

Opt-Outs in Copyright Law

Can They Save AI Training Under Fair Use and the Three-Step Test?

Darius Rostam*

The use of copyright-protected works for AI training has thrust opt-out schemes into the center of copyright discourse. Unlike conventional copyright rules requiring prior authorization, opt-out schemes permit uses by default unless rightholders affirmatively object, effectively reversing copyright's traditional logic of exclusivity. This Article provides the first comprehensive analysis of opt-out mechanisms in copyright law. Adopting a comparative perspective examining both US and EU copyright frameworks, it traces existing precedents ranging from copyright restoration to search engines, mass digitization projects, and online platform liability. The Article argues that opt-out schemes serve two principal functions: preserving rightholder autonomy through flexible „property rules“ that combine elements of property and liability rules, and controlling user infringement risk by forcing rightholders to disclose information necessary for transactions. However, these functions depend on two conditions: rightholders must be aware of both the use and their opt-out option, and the opt-out mechanism must be feasible with reasonable effort. Where these conditions are met, opt-outs can positively influence the fair use analysis, particularly factors one and four, by supporting an inference that a missing opt-out indicates no meaningful threat of substitution. Similarly, under the international three-step test, opt-outs help ensure compliance by preventing conflicts with normal exploitation and unreasonable prejudice to legitimate interests. Applying this framework to AI training under Article 4 DSM Directive and pending US litigation, this Article demonstrates that current opt-out regimes fail to meet these conditions. Rightholders often lack awareness, opt-out standards provide insufficient granularity, and many works are uploaded by third parties beyond rightholders' control. Consequently, while opt-out schemes have legitimate doctrinal foundations, their application to generative AI training is currently dysfunctional.

I. Introduction

The use of copyright-protected works for the training of artificial intelligence (AI) models has brought a peculiar instrument to the forefront of copyright discourse: opt-out schemes. In both the US as well as EU copyright law, the compilation of AI training data is either statutorily or at least factually based on a structure that seems counterintuitive to copyright law. While generally a use is forbidden unless the rightholder permits it, certain uses for AI training purposes are permitted unless the rightholder forbids them. The responsibility of users to obtain a permission prior to their use is therefore transformed into a burden of rightholders to opt-out of unwanted, otherwise lawful uses.

This opt-out logic is subject to significant controversy. While users generally favor the default of permissiveness it introduces,¹ it has been met with substantial opposition by rightholders. It is argued that opt-outs are “inconsistent with the basic principle of copyright”,

* Executive Director, Center for Transnational Intellectual Property, Media and Technology Law and Policy, Bucerius Law School; Dr. iur.; LL.M., New York University. The research was supported by a scholarship of the German Ministry for Economic Affairs and Climate Action and the German Academic Scholarship Foundation. I thank the participants of the Advanced Copyright Seminar at NYU Law for their valuable comments.

¹ See e.g. Public Knowledge, *USCO AI and Copyright Comments*, 13, <https://publicknowledge.org/policy/usco-ai-and-copyright-comments/> (last visited Jan. 30, 2026).

“antithetical to current law” and “unduly burdensome”.² These contrary views, however, are no more unsurprising than they are new.³ A line of established use cases show that opt-out schemes have long been part of copyright law. Although many situations in which copyright law demands from rightsholders to take affirmative action for their exclusive rights originated from more recent digitization developments, opt-out schemes can be found in very different contexts in both US and EU copyright law.

Scholarly discussion has so far treated these schemes separately. The debate over AI training uses now invites a more comprehensive, holistic analysis. This article examines the existing precedents and use cases of opt-out rules by looking at their requirements, connected elements and hidden background assumptions. It investigates a context-independent justification for opt-outs in copyright law, which provides a framework to evaluate if and how opt-outs may positively influence the assessment of the compilation of AI training data under the fair use analysis and the international three-step test.

The article proceeds in three steps: First, it illustrates the concept and applications of opt-outs by clarifying the terminology, tracing the existing use cases and providing an overview of the current landscape concerning AI training uses in both US and EU copyright law. In a second step, it lays out the normative case behind opt-outs from a perspective of both rightsholders as well as users. Lastly, it applies the findings to the opt-out schemes around the compilation of AI training data and investigates their influence on the outcome of the analyses under fair use and the international three-step test.

This Article argues that while there is a normative case for opt-outs in preserving the autonomy of rightsholders and in controlling the infringement risk of users, and while opt-outs may inform the analysis of both fair use and the three-step test in general, under the current conditions around AI training opt-outs have dysfunctional effects. Specifically, the failure to ensure rightsholder awareness and to provide effective, granular opt-out mechanisms means that a missing opt-out cannot reliably indicate acquiescence to AI training uses.

II. Opt-Outs: Concept and Applications

² These are examples from the excerpt of comments of stakeholders as a response to the US Copyright Office’s notice of inquiry on Copyright and AI, as well as their own recommendation, *see* U.S. Copyright Office, *Copyright and Artificial Intelligence – Part 3: Generative AI Training (Pre-Publication Version)*, 101-02, 105, <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>.

³ *See e.g.* Oren Bracha, *Standing Copyright Law on Its Head - The Googlization of Everything and the Many Faces of Property*, 85 TEX. L. REV. 1799, 1802 (2006) (giving instances of reactions to the opt-out scheme around the Google Books project, where the talk was of “standing copyright law on its head”).

Before evaluating the legitimacy of opt-outs in the context of AI training, it is essential to understand their nature and existing applications in copyright law. This part clarifies the terminology of opt-outs applied here and explores the various opt-out rules that already exist in copyright law before turning to the now controversial opt-out frameworks governing generative AI training uses.

1. Defining Opt-Outs

What opting-out means is rarely defined with precision. While the distinction between opt-out and opt-in mechanisms may seem intuitive, it involves several dimensions: the allocation of the burden to act, the direction of the desired legal change, and the object to which the mechanism applies.

Opt-out as well as opt-in schemes subscribe to the same logic. Both make some sort of affirmative step of a party necessary to move from the status quo to a different legal outcome. They differ with regards to the distribution of burdens to act: Under an opt-in rule, there is a default assumption of exclusion, meaning that a party that wants someone else to be included in a given system must persuade them to join.⁴ As the decision lies with the joining party, opt-ins are thought to preserve their personal autonomy.⁵ Opt-outs in contrast are based on a default assumption of inclusion.⁶ The burden of action falls on the included party that must communicate their diverging preference if they want to be excluded.

Relatedly, both differ regarding the direction of the option:⁷ Opt-in has an inducing notion; it relates to a change in favor of a desired legal consequence. Opt-out in contrast involves the idea of avoidance. It describes a substitution of an undesired legal consequence with a more favorable one. The object of the opt-out therefore is decisive to understand what the term means in a certain context.

In the context of intellectual property rights such as copyright, there are two of such possible objects: the existence of a property right itself, and, assuming that such protection has come into existence, its exercise against uses of protected subject matter. Concerning the existence of a property right, both directions of opt-in and opt-out play a role. Intellectual property rights that require an affirmative act other than the act of creation to come into existence operate under an

⁴ See Joy Su, *Google Book Search and Opt-Out Procedures*, 56 J. COPYRIGHT SOC'Y U.S.A. 947, 955 (2008).

⁵ See John S. Sieman, *Using the Implied License to Inject Common Sense into Digital Copyright Comment*, 85 N.C. L. REV. 885 (2006).

⁶ See Su, *supra* note 4, at 955.

⁷ Thomas Dreier, *"Opt in" and "Opt out" Mechanisms in the Internet Era – towards a Common Theory*, 26 COMPUTER L. & SECURITY REV. 144, 146 (2010) (emphasizing the difference in directionality of opt-in and opt-out mechanisms).

opt-*in* regime. The default outcome in case of inaction is that an intellectual creation is released into the public domain. To exercise exclusivity, one must actively opt into the protection by undergoing a registration procedure.

Copyright historically resembled an opt-in regime as it depended on compliance with formalities, meaning procedural conditions for an author to enjoy certain benefits from a protection of her works, such as the deposit of an artifact of the work, the registration of a renewal, or the notice of a reservation of rights.⁸ If an author did not take affirmative steps to comply with constitutive formalities, her work would not receive protection; under declarative formalities, she would not be able to enforce protection or miss out on other benefits.⁹ Such formalities had the function to discriminate between works that authors claimed protection for and works where such protection was not desired.¹⁰ By default, the latter class of works was filtered out and released to the public domain.¹¹

The situation is flipped for rights that come into existence through the act of creation without further ado. It is up to the creator then to take affirmative action to forfeit protection by opting *out* of the protection regime. Regarding copyright, formalities fell out of fashion in the second half of the 19th century, leading to their ban on international level with the Berlin Revision of the Berne Convention in 1908,¹² and an almost complete¹³ deformalization of US copyright at the turn of the 21st century.¹⁴ As no affirmative step from an author is necessary for the enjoyment of protection, automatic copyright is the norm. Conversely, it is now up to the author to opt-out of the protection by dedicating a work to the public domain in case no protection is desired.¹⁵

The terminology of “opt-out” employed here builds upon the discourse around formalities, however, relates to uses of copyright protected works instead of the existence of the property

⁸ See generally STEF VAN GOMPEL, FORMALITIES IN COPYRIGHT LAW: AN ANALYSIS OF THEIR HISTORY, RATIONALES AND POSSIBLE FUTURE 15-27 (2011).

⁹ See Stef van Gompel, *Les Formalités Sont Mortes, Vive Les Formalités! Copyright Formalities and the Reasons for Their Decline in Nineteenth Century Europe*, in PRIVILEGE AND PROPERTY 157, 164 (Ronan Deazley, Martin Kretschmer, & Lionel Bently eds., 1st ed. 2010).

¹⁰ See Christopher Sprigman, *Reform(Aliz)Ing Copyright*, 57 STAN. L. REV. 485, 502-21 (2004).

¹¹ See H.R. REP. NO. 94-1476, at 143-44.

¹² See Art. 5(2) Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, as revised at Paris on July 24, 1971 and amended in 1979, S. Treaty Doc. No. 99-27 (1986) [in the following: Berne Convention].

¹³ The US kept certain formalities in place as an incentive and not a condition for enjoyment or exercise of the rights, see e.g. 17 U.S.C. §§ 411(a), 412.

¹⁴ See Jane Ginsburg, *The U.S. Experience with Mandatory Copyright Formalities: A Love/Hate Relationship*, 33 COLUM. J. L. & ARTS 311, 333-342 (2010) (describing the development leading up to the US's Berne adherence).

¹⁵ Crucially, in copyright regimes with an author's rights tradition, such a dedication to the public domain may be hindered, effectively locking works in exclusivity, which then requires other solutions such as open licenses, see e.g. ALEXANDER PEUKERT, DIE GEMEINFREIHEIT: BEGRIFF, FUNKTION, DOGMATIK 397 (2012).

right. It concerns situations where affirmative steps of rightsholders become necessary to either permit or prevent imminent or already occurring, lawful uses. Both directions of opt-in and opt-out can play a role in this regard: Opt-*in* describes a situation where uses are generally forbidden unless the rightsholder grants a permission. A rightsholder must then take an affirmative step to allow uses. If a rightsholder remains passive and does nothing, she can be sure that no legal use can take place without her involvement, apart from existing limitations and exceptions. An opt-*out* scheme in contrast is in place when a use is generally permitted unless the rightsholder forbids it.¹⁶ In order to prevent uses, rightsholders must act. Under either an opt-in and opt-out logic as understood here, the general protection of subject matter is therefore not at stake – the question is rather whether a certain use may occur without obtaining a prior permission from the rightsholder or not.

