéki Törvényszék (Budapest Regional Court) in *Like Company v Google*¹ is seen as the first opportunity for the Court of Justice of the European Union (CJEU or Court) to rule on an issue that has surfaced at the centre of AI litigation around the world. Do AI model training and model outputs infringe exclusive rights under copyright law? Some scholars suggest that the ruling could set an important precedent for the regulation of AI systems and establish a distinct European path.²

Lower courts in many jurisdictions indeed have come to contradictory assessments on what kind of AI development can be undertaken and which generative AI applications can be brought to the market without seeking permission from copyright holders.³ The EU’s statutory

¹ Case C-250/25, *Like Company*: Request for a preliminary ruling from the Budapest Környéki Törvényszék (Hungary) lodged on 3 April 2025 – *Like Company v Google Ireland Limited*.

² See e.g. Dornis, T. W. & Lucchi, N., *Generative AI and the Scope of EU Copyright Law: A Doctrinal Analysis in Light of the Referral in Like Company v. Google*, IIC 56(10) (2025), 1800-1840 (1801) (“The case represents a pivotal juncture in the evolution of EU copyright law. ... The legal questions raised in this case transcend its specific factual context.”); Hacker, P., *Copyright, AI, and the Future of Internet Search before the CJEU: Reflections on Like Company v Google*, *VerfBlog*, 2025/7/17, <https://verfassungsblog.de/copyright-ai-cjeu/>, DOI: 10.59704/3887ddcde69e6944.

³ Key litigations in Europe include the German cases *LAION* (Hamburg Higher Regional Court, judgment of 10 December 2025 (5 U 104/24, available at <https://www.landesrecht-hamburg.de/bsha/document/NJRE001628040>))

framework, anchored in a broad definition of the text and data mining (TDM) exceptions of the CDSMD, was conceived in a pre-generative AI world. However, the AI Act (Regulation (EU) 2024/1689) explicitly asserts the applicability of TDM concepts to general-purpose AI (GPAI) model training, albeit not necessarily covering all aspects of it.⁴

In our assessment of the *Like Company* reference, we urge caution. To be clear, press publishers can have legitimate complaints against providers of AI powered services that reproduce or communicate their editorial content. This particular reference to the Court is, however, convoluted. If the Court was tempted to draw rules of wide application from narrow and murky factual settings, it needs to envisage, from a teleological and policy perspective, the increasingly infrastructural role that copyright plays in technological development.

Copyright decisions also interact with instruments developed in other policy fields. We point here specifically to interventions advanced by competition authorities, such as the French competition authority's decisions regarding Google's practices "implemented in the press sector"⁵ and the European Commission's instigation of proceedings "whether Google is distorting competition by imposing unfair terms and conditions on publishers and content creators, or by granting itself privileged access to such content".⁶

Whether exclusive rights under copyright law are the best instrument to address fundamental power imbalances in information markets remains a sensitive issue. Paradoxically, copyright rules may direct technological development towards a licensing economy in which AI systems are offered as a service by (non-European) multinationals, without solving issues such as creator consent and remuneration, which represent, in our opinion, crucial questions for a thriving EU creative sector.

and *GEMA v OpenAI* (Munich I Regional Court, judgment of 11 November 2025 (42 O 14139/24, available at <https://www.gesetze-bayern.de/Content/Document/Y-300-Z-GRURRS-B-2025-N-30204>), *Getty Images v Stability AI* in the UK ([2025] EWHC 2863 (Ch)); and *DPG Media B.V. v. Mediahuis Nederland B.V., Rechtbank Amsterdam*, C/13/737170 / HA ZA 23-690 (ECLI:NL:RBAMS:2024:6563). In the United States alone, more than 80 AI copyright suits are ongoing. See Professor Edward Lee's tracker at <https://chatgptiseatingtheworld.com/>.

⁴ See Article 53(1)(c) AI Act. As the European Copyright Society has put it in its earlier opinion, "The provisions of Arts. 3 and 4 of the CDSM Directive do not necessarily apply to all aspects or stages of the life cycle of a GPAI model or a generative AI system. Rather, these TDM exceptions cover different (but not necessarily all) aspects of the training stage of GPAI models, to the extent that those activities qualify as: (i) acts of TDM under the broad definition in Art. 2 of the CDSM Directive, and (ii) reproductions of protected subject matter of the type mentioned in Art. 3 and 4 of the CDSM Directive." See ECS Opinion: Copyright and Generative AI, January 2025, p. 5 (https://europeancopyrightsociety.org/wp-content/uploads/2025/02/ecs_opinion_genai_january2025.pdf).

