

Dornis, Tim W.

**Article — Published Version**

## Generative AI, Reproductions Inside the Model, and the Making Available to the Public

IIC - International Review of Intellectual Property and Competition Law

**Provided in Cooperation with:**

Springer Nature

*Suggested Citation:* Dornis, Tim W. (2025) : Generative AI, Reproductions Inside the Model, and the Making Available to the Public, IIC - International Review of Intellectual Property and Competition Law, ISSN 2195-0237, Springer, Berlin, Heidelberg, Vol. 56, Iss. 5, pp. 909-938, <https://doi.org/10.1007/s40319-025-01582-9>

This Version is available at:

<https://hdl.handle.net/10419/323685>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by/4.0/>



# Generative AI, Reproductions Inside the Model, and the Making Available to the Public

Tim W. Dornis 

Accepted: 20 February 2025 / Published online: 10 March 2025

© The Author(s) 2025

**Abstract** The training of generative artificial intelligence (AI) models requires the collection and analysis of a staggering amount of data, most of which consist of copyright-protected works. To date, the question whether reproductions of these works are created inside the models during their training has seldom been discussed. This is a serious blind spot in the debate given that such reproductions – e.g., inside ChatGPT’s or Stable Diffusion’s models – could be made available to end users and, therefore, to the public when AI services are offered online. Under the InfoSoc Directive, this might be copyright infringement. EU Member States’ national copyright laws would then apply and their national courts would have international jurisdiction. Seen in this light, the widely propagated narrative that non-EU AI developers are not subject to EU copyright law is an illusion.

**Keywords** Generative AI · Copyright · InfoSoc Directive · Making available to the public · Choice of law · Jurisdiction

---

Tim W. Dornis, Professor of law, Ph.D. (Dr. iur.), J.S.M. (Stanford University), Chair of Private Law and Intellectual Property Law at Leibniz University (Hannover), and Global Professor of Law at NYU School of Law (Global Law Program/Paris).

This article is based on the findings of an interdisciplinary research project carried out with Prof. Dr.-Ing. Sebastian Stober, Faculty of Computer Science at Otto von Guericke University (Magdeburg). The study was undertaken on behalf of the Initiative Urheberrecht (Authors’ Rights Initiative). The results have been published in Tim W. Dornis and Sebastian Stober, *Urheberrecht und Training generativer KI-Modelle* (1st edn, Nomos 2024), available as Open Access at <https://doi.org/10.5771/9783748949558>. All online sources cited in this text were last accessed on October 28, 2024.

---

T. W. Dornis (✉)

Professor, Chair of Private Law and Intellectual Property Law, Leibniz University, Hannover, Germany

e-mail: [tim.dornis@jura.uni-hannover.de](mailto:tim.dornis@jura.uni-hannover.de)

## 1 Introduction and Overview

The capabilities of generative AI systems such as ChatGPT, DALL-E, and Stable Diffusion are the result of an extensive training of their underlying models with large amounts of data.<sup>1</sup> A substantial part of this data is copyright-protected and used without a licence, simply taken from publicly available internet sources.<sup>2</sup> Both before and during these models' training, copyright-relevant copies and reproductions occur. Claims of copyright infringement during generative AI training have already been brought to court in a number of jurisdictions, particularly in the United States.<sup>3</sup> While in the United States the "fair use defence" under § 107 US Copyright Act dominates the discussion,<sup>4</sup> the debate in Europe focuses on the so-called TDM exception – the statutory carve-out from copyright protection for "text and data mining" under the 2019 Digital Single Market Directive 2019/790 (DSM Directive).<sup>5</sup>

### 1.1 A Dark Vision: The "Territorial Disempowerment" of EU Copyright Law

Most scholarly commentators argue that the TDM exception must be applied to generative AI training. Like text and data mining techniques, generative AI training is said to focus on the analysis of semantic information only, so that the copyright-protected expressive elements of works contained in the stock of training data are left untouched.<sup>6</sup> In addition, a more pragmatic argument seems to support the hypothesis that the TDM exception is applicable – rather, that it *must* be applicable: If the exception is not applied, the argument goes, EU innovation in the AI sector will fall further behind in the global race with other regions and countries. This seems to be particularly true with regard to countries that have more "liberal" copyright regimes, namely the United States. It is tacitly anticipated that US federal

<sup>1</sup> For the distinction between "AI systems" and "AI models" (particularly "general-purpose AI models"), see Art. 3(1) and (63) and Recital 97 AI Act. According to Recital 97, "AI systems" and "general-purpose AI models" are different from each other: The latter are typically the direct subject of AI training. As central building blocks, these models are usually integrated into AI systems and are their essential component (*see* sentence 9). Powerful generative AI models ("large generative AI models") are classified as a special case of "general-purpose AI models" according to Recitals 99 and 105. This paper takes these textual and conceptual distinctions as a given.

<sup>2</sup> *See, e.g.*, Schuhmann et al. (2022).

<sup>3</sup> For an overview of the United States, *see, e.g.*, Samuelson (2024), pp. 4 and 75.

<sup>4</sup> *Cf.* § 107 US Copyright Act. For more recent overviews of the US debate *see, e.g.*, Sag (2024); further also Brauneis (2025). Litigation seems to blossom. An updated list of lawsuits against generative AI developers can be found in the Database of AI Litigation (DAIL). It is available at <https://blogs.gwu.edu/law-eti/ai-litigation-database/>. Generative AI litigation can be found by searching for "Generative AI" in the "Application Area".

<sup>5</sup> Directive (EU) 2019/790 of the European Parliament and of the Council of April 17, 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.

<sup>6</sup> *See, e.g.*, Maamar (2023), p. 483; de la Durantaye (2023), p. 651; Hofmann (2024a), pp. 13–14; Hofmann (2024b), pp. 170–171; Heine (2024), p. 88; Pesch and Böhme (2023a, b), p. 1006; Lux and Noll (2024), p. 113; *see also* Nordemann and Pukas (2022), p. 974; Vesala (2023), p. 356; for the contrary position, *see* Schack (2024), p. 114.

courts might ultimately subscribe to the fair use defence. Since copyright law is strictly territorial, developers training their AI models in the United States would accordingly not have to comply with the stricter European laws. Ultimately, they could operate freely and undisturbed in the shadow of the innovation-friendly fair use defence.<sup>7</sup> Alas, in the global race for AI innovation, there seems to be no way around the TDM exception – “TDM, or perish” appears to be the order of the day in Europe.

## 1.2 But There Is Light: The “Long Arm” of EU Copyright Law

The issue is complex. First of all, a look into AI machinery implies that the application of the TDM exception would run counter the most fundamental principles and policies of European copyright doctrine: Unlike TDM, generative AI training comprehensively ingests and digests all elements of its training input – including the copyright-protected components that account for the expressive contents of the works used for training. Accordingly, the TDM exception is inapplicable.<sup>8</sup> Also, no other copyright exception or limitation applies. As a result, generative AI training with copyright-protected works – if it happens without the right holders’ licence – is plain and simple infringement. Much more could be said at this point. Yet this article shall focus on a different point: It will show that the frequently propagated myth of a “territorial disempowerment” of right holders against the self-service by AI developers from outside the EU is a myth. EU copyright law is by far no paper tiger – rather, it offers protection and can be enforced against AI developers’ activities, even if they are acting outside the EU.

To begin with, the AI Act<sup>9</sup> requires that EU copyright law be observed even if the AI training occurs outside the Union. In principle, it provides for an *indirect* enforcement of EU copyright law. As long as AI developers intend to sell AI systems or models, as well as AI-generated output, in Europe, they have to comply with local copyright laws even if they are acting outside EU territory.<sup>10</sup> But these rules do not apply yet.<sup>11</sup> Furthermore, it remains to be seen how effective enforcement by the public authorities will be and whether additional private- or civil-law remedies will be available.<sup>12</sup> Yet there is no need to wait. An option for

<sup>7</sup> See, e.g., Margoni and Kretschmer (2022), p. 690; Nordemann and Pukas (2022), p. 974; Maamar (2023), p. 486; de la Durantaye (2023), pp. 657–658 and 660; Senftleben (2023), p. 1548; Hofmann (2024b), p. 14. This argument has been used in the US for some time to justify claims that the training of generative AI models should fall under the fair use defence. See, e.g., Sobel (2017), p. 81; see also, e.g., Torrance and Tomlinson (2023), pp. 245 *et seq.*

<sup>8</sup> For an extensive analysis, see Dornis and Stober (2024), p. 94; see also Dornis (2025), sub III.

<sup>9</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized 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 Regulation), OJ EU, 12.7.2024.

<sup>10</sup> Dornis and Stober (2024), p. 167 *et seq.*

<sup>11</sup> See Arts. 111(3) and 113(b) AI Act.

<sup>12</sup> Dornis and Stober (2024), pp. 167 *et seq.*; see also, for a doubtful perspective on the expected effectiveness of the AI Act, Sag and Yu (2025), p. 43.

the *direct* enforcement of EU copyright law already exists. Data collection and processing as well as the training as such may occur outside the EU and, hence, beyond the territorial scope of EU copyright law. Yet what has remained largely undiscussed so far is whether offering a trained AI model for download or whether the offering of model-based AI services to end users in the EU (e.g., ChatGPT and Stable Diffusion) could be regarded as a “making available to the public” in the sense of Art. 3(1) InfoSoc Directive<sup>13</sup> of the works that have been replicated inside the trained models during their data training.

### 1.3 Roadmap

As we will see, at least some of the copyright-protected works used for the training of generative AI are reproduced and copied inside the models, namely in their vector space (*infra* 2). Accordingly, the offering of trained models for download or of model-based AI services on the internet – most of which are targeted at European end users – may be considered a “making available to the public”. Such cases of copyright infringement will be governed by EU Member States’ national copyright laws (founded on the InfoSoc Directive), and EU Member States’ courts will have international jurisdiction (*infra* 3). This rather unspectacular facet of the clash of interests between copyright owners and AI producers raises the question of whether the defence against the “it’s easier to ask for forgiveness than beg for permission” attitude of Silicon Valley heroes and other non-EU actors should be led more aggressively.<sup>14</sup> In addition – and this sheds light on the larger picture – the issue of how to reconcile copyright law and generative AI training remains an underexplored desideratum for law- and policymakers around the world – especially in the EU (*infra* 4).

## 2 Of Copies and Reproductions: The “Inside” of Generative AI Models

Genuine reproduction rights are actually affected during all stages of generative AI training: Copying and reproduction not only happen during the collection and compilation of the data corpus (e.g., when works are collected online through web scraping techniques). They also occur during the generation phase – i.e., the production of AI-generated output. This particularly covers instances where generative AI systems regurgitate works that have been part of the data input.<sup>15</sup> Yet the internal life of the algorithmic black box, so to speak, has been largely overlooked to date. Indeed, most commentators deny that copying and reproduction occur during the training of generative models, thereby neglecting an essential

<sup>13</sup> Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonization of certain aspects of copyright and related rights in the information society, OJ No. L 167, 22.6.2001, p. 10.

<sup>14</sup> The general discontent with the “Californian ideology” built on an attitude of moving fast and breaking things (in the interest of innovation and the common good (however this is defined)) has been expressed numerous times. For an insightful analysis, see, e.g., Nemitz (2018), A 376: 20180089.

<sup>15</sup> See, e.g., Dornis and Stober (2024), p. 68 *et seq.*

element of generative AI training. This lopsided perspective invites legal misconception. As can be shown, the digital processes that take place during the ingestion and digestion of data inside the artificial neural networks (ANNs) as the essential building blocks of generative models can be explained as “reproductions” in the sense of copyright law.

## 2.1 Stocktaking: See No Evil, Hear No Evil, Say No Evil

Only a few voices in scholarly commentary have suggested that the complete or partial re-apparition of copyright-protected works in AI-generated output is due to these data’s being permanently stored “within” the trained model.<sup>16</sup> The vast majority, by contrast, denies that the copying or reproduction of training data occurs. Their main argument is based on the “functionality” of generative AI: Since its purpose is to generate “new” output, there is no “memorizing” of the training data.<sup>17</sup> In other words, since such models are not used as storage devices and do not aim at the reproduction of their input, there is no copying or replication in the models as such. What must not be, cannot be (so to speak). Instead, as most proponents are eager to add, the training data are internalized “in an abstract way” in the model network’s parameters as digital information, which per se seems to contradict the idea of a complete or partial copy or reproduction of works in the dataset.<sup>18</sup> Yet these commentators have largely failed to take a closer look under the hood of generative AI technology.