2. Existing Opt-Out Schemes in Copyright

Such opt-out rules are not uncommon in copyright but can be found in diverse contexts for which AI training only represents the newest iteration. While there are isolated older instances concerning restored copyrights and a limitation concerning public performances, the advent of digital technology brought about a line of new opt-out use cases, such as search engines and digitization projects to provide access to cultural heritage, as well as the notice-and-takedown schemes that benefit service providers on the internet.

a. Copyright Restoration

A first use case concerns the restoration of copyrights under 17 U.S.C. § 104A. The provision is aimed especially at foreign rightsholders, for whom the numerous formalities of past US copyright law sometimes had harsh consequences. In case they failed to comply with a formality, for example by publishing a work without proper notice, they often inadvertently forfeited the copyright in the US and the work entered the public domain. Member states to the Berne Convention, however, are obliged to provide for the protection of works that in the country of origin have not yet entered the public domain.¹⁷ Following the US adherence to Berne and later TRIPS, § 104A was therefore enacted to automatically restore the forfeited terms of such copyrights, providing rightsholders with the full scope of remedies available under US law.¹⁸

¹⁶ Confusingly enough, this is also sometimes described as an opt-in regime, meaning that the rightsholder opts into copyright's exclusivity, *see e.g.* Tim Wu, *Tolerated Use*, 31 COLUM. J. L. & ARTS 617 (2008).

¹⁷ *Cf.* Art. 18(1), (4) Berne Convention.

¹⁸ The provision was later upheld in *Golan v. Holder*, 565 U.S. 302 (2012).

What proved to be problematic, however, are users that relied on the public domain status of such works and, after the restoration, would have infringed on the restored right. To protect their good faith, rightsholders were required to (1) announce their willingness to enforce the restored copyright by filing a notice of intent with the Copyright Office within 24 months of the date of restoration,¹⁹ or by serving such notice directly to the reliance party, and (2) to grant a grace period of twelve months of immunity.²⁰ The notice needs to contain i.a. the work title, the contact details of the rightsholder, as well as the use to which the rightsholder objects in case the notice is served directly on a reliance party.²¹ If the notice is missing or incomplete, rightsholders have no remedies against infringements of reliance parties.²²

By default, copyright's exclusive rights therefore are ineffective against reliance parties. Despite the fully restored copyright, uses without prior authorization do not trigger infringement liability unless rightsholders provide notice beforehand. It is up to the rightsholders to take the affirmative step of giving notice to enforce their copyright through injunctive or monetary relief. § 104A therefore represents an opt-out scheme, where a use is generally permitted unless the rightsholder becomes active to forbid it.

b. Performances of a Nondramatic Literary or Musical Work

A similar opt-out scheme involves performances of a nondramatic literary or musical work otherwise than in a transmission to the public under 17 U.S.C. § 110(4)(B). Such performances are not an infringement of copyright if a user does not follow a profit motive. Where performers take an admission charge and use the proceeds for charitable purposes, however, the copyright owner is granted a veto right: The limitation does not apply if the rightsholder serves a written and signed notice of objection that states a reason for the objection to the person responsible for the performance seven days before its date.²³

The idea behind this is not to compel rightsholders to indirectly contribute to donations to causes they oppose.²⁴ It is therefore up to the rightsholder to take an affirmative step to exercise

¹⁹ The date of restoration for most works was January 1, 1996 (cf. § 104A(h)(2)(A)), therefore most notices had to be filed by December 31, 1997. However, as the date of restoration for source countries that later adhere to the Berne Convention or become a WTO member country is the date of adherence (§ 104A(h)(2)(B)), later notices are possible.

²⁰ 17 U.S.C. § 104A(c), (d)(2)(A)(ii)(I).

²¹ 17 U.S.C. § 104A(e).

²² 17 U.S.C. § 104A(d)(2).

²³ 17 U.S.C. § 110(4)(B).

²⁴ See H.R. REP. NO. 94-1476, at 86, 1976 U.S.C.A.N. 5659, 5700; see also David E. Shipley, *Copyright Law and Your Neighborhood Bar and Grill: Recent Developments in Performance Rights and the Section 110(5) Exemption*, 29 ARIZ. L. REV. 475, 486 (1987).

her right of public performance by objecting to undesired public performances. As such uses are lawful even in the absence of a prior permission except where the rightsholder forbids it, § 110(4)(B) establishes an opt-out scheme.

c. Search Engines

What introduced opt-out regimes on a larger scale to copyright was the advent of digital technology and the internet. The first example in this line concerns search engines. The usefulness of web search engines is connected to the size of their index – the more websites are processed, the better are the results. Both the caching of websites and the creation and display of snippets or thumbnails as search results *prima facie* touch upon the exclusive rights of reproduction and public display.²⁵ However, in *Field v. Google* it was held that allowing internet users to access copies of works stored in the search engine’s cache was lawful, based especially on three factors:

First, the caching did not serve the same expressive function as the original website but had a “socially important” purpose in offering access to works which many other rightsholders generally choose to permit.²⁶ Second, the decision explicitly assumed that Field knew that through the upload his works would be subject to uses carried out automatically by Google’s crawlers.²⁷ Third, although Google provided technical means to easily prevent the use “in a matter of seconds”, Field chose not to make use of such means and therefore could be understood to accept the use.²⁸

The decision focused primarily on an omission by the rightsholder: Field made content available with knowledge of what uses would follow yet failed to communicate a dispreference towards the uses.²⁹ In positive terms, web site publishers therefore are expected to object to uses

²⁵ 17 U.S.C. § 106(1), (5).

²⁶ *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1118, 1121 (2006); *see also* Matthew Sag, *Copyright and Copy-Reliant Technology*, 103 NW. U. L. REV. 1607, 1618-19 (2009).

²⁷ *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1116-17 (2006) (“Field concedes that he knew that Google would automatically allow access to his works through “Cached” links when he posted them on the Internet unless he instructed otherwise. Field also knew that if an Internet user clicked on the “Cached” links to his web pages, the user would immediately download a copy of those pages from Google’s system cache.”).

²⁸ *Id.* at 1113, 1116-18.

²⁹ This is where *Field v. Google* differs from other search-related cases such as *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007). *Perfect10* concerned content not made available by the rightsholder, therefore, the rightsholder cannot be held responsible for the omission not to communicate a deviating preference concerning content she didn’t control.

by search engines in order to enforce their exclusive rights.³⁰ This also resembles an opt-out scheme as it is up to users to prevent unwanted yet otherwise lawful uses.

d. Mass Digitization and Access to Cultural Heritage

Other examples from the digital realm can be found in mass-scale digitization projects, such as the Google Books project as well as EU legislation concerning orphan and out-of-commerce works.

aa. Google Books

Google began in 2004 to create digital scans of books to create a searchable, digital corpus. Because they maintained copies of the works, made them available to project partners to download and displayed snippets to the public, the project prima facie violated the exclusive rights to reproduction, distribution, and public display,³¹ which provoked a class infringement action by the Authors Guild in 2005.³²

After long negotiations, both parties proposed a settlement³³ which would allow Google to make certain uses without prior authorization, in part against remuneration.³⁴ In return, they would have to establish a registry and use commercially reasonable efforts to locate rightsholders of unclaimed, “orphan” works to distribute payments.³⁵ However, in addition to the statutory right to request to be excluded as a member from the class,³⁶ the proposed settlement would have allowed rightsholders to exclude their books from the project.³⁷ The proposed settlement therefore established an opt-out scheme that would have put a burden to opt-out on all persons who, as of January 5, 2009, owned a US copyright interest in a book³⁸ in case they objected to such uses.

When the proposed settlement was later rejected, the opt-out structure proved to be a major obstacle. The court was concerned that the settlement would allow Google to expropriate the

³⁰ See also Oren Bracha & Patrick R. Goold, *Copyright Accidents*, 96 B.U. L. REV. 1025, 1090 (2016) (emphasizing the relevance of the opt-out possibility for the outcome).

³¹ 17 U.S.C. § 106(1), (3), (5).

³² See *Authors Guild v. Google, Inc.*, 804 F.3d 202, 208-09 (2d Cir. 2015).

³³ Amended Settlement Agreement, *Authors Guild, Inc. v. Google, Inc.*, No. 05-CV-8136-DC (S.D.N.Y. Nov. 13, 2009) [in the following: Amended Settlement Agreement].

³⁴ *Id.* at § 3.4.

³⁵ *Id.* at §§ 3.2(d), 5.2, 6.1.

³⁶ See *supra* III.; see also Michael R. Mattioli, *Opting out: Procedural Fair Use*, 12 VA. J.L. & TECH. 1, 14-17 (2007) (describing the history, benefits and underlying theory of class action opt-outs).

³⁷ See Amended Settlement Agreement §§ 1.124, 3.5(a)(i) (for books not yet digitized); § 3.5(a)(i) (for already digitized books).

³⁸ *Id.* at § 1.13.

rights of rightsholders in conflict with protections against involuntary transfers³⁹ and the statutory language of copyright's exclusive rights:⁴⁰ "As counsel for Amazon argued: '[T]he law of the United States is a copyright owner may sit back, do nothing and enjoy his property rights untrammelled by others exploiting his works without permission.' [...] Under the ASA, however, if copyright owners sit back and do nothing, they lose their rights".⁴¹ The court argued that it would contradict the purpose of copyright to expect the rightsholders to take affirmative steps to protect their works and feared that many authors would not know about the possibility to opt-out.⁴² While the Second Circuit eventually affirmed that Google engaged in fair use, the opt-out structure at least explicitly did not influence the analysis.⁴³

bb. Orphan and Out-of-Commerce Works in the EU

The EU in contrast has tackled the problem of both orphan as well as out-of-commerce literary works with legislative solutions. The Orphan Works Directive⁴⁴ from 2012 obliges member states to provide an exception or limitation for the use of certain literary works for which no rightsholder is identified or located despite a diligent search.⁴⁵ Certain public sector and cultural heritage institutions are permitted to reproduce such orphan works for the purposes of digitization and to make them available to the public.⁴⁶ Member states must ensure, however, that a rightsholder has at any time the possibility of putting an end to the work's orphan status.⁴⁷ This establishes an opt-out structure similar to the proposed Google Books settlement because uses are lawful without prior permission after carrying out a diligent search for rightsholders. They carry the burden of signaling their presence on the market *ex ante*, or of opting out by making themselves known *ex post*.

Concerning out-of-commerce works, the DSM Directive made use of a similar scheme. Works are deemed to be out of commerce when, after making a reasonable effort, it can be presumed in good faith that they are not available through customary channels of commerce.⁴⁸ Member states must enable collective management organizations (CMOs) to license certain uses

³⁹ See 17 U.S.C. § 201(e).