⁵ See Related rights: the Autorité fines Google €250 million, March 20, 2024 (<https://www.autoritedelaconcurrence.fr/en/article/related-rights-autorite-fines-google-eu250-million>).

⁶ See Commission opens investigation into possible anticompetitive conduct by Google in the use of online content for AI purposes, December 9, 2025 (https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2964).

In order to support the Court in developing a doctrine that balances rights for private and public benefit, this Opinion proceeds in three steps. We first draw attention to the uncertain factual characterisations of generative AI training and retrieval augmented generation (a technique that enables large language models to incorporate new information into outputs). These provide grounds for considering the reference inadmissible, creating space for a future, more precise, reference.

We then articulate important distinctions between the press publishers' right (PPR) under Article 15 CDSMD and claims that arise from authors' rights in original works under Article 2.a (reproduction) and 3(1) (communication to the public) of the InfoSoc Directive. Critically, the concepts of 'chatbots', 'large language models', and 'search engines' cannot be used interchangeably. They refer to distinct technologies and services, which give rise to the application of different rights under EU law.

We finally review the scope and relevance of the TDM exceptions under Article 4 CDSMD in the context of the PPR. If the chatbot in question relies on retrieval-augmented generation (RAG), the use of press publishers' data does not typically form part of the learning process (and there is no evidence in this case that it did). Its categorisation as an AI training activity is therefore doubtful.

2. Admissibility of the reference

In view of what may be at stake, we question whether the reference allows the CJEU to interpret the copyright *acquis* in a manner that is both useful for the concrete dispute and conceptually valid. References for a preliminary ruling are not meant to enable advisory opinions on general or hypothetical questions but are justified because they are necessary for the effective resolution of a dispute.⁷ In order to enable the judicial cooperation between referring courts and the Court of Justice, references under article 267 TFEU must meet certain cumulative requirements to be admissible.⁸ In particular, a request for a preliminary ruling must contain "a summary of the subject-matter of the dispute and the relevant findings of fact as determined by the referring court or tribunal, or, at least, an account of the facts on which the questions are

⁷ Judgment of 16 October 2025, *Braila Winds* (C-391/23) ECLI:EU:C:2025:799, para 28 and case-law cited.

⁸ *Ibid.*, para 32 with further references; Recommendations to national courts and tribunals in relation to the initiation of preliminary ruling proceedings OJ C/2024/6008, 9.10.2024, para 15.

based”.⁹ It is questionable whether the request of the Budapest Környéki Törvényszék meets this threshold.

The referring court justifies its request by referring to the pleading of the applicant, according to which “it is in the interests of both parties, *at the stage of determining the subject matter of the dispute*, to know how the Court of Justice interprets the law in relation to new legal questions ... so that the parties *are able to make the statements and arguments* necessary in light of that interpretation ...”.¹⁰ However, it is not the purpose of a reference for a preliminary ruling to help the parties and the referring court identify the relevant points of law. Neither does the referring court otherwise explain the subject-matter of the dispute in a sufficiently clear and detailed manner, although it rightly points out that the dispute concerns a “novel technology” that raises “new legal questions”.¹¹ The CJEU has consistently stressed that in areas “characterised by complex factual and legal situations”, there is a particular need for precision in drafting references for a preliminary ruling.¹²

The reference at issue is, however, unclear as to whether it concerns a ‘chatbot’, a ‘large language model (‘LLM’)’ and/or a ‘search engine’, and in what way these technologies, which Union law regulates separately,¹³ are interrelated in the case at hand. While the presentation of the subject matter of the request suggests that the case concerns both the legal assessment of “the responses given by the services of a chatbot, *and also* of the information communicated by means of editorial content which can be accessed through the results of a search”, the summary of the request later implies that these two uses (i.e. access through the chatbot/the results of a search) at issue are alternatives.¹⁴ In addition, the referring court distinguishes between Google Search and Google Gemini but at the same time states, without clarification, that the two services are somewhat “interconnected” and “similar”, and even calls them collectively “generative AI chatbots”.¹⁵

⁹ Art. 94(a) Rules of Procedure of the Court of Justice, OJ L 265/1.

¹⁰ Summary of the request for a preliminary ruling, Case C-250/25, para 25 (emphasis added); Budapest Környéki Törvényszék, Order of 4 April 2025, 7.P.20.629/2023/20, II paragraph 13.

¹¹ Summary of the request (n. 10), para 25; Budapest Környéki Törvényszék (n. 10), II paragraph 13 and IV.

¹² See e.g. Order of 8 October 2002, *Viacom* (C-190/02) ECLI:EU:C:2002:569, para 22.