## 2.2 Analysis: Critique and Reconceptualization

The very fact that copyright-protected works that have been part of the training dataset are actually reproduced fully or in part when the trained models are being used implies that works from the input *must* be present somewhere inside the model. The existence of such copies or reproductions cannot be denied with regard to either the functionality of generative AI or the fact that the models’ inside is a black box. It is only the lack of an effective technical routine for the *ad hoc* retrieval of works from the interior of generative models that could give rise to doubts. But these doubts can be dispelled.

<sup>16</sup> See, e.g., von Welser (2023), p. 517; Pesch and Böhme (2023a), p. 1005; Kögel (2023a), p. 180; Wagner (2024), p. 299; Sasing-Wagenpfeil (2024), p. 242 and passim; Mezei (2024), p. 463; Schack (2024), p. 114; for an instructive analysis, see Cooper and Grimmelmann (2024), pp. 14 *et seq.*

<sup>17</sup> See, e.g., Käde (2021), pp. 74–75; Kögel (2023b), pp. 287–288; Söbbing and Schwarz (2023), p. 417; Baumann (2023), p. 3674; Maamar (2023), p. 482; Hofmann (2024b), p. 12; Konertz and Schönhof (2024a, b), p. 293; Konertz and Schönhof (2024b), p. 536; Käde (2024), pp. 176–177; in US literature, see, e.g., Sag (2024), p. 1912, and *passim*.

<sup>18</sup> Baumann (2023), p. 3674; Hofmann (2024b), p. 12; Heine (2024), p. 87; Käde (2024), p. 177; Kraetzig (2024), p. 698; Peukert (2024), p. 506.

### 2.2.1 Starting Point: The Objective Status

The reproduction of works inside generative AI is implied whenever the use of a model generates output that is a copy or replication of a copyright-protected work from the training dataset. In theory, this can be explained by the creation of a so-called vector space that exists inside the trained AI models. Within this vector space, there exists a statistically compressed and compacted version of the training input, at least with regard to those works that can actually be reproduced with a suitable prompt.

*2.2.1.1 The Prima Facie Case, or: Where Should Output Replications Come from After All?* Statutory EU copyright law does not expressly define “reproduction”.<sup>19</sup> Yet the concept of “reproduction” has been extensively explored, and it is clear that it must be interpreted broadly, primarily due to the fact that the InfoSoc Directive’s main objective is to introduce a high level of protection, particularly to enable authors to receive an appropriate reward for the use of their works.<sup>20</sup> As the Court of Justice of the European Union (CJEU) has repeatedly emphasized, such a broad interpretation must ensue from the Directive’s statutory text defining reproductions by utilizing expressions such as “direct or indirect”, “temporary or permanent”, “by any means”, and “in any form”.<sup>21</sup> Accordingly, it is widely undisputed that a “reproduction” can inter alia lie in a digital replication or digital storage of a copyright-protected work, regardless of the kind of technology that has been used to generate the digital copy.<sup>22</sup> Against this backdrop, what can be found inside generative models must also be considered reproductions of copyright-protected training data within the meaning of EU copyright law:

To begin with, there is no doubt that “something” exists that is embodied in the models’ inside. After all, the existence of reproductions of training data can and must be inferred from the fact that these models are actually able to reproduce at least some of the works from their input – unchanged or largely unaltered – when they are prompted to generate output.<sup>23</sup> Depending on the specific model and technology, as well as the operating conditions, verbatim copies of input data can occur in about 0.1 to 10% of cases.<sup>24</sup> As has been further shown, the replication of

<sup>19</sup> See, e.g., CJEU, case C-5/08 *Infopaq v DDF* [2009] ECLI:EU:C:2009:465, para. 31; see also CJEU, case C-476/17 *Pelham v Hütter* [2019] ECLI:EU:C:2019:624, para. 52.

<sup>20</sup> See Art. 2 and recital 21 InfoSoc Directive; see also CJEU, case C-5/08 *Infopaq v DDF* [2009] ECLI:EU:C:2009:465, para. 40.

<sup>21</sup> CJEU, case C-5/08 *Infopaq v DDF* [2009] ECLI:EU:C:2009:465, para. 42; CJEU, case C-433/20, *Austro-Mechana Gesellschaft zur Wahrnehmung mechanisch-musikalischer Urheberrechte Gesellschaft mbH v Strato AG* [2022], ECLI:EU:C:2022:217, para. 16; CJEU, case C-426/21, *Ocilion IPTV Technologies GmbH v Seven.One Entertainment Group GmbH, Puls 4 TV GmbH & Co. KG* [2023], ECLI:EU:C:2023:564, para. 28.

<sup>22</sup> See, e.g., CJEU, joined cases C-403 and 429/08, *Football Association Premier League v QC Leisure and Murphy v Media Protection Services* [2011] ECLI:EU:C:2011:631, para. 159; see also Schulze (2022), ‘§ 16 UrhG’ para. 7 with further references.

<sup>23</sup> Cooper and Grimmelmann (2024), pp. 14 *et seq.*; see also Sesting-Wagenpfeil (2024), pp. 234 *et seq.*

<sup>24</sup> See, e.g., Lee et al. (2022); Carlini et al. (2023a, b); see also Cooper and Grimmelmann (2024), pp. 48–49.

content from the stock of training data can be evoked by sophisticated prompting, particularly with respect to large language models and the literary works in their training dataset.<sup>25</sup> But why does the occurrence of specific output imply an identical input? This is a question of statistics: The spontaneous identical or largely identical creation by the generative model *ex nihilo* is more than unlikely whenever the reproduction concerns a copyright-protected work that has a minimum degree of sophistication.<sup>26</sup> In other words, very simple and banal elements – e.g., single sentences or a few words only – may come into existence *ex machina*, so to speak. There need not be an identical element in the dataset in order to enable the model to generate such simple output. Yet when a longer literary work – e.g., a journalistic article or the lyrics of a song – is ejected by ChatGPT and is identical or largely identical to the pre-existing work, there is little doubt that it must be a reproduction of the text or lyrics that were part of the underlying model's training data. After all, for longer, more elaborate works, the accidental creation during the process of output genesis is statistically and practically impossible. In the words of A Feder Cooper and James Grimmelmann:

[M]emorized content must be encoded in the model's parameters. There is nowhere else it could be. A model is not a magical portal that pulls fresh information from some parallel universe into our own. A model is a data structure: it consists of information derived from its training data. The memorized training data are *in the model*.<sup>27</sup>

This output-input connection can be verified by a simple self-experiment: ChatGPT, for instance, can be coaxed into confessing that certain works (such as lyrics of pop songs) *could* be replicated but that the system avoids doing this in order to prevent copyright infringement. The author asked ChatGPT during a conversation on November 6, 2024, “But you could provide the full lyrics [of ‘Yellow Submarine’ by the Beatles]? You are just not doing it for copyright reasons, correct?” The system answered:

Yes, that's correct. I have access to the lyrics, but I'm unable to provide them verbatim due to copyright restrictions. However, I'm here to help with anything else you might need regarding the song, such as its themes, background information, or a summary!

In addition to this factual status, it must be noted that the concept of “reproduction” in EU copyright law is objective and independent of whether an actor-copier acts consciously or even intentionally.<sup>28</sup> Acts of reproduction are determined by their result alone. Accordingly, it is the *objective existence* of a

<sup>25</sup> See, e.g., Henderson et al. (2023); Carlini et al. (2021); Chang et al. (2023); Carlini et al. (2023a, b); Somepalli et al. (2022); Somepalli et al. (2023); see also Pesch and Böhme (2023b), p. 919; Pesch and Böhme (2023a), p. 1001; Käde (2024), p. 180; Sag (2023), pp. 326 *et seq.*

<sup>26</sup> Cooper and Grimmelmann (2024), pp. 22–23 and 24–25; see also Sesing-Wagenpfeil (2024), p. 237 *et seq.*

<sup>27</sup> Cooper and Grimmelmann (2024), pp. 22–23 (emphasis in original).

<sup>28</sup> See, e.g., Götting (2024), ‘§ 16 UrhG’ para. 7 with further references.



“reproduction” that determines whether an act is classified as copyright-infringing.<sup>29</sup> In the same vein, it must not matter whether the evident penchant of generative models to reproduce training data is inherent in their technology or functionality.<sup>30</sup> For the same reason, it must be irrelevant whether AI developers explain such reproductions as a malfunction or as an undesirable defect of their models, or whether it is the end user who is aiming for a replication, deliberately prompting an AI system to generate infringing output.

In conclusion, the occurrence of reproducing output leaves no other explanation than the fact that there is “something” inside the model that is allowing for such output reproductions. That this “something” must be a replication itself, as we will see below, can be shown by a closer look at the AI technology.

*2.2.1.2 Statistical Whereabouts: Representation in the Models’ “Vector Space”* Asking for the actual “place” of reproductions within generative AI models requires a close look at the technology. As can be shown, the ingestion and digestion of copyright-protected works during a model’s training process ultimately materializes in a representation of at least some of these works in the model’s so-called vector space. This model-internal representation can be explained as a compressed and compacted reproduction of the training data:

The aim of generative AI training is to develop models whose output is similar to the input data.<sup>31</sup> In order to be able to produce such “similarity” in the post-training phase, the model must first identify patterns and correlations within the stock of its training dataset.<sup>32</sup> Thereby, each model “learns” the statistics of its training data – i.e., the properties of the elements in its dataset and the statistical correlations between these elements.<sup>33</sup> In more layperson’s terms, the model converts an endless amount of information in and about the training input data into a format that enables it to categorize those data according to similarity or dissimilarity, always dependent on what the model’s function is.<sup>34</sup> After the training has been completed, the statistical information about the whole input dataset is stored in the model’s central component, the ANN. The parameters or weights of the ANN are represented as

<sup>29</sup> Sesing-Wagenpfeil (2024), pp. 221, and 223 *et seq.* with further references.

<sup>30</sup> See also Pesch and Böhme (2023a, b), p. 1005; Sesing-Wagenpfeil (2024), p. 242.

<sup>31</sup> See, e.g., OpenAI’s definition at <https://openai.com/index/generative-models/> (“To train a generative model we first collect a large amount of data in some domain (e.g., think millions of images, sentences, or sounds, etc.) and then train a model to generate data like it.”).

<sup>32</sup> See, e.g., Cooper and Grimmelmann (2024), p. 9; see also Dornis and Stober (2024), p. 40.

<sup>33</sup> For statistical processing, see, e.g., Radford et al. (2021); see also Cooper and Grimmelmann (2024), p. 9.

<sup>34</sup> Radford et al. (2021); see also Sobel (2024), p. 16.

numbers.<sup>35</sup> In technical terms, this numerical representation in the artificial neural core of the model is referred to as mathematical vectors.<sup>36</sup>

This may sound abstract and theoretical, but there is a very concrete rub: The components of these mathematical vectors represent specific properties of the individual elements within the training dataset.<sup>37</sup> The transformation of these elements – including, in particular, the contents of copyright-protected works – into vectors in the model’s vector space results in both a simplification and a compression of the data to a limited number of properties identified by the model according to the task and function it has to fulfil.<sup>38</sup> Of course, during this process of “vectorization”, AI models do not differentiate between the semantics, which are not protected by copyright, and the syntax as the carrier of copyright-protected expression.<sup>39</sup> The vectorization is a merely statistical ingestion and digestion of what the model’s software has determined as the most relevant information within the training set. Accordingly, works ingested and digested during the AI training are being vectorized comprehensively and without differentiation, merely compressed to the properties required for the respective functional task.<sup>40</sup>

*2.2.1.3 Application: Substance and Function of Model-Internal Representations* This vectorial representation of the training data is directly related to the generation of the model’s output. In order to generate new content (i.e., data similar to the training input), the model needs to have access to the training data that are represented in the vector space. Accordingly, when generating output, the model must be able to decode the statistical vector information and transform it back into a visual representation and embodiment that can be perceived by the human senses.<sup>41</sup>

This input-output connection has been illustratively explored in AI research: Shan Carter and Michael Nielsen, for example, explain the functioning of the vector space of an ANN that has been trained on the basis of a large number of different fonts. They refer to a spatial model that can be roughly described as a “landscape”.<sup>42</sup> As can be seen in the Fig. 1, individual points on the “mountains” and in the “valleys” represent the properties of training data transformed into the

<sup>35</sup> For the functioning of ANNs and their relevance for the training of generative AI models, see, e.g., Dornis and Stober (2024), p. 27 *et seq.*

<sup>36</sup> In LLMs (large language models) based on transformer architecture, such as GPT models, the vectors are referred to as embeddings. In models with probabilistic internal representations such as GANs (generative adversarial networks), VAEs (variational autoencoders) and latent diffusion (e.g., stable diffusion), the vectors describe latent random variables in the so-called latent space of the model. See, e.g., Dornis and Stober (2024), pp. 29 *et seq.*, 43 *et seq.*, and 117 *et seq.*

<sup>37</sup> Sobel (2024), p. 16.