⁴⁰ See 17 U.S.C. § 106(1), (3).

⁴¹ *Authors Guild v. Google, Inc.*, 770 F. Supp. 2d 666, 680-81 (S.D.N.Y. 2011).

⁴² *Id.* at 682.

⁴³ *Authors Guild v. Google, Inc.*, 804 F.3d 202, 206 (2d Cir. 2015).

⁴⁴ Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on Certain Permitted Uses of Orphan Works.

⁴⁵ Arts. 1(2), 2(1), 3, 6(1) Orphan Works Directive.

⁴⁶ Arts. 1(1), 6(2) Orphan Works Directive.

⁴⁷ Art. 5 Orphan Works Directive.

⁴⁸ Art. 8(5) DSM Directive, with certain exceptions in Art. 8(7) DSM Directive.

of such works to cultural heritage institutions – irrespective of whether the respective rightsholder is a member of the CMO.⁴⁹ It suffices that the CMO is sufficiently representative for rightsholders in the relevant type of works subject of the license;⁵⁰ in case no such CMO exists, member states must provide for an exception or limitation of the same scope.⁵¹ Crucially, however, rightsholders must be able to exclude their works from the licensing mechanism or limitation at any time, easily and effectively.⁵²

Again, this provides for an opt-out scheme because the use is permitted independently of a prior authorization by rightsholders. To enjoy exclusivity, they cannot remain passive but must either keep their work commercially available or opt-out of uses with the CMOs or cultural heritage institutions. To facilitate the opt-out, potential users must inform about future uses at least six months before on an online portal managed by the EU Intellectual Property Office.⁵³

In a similar vein, the EU legislator enabled member states to provide this extended collective licensing mechanism also for commercially available works, allowing representative CMOs to license uses for rightsholders without prior authorization.⁵⁴ While the scope of this mechanism is not limited to digitization or cultural heritage uses, it applies only in well-defined areas of use where obtaining individual authorizations are onerous and impractical.⁵⁵ Member states have to ensure that rightsholders may at any time easily exclude their works from the mechanism.⁵⁶ The onus is therefore on the rightsholders to actively prohibit undesired uses through an opt-out. Like for out-of-commerce works, there is an information obligation on the part of users: Member states must provide for appropriate measures to inform rightsholders in general, not individually, about the use and their option to exclude in a reasonable period before a use commences.⁵⁷

e. Uses on Online Platforms

Lastly, another use case of an opt-out scheme can be seen in the limitations of copyright liability for uses by online service providers in the US as well as in the EU. Under 17 U.S.C. § 512(c), the so-called DMCA safe harbor, providers of online services are not liable for infringements with respect to user material if, among other conditions, they lack knowledge of an infringement and

⁴⁹ Art. 8(1) DSM Directive.

⁵⁰ Art. 8(1) lit. a DSM Directive.

⁵¹ Art. 8(2), (3) DSM Directive.

⁵² Art. 8(4) DSM Directive.

⁵³ Art. 10(1) DSM Directive.

⁵⁴ Art. 12(1) DSM Directive.

⁵⁵ Art. 12(2) DSM Directive.

⁵⁶ Art. 12(3)(c) DSM Directive.

⁵⁷ Art. 12(3)(d) DSM Directive.

act expeditiously to hinder access to infringing material upon obtaining such knowledge. While rightsholders are free to go after the user who uploaded the material, the service providers' liability consequently hinges on affirmative action by the rightsholders: They must claim the infringement with a notice that identifies their work and the infringing material in a way that the provider is able to locate the material. They must also provide their contact details so that the service provider may contact them.⁵⁸ In order to enforce copyright through injunctive relief against service providers, rightsholders therefore have to "arm" their exclusive right through affirmative action.⁵⁹ If they remain passive, service providers are not liable for infringing uses. The EU established a similar scheme with the controversial Art. 17 DSM Directive concerning online content-sharing service providers. On the one side, it provides for a direct liability of service providers and stresses the relevance of obtaining individual authorizations.⁶⁰ However, on the other side, if despite best efforts the service providers are unable to obtain an authorization, liability for infringements because of a communication to the public vanishes in particular under two conditions: The service providers must have (1) made best efforts to prevent the availability of works for which the rightsholders have provided relevant and necessary information, and (2), reminiscent of the DMCA safe harbor, acted expeditiously to disable access to works upon receiving a notice from rightsholders.⁶¹

The crucial part is that, again, the burden is on the rightsholders to provide information to the service providers. Unless they come forward, service providers are not liable for infringements. Both schemes therefore follow an opt-out logic: Uses are lawful by default unless there is affirmative action in the form of a notice by rightsholders.

3. Opt-Outs of AI Training Uses

Copyright opt-out schemes again received increased attention due to the advent of generative AI models. Such models can generate text, images, and other content based on user prompts; their development requires access to large amounts of data, such as text, images and videos.⁶² Under the hypothesis that a model's cognitive ability scales with an increase in training dataset size, AI companies compile ever larger training datasets. In large part, such datasets

⁵⁸ 17 U.S.C. § 512(c)(3).

⁵⁹ Cf. Wu, *supra* note 16, at 621.

⁶⁰ Art. 17(1) DSM Directive.

⁶¹ Art. 17(4)(a),(b) DSM Directive.

⁶² See Recital (in the following: Rec.) 105 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act).

consist of content web-scraped from online sources.⁶³ Web content is well suited for AI training due to its diversity in viewpoints, styles, and media types, which has a positive effect on model performance.⁶⁴ It also represents a low-hanging fruit because such data is cheap to collect by making use of already compiled publicly available datasets. Other sources of training data are licensing deals with publishers and own digitization efforts.⁶⁵

Copyright issues arise throughout the various stages of generative AI model development and deployment. A significant amount of training data consists of copyright-protected works. While the analysis may differ depending on the model at stake, the compilation of training data generally requires reproductions of works that touch upon the exclusive rights of rightsholders under 17 U.S.C. § 106(1) and Art. 2 InfoSoc Directive⁶⁶.⁶⁷ Therefore, unless relevant copyright exceptions and limitations apply prior authorization by the rightsholder is required.

In the EU, while not being finally resolved, such reproductions are deemed to be covered by the exception or limitation for text and data mining that member states are required to transpose under Arts. 3, 4 DSM Directive⁶⁸.⁶⁹ The limitation under Art. 4 DSM Directive for purposes other than scientific research, however, only applies on the condition that the use has not been expressly reserved by rightsholders in an appropriate manner.⁷⁰ Unlike conventional exceptions or limitations, the provision does not exempt uses conclusively but puts their lawfulness in the

⁶³ See *Our Approach to Data and AI*, OPENAI, <https://openai.com/index/approach-to-data-and-ai/> (last visited Jan. 30, 2026) (claiming to take web crawler permissions for AI into account each time a new model is trained); Billy Perrigo, *TIME100 AI 2024: Sundar Pichai*, TIME (Sept. 5, 2024), <https://time.com/7012748/sundar-pichai-2/> (stressing that Google has “given a clear way for people to opt out of [...] AI training”); Jess Weatherbed, *Anthropic’s Crawler Is Ignoring Websites’ Anti-AI Scraping Policies*, THE VERGE (July 25, 2024), <https://www.theverge.com/2024/7/25/24205943/anthropic-ai-web-crawler-claudebot-ifixit-scraping-training-data> (observing Anthropic’s Claude Bot to engage in comprehensive scraping); *Introducing Meta Llama 3*, META AI, <https://ai.meta.com/blog/meta-llama-3/> (last visited Jan. 30, 2026) (stating that “Llama 3 is pretrained on over 15T tokens that were all collected from publicly available sources”).

⁶⁴ Cf. Wayne Xin Zhao et al., *A Survey of Large Language Models* (Mar. 11, 2025), <http://arxiv.org/abs/2303.18223> (mentioning studies that have shown that training on less diverse data degrades the generalization capability of LLMs, while increasing the data source heterogeneity is critical for improving performance).

⁶⁵ See e.g. *Partnership with Axel Springer to Deepen Beneficial Use of AI in Journalism*, OPENAI (Mar. 13, 2024), <https://openai.com/index/axel-springer-partnership/>.

⁶⁶ Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society.

⁶⁷ See e.g. Matthew Sag, *Fairness and Fair Use in Generative AI*, 92 FORDHAM L. REV. 1887, 1893 (2023) (stating that typically training involves reproductions); Katherine Lee, A. Feder Cooper & James Grimmelmann, *Talkin’ Bout AI Generation: Copyright and the Generative-AI Supply Chain (The Short Version)* (Mar. 16, 2024), <https://papers.ssrn.com/abstract=4762060> (emphasizing that every stage in the generative AI supply chain requires a reproduction).

⁶⁸ Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC.

⁶⁹ See e.g. Regional Court of Hamburg, judgment of 27 September 2024 – 310 O 227/23 – LAION, LTZ 2025, 65.

⁷⁰ Art. 4(3) DSM Directive.

discretion of the rightsholder. Uses are permitted without prior authorizations, unless the rightsholder objects. This establishes an opt-out scheme: Instead of being in the comfortable position to decide which use for what conditions to affirmatively opt into, rightsholders are forced to opt out of unwanted uses. Copyright's logic of exclusivity is therefore reversed.

In the US, at least for now the legal framework de facto follows the same logic. While the question of whether AI training uses infringe on exclusive rights of rightsholders or instead qualify as a fair use is being litigated, AI companies continue to use copyright-protected content crawled from the open web when compiling training data. Because some of the big AI companies have announced to respect certain forms of opt-outs,⁷¹ it is up to the rightsholders to either make their works unavailable for such uses in the first place or to communicate their objection, which has resulted in a significant increase of opt-out declarations.⁷²

III. The Normative Case for Opt-Outs

Some of the existing use cases have individually been subject to legal analysis.⁷³ Building on these individual discussions, the following part identifies the normative justification for opt-outs in copyright law. It asks whether and under what conditions opt-out schemes functions in the interest of both rightsholders and users. First, it identifies the essential features of the use cases: an impracticality of individual transactions and a – from a copyright perspective – special set of interests. Opt-outs can then be functionally understood as solutions to the problems these features cause. They preserve the autonomy of rightsholders by redistributing the transaction cost through an entitlement with a flexible protection level, and they make the risk of infringement manageable that users must bear by coupling the liability of users to an information disclosure from rightsholders.

1. Transaction Cost and Interests in the Use Cases

When looking at the broad picture of existing opt-out schemes, two similarities stand out.

⁷¹ See, e.g. *Overview of OpenAI Crawlers - OpenAI API*, OPENAI, <https://platform.openai.com/docs/bots> (last visited Jan. 30, 2026); *Does Anthropic Crawl Data from the Web, and How Can Site Owners Block the Crawler?* | *Anthropic Help Center*, ANTHROPIC, <https://support.anthropic.com/en/articles/8896518-does-anthropic-crawl-data-from-the-web-and-how-can-site-owners-block-the-crawler> (last visited Jan. 30, 2026); but see also Weatherbed, *supra* note 63 (giving anecdotal evidence of the ClaudeBot web crawler hitting a website that opted-out a million times in 24 hours).