¹³ Cf. recital 97 and Art. 3(1) and (63) AI Act; Art. 3(j) Regulation 2022/2065 (Digital Services Act, DSA); Commission Decision of 25.4.2023 designating Google Search as a very large online search engine in accordance with Article 33(4) of Regulation (EU) 2022/2065, C(2023) 2731 final.

¹⁴ Cf. Summary of the request (n. 10), p. 1 with para 11 (“or”). Importantly, the original Hungarian language version of the referral contains ‘illetve’, meaning ‘and’ rather than ‘or’ (hence representing a cumulative rather than alternative scenario) in Hungarian language. This translation inconsistency has to be fixed at a later stage of the proceedings.

¹⁵ Budapest Környéki Törvényszék (n. 10), I paragraph 2 and 5; Summary of the request (n. 10), paras 2, 3 and 14.

The presentation of the parties' pleadings is not helpful either. The applicant is said to plead infringements "in different ways" because its press publications are "used on an online platform (in this case, Google Search and Bard/Gemini ...)".¹⁶ What appears in this formulation as a single "platform" is separated elsewhere as "search engine results page" and "chat in the chatbot interface".¹⁷ We note that neither of these qualifies as an "online platform" in EU law.¹⁸ The applicant also appears to distinguish strictly between (implied) consent given for a display "in search engines" and the communication of content "beyond" such use,¹⁹ and no explanation is given as to why this is the case. The defendant in turn appears to plead yet another set of facts, according to which the Gemini (Bard) service is a "creative tool" that "uses the Google Search database" to "collect data".²⁰ Equally unclear is the description of the Gemini (Bard) service, which is called both "a basic model of the LLM type" and a "chatbot",²¹ although AI models and AI systems can be developed and deployed separately,²² and the four questions speak of an "LLM-based chatbot" – a notion which is solely used in the questions and not explained anywhere else.

Furthermore, the reference does not provide a definition, let alone findings or at least a sufficient account of the "process of training", which is addressed in questions 2 and 3. These questions remain unclear because the training of an LLM, which may be capable of competently performing a wide range of distinct tasks,²³ differs from the 'training' of a potentially very specific 'chatbot', which might be 'based' on an LLM, but which often involves separate techniques such as RAG.²⁴ The reference also does not distinguish between the collection of training data and their later use for the development of the LLM and/or the chatbot, although these activities may again involve different actors and different technical solutions.²⁵ The only relevant statement in the case summary, according to which the "generative AI chatbot has been trained by means of the observation and matching of patterns

¹⁶ Summary of the request (n. 10), paras 13 and 14; Budapest Környéki Törvényszék (n. 10), II paragraph 6, 7.

¹⁷ Summary of the request for a preliminary ruling, Case C-250/25, para 17; Budapest Környéki Törvényszék, Order of 4 April 2025, 7.P.20.629/2023/20, II paragraph 12.

¹⁸ Under Article 3(i) and recital 13 DSA, online platforms are only hosting service providers; whereas chatbot's activity exceeds such type of service.

¹⁹ Budapest Környéki Törvényszék (n. 10), II paragraph 8; Summary of the request (n. 10), para 16.

²⁰ Summary of the request (n. 10), para 13; Budapest Környéki Törvényszék (n. 10), II paragraph 20, 21.

²¹ Summary of the request (n. 10), para 5 ("*that* chatbot", emphasis added); Budapest Környéki Törvényszék (n. 10), I paragraph 4.

²² Recital 97 AI Act.

²³ See the definition of a general-purpose AI model in Art. 3(63) AI Act.

²⁴ Cf. European Union Intellectual Property Office, The Development of Generative Artificial Intelligence from a Copyright Perspective, 2025, DOI: 10.2814/3893780, pp. 128-153 (GenAI training process), 269-273 (RAG and its differences to GenAI model training). On the RAG technique see below section 4.

²⁵ *Ibid.*, pp. 128-153.

(something known as ‘string searching’),²⁶ is equally unclear and furthermore alludes to the operation of a search engine, which is not mentioned in questions 2 and 3.