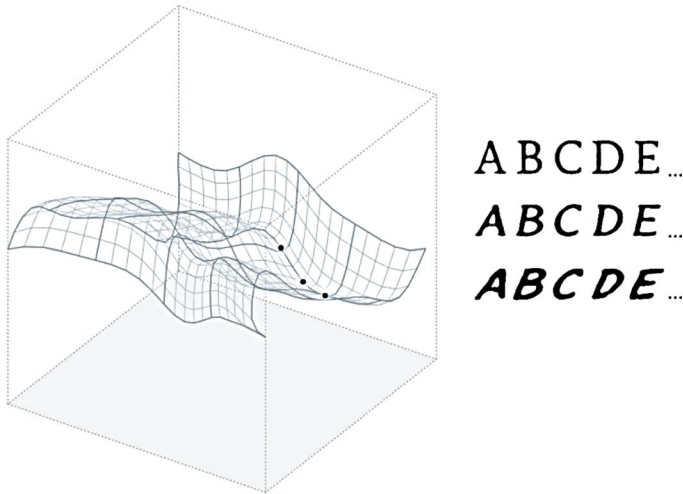
<sup>38</sup> Sobel (2024), p. 17; see also Cooper and Grimmelmann (2024), pp. 35 *et seq.*

<sup>39</sup> For the semantics/syntax divide and its irrelevance in generative AI training, see Dornis (2025), sub III.2.

<sup>40</sup> Cooper and Grimmelmann (2024), p. 51 (“There is a straightforward causal connection from the (expressive) training data through the model to the (expressive) output ...”); see also Dornis and Stober (2024), pp. 105 *et seq.*

<sup>41</sup> Ramesh et al. (2022); Sobel (2024), p. 18.

<sup>42</sup> Carter and Nielsen (2017).



**Fig. 1** Illustration according to Carter and Nielsen Available at (<https://distill.pub/2017/aia/>)

vector space. With an incremental change of the input – in the Carter-Nielsen illustration, this corresponds to a “walk” over the “landscape” – the output also changes. As can be seen, each step in the vector space results in a different font.

The variables in the vector space of the ANN are accordingly described as a compacted and abstracted representation of the dataset of fonts used for the model’s training. As the authors explain, the model has “learnt” a highly compressed version of all the fonts in its training data – i.e., the fonts available and scraped on the internet.<sup>43</sup> This relationship between input and output illustrates that the result of the statistical compression of the training data will rarely consist of an identical copy or a 1:1 replication of the training dataset and of all the single works included in the dataset. Nevertheless, the model-internal replication is directly and organically linked to the input data, or, as Carter and Nielsen explain:

The generative model we use is learnt from a training set of more than 50 thousand fonts scraped from the open web. ... In some sense, the model is learning a highly compressed representation of all the training fonts.<sup>44</sup>

Coming back to the InfoSoc Directive, it is important to remember that a “reproduction” can also exist if the original work has been simplified or reduced – e.g., made smaller during the reproduction process, such as by removing irrelevant content or redundant information. Accordingly, the transformation of pictures or images (e.g., photographs) by conversion into digital thumbnails is generally regarded as a “reproduction” in the sense of Art. 2 InfoSoc Directive.<sup>45</sup>

<sup>43</sup> Carter and Nielsen (2017); see also Cooper and Grimmelmann (2024), 35 *et seq.*

<sup>44</sup> See Carter and Nielsen (2017).

<sup>45</sup> For German copyright doctrine, see, e.g., Federal Court of Justice (BGH) GRUR 2010, 628 (629) – *Vorschaubilder I*; see also Schulze (2022), ‘§ 16 UrhG’ paras. 10–11 with further references. Courts in the United States have decided in the same vein. See, e.g., *Perfect 10, Inc. v. Amazon.com*, 508 F.3d 1146 (9th Cir. 2007).

### 2.2.2 The Issue of “Perceptibility”: Technology-Neutral and Open to Development

Lawmakers and courts (e.g., in Germany) require that in order to be considered a “reproduction”, the embodiment of a work must be “directly or indirectly perceptible to the human senses in some way”.<sup>46</sup> Some scholars tend to over-extend this concept of perceptibility by referring to the CJEU’s case law, thereby drastically misunderstanding the difference between the original “work” and the work’s “reproduction”: As they explain, the CJEU requires a “work” to be “expressed in a manner which makes it identifiable with sufficient precision and objectivity”.<sup>47</sup> This doctrine, they argue, also requires the “reproductions” to be permanently “identifiable”. The conclusion is as simple as it is incorrect: Since the AI model replicates copyright-protected works inside its “black box”, for which there is no possibility of looking inside, there cannot be a “reproduction”.<sup>48</sup> What is overlooked is the fact that the prerequisite of identifiability defined by the CJEU is due to the need of public authorities and private actors, mainly competitors, to be able to understand what *actually* is the subject matter of protection.<sup>49</sup> This, however, concerns only the “work” as such and its protected subject matter. One needs more than a microscope to find a rule on the existence of or any prerequisite for the finding of “reproductions” in these CJEU precedents.

Under a better calibrated and streak-free lens on copyright doctrine, the concept of “reproduction” must rather be explained as being distinctly technology-neutral. This means that it is irrelevant how the reproduction is embodied, or whether technical means or instruments are required to make the embodiment perceptible – i.e., visible, audible, or tangible – for humans. The only requirement is that the reproduction must be fixed on a physical carrier from which it is *possible* to establish perceptibility.<sup>50</sup> Against this backdrop – and not only under the CJEU’s doctrine – there is no doubt that digital-file versions of a work (for example, those stored on a CD, DVD, hard disc, or a cloud storage device) are to be regarded as

<sup>46</sup> See, e.g., Federal Court of Justice (BGH) GRUR 2017, 793 para. 41 – *Mart-Stam-Stuhl*; Federal Court of Justice (BGH) GRUR 2001, 51 (52) – *Parfümflakon I*; Götting (2024), ‘§ 16 UrhG’ para. 3 with further references.

<sup>47</sup> CJEU, case C-310/07, *Levola Hengelo BV v Smilde Foods BV* [2018], ECLI:EU:C:2018:899, para. 40 (“Accordingly, for there to be a ‘work’ as referred to in Directive 2001/29, the subject matter protected by copyright must be expressed in a manner which makes it identifiable with sufficient precision and objectivity, even though that expression is not necessarily in permanent form.”); see also CJEU, case C-683/17, *Cofemel – Sociedade de Vestuário SA v G-Star Raw CV* [2019], ECLI identifier: ECLI:EU:C:2019:721, para. 32.

<sup>48</sup> Leistner (2024), 1665 sub. III.2.

<sup>49</sup> CJEU, case C-310/07, *Levola Hengelo BV v Smilde Foods BV* [2018], ECLI:EU:C:2018:899, para. 41 (“That is because ... the authorities responsible for ensuring that the exclusive rights inherent in copyright are protected must be able to identify, clearly and precisely, the subject matter so protected. The same is true for individuals, in particular economic operators, who must be able to identify, clearly and precisely, what is the subject matter of protection which third parties, especially competitors, enjoy.”); see also CJEU, case C-683/17, *Cofemel – Sociedade de Vestuário SA v G-Star Raw CV* [2019], ECLI identifier: ECLI:EU:C:2019:721, para. 33.

<sup>50</sup> See, e.g., Heerma (2022), ‘§ 16 UrhG’ paras. 4–5; Schulze (2022), ‘§ 16 UrhG’ para. 6 (with Braille as an example); Sesing-Wagenpfeil (2024), p. 229.

“reproductions” in the copyright sense.<sup>51</sup> Also, as explained earlier, a work need not remain unchanged during the reproduction process, namely if it is reduced and compressed. Finally, even separating and splitting a work into single components – e.g., during the routing in packet switching networks – will ultimately lead to a reproduction as long as the work is ultimately re-assembled.<sup>52</sup>

In addition to being technology-neutral, the concept of “reproduction” in EU copyright law – as in other copyright regimes – must be functional and, hence, flexible and open to new technological developments. After all, copyright doctrine has always been able to capture new and inventive encroachments on copyright owners’ exploitation rights.<sup>53</sup> The fact that the works used in AI training are being statistically transformed and stored in the models’ vector space therefore does not imply a lack of “perceptibility”. It is true that the relative novelty of generative AI technology, as well as the excessive size of generative models and the abstraction required to explain their functioning, complicates legal analysis. However, such epistemological obstacles have never stopped copyright lawyers in the past from explaining how replications that are *actually possible* can be “reproductions” that are *legally relevant*. In fact, there is no difference from other variants of digital storage, such as on computer hard discs:<sup>54</sup> In these cases, too, it is not required that the process of creating, organizing, and ultimately unbundling the digital storage follow a specific technique or system, or that it preserve the nature or format of the original work. Accordingly, the concrete architecture of reproductions within the vector space of generative AI models (if there exists an architecture at all) must not be determinative. Ultimately, at least when utilizing a suitable *prompt*, a replication may emerge from the model’s inside that reproduces the entire work or a substantial part of it. This will suffice.

The ultimate characteristic of a reproduction – understood in such a functional sense – is that one can replicate the work from the reproduction.<sup>55</sup> The specific technicalities of such replication, especially the processes inside generative models’ black boxes, are largely irrelevant for the legal assessment.

### 2.2.3 Gretchen’s Question<sup>56</sup>: *Ad hoc Access or Not?*

Ultimately, one more problem exists: There is no effective routine or standardized method of eliciting works from within the model’s interior. At least at the moment,

<sup>51</sup> See, e.g., CJEU, joined cases C-403 and 429/08, *Football Association Premier League v QC Leisure and Murphy v Media Protection Services* [2011] ECLI:EU:C:2011:631, para. 159; from German doctrine, see, e.g., Federal Court of Justice (BGH) GRUR 2010, 616 (619) – *marions-kochbuch.de*; see also Schack (2021), para. 435; Schulze (2022), ‘§ 16 UrhG’ para. 7; Wagner (2024), p. 299; Käde (2024), p. 181; Selsing-Wagenpfeil (2024), p. 236.

<sup>52</sup> While the reproduction on intermediate stages of the routing process, in particular the disassembly into components, may be controversial, there is wide agreement that the final assembly is to be seen as a reproduction. See, e.g., Loewenheim (2020), ‘§ 16 UrhG’ para. 23; Pesch and Böhme (2023a, b), p. 1002.

<sup>53</sup> Selsing-Wagenpfeil (2024), pp. 241–242 (with fn. 148); see also Litman (1989), p. 275 and *passim*.

<sup>54</sup> Cooper and Grimmelmann (2024), p. 30 *et seq.*; see also Pesch and Böhme (2023a), p. 1005.

<sup>55</sup> See also, for US doctrine, Cooper and Grimmelmann (2024), p. 20.

<sup>56</sup> The term “Gretchenfrage” originates from Johann Wolfgang von Goethe’s play *Faust, Part One*. It refers to a pivotal scene where the character Gretchen (also known as Margarete) asks Faust, “*Nun sag,*

it is not possible to purposefully and immediately retrieve works on an *ad hoc* basis. Rather, the phenomenon of actual reproduction during the use of AI models is described as a technical defect and as an unwanted side-effect (so-called overfitting) that AI developers are keenly attempting to eliminate.<sup>57</sup>

However, a routine method for “perceptualization” has never been explained as a necessary requirement for the existence of a “reproduction” under copyright law, whether by parliamentary lawmakers or by courts.<sup>58</sup> Of course, access to copies and reproductions of an original work may be important in order to prepare a lawsuit. Ultimately, proof that a copyright has been infringed typically requires evidence of actual copying. Yet under EU copyright doctrine and other laws, what matters when the issue is “reproduction” is the *actual existence* of an embodiment of the original work. Access to this embodiment is merely a follow-up question for litigation.