⁷² Shayne Longpre et al., *Consent in Crisis: The Rapid Decline of the AI Data Commons* (July 24, 2024), <http://arxiv.org/abs/2407.14933>.

⁷³ See e.g. Bracha, *supra* note 3 (concerning Google Books and digital libraries); Mattioli, *supra* note 36 (concerning Google Books); Su, *supra* note 4 (concerning the Google Books project); Sag, *supra* note 26 (concerning so-called copy-reliant technologies such as search engines and the Google Books project).

First, almost all use cases concern situations where individual transactions are impractical: The group of opt-outs concerning mass digitization and access to cultural heritage involves uses where the amount of the works to be used and the corresponding rightsholders is so high that the cumulative transaction cost become staggering. The Google Books Project was estimated to involve clearance cost of above \$10 billion, not including the following royalties that would be paid to authors.⁷⁴

The provisions concerning orphan and out-of-commerce works are also intended to facilitate mass digitization projects, where due to the unavailability of rightsholders and their general disinterest in an economic exploitation, it becomes at least very resource-intensive if not impractical to acquire permissions in individual transactions.⁷⁵

The same applies to uses by search engines that involve a very large amount of works from the open web.⁷⁶ Concerning the uses by online platforms, it is not the platforms but the users that decide which content to upload which makes it difficult to obtain authorizations prior to the display of works; in addition, the platforms also engage in mass uses of many works with many rightsholders.

For the restoration of copyrights, the problem is not that transactions are impractical due to the high transaction cost because such uses not necessarily involve a large amount of works. Rather, users have no reason to initiate a transaction in the first place as the works were part of the public domain before the restoration.

The limitation for public performances under 17 U.S.C. § 110(4)(B) in contrast stands out in this regard. It may be unreasonable to expect performers that stand not to gain in any way from the spontaneous recitation of a poem, for example, to obtain an authorization – however, this relates to the quality and not, as for the other use cases, the quantitative aspects of the use. The impracticality of transactions here also arises from a disproportionality between the transaction cost and the anticipated benefits of the use for an individual user.⁷⁷ The public performances that § 110(4)(B) allows for are often dispensable for the user and therefore easily discouraged by even small transaction costs. Furthermore, any positive effects of the performance on others are not

⁷⁴ See Sag, , 26, at 1662.

⁷⁵ Cf. Rec. 7 Orphan Works Directive, Rec. 30 DSM Directive.

⁷⁶ While limited amounts of high-quality content are available for licensing, so-called vernacular works, meaning the vast amount of online content created by the general public, generally is not, see U.S. Copyright Office, *supra* note 2, at 70.

⁷⁷ See Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 COLUM. L. REV. 1600, 1628-30 (1982) (describing different instances of market breakdown that may justify a fair use finding).

internalized by the performers. What makes transactions therefore impractical here is the often spontaneous, opportunistic nature of the use.

Second, the use cases reflect specific objective interests of rightsholders and users: In general, for most of the use cases, on the one side, users have a qualified interest in the use, while on the other side the interest of rightsholders in prohibiting the use usually is decreased.

This is particularly evident for uses by search engines: There is a large public interest in enhancing information-gathering techniques on the internet, and many website publishers invite uses by search engines to profit from traffic that they direct to their site.⁷⁸ Concerning the mass digitization use cases, there is a large common interest in the preservation and dissemination of cultural heritage for future generations.⁷⁹

The status of a work as orphan or out-of-commerce, which often involves older works where an individual exploitation is not promising, the possibility for rightsholders to profit from the use, and the special user base of public institutions acting for non-commercial purposes indicate also a diminished interest of rightsholders in exclusivity.

For content-sharing uses on online platforms, there also is a qualified public interest in the wide access to cultural and creative works that the platforms provide while rightsholders in general have a low interest in preventing such uses: The debate around the controversial Art. 17 DSM Directive, for example, was marked by demands of closing a “value gap” between rightsholders and the platforms, not of preventing the platforms from using works at all.⁸⁰

The limitation for performances of a nondramatic literary or musical work ensures through several requirements that the performing party will not profit commercially from the use and excludes transmissions to the public which lowers the interest of rightsholders in insisting on an individual transaction. Also, there is a large public interest in allowing for a spontaneous and permission-less performance of works in the public as a part of the cultural and social sphere.

A special case in this regard is the restoration of copyrights. The interests of users are qualified in that they relied in good faith on the public domain status of the works. For a decreased interest in exclusivity by rightsholders, however, the only evidence is the failure to

⁷⁸ It remains to be seen whether website publishers will continue to invite such uses when search engines stop sending traffic outside of their own ecosystem, *see e.g.* Nilay Patel, *Google Zero Is Here — Now What?*, THE VERGE (May 30, 2024), <https://www.theverge.com/24167865/google-zero-search-crash-housefresh-ai-overviews-traffic-data-audience>.

⁷⁹ *Cf.* Rec. 1 Orphan Works Directive; Rec. 25 DSM Directive.

⁸⁰ This is demonstrated for example by the fact that between July and December 2023 rightsholders in over 90% of all Content ID claims choose to earn revenue from the ads run against a video instead of blocking it, *see YouTube Copyright Transparency Report*, GOOGLE TRANSPARENCY REPORT, <https://transparencyreport.google.com/youtube-copyright/intro> (last visited Jan. 30, 2026).

comply with US formalities which may rather be indicative of the complicated nature of the formalities.

2. Preserving the Autonomy of Rightsholders

With regards to this distribution of the transaction cost and general interests concerning a use, in the use cases the opt-out schemes function to preserve the autonomy of rightsholders. They do so by redistributing the transaction costs between rightsholder and user which – under certain conditions laid out below – helps to realize the diverging interests of different groups of rightsholders.

a. Market Failure under a Property Rule

Following the taxonomy of *Calabresi/Melamed*, we can distinguish between entitlements protected by a property rule, meaning in the copyright context that users must obtain a permission for a use through voluntary transactions, and a liability rule, where users are free to use a work yet have to compensate the rightsholder for it.⁸¹ When rightsholders may not prevent others from using a work but users also owe no compensation for the use, the protection level can be described as a zero-price liability rule.⁸²

Copyright is usually protected with a property rule. This is built on the assumption that a significant amount of rightsholders will want to exploit their works in an exclusive mode, for which the property rule not only saves them from wasteful effort of protecting their works and policing for unauthorized uses but also provides them with a mechanism to enforce negotiations through injunctive relief. Rightsholders that in contrast want to allow for uses of their works without individual transactions are free to do so yet carry the burden of signifying their permission to others.⁸³ Copyright's entitlement design therefore structurally prefers exclusivity; however, this imbalance is sound as long as the entitlement may change hands following the decision of the rightsholder.

What the similarities regarding the interests of both parties and the transaction costs show, though, is that in the opt-out use cases these assumptions are unfounded. Because of the high number of rightsholders and works to be used as well as the disproportion between the use and

⁸¹ See Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1105-06 (1971).

⁸² See Mark A. Lemley & Philip J. Weiser, *Should Property or Liability Rules Govern Information*, 85 TEX. L. REV. 783, 786 (2006); Abraham Bell & Gideon Parchomovsky, *Liability Rules*, 101 MICH. L. REV. 1, 6 (2002) (describing such rules as “zero-order”).

⁸³ Cf. Bracha, *supra* note 3, at 1827 (emphasizing the substantial high cost of such a rule concerning the access to websites).

its benefit, transactions are impractical and will mostly fail already at the stage of searching and identifying the rightsholder as the transaction partner. A property rule therefore leads to a market failure: The group of rightsholders that want to exploit their works exclusively cannot do so because they have no transaction partner, and the group of rightsholders that want to allow for uses may signify their permission yet to no avail as their signal will often not be found by potential users. The protection by a property rule therefore only serves the group of rightsholders that want to prevent all uses of their works permanently; all other rightsholders are stuck with an entitlement they (and others) cannot make use of.

This is problematic for copyright especially in two ways: For one, it interferes with the incentive of authors to create works in the first place because the prospect of amortizing their production cost is curtailed. More importantly, however, it constrains the autonomy of rightsholders. The freedom to decide on the use of one's creation free from the influence of others serves not only to optimize the allocation of entitlements. It also carries a normative dimension of self-determination that sets a property rights-based copyright system apart from other regimes capable of incentivizing creativity, such as private or state patronage.⁸⁴ It is therefore not enough for copyright to seclude to a paradigm of exclusivity – what needs to be justified in light of its purpose is not only whether the entitlement is granted but also what protection level is chosen.⁸⁵ Because of the orientation of copyright in the US towards promoting the dissemination of works to promote innovation in the public interest, it is not viable to one-sidedly favor exclusivity in the use cases because this would leave the idea of autonomy incomplete.

b. Property Rules as a Solution

Opt-out schemes solve this problem by breaking up the conventional taxonomy of protection levels. Under an opt-out scheme, as a default users may make use of a work without prior authorization by rightsholders, meaning that the right is protected under a liability rule or a zero-price liability rule.⁸⁶ What makes the entitlement special, however, is that it exhibits a second state, which is triggered by the opt-out declaration.⁸⁷ As soon as rightsholders opt-out, voluntary transactions become necessary and users must obtain a permission prior to any use, which flips

⁸⁴ See Sag, *supra* note 26, at 1675 (stressing also the relevance of autonomy for the institutional design of copyright).

⁸⁵ *Id.* at 1673.

⁸⁶ There is a liability rule in place concerning orphan works in the EU, *cf.* Art. 6(5) Orphan Works Directive, as well as some transpositions of Art. 17 DSM Directive. In all other cases, there are zero-price liability rules.

⁸⁷ See Bracha, *supra* note 3, at 1827.

the protection level to a property rule. Opt-out regimes therefore provide an entitlement whose protection level by design features both a liability and a property rule state, for which *Bell/Parchomovsky* coined the portmanteau “loproperty rule”.⁸⁸

Loperty rules complement the taxonomy of property and liability rules with a temporal dimension.⁸⁹ The protection level can be differentiated not only along a binary of whether to require voluntary transactions but also *when* to require them. Regarding the context of the opt-out use cases, loperty rules provide for the necessary flexibility that a static property or liability rule would lack. Under a static property rule, copyright would produce a market failure because transactions fail structurally due to the high transaction cost; a static liability rule on the other side would either come with the Achilles’ heel of out-of-market pricing that may cause windfall profits or undercompensation, or, when in form of a zero-price liability rule, provides for no compensation at all.⁹⁰ What the loperty rule allows is to combine the advantages of both protection levels in a way that offsets the disadvantages of both.⁹¹ In the initial state, the entitlement is protected under a (zero-price) liability rule which solves the transaction cost problem because uses are generally lawful without prior transactions. However, rightsholders are always free to substitute the missing or inadequate compensation with a market-based one by switching to the secondary state of the property rule. Where one protection level would fail, the other steps in.

c. Conditions for an Autonomy-Preserving Effect

Loperty rules have a preserving function for the autonomy of rightsholders as the entitlement does not remain stuck in a property rule but allows for voluntary transactions (only) in case they are demanded. The entitlement is therefore at least formally capable of realizing the interests of all three groups from above: Rightsholders that want to allow for uses without individual transactions can remain passive; rightsholders that want to exploit their works through voluntary transactions may do so by opting out into the property rule; and the same goes for rightsholders that want to prevent any use of their works permanently.