Last but not least, the lack of a description of the “process of training” renders these questions 2 and 3 hypothetical because it has not been established whether the aggrieved infringements are subject to Hungarian and thereby Union copyright law. The parties apparently disagree. Whereas the applicant argues that “the chatbot used works by Hungarian authors and, during its training, ... infringed the applicant’s right of reproduction”, the defendant contends that it “did not use any Hungarian hardware infrastructure to train the chatbot; that training did not take place in Hungary and, therefore, Hungarian law is not applicable.”²⁷ The referring court neither expresses its position regarding the applicability of Hungarian law nor does it pose a respective question, although there is scholarly debate about which law is applicable to the training of AI models and systems.²⁸ If one follows the conventional, territorial approach, the application of Hungarian and EU copyright law to the case at hand depends upon the locus of the reproductions during the “process of training”.²⁹ If these reproductions do not occur within Hungary, questions 2 and 3 are entirely hypothetical and should therefore not be considered.³⁰

²⁶ Summary of the request (n. 10), para 4; Budapest Körményi Törvényszék (n. 10), I paragraph 4.

²⁷ Summary of the request (n. 10), paras 17 and 19; Budapest Körményi Törvényszék (n. 10), II paragraphs 2, 16.

²⁸ See Quintais, J. P., Copyright, the AI Act and Extraterritoriality (June 19, 2025), available at <https://ssrn.com/abstract=5316132>; Metzger, A., Krise des Journalismus - Lösungsansätze des Urheberrechts, *Juristenzeitung* 2025, 328-333 (332); Schack, H., Auslesen von Webseiten zu KI-Trainingszwecken als Urheberrechtsverletzung de lege lata et ferenda, *Neue Juristische Wochenschrift* 77(3) (2024), 113-117 (115) (domicile of the company responsible for the training); Rosati, E., Infringing AI: Liability for AI-generated outputs under international, EU, and UK copyright law, *European Journal of Risk Regulation* 16(2) (2025), 603-627 (615) (“if the acts of extraction and reproduction undertaken during TDM processes are functional ... 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”); Azzi, T., Conseil supérieur de la propriété littéraire et artistique (CSPLA): Publication du rapport de mission sur la loi applicable aux modèles d'IA générative commercialisés dans l'Union européenne, 18 December 2025 (<https://www.culture.gouv.fr/nous-connaître/organisation-du-ministere/conseil-superieur-de-la-propriete-litteraire-et-artistique-cspla/travaux-et-publications-du-cspla/missions-du-cspla/publication-du-rapport-de-mission-sur-la-loi-applicable-aux-modeles-d-ia-generative-commercialises-dans-l-union-europeenne>).

²⁹ Code of Practice for General-Purpose AI Models, Copyright Chapter, 10.7.2025, available at <https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>, Commitment 1(2) („before carrying out any copyright-relevant act in the territory of the relevant Member State“); Stieper M. & Denga, M., Die Reichweite des EU-Urheberrechts nach der KI-VO. Gewerblicher Rechtsschutz und Urheberrecht 126(20) (2024), 1473-1483 (1477); Peukert, A., Copyright in the Artificial Intelligence Act – A Primer. *Gewerblicher Rechtsschutz und Urheberrecht International* 126(6), 497-509 (505-506); Peukert A. & Castets-Renard, C., The Copyright Chapter of the EU’s Code of Practice for General-Purpose AI Models: A Commentary, *International Review of Intellectual Property and Competition Law* 56 (2025), 1950-1993 (1961); Quintais (n. 28).

³⁰ Cf. Judgment of 24 October 2024, *Kwantum Nederland BV and Others v Vitra Collections AG*, (C-227/23) ECLI:EU:C:2024:914, para 45; *Getty Images v Stability AI* [2025] EWHC 2863 (Ch), para 9 (“it is now acknowledged by Getty Images that (i) there is no evidence that the training and development of Stable Diffusion took place in the United Kingdom (such that what has been called “the Training and Development Claim” has been abandoned”).

3. The press publishers’ right as the sole subject matter of the dispute and the scope of that right under the CDSM Directive

Even if the CJEU were to find all or some of the questions to be admissible, the *Like Company* reference is not the “big case” that would allow the Court to clarify under which conditions generative AI training and deployment is in conformity with EU copyright law in general. The reason is that the applicant based its claims solely on its PPR under the Hungarian Copyright Act and not on authors’ rights in works of a journalistic nature.³¹

The PPR under Article 15 CDSMD represents a new ‘related right’ for the benefit of press publishers established in a Member State in respect of the online reproduction and making available to the public of their press publications (as defined by Article 2(4) CDSMD) by information society service providers (ISSPs, as defined by Directive (EU) 2015/1535).