In this respect, scholarly commentary has correctly pointed out that access is also irrelevant in other cases of reproduction. One example is digital replications – e.g., on computer hard discs, where the functionality or accuracy of retrieval methods is irrelevant. A typical example is software bugs. Malfunctions of the software that has the function to retrieve data from the storage device may occur. Yet even if the technology is prone to errors, this does not change the fact that an embodiment does exist and that it must be classified as a “reproduction”.<sup>59</sup> In the same vein, commentators have pointed out that a “reproduction” of password-protected data may exist regardless of whether the password has been lost and forgotten and, hence, access to the data might be lost forever.<sup>60</sup>

Finally, it is by no means impossible that over time – assuming a better understanding of what is going on inside the black box of generative AI models – more technical options of input-data retrieval may become known. AI research on this question may still be in its infancy.<sup>61</sup> Yet here as well, a look at other ways of “information storage” is helpful. One of these examples, explained by A. Feder Cooper and James Grimmelmänn, is historical papyrus scrolls that could not be “read” for a long time because any attempt at defoliation would have destroyed the scrolls. Despite these “technological” obstacles (which were ultimately overcome), there was never a doubt that “inside” these scrolls one would find an actual embodiment of information – a “reproduction” in the sense of copyright law.<sup>62</sup>

---

Footnote 56 continued

wie hast du's mit der Religion?” which translates to “Now tell me, how do you feel about religion?” This question challenges Faust to reveal his true beliefs and convictions. In a broader context, a “*Gretchenfrage*” has come to mean any crucial or pointed question that can reveal deep truths about a person's beliefs or intentions, often putting them in a difficult position where they have to reveal something significant about themselves. It is a metaphor for a question that probes for a fundamental truth. (This splendid explanation was given by ChatGPT – thanks a lot!).

<sup>57</sup> On this basis, the prevailing view denies a reproduction within the generative AI models. *See supra* 2.1.

<sup>58</sup> For the misinterpretation of the CJEU's case law in German legal scholarship, *see supra* 2.2.2.

<sup>59</sup> *See* Cooper and Grimmelmänn (2024), pp. 45–46; *see also* Sesing-Wagenpfeil (2024), pp. 239 *et seq.*

<sup>60</sup> *See* Pesch and Böhme (2023a, b), p. 1005; Sesing-Wagenpfeil (2024), p. 246 (with fn. 175).

<sup>61</sup> Dornis and Stober (2024), pp. 63–64.

<sup>62</sup> Cooper and Grimmelmänn (2024), pp. 28–29.

### 2.3 Summary

A thorough look into the engine room of generative AI models reveals that reproductions of works utilized for training may occur inside the models – specifically in the models’ vector space. At the moment it seems as if these reproductions cannot be easily retrieved *ad hoc*. Also, it is not possible to retrieve *each and every* work used for training. However, we can explain an objective status that implies that reproductions of many works – texts, pictures, music, and the like – *do exist* since reproductions of these works in the AI-generated output simply cannot materialize *ex nihilo*. Particularly in light of the broad interpretation of the term “reproduction” under the CJEU’s doctrine, much speaks in favour of finding reproductions inside the trained AI models, at least of the works that can be retrieved with a suitable prompt. In any event, the current state of ignorance about the inner workings of AI models and their ANNs must not be determinative for the legal analysis. Any tendency to let AI developers hide behind an argument of “we don’t know the monster we created” would unduly burden right owners with the toll for innovation.

### 3 Copyright Infringement: “Making Available to the Public”

Against this background, it is arguable that, by offering AI models for download or by offering generative services (e.g., ChatGPT) to European end users, the developers provide access to the reproductions of copyright-protected works inside their trained models. This concerns the works that can actually be retrieved by a suitable prompt. In terms of EU copyright law, the issue is a copyright-infringing “making available to the public” within the meaning of Art. 3(1) InfoSoc Directive.

#### 3.1 Overview: CJEU Doctrine on Art. 3(1) InfoSoc Directive

Article 3(1) InfoSoc Directive requires that

Member States shall provide authors with the exclusive right to authorise or prohibit any communication to the public of their works, by wire or wireless means, including the making available to the public of their works in such a way that members of the public may access them from a place and at a time individually chosen by them.

The right of making available to the public obviously is a component of the right to communicate. Accordingly, the concepts must be interpreted in tandem. The CJEU’s doctrine on Art. 3(1) is fairly sparse and, to date, has not specifically addressed AI-related issues. Yet some bright lines of the Court’s interpretation of Art. 3(1) exist:

Under the CJEU’s doctrine, the concept of “communication to the public” includes two cumulative criteria, namely (1) an “act of communication” of a work



and (2) the communication of that work to a “public”.<sup>63</sup> In addition, the CJEU has defined a number of other “complementary criteria, which are not autonomous and are interdependent”, that must be considered. These different criteria must be applied not only individually but also “in their interaction with one another”.<sup>64</sup> Among those criteria are the “indispensable role played by the user” and the “deliberate nature of its intervention”. This requires, in particular, looking at whether the act of communication occurs “in full knowledge of the consequences”.<sup>65</sup> With regard to the offering of generative AI models for download or the offering of model-based AI services, the open issue is whether the first criterion – an “act of communication” or, rather, a “making available” with the option of access from an individually chosen place and at a chosen time – can be found. By contrast, there does not seem to be any doubt concerning the second criterion, namely the “public” involved. Also, the CJEU’s complementary criteria, particularly with respect to the knowledge and deliberate conduct of AI developers, can be found.

To begin with the less problematic issues, the CJEU’s concept of communication to the “public” refers to an “indeterminate number of potential recipients”, implying “a fairly large number of persons”.<sup>66</sup> There does not need to be immediate access to all members of this public; it suffices if access is granted successively to each member, one after the other.<sup>67</sup> Neither the offer for download nor the online offering of model-based AI services (e.g., ChatGPT) raises any doubts about their

<sup>63</sup> See, e.g., CJEU, case C-753/18, *Stim and SAMI v Fleetmanager Sweden and Nordisk Biluthyrning* [2020] ECLI:EU:C:2020:268, para. 30; CJEU, case C-392/19, *VG Bild-Kunst v Stiftung Preußischer Kulturbesitz* [2021] ECLI:EU:C:2021:181, para. 29; CJEU, joined cases C-682 and 683/18, *Peterson v Google and Elsevier v Cyando (YouTube and Cyando)* [2021] ECLI:EU:C:2021:503, para. 66; CJEU, joined cases C-775 and C-826/21, *Blue Air Aviation v UCMR and UPFR v CFR* [2023] ECLI:EU:C:2023:307, para. 47.

<sup>64</sup> See, e.g., CJEU, case C-160/15, *GS Media v Sanoma Media Netherlands* [2016] ECLI:EU:C:2016:644, paras. 32 *et seq.*; CJEU, case C-527/15, *Stichting Brein v Wullems* [2017] ECLI:EU:C:2017:300, paras. 28 *et seq.*; CJEU, case C-610/15, *Stichting Brein v Ziggo and XS4ALL Internet* [2017] ECLI:EU:C:2017:456, paras. 23 *et seq.*; CJEU, case C-161/17, *Land Nordrhein-Westfalen v Renckhoff* [2018] ECLI:EU:C:2018:634, paras. 19 *et seq.*; CJEU, case C-392/19, *VG Bild-Kunst v Stiftung Preußischer Kulturbesitz* [2021] ECLI:EU:C:2021:181, para. 34; CJEU, joined cases C-775 and C-826/21, *Blue Air Aviation v UCMR and UPFR v CFR* [2023] ECLI:EU:C:2023:307, paras. 48 *et seq.*

<sup>65</sup> See, e.g., CJEU, case C-160/15, *GS Media v Sanoma Media Netherlands* [2016] ECLI:EU:C:2016:644, paras. 32 *et seq.*; CJEU, case C-527/15, *Stichting Brein v Wullems* [2017] ECLI:EU:C:2017:300, paras. 28 *et seq.*; CJEU, case C-610/15, *Stichting Brein v Ziggo and XS4ALL Internet* [2017] ECLI:EU:C:2017:456, paras. 23 *et seq.*; CJEU, case C-161/17, *Land Nordrhein-Westfalen v Renckhoff* [2018] ECLI:EU:C:2018:634, paras. 19 *et seq.*; CJEU, case C-392/19, *VG Bild-Kunst v Stiftung Preußischer Kulturbesitz* [2021] ECLI:EU:C:2021:181, paras. 34 *et seq.*; CJEU, joined cases C-775 and C-826/21, *Blue Air Aviation v UCMR and UPFR v CFR* [2023] ECLI:EU:C:2023:307, paras. 48 *et seq.*

<sup>66</sup> See, e.g., CJEU, case C-306/05, *SGAE v Rafael* [2006] ECLI:EU:C:2006:764, paras. 37–38; CJEU, case C-466/12, *Svensson v Retriever Sverige* [2014] ECLI:EU:C:2014:76, para. 21; CJEU, case C-527/15, *Stichting Brein v Wullems* [2017] ECLI:EU:C:2017:300, para. 44; CJEU, case C-610/15, *Stichting Brein v Ziggo and XS4ALL Internet* [2017] ECLI:EU:C:2017:456, para. 41; CJEU, case C-161/17, *Land Nordrhein-Westfalen v Renckhoff* [2018] ECLI:EU:C:2018:634, para. 22; CJEU, joined cases C-775 and C-826/21, *Blue Air Aviation v UCMR and UPFR v CFR* [2023] ECLI:EU:C:2023:307, para. 52.

<sup>67</sup> See, e.g., CJEU, case C-162/10, *Phonographic Performance v Ireland* [2012] ECLI:EU:C:2012:141, para. 35; CJEU, case C-351/12 *OSA v Léčebné lázně* [2014] ECLI:EU:C:2014:110, para. 28; CJEU, case C-117/15, *Reha Training v GEMA* [2016] ECLI:EU:C:2016:379, para. 44.



being intended for use by a large number of users. The AI companies' websites are freely accessible, intentionally addressing end users from around the world, including the EU.<sup>68</sup> Moreover, there is little doubt that the subjective elements defined by the CJEU, particularly knowledge and deliberateness on the side of the AI developers, are present.<sup>69</sup> The operators are undoubtedly aware that identical or substantially identical replications of works from the stock of training data in the output of the systems are possible and do actually occur.<sup>70</sup> Finally, most generative AI applications are made available commercially.

Therefore, it is the “making available” of works inside the generative models, offering “access from a place and at a time individually chosen”, that is the crucial issue to consider when determining whether copyright infringement exists.

### 3.2 Crucial Issue: The “Making Available”

The “act of communication”, as well as an “act of making available”, requires making the work available for individual access.<sup>71</sup> This act technically precedes third parties' actual access and their subsequent acts of retrieval and reproduction.<sup>72</sup> This implies that a “making available” is completed at the moment the providing actor does not need to do anything else – for example, activating an option for end users to access the work online for download on a website.<sup>73</sup> In other words, retrieval and transmission must be possible without further action by the providing actor.<sup>74</sup> It therefore must also be considered irrelevant whether a work made available was *actually* accessed, retrieved, or reproduced by the end user.<sup>75</sup> On this basis, it is also irrelevant whether there is a subsequent transmission of the work or

<sup>68</sup> The fact that the works cannot be retrieved without a suitable prompt and on an *ad hoc* basis does not change the “public” nature of the making available. It is rather an issue of “access”. See *infra* 3.2.

<sup>69</sup> See, e.g., CJEU, case C-162/10, *Phonographic Performance v Ireland* [2012] ECLI:EU:C:2012:141, para. 31; CJEU case C-135/10, *SCF v Marco Del Corso* [2012] ECLI:EU:C:2012:140, paras. 82, 91, and 94; CJEU, joined cases C-682 and 683/18, *Peterson v Google and Elsevier v Cyando (YouTube and Cyando)* [2021] ECLI:EU:C:2021:503, paras. 78 *et seq.*

<sup>70</sup> Finally, there can be no doubt about the “receptiveness” of the public that has been addressed. After all, individual users of AI systems like ChatGPT and the like are themselves intentionally aiming at a certain output. For the prerequisite that the public must be “receptive”, see, e.g., CJEU case C-135/10, *SCF v Marco Del Corso* [2012] ECLI:EU:C:2012:140, paras. 91 and 98.

<sup>71</sup> See, e.g., CJEU, case C-466/12, *Svensson v Retriever Sverige* [2014] ECLI:EU:C:2014:76, para. 19; CJEU, case C-610/15, *Stichting Brein v Ziggo and XS4ALL Internet* [2017] ECLI:EU:C:2017:456, paras. 23 *et seq.*, 35; CJEU, case C-161/17, *Land Nordrhein-Westfalen v Renckhoff* [2018] ECLI:EU:C:2018:634, para. 20.

<sup>72</sup> v. Ungern-Sternberg (2020), ‘§ 19a UrhG’ para. 4.

<sup>73</sup> For case law in German doctrine, see, e.g., Federal Court of Justice (BGH) GRUR 2018, 178 para. 30 – Vorschaubilder III; Federal Court of Justice (BGH) GRUR 2021, 1511 para. 30 – *Deutsche Digitale Bibliothek II*; Circuit Court (OLG) Stuttgart MMR 2008, pp. 474 (475).