However, whether all interests are in fact realized depends on the context of the use at stake. Just as high transaction costs under an opt-in logic keep otherwise willing rightsholders in a pointless property rule, an opt-out scheme may bind unwilling rightsholders to a similarly

⁸⁸ See Bell and Parchomovsky, *supra* note 82, at 31, 53-54.

⁸⁹ See *id.* at 5-6, 65.

⁹⁰ See Lemley and Weiser, *supra* note 82, at 820 (identifying the challenge of setting the terms of access under uncertainty as the Achilles’ heel of the case for liability rules).

⁹¹ See *generally* Bell and Parchomovsky, *supra* note 82, at 7 (describing pliability rules).

detrimental (zero-price) liability rule. To prevent the lopertry rule from being just as sticky in its initial state as a static property rule, two factors are important: the rightsholder's awareness of the use and the opt-out possibility, as well as the ease of the opt-out mechanism.

The awareness of the rightsholder must relate to two circumstances in order for the scheme to work. First, it must involve the fact that a copyright-relevant use of a work of the rightsholder is taking place. If a rightsholder is not aware of the use, she will be unable to take affirmative steps to forbid it, even if she disprefers the use. Moreover, the awareness must also extend to the fact that there is a possibility for the rightsholder to opt-out.⁹² Even if a rightsholder knows that an unwanted use is underway, she will not opt-out if she thinks that she has no other option than to tolerate the use which is usually the case under conventional limitations or exceptions in copyright law. The rightsholder's awareness of both the use as well as its opt-out character is therefore a minimum requirement for the scheme to have an autonomy-preserving effect.

To be sure, this does not necessarily mean that rightsholders always must be positively informed about both circumstances. It is sufficient that rightsholders are put in a position which makes it likely that unwilling rightsholders will find out about uses and their ability to do something about it, and will obtain the necessary information themselves.⁹³ This brings information obligations of users into focus. Some of the existing use cases already demand a passive or even active information of the rightsholder by the user: The crawling of web pages by search engines, for example, is usually traceable through server logs; uses of orphan works and out-of-commerce works both require an active information about the use, however, not to the user but an intermediary database provider.

The effort that can reasonably be expected from rightsholders to become aware of the right to opt-out follows from the interplay of the level of information provided by users and the rightsholders' ability to monitor uses and their options. If both are optimized, the activation of rightsholders in order to initialize transactions under the lopertry rule will lower the transaction cost when compared with a property rule.⁹⁴

⁹² See also Kevin J. Hickey, *Consent, User Reliance, and Fair Use*, 16 YALE J.L. & TECH. 397, 429 (2013) (identifying the awareness of the proposed use as well as of the opt-out option as characteristics of a type of consent-seeking interactions).

⁹³ This addresses only the information cost that necessarily prevent an opt-out. It does not, however, tell rightsholders how an opt-out compares to an exclusive exploitation, and therefore does not prevent uninformed opt-outs. This is because while uninformed opt-outs might be unfavorable from a public policy perspective, copyright law generally does not discriminate on the basis of the reason rightsholders object to a use. See generally Oren Bar-Gill & Omri Ben-Shahar, *Rethinking Nudge: An Information-Costs Theory of Default Rules*, 88 U. CHI. L. REV. 531, 536 f. (2021) (pointing out that high information costs do not prevent people from opting out, resulting in uninformed opt-outs).

⁹⁴ Cf. Bracha, *supra* note 3, at 1836 (assuming in the context of a digital library that given adequate disclosure, a lopertry rule will lower the search cost compared to a property rule).

The second condition that needs to be fulfilled to respect the preferences of all groups of rightsholders concerns the difficulty of the opt-out mechanism.⁹⁵ Even if a rightsholder knows of the use and her option to opt-out, if the opt-out is too burdensome she will refrain from doing so. The ease of the means to switch to the property rule therefore plays a decisive role for realizing the autonomy of rightsholders.⁹⁶ Compared with the cumulative transaction cost under a property rule, the property rule has the advantage that rightsholders only need to become active concerning works they want to prevent uses of, unlike under a property rule where users need to find out a rightsholder's preference towards their use concerning all works.⁹⁷

Again, information delivered by users influences the rightsholders' effort. If rightsholders are required to opt-out vis-à-vis single users but such users provide rightsholders with the necessary details to easily address them, the opt-out becomes easier. Another important aspect in this regard is the nature of the opt-out. Digital opt-outs can be cheap and signal dispreference to every prospective user once and for all, which will often result in lower transaction cost compared to single users initiating individual transactions with rightsholders under a property rule.

This is exemplified by opt-outs from search engines. In *Field v. Google*, the court emphasized that the opt-out mechanism, either meta-tags or robots.txt files, were highly publicized and well-known industry standards that could be employed "in a matter of seconds".⁹⁸ However, the difficulty can vary depending on the skills of the person that is required to make use of the mechanism.⁹⁹ The DMCA safe harbor, for example, requires a distinct form of notice which encompasses information that will always be present with the rightsholder.¹⁰⁰ Service providers must designate an agent that receives the notifications, whose information, including an email address, has to be made available on its website as well as to the Copyright Office, which then

⁹⁵ This concerns so-called "mechanical costs", see Bar-Gill and Ben-Shahar, *supra* note 93, at 534 (describing mechanical costs as the actual costs of opting-out).

⁹⁶ See also Sag, *supra* note 26, at 1607 (emphasizing that the costs need to be sufficiently small such that autonomy is preserved); Bracha, *supra* note 3, at 1836 (describing that an opt-out regime will be more cost-efficient i.e. to the extent the opt-out mechanism is cheap, accessible and easy); Hickey, *supra* note 92, at 420, 423 (assuming that the communication costs of rightsholder and user, which entail the cost of rightsholders to announce their preference and to opt-out, must not exceed the expected social benefits).

⁹⁷ See Bracha, *supra* note 3, at 1835; see also U.S. Copyright Office, *supra* note 2, at 105 (arguing concerning AI opt-outs that they should be straightforward enough for rightsholders without legal or technical knowledge to use them).

⁹⁸ *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1116 f., 1119 (2006); see also Sag, *supra* note 26, at 1607 (describing the opt-outs as "virtually costless").

⁹⁹ In *Field v. Google Inc.*, the rightsholder that had to opt-out was an author and attorney, that had implemented a robots.txt file to invite search engines to index his site, *cf.* *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1117 (2006).

¹⁰⁰ *Cf.* 17 U.S.C. § 512(c)(3)(A).

publishes a directory.¹⁰¹ This lowers the effort to opt-out immensely. The European counterpart, Art. 17 DSM Directive, in contrast specifies neither the scope nor the way to provide information to opt-out of uses, which allows for flexibility in the transposition yet may aggravate the burden for rightsholders. The transposition in Germany, for example, requires the appointment of a person authorized to receive service that must be pointed out in an easily recognizable and directly accessible manner on the service provider's website.¹⁰² The enforcement of a restored copyright will take a similar low effort as the required information is trivial for rightsholders to give,¹⁰³ and it is sufficient to send the notice to the Copyright Office, which makes it obsolete to identify users.

The notice of objection concerning unwanted public performances under § 110(4)(B), in contrast, is rather complex: While rightsholders must only state any reason for their objection, they must do so in writing, and the notice must be signed and be served on the person responsible for the performance at least 7 days before. Crucially, there are no obligations for users to inform about a planned performance.¹⁰⁴ Therefore, this will allow rightsholders to prevent performances only in exceptional cases where they are aware beforehand.

d. Reliance Interests

The flexible ontology of property rules comprising of property and liability rule elements does not speak to the conditions under which the switch from one level of protection to another should be available to rightsholders. A limiting factor in this regard could be reliance interests by users.¹⁰⁵ Users might rely on the possibility to use works permissionless in case of an initial zero-price liability rule; under an initial liability rule their reliance interest might be even higher because they compensated rightsholders for their use (even though their compensation only covers uses that have already taken place and does not secure a future right to use). Furthermore, there will most likely be situations where users have made investments in reliance on the possibility to use

¹⁰¹ Cf. 17 U.S.C. § 512(c)(2).

¹⁰² Gesetz über die urheberrechtliche Verantwortlichkeit von Diensteanbietern [UrhDaG] [Act on Copyright Liability of Online Content Sharing Service Providers], May 31, 2021, BGBl. I at 1204, 1215, § 20 (Ger.).

¹⁰³ See 17 U.S.C. § 104A(e), which requires the title of the work, an address and telephone number at which the owner may be contacted, and for a notice served directly on a reliance party, in addition the use to which the owner objects.

¹⁰⁴ See 2 Nimmer on Copyright § 8.15 (2025) (criticizing the missing information of the rightsholder and suggesting that courts may hold that a performer may impose an admission only on the condition that a reasonable effort has been made to give the copyright owner notice of a performance).

¹⁰⁵ See generally Bo S. L. Kim, *Copyright's Public Reliance Interests*, 99 WASH. L. REV. 107, 126-29 (2024) (describing reliance interests generally as arising from various uses by persons even lying outside a bilateral infringement context, however, narrowing it down applied to the scenes a faire doctrine to uses that in the aggregate give rise to audience expectations through for example cultural recognizability).

works and are therefore “locked-in” to a use.¹⁰⁶ If rightsholders can deprive them of the possibility to use without further ado, these interests could easily be frustrated.

When users have something to lose, it also invites strategic behavior by rightsholders to extract rents by holding out their position. *Sag* identifies such conduct at play for example in *Field v. Google* where the rightsholder was found to have no genuine objection to the opt-out logic but deliberately tried to exploit Google’s standard practice of caching websites to extract surplus value.¹⁰⁷ Crucially, however, the problem there was that the plaintiff did not conform to the opt-out-scheme when objecting to the use, not that he objected at all.¹⁰⁸ What makes dealing with strategic behavior so difficult – both in the case of an opt-out as well as a conventional exploitation based on an exclusive right – is that copyright is precisely intended to enable rightsholders to exploit their position.¹⁰⁹ The statutory monopoly granted to rightsholders entails the risk that the value they extract derives not from characteristics of their work but from the permission cost that users would have to bear if the opt-out scheme would be held unlawful (like in the *Field* case, based on the conventional exclusive right), or the then sunk cost the user incurred in expectation of the possibility to use (like in the opt-out use cases, on the basis of a successful opt-out).

Concerning opt-outs, there are different reasons for giving rightsholders such an ability. For one, it can be argued that users are no worse off than they would be under a conventional, unconditional exclusive right. Rather, the opt-out structure puts them in a favorable position which the rightsholder can then withdraw again without putting them at a disadvantage compared to the conventional state. However, such an argument must logically subscribe to the exclusivity paradigm from which an opt-out scheme then deviates. What the normative case for property rules though has shown is that the level of protection must be justified considering each use. Where the conditions for exclusivity are not given, such as in the cases where opt-out schemes apply, it is inconsequent to argue that users are in a better position compared to a protection level that does not apply in the first place.