On the one hand, this new related right builds heavily on the *acquis*, as visible by the reference to Articles 2 and 3(2) InfoSoc Directive in Article 15 CDSMD.³² Hence terms like ‘reproduction’, ‘making available to the public’ or ‘information society service provider’ shall be interpreted in light of the formerly existing secondary legislation and CJEU case law. On the other hand, Articles 2(4) and 15 CDSMD narrow down the scope of the PPR in multiple ways.³³

First, press publishers are granted this right solely to protect their “organisational and financial contribution” or investments. Press publications are not ‘works’, but mere ‘collections’ of publications, including but not limited to ‘works or other subject matter’ protected by copyright or other related rights.³⁴ The new right was introduced as a tool to strengthen press publishers’ bargaining power against ISSPs.³⁵ Accordingly, Member States may transpose Article 15 CDSMD into national legislation in ways that include specific obligations on ISSPs to conduct negotiations to authorize the use of press publications,³⁶ as long as “those provisions do not deprive publishers of the possibility of refusing to grant such authorisation or of granting it free

³¹ See Budapest Környéki Törvényszék (n. 10), II paragraphs 2-7 and CJEU, Judgment of 9 February 2012, Luksan (C-277/10), ECLI:EU:C:2012:65.

³² See recital 57 CDSMD.

³³ AG Opinion of 10 July 2025, Meta Platforms (C-797/23), ECLI:EU:C:2025:552, para. 24.

³⁴ See recital 56 CDSMD.

³⁵ See recital 55 CDSMD; see further AG Opinion, Meta Platforms (n. 33), para. 25.

³⁶ AG Opinion, Meta Platforms (n. 33), para. 41.

of charge, that they do not impose on ISSPs any payment obligation unrelated to the actual or intended use of such publications, and that they do not restrict the contractual freedom of the parties in a binding manner”.³⁷ Furthermore, the PPR leaves intact the copyright of authors and other rightsholders whose works are included in the press publications as established in Article 15(2) CDSMD. As such, these other rightsholders might independently exercise and enforce their copyright (or related rights) against the users of those works or other subject matters.

Secondly, the scope of PPR is limited to press publications that are published under the initiative, editorial responsibility and control of the press publisher, and that are first published on or after 6 June 2019. Consequently, user generated content is excluded from the scope of this related right.³⁸

Thirdly, the PPR protects press publishers only against the ‘online use’ of their press publications by information society service providers (ISSPs), as mentioned by Article 15(1) and reaffirmed by various recitals. As such, the ‘online’ use does not equal with every possible ‘digital’ use as originally proposed by the European Commission;³⁹ and it is expressly bound to acts of reproductions and making available to the public carried out via the Internet by ISSPs, that is via services “normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services”.⁴⁰ The interpretation of what constitutes an ‘online use’ is thus also determinative for identifying the services covered by the right. Recital 54 CDSMD refers to “online services, such as news aggregators or media monitoring services, for which the reuse of press publications constitutes an important part of their business models and a source of revenue”. It is also questionable whether providers of AI chatbots can be equated with such services, inter alia because it may be impossible to establish a clear link between an AI summary and a particular press publication - think about the case that a journalistic work is published in parallel by numerous press publishers. Article 15(2) CDSMD actively enables such parallel publications of journalistic content by providing that the PPR “shall not deprive” journalists from exploiting their works “independently from the press publication in which they are incorporated” and that it “shall not be invoked to prohibit the use by other authorised users”.

³⁷ AG Opinion, *Meta Platforms* (n. 33), para 62. See Kowala, M., *Publishers’ rights and copyright law. Safeguarding access to information and media pluralism* (Routledge, 2025) 202.

³⁸ AG Opinion, *Meta Platforms* (n. 33), para. 28.

³⁹ Cf. Art. 11(1) Proposal for the CDSM Directive, COM/2016/593 final.

⁴⁰ Article 1 paragraph 1 point b of Directive 2015/1535.

Fourthly, the Directive includes numerous exceptions and limitations to the PPR. Three of them are expressly listed in Article 15(1) CDSMD, namely: private or non-commercial uses by individuals; hyperlinking; and, seemingly the most important one for the referred case, the use of individual words or very short extracts of press publications.⁴¹ Some further limitations are prescribed in Article 15(2) and (3) CDSMD, namely: neither authorised uses of third parties under a non-exclusive license nor the use of public domain material shall be prohibited; and exceptions and limitations under Article 5 InfoSoc Directive shall apply *mutatis mutandis* to the PPR.

Finally, the term of protection of the PPR is limited to two years from 1 January of the year following the publication of the press publication (Art. 15(4) CDSMD). Press publishers are also required to share their revenues with authors and related rights holders whose works and other subject matter were included in the press publications (Art. 15(5) CDSMD).

In sum, since all the questions center around the protection of press publishers rather than authors, Articles 2 and 3 of the InfoSoc Directive and Article 4 CDSMD are not at stake per se for all the subject matter they cover *but only in connection with Article 15 CDSMD*. Accordingly, if the CJEU were to give a ruling on the substance of the case, such ruling must be strictly limited to the scope of the PPR as codified in Articles 2(4) and 15 as well as the recitals of the CDSMD.