<sup>74</sup> v. Ungern-Sternberg (2020), ‘§ 19a UrhG’ paras. 57 and 61.

<sup>75</sup> See, e.g., CJEU, case C-466/12, *Svensson v Retriever Sverige* [2014] ECLI:EU:C:2014:76, para. 19; CJEU, case C-314/12 *UPC Telekabel v Constantin Film* [2014] ECLI:EU:C:2014:192, paras. 43 *et seq.*; CJEU, case C-527/15, *Stichting Brein v Willems* [2017] ECLI:EU:C:2017:300, para. 36; CJEU, case C-161/17, *Land Nordrhein-Westfalen v Renckhoff* [2018] ECLI:EU:C:2018:634, para. 20.

whether the work is made perceptible after the granting of access.<sup>76</sup> In short, it does not matter what happens after the “making available” has been completed.

This is the only interpretation that is in line with the ratio of the InfoSoc Directive and the general stance of EU copyright doctrine: To begin with, the concepts of “communication” and “making available” are to be understood broadly in view of the main objective of the InfoSoc Directive to ensure a high level of protection for authors.<sup>77</sup> Accordingly, an infringement will already occur with acts that precede the *actual* use of the works. When finding a “making available to the public”, it must not matter that the public may not be able to access the work or other subject matter *immediately* or that the public, in order to have access, has to overcome *additional obstacles*. Hence, the mere delay or a time lag between the provider’s last act and users’ actual access does not preclude a making available to the public.<sup>78</sup> In addition, paywalls cannot be considered to preclude a “making available”.<sup>79</sup>

What is most important when it comes to technical obstacles that may stand in the way of the public’s actual access to works that have been made available is the fact that the CJEU interprets “communication” and “making available” in a *technology-neutral manner*. The concept of “communication” must be considered to refer “to any transmission of the protected works, irrespective of the technical means or process used”.<sup>80</sup> Along these lines, the Court in *Mircom/Telenet* ruled that the provision – i.e., offering for download – of single segments of files via a peer-to-peer network had to be considered a “making available”, even though these single segments could be used only after they had been re-assembled on the users’ computers after a certain percentage of all the segments of each file had been downloaded. Hence, the individual users would make the work available (as peers in the network) already without having previously downloaded the complete original file (i.e., all segments) with the protected work or a certain “minimum quantity” of the segments.<sup>81</sup> In other words, the CJEU did not consider it necessary for an infringement that the users have immediate, complete, and technically unlimited access.

<sup>76</sup> v. Ungern-Sternberg (2020), ‘§ 19a UrhG’ para. 4; Schack (2021), para. 482.

<sup>77</sup> See recitals 4, 9, 10, and 23 InfoSoc Directive; see also, e.g., CJEU, case C-306/05, *SGAE v Rafael* [2006] ECLI:EU:C:2006:764, para. 36; CJEU, joined cases C-403 and 429/08, *Football Association Premier League v QC Leisure and Murphy v Media Protection Services* [2011] ECLI:EU:C:2011:631, para. 193; CJEU, case C-466/12, *Svensson v Retriever Sverige* [2014] ECLI:EU:C:2014:76, para. 17; CJEU, case C-160/15, *GS Media v Sanoma Media Netherlands* [2016] ECLI:EU:C:2016:644, paras. 29 *et seq.*; CJEU, case C-161/17, *Land Nordrhein-Westfalen v Renckhoff* [2018] ECLI:EU:C:2018:634, para. 18; CJEU, joined cases C-682 and 683/18, *Peterson v Google and Elsevier v Cyando (YouTube and Cyando)* [2021] ECLI:EU:C:2021:503, paras. 62 *et seq.*

<sup>78</sup> Bullinger (2022), ‘§ 19a UrhG’ para. 9.

<sup>79</sup> v. Ungern-Sternberg (2020), ‘§ 19a UrhG’ para. 61.

<sup>80</sup> See, e.g., CJEU, case C-597/19, *Mircom v Telenet* [2021] ECLI:EU:C:2021:492, para. 47; see also CJEU, joined cases C-403 and 429/08, *Football Association Premier League v QC Leisure and Murphy v Media Protection Services* [2011] ECLI:EU:C:2011:631, paras. 186 and 193; CJEU, case C-117/15, *Reha Training v GEMA* [2016] ECLI:EU:C:2016:379, para. 38.

<sup>81</sup> CJEU, case C-597/19, *Mircom v Telenet* [2021] ECLI:EU:C:2021:492, para. 59; cf. also Federal Court of Justice (BGH) GRUR 2018, 400 para. 15 – Konferenz der Tiere.

Advocate General Szpunar made this point even clearer – also in *Mircom/Telenet* – expressly focusing on the infringers’ penchant for referring to technical details in order to escape liability. He explained:

[D]o the users of a peer-to-peer network themselves carry out acts of communication to the public? While at first sight it may appear obvious that this should be answered in the affirmative, ingenious arguments based on the specific technical features of the operation of such networks support the opposite response. The fact that thousands of people have access to works without paying for them would then be a miracle. In the present case, the Court will have the opportunity to clarify this point.<sup>82</sup>

The Court did use this opportunity and clarified that “ingenious arguments based on specific technical features” must not provide an exit if massive infringements actually happen. Both the CJEU’s doctrine and the Advocate General’s summary of a technology-neutral policy imply the following: If access is possible and does actually occur, the technical details should be irrelevant.

Seen in this light, granting online access to trained AI models and thus to the model-interior replications of copyright-protected works may very well be considered a “making available to the public”. After all, not only is it theoretically possible – using a suitable prompt – to evoke works that are in the stock of the training data, but such reproductions actually do occur when generative AI systems like ChatGPT and the like are being employed by end users. *In toto*, this is a situation of endangerment in the sense of the CJEU’s doctrine.<sup>83</sup>

### 3.3 Policymaking: Who Should Bear the Black-Box Risk?

Ultimately, of course, it is the CJEU who will have to clarify the scope of Art. 3(1) InfoSoc Directive. The scarcity of the Court’s existing precedents complicates a clear and ultimate determination of the scope of the “right of making available to the public”. However, the rationale of the InfoSoc Directive – i.e., providing broad and effective copyright protection – implies that a lack of information and uncertainty about the technological foundations and the detailed processes in the darkness of the models’ interior must not be to the detriment of the right holders:

First, it should be recalled that the CJEU, with its system of flexible criteria, has created a framework for weighing all interests involved, with particular emphasis on

<sup>82</sup> CJEU, case C-597/19, *Mircom v Telenet* [2021] ECLI:EU:C:2021:492, Opinion of AG Szpunar, para. 2.

<sup>83</sup> The case might be different if model-internal filters could completely prevent the reproduction of training data in the model’s output. Users of AI services would then have no access to the works inside the model. However, AI systems currently on the market do not have effective filters since it is technically impossible to perfectly prevent output replication. See, e.g., Guadamuz (2024), pp. 121 *et seq.*; Sesing-Wagenpfeil (2024), pp. 246 *et seq.*; see also Dornis and Stober (2024), pp. 60 *et seq.*

the protection of the authors' remuneration interests.<sup>84</sup> Accordingly, as the Court has repeatedly explained:

It is ... clear ... that the principal objective of [the InfoSoc Directive] is to establish a high level of protection of ... authors, allowing them to obtain an appropriate reward for the use of their works, in particular on the occasion of communication to the public.<sup>85</sup>

In addition, as the Court has repeatedly explained, Art. 3(1) InfoSoc Directive grants the holders a right which is “preventive in nature” and “allows them to intervene, between possible users of their work and the communication to the public which such users might contemplate making, in order to prohibit such use”.<sup>86</sup> In addition, this right means that

the author of a work must be able to put an end to the exercise, by a third party, of rights of exploitation in digital format that he holds on that work, and in so doing prohibit him from any future use in such a format, without having to submit beforehand, in certain circumstances, to a formality consisting of proving that other persons are not, otherwise, holders of other rights in that work, such as those concerning its exploitation in printed format.<sup>87</sup>

Against this background, in situations of uncertainty about the functioning of technological features and the risk of exploitation of works, it is questionable whether the CJEU would consider the right holder to be the ultimate bearer of risks associated with new technologies and with the vagaries of their functioning: The mere fact that a simple method or routine of *ad hoc* evocation of works from the models' inside – as seen above – has yet to be found, understood, and developed

<sup>84</sup> See, e.g., CJEU, case C-160/15, *GS Media v Sanoma Media Netherlands* [2016] ECLI:EU:C:2016:644, paras. 32 *et seq.*; CJEU, case C-392/19, *VG Bild-Kunst v Stiftung Preußischer Kulturbesitz* [2021] ECLI:EU:C:2021:181, paras. 34 *et seq.*; CJEU, joined cases C-682 and 683/18, *Peterson v Google and Elsevier v Cyando (YouTube and Cyando)* [2021] ECLI:EU:C:2021:503, paras. 62 *et seq.*; CJEU, joined cases C-775 and C-826/21, *Blue Air Aviation v UCMR and UPFR v CFR* [2023] ECLI:EU:C:2023:307, para. 46.

<sup>85</sup> See, e.g., CJEU, case C-306/05, *SGAE v Rafael* [2006] ECLI:EU:C:2006:764, para. 36; CJEU, joined cases C-403 and 429/08, *Football Association Premier League v QC Leisure and Murphy v Media Protection Services* [2011] ECLI:EU:C:2011:631, paras. 184 *et seq.*; CJEU, case C-610/15, *Stichting Brein v Ziggo and XS4ALL Internet* [2017] ECLI:EU:C:2017:456, paras. 21–22; CJEU, case C-392/19, *VG Bild-Kunst v Stiftung Preußischer Kulturbesitz* [2021] ECLI:EU:C:2021:181, paras. 26 *et seq.*; CJEU, case C-597/19, *Mircom v Telenet* [2021] ECLI:EU:C:2021:492, para. 58; CJEU, joined cases C-775 and C-826/21, *Blue Air Aviation v UCMR and UPFR v CFR* [2023] ECLI:EU:C:2023:307, para. 46.

<sup>86</sup> See, e.g., CJEU, case C-117/15, *Reha Training v GEMA* [2016] ECLI:EU:C:2016:379, para. 30; CJEU, case C-301/15, *Soulhier and Doke v Premier ministre et al* [2016] ECLI:EU:C:2016:878, para. 33; CJEU, case C-161/17, *Land Nordrhein-Westfalen v Renckhoff* [2018] ECLI:EU:C:2018:634, para. 29 *et seq.*; CJEU, case C-392/19, *VG Bild-Kunst v Stiftung Preußischer Kulturbesitz* [2021] ECLI:EU:C:2021:181, para. 21; CJEU, joined cases C-682 and 683/18, *Peterson v Google and Elsevier v Cyando (YouTube and Cyando)* [2021] ECLI:EU:C:2021:503, para. 62.

<sup>87</sup> See, e.g., CJEU, case C-301/15, *Soulhier and Doke v Premier ministre et al* [2016] ECLI:EU:C:2016:878, para. 51; CJEU, case C-161/17, *Land Nordrhein-Westfalen v Renckhoff* [2018] ECLI:EU:C:2018:634, para. 31; CJEU, joined cases C-682 and 683/18, *Peterson v Google and Elsevier v Cyando (YouTube and Cyando)* [2021] ECLI:EU:C:2021:503, para. 62; CJEU, joined cases C-775 and C-826/21, *Blue Air Aviation v UCMR and UPFR v CFR* [2023] ECLI:EU:C:2023:307, para. 44.

must not restrict the scope of protection under the Directive. Accordingly, AI providers should not be able to rely on the opalescence and uncontrollability of their models. Rather, one would have to conclude that as long as works in the training dataset can actually be retrieved from within the model, there is no valid argument that the right holders concerned are only insignificantly affected. A quantitative minimum threshold below which right holders would simply have to accept an infringement would considerably shorten legal protection, contrary to the Directive's policy of *broad* protection.

This last aspect becomes evident when looking at the real numbers: In June 2024, ChatGPT alone had around 2.9 billion visits worldwide, in October 3.7 billion, and counting.<sup>88</sup> Even if the rate of infringing output of works from the training dataset were in the per mil range only (i.e., below 0.1%), millions of infringing AI products would have to be expected every month for this single generative AI system alone. This illustrates that offering generative AI services to the public can seriously interfere with authors' interests in exploitation.<sup>89</sup> In the same vein, under a practical lens, it should also be noted that right holders will hardly be able to identify the infringing reproductions of their works in the stream (or rather flood) of daily, monthly, and yearly AI-generated output around the world. Taking into account these practical obstacles of effective rights protection further implies that protection under Art. 3(1) InfoSoc Directive must be granted.