Another argument in favor of an unrestricted opt-out right would go along the lines of “you should have known better”, pointing out that it should have been clear to users from the start

¹⁰⁶ See e.g. Hickey, *supra* note 92, at 422 (describing such a scenario abstractly).

¹⁰⁷ *Sag*, *supra* note 26, at 1679-80; see also *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1117 (2006) (assuming that *Field* decided to manufacture a claim for copyright infringement against Google in the hopes of making money from Google’s standard practice).

¹⁰⁸ See *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1117 (2006) (stressing that *Field* could have informed Google not to provide cached links but choose to remain silent).

¹⁰⁹ See *Sag*, *supra* note 26, at 1680 (stating that ordinarily, a copyright owner should be allowed to hold-out insofar as to realize the value she places in her works).

that rightsholders may opt-out of the use. However, this line of reasoning would devalue the initial state of the property rule. Who would make use of it for investment-intense uses such as complex digitization projects under the constant risk that the rightsholder is equipped with the option to extract any surplus value or prohibit any further use?

A normative case for opt-outs therefore must consider both the interests of rightsholders in the opt-out possibility as well as the reliance interests of users. An opt-out could for example require rightsholders to compensate users in exchange for the gained property rule protection. This would also ensure that rightsholders only switch if the exclusive right is worthwhile. Yet, in case the opt-out is intended to take effect erga omnes, there is no concrete user as a counterpart to compensate. Also, under a property rule the rightsholder is already burdened with the cost of monitoring uses, giving notice and negotiating the use.¹¹⁰

Instead, inspiration for an adequate balancing can be found where the reliance interests of users are probably the greatest: an opt-out that involves public domain works, which is the case for the restoration of copyright under § 104A. There, the reliance interests are protected through three instruments: First, the time frame for the rightsholder to opt-out is limited to 24 months after the date of restoration.¹¹¹ This not only provides for repose in the interest of the user, it is also consistent with the generally short exploitation period of most works.¹¹² Opt-outs after this period are more likely not to be connected to legitimate exploitation or dissemination motives of the rightsholder. Second, rightsholders are required to give notice of their intent to enforce exclusive rights which prevents inadvertent infringements.¹¹³ Third, remedies are barred for infringements that ended before a grace period of 12 months after giving notice.¹¹⁴ Users therefore have planning security and are provided with a use-up term, which enables them to control the risk of opportunistic opt-outs. The grace period also ensures that users can prevent rightsholders from getting to a troublesome post-use bargaining position, where users have no leverage other than taking their chances with a fair use defense and forcing the rightsholder to sue.¹¹⁵

3. Controlling the Infringement Risk of Users

¹¹⁰ See Bracha, *supra* note 3, at 1833.

¹¹¹ 17 U.S.C. § 104A(d)(2)(A)(i).

¹¹² See e.g. Kristelia García, James Hicks & Justin McCrary, *Copyright and Economic Viability: Evidence from the Music Industry*, 17 J. EMPIRICAL LEGAL STUD. 696 (2020) (suggesting empirically a quick depreciation of commercial music).

¹¹³ Cf. 17 U.S.C. § 104A(e).

¹¹⁴ 17 U.S.C. § 104A(d)(2)(A)(ii), (B)(ii).

¹¹⁵ See Hickey, *supra* note 92, at 419-20 (describing the situation where rightsholders are aware of the use but decline to announce their preferences beforehand, leading to a favorable bargaining position).

Another way to look at the normative advantages of opt-out schemes is focusing on users through the lens of the risk of infringement. While generally users must bear the infringement risk, under an opt-out scheme the risk-inducing exclusivity of copyright is subject to a procedure. This proceduralization serves as an information-forcing mechanism that forces rightsholders to provide users with the information they lack to initiate a transaction.

a. Proceduralization

What makes the use cases problematic for users are information deficits.¹¹⁶ Because of the high transaction cost, users lack the information whether a certain work is subject to copyright protection, who the rightsholder is and whether she permits the use or not. However, under a property rule, they must bear the risk of infringement.¹¹⁷ Rational users therefore refrain from a use as they are unable to control the risk with reasonable effort. The infringement risk consequently has a chilling effect on lawful uses of works that are not (anymore) subject to protection or whose rightsholders are willing to permit uses without explicit transactions.

Under an opt-out scheme, this undersupply of information on the part of users is remedied by a proceduralization of the exclusive right. In a broader sense, proceduralization describes a reinforcement or replacement of substantive criteria with formal procedures.¹¹⁸ It relieves legislators, for example, from having to make substantial decisions under uncertainty about their effects. Regulated is not the “what”, meaning a certain substantive result, but the “how”, the procedural way to get there.¹¹⁹ In a narrow sense, a copyright scheme that subscribes to an opt-out logic procedurally restructures the exclusive right.¹²⁰ Only after a rightsholder has successfully complied with a defined procedure, the opt-out, property-rule protection takes effect.¹²¹ Therefore, instead of a substantive determination of the protection level to govern a certain use case, only a binding formal procedure is put in place that delegates the decision-making authority but also redistributes the burden of action to the rightsholder.

¹¹⁶ See Bracha, *supra* note 3, at 1839.

¹¹⁷ Cf. 17 U.S.C. §§ 501(a), 502.

¹¹⁸ Cf. SEBASTIAN HIRSCHBERGER, PROZEDURALISIERUNG IM EUROPÄISCHEN BINNENMARKTRECHT: DIE VERFAHRENDIMENSION DER GRUNDFREIHEITEN UND DES BEIHILFENRECHTS [PROZEDURALIZATION IN EUROPEAN INTERNAL MARKET LAW] 27 (2010).

¹¹⁹ See e.g. Tatjana Sheplyakova, *Prozeduralisierung des Rechts – Tema con Variazioni*, in PROZEDURALISIERUNG DES RECHTS [PROZEDURALIZATION OF LAW] 1, 1, 6 (Tatjana Sheplyakova ed., 2018); GERALF-PETER CALLIESS, PROZEDURALES RECHT [PROZEDURAL LAW] 178 (1999).

¹²⁰ Dan Wielsch, *Die Zugangsregeln Der Intermediäre: Prozeduralisierung von Schutzrechten*, GRUR 665, 670 (2011) (coining the term of a procedural re-structuring of property rights).

¹²¹ For an example of proceduralization of the exclusive right in patent law, see Case C-170/13, Huawei Techs. Co. v. ZTE Corp., ECLI:EU:C:2015:477, ¶¶ 60–65 (July 16, 2015).

b. Information-Forcing Mechanism

The proceduralization of copyright has an influence on the distribution of information between rightsholders and users. The cheapest information provider for the information that users lack are rightsholders.¹²² They are aware of their ownership, their preference towards uses and ways how users can get in contact. With exclusivity as a default under an opt-in logic, however, there is no incentive for rightsholders to come forward with this information.¹²³ They can “sit back, do nothing and enjoy [their] property rights untrammelled by others exploiting [their] works without permission”¹²⁴.

A property rule changes this. Like a default rule in contract law, however, here applied in a non-negotiation setting,¹²⁵ it defines a default state of liability rule protection that may be substituted for a second state of property rule protection. As there generally is a qualified user interest in the use while the interest of rightsholders in exclusivity is usually diminished,¹²⁶ this default rule conforms with the predominant preferences of the parties. In terms of the economic analysis of default rules, it therefore represents a majoritarian default rule. A majoritarian default rule serves to facilitate an efficient allocation of resources by providing regulatory patterns as a default that are based on typical, i.e. generally majority, preferences.¹²⁷ For the remaining minority with diverging preferences, however, the rule has a penalizing effect:¹²⁸ They must derogate the default rule to have the scheme reflect their interests. Rightsholders that want to exploit their works exclusively therefore are incentivized to take affirmative steps and opt-out into a property rule.¹²⁹

The crucial point about this arrangement is that the opt-out mechanism has an information-disclosing effect. The opt-out marks not only the switch to the property rule but also signals to

¹²² See Bracha, *supra* note 3, at 1839.

¹²³ *Id.*

¹²⁴ Authors Guild v. Google, Inc., 770 F. Supp. 2d 666, 681 (S.D.N.Y. 2011).

¹²⁵ See Bracha, *supra* note 3, at 1840.

¹²⁶ See *supra* III.3.a.

¹²⁷ Ian Ayres & Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 99 YALE L.J. 87, 89 (1989) (describing this as an “would have wanted” approach).

¹²⁸ See Bracha, *supra* note 3, at 1839. The rule here does not resemble, however, a penalty default rule. Penalty default rules are intentionally designed not to reflect the preference of a party to incentivize it to disclose information, which in certain contexts promises a higher efficiency than a majoritarian default rule, see Ayres and Gertner, *supra* note 127, at 95-100. In other words, they require a derogation of the default rule to be efficient. The use cases of opt-outs in copyright, however, not necessarily have to result in a derogation – it will often be in the interest of both parties for the (zero-price) liability rule to stay in place because it allows for permission-less uses.

¹²⁹ But see Hickey, *supra* note 92, at 448-451 (pointing out that this also incentivizes rightsholders to preemptively claim protection for all their works out of fear to lose their rights and thereby discourage “tolerated uses”, meaning uses that are “technically infringing, but often permitted in practice”); for the notion of tolerated uses, see Wu, *supra* note 16.

users that the party that opts out assumes that the work is copyright-protected, that they claim to be the owner of it, and that they prefer no use to occur without prior individual transactions.

For one, forcing rightsholders to reveal their dispreference in advance before a use prevents harmful opportunism that would create social cost as it deters socially meritorious uses.¹³⁰ Furthermore, while this information will not solve all problems for users concerning their infringement risk, it will at least put them in a position to manage the risk by obtaining the necessary information. In the best case, rightsholders are forced to disclose the information that was missing for a transaction to occur in the first place, especially their identity and availability.

This is given in most use cases: To enforce a restored copyright under § 104(A), for example, rightsholders must provide an address and telephone number where they can be contacted.¹³¹ In order to object to public performances under § 110(4)(B), rightsholders must serve users with a notice that reveals their identity. The same is required for complaints to service providers under the DMCA safe harbor which demands that a notification includes information reasonably sufficient to permit the service provider to contact the complaining party.¹³² The exclusion of an out-of-commerce work as well as the provision of information to service providers under Art. 17(4)(b), (c) DSM Directive also generally will reveal the rightsholder's identity. Moreover, the status as an orphan work only ends if rightsholders signal their presence on the market again.¹³³ Only concerning search engines the mechanisms to opt-out of uses, e.g. a declaration under the robots.txt protocol, not necessarily provide information on how to contact the rightsholder.