4. On the scope and relevance of Article 4 CDSM Directive

In the context of the *Like Company* reference, Article 4 CDSMD provides for an exception to the rights laid down in Article 15(1) of that directive in respect of reproductions and extractions of lawfully accessible press publications carried out for the purposes of TDM. Question 3 of the reference seeks to ascertain whether Article 4 is applicable to the factual circumstances described therein, on the assumption that Question 2 is answered in the affirmative. As already noted, the description of those circumstances is not always consistent or sufficiently precise, which restricts the possibility of undertaking a detailed assessment. Nevertheless, some observations on the validity and applicability of Article 4 CDSMD in connection with Article 15(1) CDSMD can be developed.

⁴¹ See recitals 55, 57 and 58 CDSMD.

As a preliminary matter, it is necessary to determine whether the TDM activities required in order to substantiate the alleged facts are, in principle, subject to authorisation by rightholders. While that question can be answered relatively straightforwardly in relation to works of authorship, the position is less clear with regard to the PPR, given the more limited scope of protection afforded by this related right. Accordingly, the examination of the applicability of Article 4 CDSMD must begin with an assessment of whether the material scope of Article 15 CDSMD, namely acts of reproduction and acts of making available to the public for the online use of press publications by ISSPs, is capable of encompassing activities connected with the training of AI models by such providers, where those providers do not operate, inter alia, as news aggregators or media monitoring services within the meaning of recital 54 CDSMD.

Only if it is established that the factual situation underlying the reference is capable of falling within the scope of Article 15 CDSMD can the applicability of Article 4 be examined. As indicated above, the subject matter protected under Article 15, namely press publications, is not excluded from the scope of the exception laid down in Article 4. Consequently, acts of reproduction of qualifying press publications carried out for the purposes of TDM fall, in principle, within that exception, provided that the conditions laid down therein are satisfied and that no valid reservation of TDM uses by rightholders was expressed at the relevant time.⁴²

A further issue of importance concerns the question whether the exception provided for in Article 4 CDSMD is capable of covering the training of a model for the purpose of developing AI applications of the type described in the reference. That issue can, however, be addressed only on the basis of a sufficiently precise delineation of the relevant facts. As previously observed, the reference exhibits a degree of terminological inconsistency, in particular in its interchangeable use of concepts such as ‘chatbots’, ‘large language models’, and ‘search engines’. Those terms refer to distinct technologies and services, which may give rise to the application of different rights under EU law. Their precise definition is therefore indispensable for a proper assessment of the applicability of Article 4 CDSMD to the various technological stages involved.

In general terms, Article 4 CDSMD is capable of applying to situations involving reproductions carried out for the purpose of training AI models.⁴³ That conclusion follows both from the wording of the CDSMD, in particular the broad and non-exhaustive definition of TDM set out

⁴² Regarding the three-step test see Senftleben M., Text and Data Mining, Generative AI, and the Copyright Three-Step Test, IIC 57 (2026), 1-41, <https://doi.org/10.1007/s40319-026-01680-2>.

⁴³ See ECS Opinion (n. 4); Kretschmer, M., Margoni, T. & Oruç, P. Copyright Law and the Lifecycle of Machine Learning Models, IIC 55(1) (2024), 110-138, <https://doi.org/10.1007/s40319-023-01419-3>.

in Article 2(2) CDSMD, and from the subsequent adoption, in 2024, of the AI Act. Recital 105 of the AI Act states that, where “the rights (sic!) to opt out has been expressly reserved in an appropriate manner, providers of general-purpose AI models need to obtain an authorisation from rightholders if they want to carry out text and data mining over such works”. That interpretation is confirmed by Article 53(1)(c) of the AI Act and is further elaborated in the GPAI Code of Practice assessed as adequate by the European Commission and the AI Board.⁴⁴

Finally, although Question 3 is again unclear in this respect, the description of the facts in the reference appears to relate to the use of up-to-date information at inference time by Gemini (Bard), that is to say during the generation of outputs, a process or technique commonly referred to as RAG. RAG is a deployment-stage technique whereby a generative AI system retrieves external content and provides it to the model as additional context when generating a response. RAG is generally distinguished from model training, since it does not rely exclusively on (or affect) model weights but instead dynamically accesses external sources such as search engines, proprietary databases, or internal document repositories to optimize outputs.⁴⁵ Well-known applications of RAG include so-called ‘answer engines’, which provide users with concise and contextually relevant responses, as appears to be the case with Gemini (Bard).⁴⁶