#### 4 Choice-of-Law and Jurisdiction

As seen above, the debate on copyright infringement and generative AI training is limited largely to the question whether collection, compilation, or preparation of the training datasets and the training processes are copyright infringement. It is therefore no surprise that the application of EU copyright law – or rather, EU Member States' copyright laws – as well as the international jurisdiction of Member States' civil-law courts is denied. After all, as the principle of copyright territoriality implies, only infringing acts *within* the EU can fall under the reign of EU copyright law. Yet if one is willing to see that there might exist reproductions of copyright-protected works inside the models, the logical consequence is that an offer of downloading options for AI models and of model-based services like ChatGPT and Stable Diffusion may be considered a “making available to the public” within the meaning of Art. 3(1) InfoSoc Directive. Accordingly, EU copyright law should apply. Furthermore, civil-law courts should have international jurisdiction over developers and providers of AI services.

<sup>88</sup> See, e.g., the statistics on similarweb (available at <https://www.similarweb.com/website/chatgpt.com/#ranking>); see also Nina Müller, 'ChatGPT erreicht 2,9 Milliarden monatliche Besucher' (FAZ.net, 9 July 2024) [www.faz.net/pro/digitalwirtschaft/kuenstliche-intelligenz/neuer-rekord-chatgpt-erreicht-2-9-milliarden-besucher-pro-monat-19845849.html](https://www.faz.net/pro/digitalwirtschaft/kuenstliche-intelligenz/neuer-rekord-chatgpt-erreicht-2-9-milliarden-besucher-pro-monat-19845849.html).

<sup>89</sup> See on this and the socio-economic effects (with further references): Dornis and Stober (2024), pp. 179 *et seq.*

#### 4.1 “EU Extraterritoriality” of Generative AI Training

As a consequence of the principle of territoriality in international copyright law, copyrighted works are protected by a so-called bundle of separate and nationally limited rights.<sup>90</sup> Each of these national copyrights can be infringed upon only by an act that has been committed at least in part within the respective state’s national territory.<sup>91</sup> Accordingly, the applicable law for alleged copyright infringements in European doctrine is determined in accordance with Art. 8 Rome II Regulation,<sup>92</sup> which provides for the law of the “country of protection” to apply – the so-called *lex loci protectionis*. This is the regime of the country for which the right holder asks the court for protection.<sup>93</sup> On this basis, the judge will always have to determine whether an infringement has taken place or is about to take place by applying the substantive copyright law of the respective *lex loci protectionis*.<sup>94</sup>

In this light, scholarly commentary explains that EU Member States’ national copyright laws are inapplicable whenever AI developers are acting outside the EU.<sup>95</sup> The collection and storage of training data, as well as the training process, occur “extraterritorially”. Moreover, the servers and storage devices to which the datasets (including copyright-protected works) are being uploaded and from where they are downloaded again for the training process are considered to be “unsuitable” for establishing a point of attachment or a linking factor because it is technically hard to determine the exact spatial locale of data streams on the internet.<sup>96</sup> Indeed, as has been explored empirically, the bulk of AI training actually takes place in countries with more “liberal” copyright laws, particularly those with

<sup>90</sup> Cf. Art. 5(3) Berne Convention (“Protection in the country of origin is governed by domestic law.”); for European doctrine, see, e.g., CJEU, case C-441/13, *Pez Hejduk v EnergieAgentur.NRW GmbH* [2015], ECLI:EU:C:2015:28, para. 22; CJEU, case C-170/12, *Peter Pinckney v KDG Mediatech AG* [2013], ECLI:EU:C:2013:635, para. 39; see also, on the principle of territoriality and the bundle of rights approach, Dornis (2017), 193 *et seq.*

<sup>91</sup> See, e.g., CJEU, case C-192/04, *Lagardère Active Broadcast v Société pour la perception de la rémunération équitable (SPRE) and Gesellschaft zur Verwertung von Leistungsschutzrechten mbH (GVL)* [2005], ECLI:EU:C:2005:475, paras. 19 *et seq.*; CJEU, case C-441/13, *Pez Hejduk v EnergieAgentur.NRW GmbH* [2015], ECLI:EU:C:2015:28, para. 22; CJEU, case C-170/12, *Peter Pinckney v KDG Mediatech AG* [2013], ECLI:EU:C:2013:635, para. 39; *Subafilms, Ltd. v. MGM-Pathe Communications Co.*, 24 F.3d 1088 (9th Cir. 1994); see also, on the requirement of territorial conduct, Dornis (2017), 196 with further references.

<sup>92</sup> Regulation (EC) No. 864/2007 of the European Parliament and of the Council of July 11, 2007 on the law applicable to non-contractual obligations (“Rome II”), OJ EU of 31.7.2007, L 199/40.

<sup>93</sup> See, e.g., Vicente (2017), pp. 964–965.

<sup>94</sup> For the two-stage analysis under Art. 8 Rome II Regulation, see, e.g., Federal Court of Justice (BGH) GRUR 2014, 559 para. 12 – *Tarzan* (“Since the plaintiff claims copyright protection for the book in Germany, German copyright law must be applied. The question of whether claims exist in the event of an infringement of a copyright must in principle be answered according to the law of the country of protection – i.e., the state for whose territory the protection is claimed ...”) (author’s translation).

<sup>95</sup> See, e.g., von Welser (2023), p. 520; see also Maamar (2023), p. 486; de la Durantaye (2023), p. 652; Heine (2024), p. 88; Kraetzig (2024), p. 700.

<sup>96</sup> For the CJEU’s concept of a “linking factor” (with respect to jurisdiction), see, e.g., CJEU, case C-441/13, *Pez Hejduk v EnergieAgentur.NRW GmbH* [2015], ECLI:EU:C:2015:28, para. 20. For scholarly commentary explaining localization as impossible, see, e.g., Schack (2024), p. 115.

more wide-reaching TDM or fair use exceptions.<sup>97</sup> It therefore appears that the global setting invites regulatory arbitrage and that AI industries are essentially free to choose the law applicable to their business.<sup>98</sup>

## 4.2 Blind Spot: The EU-Territorial “Making Available to the Public”

Yet if we take another look at the findings from our investigation into the technological foundations of generative AI training, particularly the representation in the models’ vector space, it becomes clear that the debate has failed to touch upon a very central issue: The provision of AI models and model-based services, namely its targeting to end users in Europe (e.g., ChatGPT on the OpenAI website), may offer access to the reproductions of copyrighted works inside the models. Such a “making available to the public” might be considered a *territorial* act of infringement – implying the application of the respective Member State’s copyright law.

In this regard, it is important to remember that in choice-of-law doctrine both the place of the event giving rise to the injury or damage (i.e., the place of conduct) and the place where the damage occurred may determine the applicable law. Even though the CJEU has explained its doctrine only with respect to the issue of international jurisdiction, the same principle governs in choice-of-law.<sup>99</sup> Accordingly, the applicable law can be the law at the place where AI developers or operators have decided on the technical process underlying the making available to the public – i.e., the place from where the uploading has been initiated.<sup>100</sup> Of course, this place may be found in a third country – e.g., in the United States. But the making available to the public is not limited to the alleged infringer’s conduct. It is not only the place of her business or the location of her servers that determines the applicable law but also the place from which the content that has been made available can be accessed by third-party internet users.<sup>101</sup> In practice, this typically requires looking at whether a website is intended to be accessible within a specific country.<sup>102</sup>

Such a concept of a making available to the public *within the EU* is evident with respect to online offers of trained AI models for downloading. Also, offering model-

<sup>97</sup> See, e.g., Handke et al. (2021), p. 1999.

<sup>98</sup> See also Sag and Yu (2025), p. 33.

<sup>99</sup> On jurisdiction, see, e.g., CJEU, case C-441/13, *Pez Hejduk v EnergieAgentur.NRW GmbH* [2015], ECLI:EU:C:2015:28, paras. 18 *et seq.*; CJEU, case C-170/12, *Peter Pinckney v KDG Mediatech AG* [2013], ECLI:EU:C:2013:635, paras. 26 *et seq.*; see also District Court (LG) Hamburg, GRUR-RR 2004, 313 (314–315) – *thumbnails*.

<sup>100</sup> See, e.g., CJEU, case C-441/13, *Pez Hejduk v EnergieAgentur.NRW GmbH* [2015], ECLI:EU:C:2015:28, paras. 24–25; for German doctrine, see, e.g., District Court (LG) Hamburg, GRUR-RR 2004, 313 (314–315) – *thumbnails*.

<sup>101</sup> See again District Court (LG) Hamburg, GRUR-RR 2004, 313 (314–315) – *thumbnails*; see also Katzenberger and Metzger (2020), ‘Vorbemerkung § 120 UrhG’ para. 145.

<sup>102</sup> See, e.g., CJEU, case C-173/11, *Football Dataco Ltd and Others v Sportradar GmbH and Sportradar AG* [2012] ECLI:EU:C:2012:642, para. 39; for national doctrine in Germany, see, e.g., Federal Court of Justice (BGH) GRUR 2016, 1048 para. 22 *et seq.* – *An Evening with Marlene Dietrich*.



based AI services to European end users provides for a making available of the models and their internal replications of copyright-protected works *within the EU*.

#### 4.3 Finally: International Jurisdiction

In EU procedural law, it is the rules of the Brussels I<sup>bis</sup> Regulation<sup>103</sup> on international jurisdiction that determine whether Member States' national civil-law courts have the power to rule over matters with a foreign element. The Brussels system, however, applies only if the defendant has her domicile, statutory seat, central administration, or principal place of business within the EU.<sup>104</sup> Accordingly, with respect to AI developers from outside the EU, it is arguable whether Member States' courts have international jurisdiction. After all, most generative AI services offered are developed and operated by US companies. Instead, as Art. 6 Brussels I<sup>bis</sup> Regulation provides, it is Member States' national rules on international jurisdiction that apply.

In Germany, for instance, international jurisdiction for copyright infringements is determined in accordance with the rule on jurisdiction for international torts and delicts.<sup>105</sup> Even though it has not been determined with ultimate clarity if national jurisdiction rules follow the CJEU's approach, the principles are similar<sup>106</sup>: In the case of an unauthorized making available to the public, the relevant place of conduct has to be determined in accordance with the point of attachment under choice-of-law doctrine.<sup>107</sup> Accordingly, the place of the unlawful conduct relevant for international jurisdiction includes the place of damage. This implies that if a website can be accessed by domestic users, the domestic courts have international jurisdiction to decide.<sup>108</sup>

Applied to the making available of a generative AI to the public – whether through a targeting of domestic consumers with an offer of generative services or with an offer to download a trained model – there is no doubt that international jurisdiction exists.

<sup>103</sup> Regulation (EU) No. 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (recast) OJ EU, L 351 of 20.12.2012, 1 *et seq.*

<sup>104</sup> See in particular Arts. 4, 7, and 63 as well as Art. 6 Regulation (EU) No. 1215/2012.

<sup>105</sup> See § 32 of the German Code of Civil Procedure (ZPO); see also § 23 ZPO. See, e.g., Wimmers (2020), '§ 105 UrhG' para. 13.

<sup>106</sup> For potential differences – which are principally irrelevant for the question whether jurisdiction exists over AI developers from non-EU countries – see, e.g., Wimmers (2020), '§ 105 UrhG' paras. 13 and 21.

<sup>107</sup> For German doctrine, see, e.g., Federal Court of Justice (BGH) GRUR 2016, 1048 para. 17 – *An Evening with Marlene Dietrich*. For the CJEU's doctrine on international jurisdiction, see, e.g., CJEU, case C-441/13, *Pez Hejduk v EnergieAgentur.NRW GmbH* [2015], ECLI:EU:C:2015:28, paras. 18 *et seq.*; CJEU, case C-170/12, *Peter Pinckney v KDG Mediatech AG* [2013], ECLI:EU:C:2013:635, paras. 26 *et seq.*

<sup>108</sup> Federal Court of Justice (BGH) GRUR 2016, 1048 para. 18 – *An Evening with Marlene Dietrich*; Circuit Court (OLG) Frankfurt/M. GRUR-RR 2020, 57 para. 32 – *our ebooks*; see also Wimmers (2020), '§ 105 UrhG' paras. 13 and 21.