From the user perspective, the function of a property rule therefore is to equalize the information differential between rightsholder and user. An opt-out scheme enforces a synchronicity between the user's risk of infringement with their level of information: If users cannot be expected to have sufficient information to initiate a transaction, they carry no infringement risk under the initial liability rule state. When rightsholders opt-out into the property rule, it comes at the prize of revealing information to users that, in the least, give them a reason and direction for further investigations, and at best, will lower transaction cost so much that they can initiate a transaction. As soon as users are provided with the information, they carry the risk of infringement again yet now can manage it properly. What this ensures is a coincidence

¹³⁰ See Hickey, *supra* note 92, at 420 ff. (describing the cost of rightsholders' opportunism through the deterrence of socially-valuable uses).

¹³¹ Cf. 17 U.S.C. § 104(A)(e)(1)(A)(i), (2)(B).

¹³² Cf. 17 U.S.C. § 512(c)(3)(A)(iv).

¹³³ Cf. Art. 5 Orphan Works Directive.

of the legal necessity and the economic possibility of individual transactions – the latter must be present for the former to arise.

IV. Implications for Fair Use and the Three-Step Test

Now that the abstract normative justification for opt-outs in copyright has been laid out, the following part applies this framework to a particularly contested area, the training of generative AI systems. Whether the opt-out schemes that govern reproductions for this purpose hold up must be assessed against two key instruments: fair use and the three-step test. In the US, the lawfulness of AI training uses that operate under an opt-out logic hinges on whether they qualify as a fair use; in the same vein, the destiny of the statutory opt-out scheme in the EU depends on its conformity with the three-step test as it must be in accordance with the test's requirements. Two questions must be differentiated: first, the general influence of opt-outs on the analysis under each instrument, and second the concrete effects in relation to the compilation of AI training data.

1. Fair Use

A use is not infringing if it qualifies as fair under 17 U.S.C. § 107. The fair use doctrine, originally judicially developed as an “equitable rule of reason”¹³⁴ and later codified, is built on the premise that the utilitarian goal of US copyright law does not justify broad exclusive rights concerning uses that do not interfere with the incentive to create new works. If such uses would be subject to prior permission, copyright would stifle the very creativity it serves to promote.¹³⁵ The analysis is thus oriented towards the prospect of substitution of protected expression: It supports a finding of fair use if a use is unlikely to substitute for a use of the original work by the rightsholder.¹³⁶ Four statutory factors are particularly relevant yet neither determinative nor exclusive for the inquiry: (1) the purpose and character of the use, (2) the nature of the copyrighted work, (3) the amount and substantiality of the portion used in relation to the copyrighted work, and (4) the effect of the use upon the potential market for or value of the copyrighted work.¹³⁷

a. The General Influence of Opt-Outs

¹³⁴ See e.g. *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 448 (1984).

¹³⁵ See e.g. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 577 (1994).

¹³⁶ See e.g. *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 528 (2023) (describing the problem of substitution as “copyright’s *bête noire*”).

¹³⁷ 17 U.S.C. § 107.

Though fair use and opt-out schemes are functionally equivalent in that they both help to realize uses that would otherwise fail due to high transaction cost,¹³⁸ opt-out schemes generally do not fit squarely into the conventional pillars of fair use analysis. In the abstract, opt-outs provide for fairness in a procedural sense.¹³⁹ Neither does an opt-out scheme ensure that uses that are permitted by default qualify as fair in substance, nor does it conversely allow the conclusion that when a rightsholder chooses to opt out of a certain use, such use is particularly unfair. Fair use in contrast aims to ensure substantive fairness which is illustrated for instance by the four statutory factors that aim at a substantial quality. What also complicates introducing opt-outs to fair use is the user-focused perspective of the inquiry: The factors generally investigate the effects of user action on the exploitation of the original work; only factor two considers the rightsholders' sphere.¹⁴⁰

A successful integration into the analysis therefore raises the question under what circumstances the procedural fairness of an opt-out scheme ensures that the substantive conditions for a finding of fair use are given.

aa. The Purpose and Character of the Use and the Effect upon the Market

Opt-outs can become particularly relevant under factor one. They can inform the underlying inquiry of factor one, even though they bear no direct relationship to its surface-level concepts. The first factor is dominated by the ideas of transformativeness, justification and commerciality. While the Supreme Court originally described transformativeness in *Campbell* as an investigation in whether a new work “adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message”¹⁴¹, the *Warhol* Court narrowed the analysis down to the degree of which “the use at issue has a purpose or character different from the original”¹⁴². Justification relates to “the reasons for, and nature of, the copier’s use of an original work”.¹⁴³ Commerciality points to whether the user seeks to profit directly from the use of a plaintiff’s work.¹⁴⁴

With all three concepts, however, opt-outs stand in no direct relationship. The mere presence of an opt-out option is unrelated to the quality of a use; neither does it ensure that the

¹³⁸ See e.g. Gordon, *supra* note 77 (describing fair use as a mechanism in case of market breakdown).

¹³⁹ Mattioli, *supra* note 36, at 28.

¹⁴⁰ See also Hickey, *supra* note 92, at 409.

¹⁴¹ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994).

¹⁴² *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 528 (2023).

¹⁴³ *Id.* at 509; see also *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 580-81 (1994).

¹⁴⁴ Cf. e.g. *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 818 (9th Cir. 2003).

use has a different purpose than the original use, nor does it implicate a good positive¹⁴⁵ reason for the use or that a use is less commercial.

Yet the concepts under factor 1 are not an end in themselves. They function as a proxy for the underlying inquiry on the likelihood of substitution of protected expression.¹⁴⁶ The other influential and relevant statutory factor for opt-outs, factor four, follows the same north star.¹⁴⁷ It sheds light on the market and value effects of a use, however, focuses not on a difference in purpose but on the prospect of market competition through a substitutive copy.¹⁴⁸ The right criterion to measure the influence of opt-outs on the analysis against therefore is the superordinate prospect of substitution.

In this regard, opt-outs allow a simple but striking inference: If a rightsholder does not opt-out, it is improbable that there is any relevant substitution based on a similar purpose or a meaningful effect on the market or the value of the work,¹⁴⁹ meaning that the incentive to create is not stifled.¹⁵⁰ This notion is already hinted at in *Field v. Google* where the court held that “the fact that the owners of billions of Web pages choose to permit these links to remain is further evidence that they do not view Google's cache as a substitute for their own pages”.¹⁵¹

However, for the inference to inform the analysis the two critical conditions that influence the normative justification for opt-outs must be taken into account: Tolerating a use and refraining from opting out can only be attributed to a conscious and autonomous decision of the rightsholder if she was aware of both the use and the possibility to opt-out, and the means of opting-out are not too burdensome. This is because only under these conditions it is effectively the rightsholder and not the user that controls the use.¹⁵² If the use would pose the danger of market substitution to a relevant degree, the rightsholder would simply not tolerate it.

¹⁴⁵ Other than the negative „the rightsholder did not object“.

¹⁴⁶ Concerning justification, see *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 581 n.14 (1994); see also *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 511, 532 (2023) (introducing the notion of „independent justification“ which seems to overlap with the presence of a purpose different from the rightsholder's purpose).

¹⁴⁷ Alternatively, another possible avenue for opt-outs into fair use analysis is the concept of the source of loss under factor 4, see e.g. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 591-92 (1994) (exploring harm that is not „cognizable under the copyright act“).

¹⁴⁸ See e.g. *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 536 n.12 (2023); *Authors Guild v. Google, Inc.*, 804 F.3d 202, 223 (2d Cir. 2015); but see U.S. Copyright Office, *supra* note 2, at 65 f. (advancing an unprecedented theory of “market dilution”, meaning that market harm may also stem from “stylistically similar” material, effectively introducing market harms from sources other than substitution of expression into factor 4 analysis).

¹⁴⁹ Or alternatively: there is a (potential) loss in the market, however, the source of the loss is not the use but the failure of the rightsholder to opt-out.

¹⁵⁰ See also Hickey, *supra* note 92, at 414, 427.

¹⁵¹ *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1119 (2006).

¹⁵² Cf. *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1119 (2006); see also *Sag*, *supra* note 26, at 1677.

In a way, opt-out schemes therefore allow to shortcut the different heuristics that have been developed under the fair use analysis by substituting them with the account of the actually affected rightsholder. Under an opt-out regime, we may just take the word of the person whose incentive is supposed to be stifled (if the copyright owner is the author) or who is supposed to profit from the exploitation of the work (if the copyright owner is a derivative or translative rightsholder) to assess whether the use is fair.¹⁵³ To consider opt-outs under fair use analysis therefore entertains a different avenue than the traditional concepts but reaches the same goal: cutting back the exclusive rights where they work against copyright's goal of promoting the creation of new expressive works. That opt-outs do not fit seamlessly into the conventional fair use concepts presents no obstacle as the wording of 17 U.S.C. § 107 allows for the consideration of additional factors.

bb. Good and Bad Faith

Another doctrinal anchor for opt-outs can be found in the considerations around good and bad faith. Relating to Ninth Circuit precedents, the *Field* decision – that explicitly dealt with an opt-out scheme – addressed Google's good faith under an additional factor to the analysis.¹⁵⁴ It stressed that Google explained rightsholders how to opt-out, honored opt-outs when they occurred through automated means, and mitigated the impact on the original webpage.¹⁵⁵ Particularly relevant was also that Google only later learned through the complaint about Field's objection.¹⁵⁶

The relevance of good or bad faith to fair use analysis, however, currently is precarious at best. While in *Harper & Row*, the Supreme Court stressed the importance of “good faith and fair dealing” for fair use,¹⁵⁷ in *Campbell* the Court expressed its doubts whether subjective factors were relevant to the inquiry.¹⁵⁸ The Court revisited the issue in *Google* and found the skepticism reasonable as copyright “is not a privilege reserved for the well-behaved”.¹⁵⁹ Yet, this line of reasoning in both cases dealt primarily with the negative weight of bad faith – in *Campbell* the

¹⁵³ It is important not to overstate the advantages in this regard. Opt-outs only allow for a negative selection in that they help to identify situations where there is no or a low likelihood of meaningful substitution. This does not mean, however, that if a rightsholder opts out the likelihood of competition by the use is high. It only shows an interest in exclusivity.

¹⁵⁴ See also Nari Na, *Testing the Boundaries of Copyright Protection: The Google Books Library Project and the Fair Use Doctrine Note*, 16 CORNELL J. L. & PUB. POL'Y 417, 439-40 (2006).

¹⁵⁵ *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1122 f. (2006).

¹⁵⁶ *Id.* at 1123.

¹⁵⁷ *Harper & Row, Pub., Inc. v. Nation Enters.*, 471 U.S. 539, 562 (1985); see also e.g. *Maxtone-Graham v. Burtchae*, 803 F.2d 1253, 1261 (2d Cir. 1986); *Williamson v. Pearson Educ., Inc.*, 2001 WL 1262964, at *5 (S.D.N.Y. 2001).

¹⁵⁸ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 586 n.18 (1994).