In the context of the copyright-related obligations laid down by the AI Act, a similar distinction is drawn in the Explanatory Notice and Template for the Public Summary of Training Content for General-Purpose AI Models required under Article 53(1)(d). That document states that, since Article 53(1)(d) expressly refers to ‘training’, other input data used during the operation of the model, for example through RAG techniques, need not be included in the mandatory sections of the template, unless the model actively learns from such input data.⁴⁷ In other words,

⁴⁴ ECS Opinion (n. 4); European Commission, Commission Opinion of 1.8.2025 on the assessment of the General-Purpose AI Code of Practice within the meaning of Article 56 of Regulation (EU) 2024/1689, COM(2025) 5361 final, para. 2; AI Board, Conclusion on the Assessment of the General-Purpose AI Code of Practice pursuant to Article 56 of Regulation 2024/1689, (<https://ec.europa.eu/newsroom/dae/redirection/document/118687>); Quintais, J. P., Generative AI, Copyright and the AI Act. *Computer Law and Security Review* 59 (2025), Article 106107.

⁴⁵ U.S. Copyright Office, *Copyright and Artificial Intelligence: Part 3: Generative AI Training* (Pre-publication version, May 9 2025, <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>), pp. 22-23, 30-31; see also EUIPO Study (n. 24), pp. 272-273, distinguishing between two main types of RAG systems, namely static RAG and dynamic RAG, and distinguishing RAG from other approaches, such as fetching and training.

⁴⁶ EUIPO Study (n. 24), p. 269.

⁴⁷ European Commission, Annex to the Communication to the Commission: Explanatory Notice and Template for the Public Summary of Training Content for General-Purpose AI Models Required by Article 53(1)(d) of Regulation (EU) 2024/1689 (AI Act), C(2025) 5235 final (24 July 2025), pp. 3-4. See also Veale, M. & Quintais, J. P., The Obligations of Providers of General-Purpose AI Models (Articles 53, 54 and 56) in Malgieri, G., González Fuster, G., Mantelero, A. & Zanfir-Fortuna, G. (eds), *The EU Artificial Intelligence Act: A Thematic*

where the use of data in RAG does not form part of the learning process, it is not regarded as a training activity.

Since Question 3 of the reference expressly refers back to Question 2 and to the “process of training an LLM-based chatbot”, and seeks clarification as to the applicability of Article 4 CDSMD to that process, the referring court would be required to establish and clarify those factual elements in the present case. As the reference currently stands, it does not do so. Consequently, should the CJEU decide to answer the reference, it would, in our view, be required to confine its analysis to the question whether Article 4 CDSMD applies to model-training activities carried out prior to the placing of the model on the market, to the exclusion of RAG techniques that do not form part of the learning process. Subject to any further factual clarification to the contrary, it is our view that the training of AI models as described in the reference falls within the definition of TDM and is therefore covered by the exception laid down in Article 4 CDSMD.

5. Proposed approach to the referred questions

In light of the foregoing analysis, it is submitted that the CJEU should reorder and regroup the questions referred for a preliminary ruling as follows.

In essence, the questions referred concern exclusively the use of press publications owned by press publishers in the context of the training of a specific category of LLM and the subsequent display of content in the context of an LLM-based chatbot.

The Court should first examine Questions 2 and 3. As regards the training phase, the reference is limited to the exclusive right of reproduction of press publishers under Article 15 CDSMD, that right being linked to the general right of reproduction laid down in Article 2 InfoSoc Directive. No other reproduction right, nor any other category of rightholder, is implicated, in particular as regards authorial works.

Question 2 seeks to ascertain whether the process of training an LLM-based chatbot, as described in the order for reference, entails an act of reproduction falling within the scope of the PPR provided for in Article 15(1) CDSMD. This question can only be answered in the

Commentary (Hart Publishing 2026); pre-print available at [10.2139/ssrn.5744602](https://ssrn.com/abstract=5744602); Margoni, T., TDM and generative AI: Lawful access and opt-outs, *Auteurs & Media* 2024/2.

affirmative if the Court finds that relevant reproductions during the training can be qualified as reproductions “for the online use of (...) press publications by information society service providers” (cf. Article 15(1) CDSMD).