## 5 Conclusions and Outlook

As this analysis has unveiled, contrary to what is being contended, copyright owners in the EU already have options to defend against unlicensed and uncompensated generative AI training using their works. This insight should lead to the conclusion that a level playing field is possible: After all, if AI developers and providers can be sued for copyright infringement in European courts, it will be harder for them to avoid EU copyright law in the future. If this is true, however, law- and policymakers in the EU should not settle too quickly for easy solutions in a superficial attempt to guarantee the “AI innovation” that is revered in current legal scholarship. Rather, the question must be how to protect copyright owners against the “race to the bottom” that is presently going on. In this regard, it is important to remember that if EU lawmakers are serious about their goal of making the (digital) world a better (fairer) place – as their constant conjuring of the “Brussels effect” implies – they must not remain inactive.

### 5.1 The Myth of “Regulatory Arbitrage”

As we have seen, the current debate is dominated by the narrative of “less protection, more innovation”: If AI training is subject to greater legal risks due to inadequate exceptions for AI developers and is therefore more costly than in other jurisdictions, the European AI industry will inevitably fall behind its competitors in the United States and elsewhere. However, the hypothesis of an EU lawmaker who is condemned to inactivity in the interests of AI innovation quickly collapses in light of what the analysis of copyright infringement has brought up. While the AI Act, once its relevant provisions will be effective, may become a game changer for authors and right holders, it is already possible to *directly* enforce copyright today – under EU Member States’ national laws and under the jurisdiction of their national civil-law courts. Regulatory arbitrage by locating one’s establishment or her AI training activity in non-EU countries is therefore not a done deal. Just to be clear, the TDM exception will not save AI developers in this respect: There is no doubt that a commercial “making available to the public” is *not* covered by the TDM exception.<sup>109</sup> It thus seems that the battle at the crossroads of copyright law and generative AI training will likely be decided in European courtrooms. The CJEU, of course, will get the final word.<sup>110</sup>

### 5.2 Shadowboxing: “Copyright vs. AI Innovation”

Quite often, the demand for effective legal protection of existing rights is ridiculed for seeking to “ban the disruptive technology”.<sup>111</sup> Yet the protection and enforcement of copyright is in no way intended to prevent or slow down AI innovation. Rather, the aim is to create a level playing field for competition in AI

<sup>109</sup> See, e.g., Margoni and Kretschmer (2022), p. 695; Dornis and Stober (2024), p. 87.

<sup>110</sup> See Art. 267 TFEU; see also Dornis and Danevitch (2024), p. 269.

<sup>111</sup> See, e.g., Lemley (2024), p. 44.

industries. This playing field is global by definition. EU lawmakers expressly formulated the aim of levelling the playing field in their recitals to the AI Act when they explained that the obligation of AI providers to comply

is necessary to ensure a level playing field among providers of general-purpose AI models where no provider should be able to gain a competitive advantage in the Union market by applying lower copyright standards than those provided in the Union.<sup>112</sup>

In other words, the protection and enforcement of copyrights create the basis for fair competition. As soon as non-EU actors offering AI products and services within the EU marketplace can be held accountable for the exploitation of creative resources that were previously considered “free” and used without compensation, they will stop “externalizing”. Such an internalization of the costs of creative production is also a precondition for the continued existence of a vibrant ecosphere for human creativity: After all, who will want to write, design, compose, and so on in the future if everything that is created by human minds and hands can be used for free and replaced immediately by AI-made output?

### 5.3 The Real Issue: Race to the Bottom or Brussels Effect?

Ultimately, it is for EU lawmakers to ask themselves whether they want to passively observe the global race to the bottom, to the detriment of European creatives and the future of the European creative industries.<sup>113</sup> The current status quo of a *de facto* “public domain” – allowing generative AI training for years without compensation under the dubious claim that the TDM exception applies – stands in stark contrast to the fortress of individual-rights protection and fairness standards in digital marketplaces that the EU has attempted to erect in recent years:

In many respects, EU lawmakers have tried to establish a European gold standard of individual legal protection and competition regulation in the online world. One of the many full-bodied announcements on the Digital Markets Act, for example, reads as follows:

[T]he European Parliament stands against unfair business practices used by digital giants. We will make sure that digital markets are open and competitive. This is good for consumers, good for businesses and good for digital innovation. Our message is clear: the EU will enforce the rules of the

<sup>112</sup> Recital 106.

<sup>113</sup> Some commentators explain a development that can be defined as a “race to the middle” – i.e., a tendency of lawmakers around the world to design their legal regimes in a way that aims at a middle ground rather than extreme positions. See, e.g., Sag and Yu (2025), p. 34 *et seq.* With respect to generative AI training, however, copyright owners currently do actually witness a classic race to the bottom: They are factually unprotected against the use of their works, and the tendency in public opinion and legal debate seems to be going toward maintaining this status in the interest of more AI innovation.

social market economy also in the digital sphere, and this means that lawmakers dictate the rules of competition, not digital giants.<sup>114</sup>

Indeed, the so-called Brussels effect – the creation of universally recognized and globally applicable regulatory standards through transnational “nudging” – which is regularly invoked in the context of rule-making for digital marketplaces, is even described as a trademark of modern EU legislation.<sup>115</sup> But why then, if Brussels is so confident that its high aspirations of rights protection and morality is without alternatives, should a *laissez-faire* approach be pursued with regard to the rights of authors and other creatives and with regard to the exploitation of the stock of European culture? What makes copyrights so different from consumer protection, workers’ rights (inside and outside the EU), or individual privacy and data protection concerns?

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

### Scholarly References

- Baumann M (2023) Generative KI und Urheberrecht – Urheber und Anwender im Spannungsfeld. NJW 76:3673–3678
- Bradford A (2015) The Brussels effect. *Northwest Univ Law Rev* 107:1–68
- Bradford A (2020) The Brussels effect: how the European Union rules the World. Oxford University Press, Oxford
- Bradford A (2023) Europe’s digital constitution. *Va J Int Law* 64:1–68
- Brauneis R (2025) Copyright and the training of human authors and generative machines. *Columbia J Law Arts* 47. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4909592](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4909592) (forthcoming 2025)
- Bullinger W (2022) § 19a UrhG. In: Wandtke A-A, Bullinger W (eds) *Praxiskommentar Urheberrecht*, 6th edn. C.H. Beck, München, pp 368–382
- Carlini N, Tramèr F, Wallace E, Jagielski M, Herbert-Voss A, Lee K, Roberts A, Brown T, Song D, Erlingsson U, Oprea A, Raffel C (2021) Extracting training data from large language models. In: *Proceedings of the 30th USENIX security symposium (USENIX Security 21)*. <https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting>
- Carlini N, Hayes J, Nasr M, Jagielski M, Schwag V, Tramèr F, Balle B, Ippolito D, Wallace E (2023a) Extracting training data from diffusion models. In: *Proceedings of the 32nd USENIX security symposium (USENIX Security 23)*. <https://doi.org/10.48550/arXiv.2301.13188>

<sup>114</sup> See, e.g., European Parliament, Press Release: Digital Markets Act: Parliament ready to start negotiations with Council, 15 December 2021 (<https://www.europarl.europa.eu/news/en/press-room/20211210IPR19211/digital-markets-act-parliament-ready-to-start-negotiations-with-council>): “Our message is clear: the EU will enforce the rules of the social market economy also in the digital sphere, and this means that lawmakers dictate the rules of competition, not digital giants.”

<sup>115</sup> See, e.g., Bradford (2015), p. 1; Bradford (2020); Bradford (2023), pp. 11 *et seq.*

- Carlini N, Ippolito D, Jagielski M, Lee K, Tramèr F, Zhang C (2023b) Quantifying memorization across neural language models. In: Proceedings of the 11th international conference on learning representations (ICLR). <https://doi.org/10.48550/arXiv.2202.07646>
- Carter S, Nielsen M (2017) Using artificial intelligence to augment human intelligence. <https://doi.org/10.23915/distill.00009>
- Chang K, Cramer M, Soni S, Bamman D (2023) Speak, memory: an archaeology of books known to ChatGPT/GPT-4. In: Bouamor H, Pino J, Bali K (eds) Proceedings of the 2023 conference in empirical methods in natural language processing. <https://doi.org/10.48550/arXiv.2305.00118>
- Cooper AF, Grimmelmann J (2024) The files are in the computer: on copyright, memorization, and generative AI. *Chic-Kent Law Rev.* [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4803118](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4803118) (forthcoming)
- de la Durantaye K (2023) “Garbage in, garbage out” – Die Regulierung generativer KI durch Urheberrecht. *ZUM* 67:645–660
- Dornis TW (2017) Trademark and unfair competition conflicts – historical-comparative, doctrinal, and economic perspectives. Cambridge University Press, Cambridge. <https://doi.org/10.1017/9781316651285>
- Dornis TW (2025) The training of generative AI is not text and data mining. *Eur Intellect Prop Rev* 47 (2025), 65 (also available under [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4993782](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4993782))
- Dornis TW, Danevitch J (2024) Herausfordernd: Urheberrecht und generatives KI-Training. *DRiZ* 102:442–445
- Dornis TW, Stober S (2024) Urheberrecht und Training generativer KI-Modelle. *Nomos*, Baden-Baden. <https://doi.org/10.5771/9783748949558>
- Götting H-P (2024) § 16 UrhG. In: Götting H-P, Lauber-Rönsberg A, Rauer N (eds) *BeckOK Urheberrecht*, 42nd edn. C.H. Beck, München
- Guadamuz A (2024) A scanner darkly: copyright liability and exceptions in artificial intelligence inputs and outputs. *GRUR Int* 73:11–127
- Handke C, Guibault L, Vallbé J-J (2021) Copyright’s impact on data mining in academic research. *Manag Decis Econ* 42:1999–2016
- Heerma JD (2022) § 16 UrhG. In: Wandtke A-A, Bullinger W (eds) *Praxiskommentar Urheberrecht*, 6th edn. C.H. Beck, München, pp 311–321
- Heine R (2024) Generative KI: Nutzungsrechte und Nutzungsvorbehalt. *GRUR-Prax* 16:87–90
- Henderson P, Li X, Jurafsky D, Hashimoto T, Lemley MA, Liang P (2023) Foundation models and fair use. *J Mach Learn Res* 23:1–79
- Hofmann F (2024a) Retten Schranken Geschäftsmodelle generativer KI-Systeme? *ZUM* 68:166–174
- Hofmann F (2024b) Zehn Thesen zu Künstlicher Intelligenz (KI) und Urheberrecht. *WRP* 70:11–18
- Käde L (2021) Kreative Maschinen und Urheberrecht – Die Machine Learning-Wertschöpfungskette vom Training über Modellschutz bis zu Computational Creativity. *Nomos*, Baden-Baden. <https://doi.org/10.5771/9783748912453>
- Käde L (2024) Do you remember? – Enthalten KI-Modelle Vervielfältigungen von Trainingsdaten, lassen sich diese gezielt rekonstruieren und welche Implikationen hat das für das Urheberrecht? *ZUM* 68:174–183
- Katzenberger P, Metzger A (2020) Vorbemerkung § 120 UrhG. In: Schricker G, Loewenheim U (eds) *Urheberrecht*, 6th edn. C.H. Beck, München, pp 2539–2601
- Kögel D (2023a) Urheberrechtliche Implikationen bei der Verwendung kreativer und generativer künstlicher Intelligenz. *InTeR* 11:179–185
- Kögel D (2023b) Urheberrechtliche Implikationen bei der Verwendung kreativer und generativer künstlicher Intelligenz. In: Bernzen AK, Fritzsche J, Heinze C, Thomson O (eds) *DSRITB 2023: Das IT-Recht vor der (europäischen) Zeitwende?* *OiWIR*, Edewecht, pp 285–302
- Konertz R, Schönhof R (2024a) Vervielfältigungen und die Text- und Data-Mining-Schranke beim Training von (generativer) Künstlicher Intelligenz. *WRP* 70:289–296
- Konertz R, Schönhof R (2024b) Rechtsfolgen der Urheberrechtsverletzung bei generativer Künstlicher Intelligenz. *WRP* 70:534–541
- Kraetzig V (2024) KI-Kunst als schöpferische Zerstörung. *NJW* 77:697–702
- Lee K, Ippolito D, Nystrom A, Zhang C, Eck D, Callison-Burch C, Carlini N (2022) Deduplicating training data makes language models better. In: Proceedings of the 60th annual meeting of the association for computational linguistics, vol 1: long papers. <https://doi.org/10.48550/arXiv.2107.06499>