Should that question be answered in the affirmative, it then becomes necessary to address Question 3, which concerns whether such reproduction of lawfully accessible press publications (the reference incorrectly speaks of ‘works’) is covered by the exception for TDM laid down in Article 4 CDSMD.⁴⁸

In our view, if the CJEU were to find that the training of an LLM-based chatbot in the present context gives rise to a reproduction that is relevant under the PPR, such reproduction is, in principle, covered by the TDM exception laid down in Article 4 CDSMD, provided that the remaining conditions for the application of that exception are satisfied.⁴⁹

Distinct from the training phase, the Court should then address jointly Questions 4 and 1, which relate to the legal characterisation of the use of press publications in the display, by an LLM-based chatbot, of excerpts from press publications in its responses. Here again, only the related rights of press publishers, namely the rights of reproduction and of making available to the public for the online use of a press publication, are at issue. In our view, Question 4 can only be answered in the negative in light of the facts of the case involved, as presented in section 2 above, and the limited scope of the PPR, as presented in section 3 above.

6. Conclusion

While it is tempting to encourage the Court to develop a test “where user-facing responses reproduce protected material in a manner that substitutes for access to the source”,⁵⁰ getting the scope of exclusive rights wrong on the basis of an ambiguous characterisation of a fast moving technology may result in a failure to realise the societal benefits of AI as a potential general-purpose technology.⁵¹ Furthermore, a reference focussing on the PPR, rather than authorial

⁴⁸ In that regard, it should be noted that the question refers to ‘works’ because the definition of ‘press publications’ in Article 2(4) of that directive encompasses works. The decisive point is, however, that the right at issue in the present reference is a related right, namely the PPR.

⁴⁹ See ECS Opinion (n 4).

⁵⁰ Dornis & Lucchi (n. 2), 1835.

⁵¹ General-purpose technologies, such as the steam engine, the electric motor, and semiconductors “are characterised by pervasiveness, inherent potential for technical improvements, and ‘innovational

works, is not a suitable vehicle for fixing fundamental issues stemming from both AI model training and content generation practices, nor the crucial issue of authors' consent and remuneration.⁵²

Drafting committee: Péter Mezei, Martin Kretschmer, Thomas Margoni, Alexander Peukert, João Pedro Quintais.

Signatories

Prof. Valérie Laure Benabou, Professor, Université Paris Saclay / UVSQ, France

Prof. Lionel Bently, Professor of Intellectual Property Law, University of Cambridge, United Kingdom

Prof. Estelle Derclaye, Professor of Intellectual Property Law, University of Nottingham, United Kingdom

Prof. Séverine Dusollier, Professor, Sciences Po Paris, France

Prof. Jonathan Griffiths, Professor of Intellectual Property Law, Queen Mary, University of London, United Kingdom

Prof. Martin Husovec, Associate Professor, LSE Law School, London School of Economics and Political Science (LSE), United Kingdom

Prof. Martin Kretschmer, Professor of Intellectual Property Law, University of Glasgow and Director, CREATE, United Kingdom

Prof. Thomas Margoni, Research Professor of Intellectual Property Law, Centre for IT & IP Law (CiTiP), KU Leuven, Belgium

Prof. Axel Metzger, Professor of Civil and Intellectual Property Law, Humboldt-Universität, Berlin, Germany

Prof. Péter Mezei, Professor, University of Szeged, Hungary; Adjunct Professor (dosentti), University of Turku, Finland

Prof. Alexander Peukert, Professor of Civil, Commercial and Information Law, Goethe-University Frankfurt am Main, Germany⁵³

Prof. João Pedro Quintais, Associate Professor, University of Amsterdam, Institute for Information Law (IViR), Netherlands

Prof. Ole-Andreas Rognstad, Professor of Law, Department of Private Law/Centre for European Law, University of Oslo, Norway

Prof. Martin Senftleben, Professor of Intellectual Property Law and Director of the Institute for Information Law (IViR), University of Amsterdam, Netherlands

complementarities". See Bresnahan, T. F. & Trajtenberg, M., General purpose technologies 'Engines of growth'?, *Journal of Econometrics* 65(1) (1995), 83-108.

⁵² On the problematic interface between the PPR and journalistic works incorporated into press publications, cf. recital 59 CDSMD.

⁵³ From September 2024 to July 2025 Prof. Peukert served as a co-chair of a working group responsible for the drafting of the Copyright Chapter of the First General-Purpose AI Code of Practice under the AI Act.

Prof. Caterina Sganga, Professor of Comparative Private Law, Scuola Superiore Sant'Anna (Pisa), Italy

Prof. Alain Strowel, Professor, Saint-Louis University and UCLouvain, Belgium

Prof. Tatiana Eleni Synodinou, Professor of Private and Commercial Law, University of Cyprus, Cyprus

Prof. Mireille van Eechoud, Professor of Information Law, Institute for Information Law (IViR), University of Amsterdam, Netherlands