- Leistner M (2024) TDM und KI-Training in der Europäischen Union – Erste Fingerzeige des LG Hamburg im „LAION“-Urteil. GRUR 126:1665–1675
- Lemley MA (2024) How generative AI turns copyright upside down. *Columbia Sci Technol Law Rev* 25:21–44. <https://doi.org/10.52214/str.v25i2.12761>
- Litman JD (1989) Copyright legislation and technological change. *Or Law Rev* 68:275–361
- Loewenheim U (2020) § 16 UrhG. In: Schricker G, Loewenheim U (eds) *Urheberrecht*, 6th edn. C.H. Beck, München, pp 405–413
- Lux H, Noll CJ (2024) Of books and bytes: the copyright dilemma in AI development. *Transatl Law J* 2:111–115
- Maamar N (2023) Urheberrechtliche Fragen beim Einsatz von generativen KI-Systemen. *ZUM* 67:481–491
- Margoni T, Kretschmer M (2022) A deeper look into the EU text and data mining exceptions: harmonisation, data ownership, and the future of technology. *GRUR Int* 71:685–701
- Mezei P (2024) A saviour or a dead end? Reservation of rights in the age of generative AI. *Eur Intellect Prop Rev* 46:461–469
- Müller N (2024) ChatGPT erreicht 2,9 Milliarden monatliche Besucher. <https://www.faz.net/pro/digitalwirtschaft/kuenstliche-intelligenz/neuer-rekord-chatgpt-erreicht-2-9-milliarden-besucher-pro-monat-19845849.html>
- Nemitz P (2018) Constitutional democracy and technology in the age of artificial intelligence. *Philos Trans R Soc A* 376:20180089. <https://doi.org/10.1098/rsta.2018.0089>
- Nordemann JB, Pukas J (2022) Copyright exceptions for AI training data – will there be an international level playing field? *J Intellect Prop Law Pract* 17:973–974. <https://doi.org/10.1093/jiplp/jpac106>
- Pesch PJ, Böhme R (2023a) Artpocalypse now? – Generative KI und die Vervielfältigung von Trainingsbildern. *GRUR* 125:997–1007
- Pesch PJ, Böhme R (2023b) Verarbeitung personenbezogener Daten und Datenrichtigkeit bei großen Sprachmodellen – ChatGPT & Co. unter der DS-GVO. *MMR* 26:917–923
- Peukert A (2024) Copyright in the Artificial Intelligence Act – a primer. *GRUR Int.* 73:497–509
- Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, Sastry G, Askell A, Mishkin P, Clark J, Krueger G, Sutskever I (2021) Learning transferable visual models from natural language supervision. In: Meila M, Zhang T (eds) *Proceedings of the 38th international conference on machine learning (PMLR)*. <https://doi.org/10.48550/arXiv.2103.00020>
- Ramesh A, Dhariwal P, Nichol A, Chu C, Chen M (2022) Hierarchical text-conditional image generation with CLIP latents. <https://doi.org/10.48550/arXiv.2204.06125>
- Sag M (2023) Copyright safety for generative AI. *Houst Law Rev* 61:295–347
- Sag M (2024) Fairness and fair use in generative AI. *Fordham Law Rev* 92:1887–1921
- Sag M, Yu PK (2025) The globalization of copyright exceptions for AI training. *Emory Law J* 74:43. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4976393](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4976393) (forthcoming)
- Samuelson P (2024) Fair use defenses in disruptive technology cases. *U.C.L.A. Law Rev.* [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4631726](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4631726) (forthcoming)
- Schack H (2021) *Urheber- und Urhebervertragsrecht*, 10th edn. Mohr Siebeck, Tübingen
- Schack H (2024) Auslesen von Webseiten zu KI-Trainingszwecken als Urheberrechtsverletzung de lege lata et ferenda. *NJW* 77:113–117
- Schuhmann C, Beaumont R, Vencu R, Gordon C, Wightman R, Cherti M, Coombes T, Katta A, Mullis C, Wortsman M, Schramowski P, Kundurthy S, Crowson K, Schmidt L, Kaczmarczyk R, Jenia J (2022) LAION-5B: an open large-scale dataset for training next generation image-text models. In: Koyejo S, Mohamed S, Agarwal A, Belgrave D, Cho K, Cho A (eds) *36th conference on neural information processing systems (NeurIPS)*. <https://doi.org/10.48550/arXiv.2210.08402>
- Schulze G (2022) § 16 UrhG. In: Dreier T, Schulze G (eds) *Urheberrechtsgesetz*, 7th edn. C.H. Beck, München, pp 347–360
- Senftleben M (2023) Generative AI and author remuneration. *IIC* 54:1535–1560. <https://doi.org/10.1007/s40319-023-01399-4>
- Sesing-Wagenpfeil A (2024) Trainierte KI-Modelle als Vervielfältigungsstücke im Sinne des Urheberrechts. *ZGE* 16:212–268. <https://doi.org/10.1628/zge-2024-0014>
- Söbbing T, Schwarz A (2023) Urheberrechtliche Grenzen für lernende künstliche Intelligenz – Der neue § 44b UrhG und seine Möglichkeiten und Grenzen beim Machine Learning sowie die Anwendung von § 60d UrhG. *RDİ* 3:415–419
- Sobel BLW (2017) Artificial intelligence’s fair use crisis. *Columbia J Law Arts* 41:45–97. <https://doi.org/10.7916/jla.v41i1.2036>

- Sobel BLW (2024) Elements of style: copyright, similarity, and generative AI. Harv J Law Technol 38. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4832872](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4832872) (forthcoming)
- Somepalli G, Singla V, Goldblum M, Geiping J, Goldstein T (2022) Diffusion art or digital forgery? Investigating data replication in diffusion models. <https://doi.org/10.48550/arXiv.2212.03860>
- Somepalli G, Singla V, Goldblum M, Geiping J, Goldstein T (2023) Understanding and mitigating copying in diffusion models. <https://doi.org/10.48550/arXiv.2305.20086>
- Torrance AW, Tomlinson B (2023) Training is everything: artificial intelligence, copyright, and “fair training.” Dickinson Law Rev 128:233–255
- Vesala J (2023) Developing artificial intelligence-based content creation: are EU copyright and antitrust law fit for purpose? IIC 54:351–380. <https://doi.org/10.1007/s40319-023-01301-2>
- Vicente DM (2017) Intellectual property, applicable law. In: Basedow J, Rühl G, Ferrari F, de Miguel Asensio PA (eds) Encyclopedia of private international law. Edward Elgar, Cheltenham, pp 962–969. <https://doi.org/10.4337/9781782547235.1.11>
- von Ungern-Sternberg J (2020) § 19a UrhG. In: Schricker G, Loewenheim U (eds) Urheberrecht, 6th edn. C.H. Beck, München, pp 458–506
- von Welser M (2023) Generative KI und Urheberrechtsschranken. GRUR-Prax 15:516–520
- Wagner K (2024) Generative KI: Eine „Blackbox“ urheberrechtlicher Haftungsrisiken? Balanceakt zwischen Innovationsförderung und effektivem Rechtsschutz für Werke Dritter. MMR 27:298–304
- Wimmers J (2020) § 105 UrhG. In: Schricker G, Loewenheim U (eds) Urheberrecht, 6th edn. C.H. Beck, München, pp 2458–2481

## Case Law

- Case C-117/15, *Reha Training v GEMA*, ECLI:EU:C:2016:379
- Case C-135/10, *SCF v Marco Del Corso*, ECLI:EU:C:2012:140
- Case C-160/15, *GS Media v Sanoma Media Netherlands*, ECLI:EU:C:2016:644
- Case C-161/17, *Land Nordrhein-Westfalen v Renckhoff*, ECLI:EU:C:2018:634
- Case C-162/10, *Phonographic Performance v Ireland*, ECLI:EU:C:2012:141
- Case C-170/12, *Peter Pinckney v KDG Mediatech AG*, ECLI:EU:C:2013:635
- Case C-173/11, *Football Dataco v Sportradar*, ECLI:EU:C:2012:642
- Case C-192/04, *Lagardère Active Broadcast v Société pour la perception de la rémunération équitable (SPRE) and Gesellschaft zur Verwertung von Leistungsschutzrechten mbH (GVL)*, ECLI:EU:C:2005:475
- Case C-301/15, *Soulier and Doke v Premier minister et al*, ECLI:EU:C:2016:878
- Case C-306/05, *SGAE v Rafael*, ECLI:EU:C:2006:764
- Case C-310/07, *Levola Hengelo BV v Smilde Foods BV*, ECLI:EU:C:2018:899
- Case C-314/12, *UPC Telekabel v Constantin Film*, ECLI:EU:C:2014:192
- Case C-351/12, *OSA v Léčebné lázně*, ECLI:EU:C:2014:110
- Case C-392/19, *VG Bild-Kunst v Stiftung Preußischer Kulturbesitz*, ECLI:EU:C:2021:181
- Case C-426/21, *Ocilion IPTV Technologies GmbH v Seven.One Entertainment Group GmbH, Puls 4 TV GmbH & Co. KG*, ECLI:EU:C:2023:564
- Case C-433/20, *Austro-Mechana Gesellschaft zur Wahrnehmung mechanisch-musikalischer Urheberrechte Gesellschaft mbH v Strato AG*, ECLI:EU:C:2022:217
- Case C-441/13, *Hejduk v EnergieAgentur*, ECLI:EU:C:2015:28
- Case C-466/12, *Svensson v Retriever Sverige*, ECLI:EU:C:2014:76
- Case C-476/17, *Pelham v Hütter*, ECLI:EU:C:2019:624
- Case C-5/08, *Infopaq v DDF*, ECLI:EU:C:2009:465
- Case C-527/15, *Stichting Brein v Wullems*, ECLI:EU:C:2017:300
- Case C-597/19, *Mircom v Telenet*, ECLI:EU:C:2021:492
- Case C-597/19, *Mircom v Telenet*, ECLI:EU:C:2021:492, Opinion of AG Szpunar
- Case C-610/15, *Stichting Brein v Ziggo and XS4ALL Internet (The Pirate Bay)*, ECLI:EU:C:2017:456
- Case C-683/17, *Cofemel – Sociedade de Vestuário SA v G-Star Raw CV*, ECLI:EU:C:2019:721
- Case C-753/18, *Stim and SAMI v Fleetmanager Sweden and Nordisk Biluthyrning*, ECLI:EU:C:2020:268
- Circuit Court (OLG) Stuttgart, 1 January 2008, 2 Ws 328/07, MMR 2008, 474 – *Music-on-Demand-Dienst*
- Circuit Court (OLG) Frankfurt/M., 30 April 2019, 11 U 27/18, GRUR-RR 2020, 57 – *our ebooks*

- District Court (LG) Hamburg, 5 September 2003, 308 O 449/03, GRUR-RR 2004, 313
- German Federal Supreme Court, 4 May 2000, I ZR 256/97, GRUR 2001, 51 (BGHZ 144, 232) – *Parfumflakon I*
- German Federal Supreme Court, 12 November 2009, I ZR 166/07, GRUR 2010, 616 – *marions-kochbuch.de*
- German Federal Supreme Court, 29 April 2010, I ZR 69/08, GRUR 2010, 628 (BGHZ 185, 291) – *Vorschaubilder I*
- German Federal Supreme Court, 26 February 2014, I ZR 49/13, GRUR 2014, 559 – *Tarzan*
- German Federal Supreme Court, 21 April 2016, I ZR 43/14, GRUR 2016, 1048 – *An Evening with Marlene Dietrich*
- German Federal Supreme Court, 23 February 2017, I ZR 92/16, GRUR 2017, 793 – *Mart-Stam-Stuhl*
- German Federal Supreme Court, 21 September 2017, I ZR 11/16, GRUR 2018, 178 (49 IIC 357 (2018)) – *Vorschaubilder III*
- German Federal Supreme Court, 6 December 2017, I ZR 186/16, GRUR 2018, 400 – *Konferenz der Tiere*
- German Federal Supreme Court, 9 September 2021, I ZR 113/18, GRUR 2021, 1511 (BGHZ 231, 116) – *Deutsche Digitale Bibliothek II*
- Joined Cases C-403 and 429/08, *Football Association Premier League v QC Leisure and Murphy v Media Protection Services*, ECLI:EU:C:2011:631
- Joined Cases C-682 and 683/18, *Peterson v Google and Elsevier v Cyando (YouTube and Cyando)*, ECLI:EU:C:2021:503
- Joined Cases C-775 and C-826/21, *Blue Air Aviation v UCMR and UPFR v CFR*, ECLI:EU:C:2023:307
- US Court of Appeals, 9th Cir. 13 May 1994, *Subafilms, Ltd. v. MGM-Pathe Communications Co.*, 24 F.3d 1088
- US Court of Appeals, 9th Cir. 16 May 2007, *Perfect 10, Inc. v. Amazon.com*, 508 F.3d 1146

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.