¹⁵⁹ *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 32 (2021).

question arose because the Defendants asked for a license and were denied permission,¹⁶⁰ in *Google* the Court explicitly only addressed bad faith: “We have no occasion here to say whether good faith is as a general matter a helpful inquiry”.¹⁶¹ The precedents therefore do not necessarily support the opposite conclusion that good faith must never have any weight in favor of a fair use finding.¹⁶²

In the opt-out use cases, taking good faith into account brings characteristics of a use into the analysis which speak to its character, that would otherwise be effectively ignored. Primarily, by offering an opt-out, users effectively choose a smaller interference with the rightsholder’s position than without, where rightsholders would have to start an infringement action against unwanted uses. Users instead provide a mechanism to mitigate (typically ex ante, sometimes at least ex post) any infringements of exclusive rights and thereby show that they want to respect the autonomous decision of the rightsholder.

Even further, opt-outs also give reason to reconsider the weight of bad faith in fair use analysis. If a user allows for or even advertises with an opt-out possibility yet ignores opt-outs, the explicit autonomous decisions of rightsholders are undermined. At first sight, *Campbell* and *Google* seem to preclude the relevance of such behavior for the analysis as it was precisely a denial of permission prior to the use, similar to an opt-out, that was held not to weigh against fair use in *Campbell*. However, the situation under an opt-out scheme is different.

The use in *Campbell* was open and rather easily detectable – that the band 2 Live Crew used the song “Pretty Woman” for their rap despite the denied permission would have come to the attention of the rightsholders without large efforts anyway. Therefore, from the rightsholder’s perspective, it did not carry much weight that the user asked for permission beforehand and it was denied because there still was a realistic chance for the rightsholders to have the issue resolved in court if a use occurs without authorization (as was ultimately the case). Rightsholders were thus not left empty-handed, even though the Court held that a use may be fair despite a denied permission.

The opt-out use cases in contrast involve uses that are not easily detectable. Often, without affirmative information by the user, rightsholders cannot effectively find out whether a use is for a purpose they would object to. If ignoring opt-outs never plays a role for the fair use finding, there is no incentive for users to honor the rightsholder’s objection, and rightsholders have no

¹⁶⁰ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 586 n.18 (1994).

¹⁶¹ *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 32 (2021).

¹⁶² See also Hickey, *supra* note 92, at 413 (stressing that *Campbell* does not foreclose the proposition that an ex ante request for permission should weigh in favor of fair use).

real chance to defend themselves against unwanted uses they believe they already have successfully objected to.¹⁶³ Similar types of conduct that would weigh against fair use in this context are to actively circumvent the information of rightsholders about uses or to make it unduly difficult for the rightsholder to opt out. At the same time, taking bad faith into account does not lead to an asymmetrical grave burden on users, as the bad faith of rightsholders, for example through opportunistic behavior, should consequently also be included in the analysis.¹⁶⁴

cc. Supra-Individual Beneficial Effects

A complication arises, however, from the typical litigation scenario. Most likely, an infringement action is brought by someone that perceives the use to be unfair yet chooses not to opt-out, such as in *Field v. Google*. For the scheme to prevail then as fair, the analysis must allow the harm for this individual rightsholder to be weighed against the collective advantages, not of the concrete use but of the opt-out design as a whole.¹⁶⁵ The wording of factor four, however, limits the analysis to the effect of “the use” upon “the copyrighted work”. This confines the analysis to the effects on the litigation parties.

Yet, for example the *Google* decision shows that supra-individual positive effects may influence the analysis to the detriment of the rightsholder under this factor.¹⁶⁶ In *Google*, the Court explicitly balanced the public benefits of the use with the losses to the copyright owner.¹⁶⁷ From this point, it is only a small step to abstract from the positive effects of a concrete use to the advantages of the underlying institutional design which enabled the use in the first place.¹⁶⁸ Similarly, supra-individual negative effects are sometimes included in the analysis under factor four when it comes to market harm through a dilution of the market by indirect substitution.¹⁶⁹ It is only reasonable then to also take into account supra-individual positive effects on the market.

¹⁶³ This also explains why putting rightsholder in a better position with the opt-out option may effectively create a worse fair use case for users: Rightsholder will reasonably rely on the users to honor the opt-out.

¹⁶⁴ See *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1123 (2006) (taking into account that the rightsholder took affirmative steps to provoke the use and weighing it against him in favor of fair use); see generally Hickey, *supra* note 92, at 414-15, 444 (arguing in favor of including bad faith of rightsholders as a factor in fair use analysis). See *supra* III.3.b.dd. on opportunistic behavior.

¹⁶⁵ See *supra* III.3.b., c (naming the normative advantages).

¹⁶⁶ In *Campbell*, in contrast, the analysis evolved around positive effects of the use on the market *for the work*, see *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590 n.21 (1994).

¹⁶⁷ See *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 35-36 (2021).

¹⁶⁸ See also Sag, *supra* note 26, at 1675 (arguing in favor of considering the institutional design of copyright under fair use).

¹⁶⁹ See U.S. Copyright Office, *supra* note 2, at 65 (advancing a theory under which factor 4 should be read to encompass effects through the dilution of markets for works of the same kind as the training data of generative AI systems). This theory was also adopted by Judge Chhabria in another generative AI copyright case, see *Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417-VC, slip op. at 4 (N.D. Cal. June 25, 2025).

b. Application to the Compilation of Generative AI Training Data

When applying the general findings concerning opt-outs to the compilation of AI training data, after the *Warhol* decision it must be clear to what particular use the analysis relates.¹⁷⁰ The use at stake here is defined as a currently prototypical use in the AI context: Copyright-protected works available on the open web without an opt-out but not necessarily uploaded by a rightsholder are compiled to serve as training data for a generative AI model that is capable of generating outputs which could substitute for the original works.¹⁷¹

aa. The Purpose and Character of the Use and the Effect upon the Market

Based on the standard developed above,¹⁷² the decisive question is whether for this use one can infer that the likelihood of a substitutive effect or meaningful market harm is low. As a reminder, in the abstract this is the case if the rightsholder was aware of both the use as well as her right to opt-out and had an effective possibility to opt-out with reasonable effort, yet did not do so.

Concerning the awareness of rightsholders, we can distinguish between the point of time and the depth of knowledge. For the point of time, there appears to be a turning point with the advent of consumer-ready generative AI models, starting with the release of ChatGPT by OpenAI in late 2022. Beforehand, while there was a general awareness of web scraping uses for the purposes of search indexing, web archiving and research,¹⁷³ data compilation for the purposes of AI training was relatively unknown. Because AI models and their use cases were discussed predominantly among experts, the vast majority of rightsholders that made works available online likely was not aware of such uses. The proliferation of opt-outs that rightsholders have declared since then, however, demonstrates that this has changed. From 2023 to 2024 alone, for example, almost 25% of tokens from the most critical domains, and 5% of all tokens, from the entire corpora of three data sources that are widely used (derivatively) in modern foundation models, have become subject to an opt-out restriction.¹⁷⁴ Also, during Google's recent search antitrust trial, it was revealed that Google removed 80 billion of 160 billion tokens from AI training by

¹⁷⁰ See *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 533 (2023).

¹⁷¹ Including potential substitution in the use makes it a particular interesting case because that is where conflicts are the strongest.

¹⁷² See *supra* IV.1.a.aa.

¹⁷³ See e.g. Andrew Sellars, *Twenty Years of Web Scraping and the Computer Fraud and Abuse Act*, 24 B.U. J. SCI. & TECH. L. 372, 374 (2018) (naming different applications of web scraping); Christopher Olston & Marc Najork, *Web Crawling*, 4 FOUNDATIONS AND TRENDS IN INFORMATION RETRIEVAL 175, 180 (2010) (giving an early history of crawling).

¹⁷⁴ Longpre et al., *supra* note 72 (describing a proliferation of restrictions concerning the corpora of C4, RefinedWeb, and Dolma through robots.txt).

Google DeepMind on the basis of opt-outs by publishers.¹⁷⁵ This shows a significantly increased, broad awareness for AI training uses and opt-outs by rightsholders.

For current compilation efforts, the knowledge requirement therefore is fulfilled – rightsholders that want to object to such uses have had enough reason to obtain information about their opt-out possibilities. However, this also means that datasets that were scraped before such general knowledge existed do not support the conclusion that rightsholders consciously refrained from opting out, which forecloses the inference that there is no meaningful prospect of substitution. Generative AI system developers are therefore required to update old datasets to reflect the current autonomous decision of rightsholders towards such uses for the opt-out to have a positive influence on their fair use case.

The depth of knowledge of rightsholders makes no difference in this regard. In some cases, vigilant rightsholders will have concrete knowledge of which bot is scraping their works.¹⁷⁶ Yet abstract awareness of the phenomenon of web scraping for purposes of AI training data compilation is sufficient for the opt-out scheme to function. It provides enough reason for unwilling rightsholders to find out about potential uses and how to prevent them through opt-outs.

Concerning the ease of the opt-out mechanism, the situation looks different. Since generative AI models became publicly available, a plethora of standards have emerged that provide different mechanisms.¹⁷⁷ However, there is no generally recognized standard that would allow rightsholders to ensure that no training uses will occur. What comes closest is the Robots Exclusion Protocol (REP), a voluntary¹⁷⁸ compliance protocol to indicate a website owner's

¹⁷⁵ See *Google Can Train Search AI With Web Content Even After Opt-Out*, BLOOMBERG.COM, May 3, 2025, <https://www.bloomberg.com/news/articles/2025-05-03/google-can-train-search-ai-with-web-content-even-after-opt-out>.

¹⁷⁶ See e.g. Weatherbed, *supra* note 63 (describing an instance where the ClaudeBot web crawler, a bot by the AI company Anthropic, requested the website ifixit.com almost a million times in 24 hours which was noticed by the iFixit as the rightsholder).

¹⁷⁷ Location-based approaches, meaning declarations that apply to all works hosted on a specific location, encompass for example robots.txt, ai.txt, the TDM Reservation protocol, noai meta-tags, domain registrations in do-not-train registries and HTTP headers; unit-based approaches in contrast are designed to manage opt-outs at the level of a copy, such as do not train-declarations on haveibeentrained.com or the planned credential system of the Adobe-led Content Authenticity Initiative, see Paul Keller, *Open Future Policy Brief #6, Considerations for Opt-Out Compliance Policies by AI Model Developers*, 2 f., https://openfuture.eu/wp-content/uploads/2024/05/240516considerations_of_opt-out_compliance_policies.pdf; see Patrick Kulp, *Adobe's New Tool Aims to Let Artists Opt out of AI Training*, TECH BREW, <https://www.emergingtechbrew.com/stories/2024/10/11/adobe-tool-artists-opt-out-ai-training> (last visited Jan. 30, 2026) (describing the planned credential system of the Content Authenticity Initiative).

¹⁷⁸ This does not mean that there are not ways to enforce compliance with the protocol on a technical level, see e.g. *Robotcop: Enforcing Your Robots.Txt Policies and Stopping Bots before They Reach Your Website*, THE CLOUDFLARE BLOG (Dec. 10, 2024), <https://blog.cloudflare.com/ai-audit-enforcing-robots-txt/> (announcing a tool that allows website owners to audit compliance with the protocol by AI service providers and to enforce compliance at the network level).

preference on which content web crawlers are allowed to visit, that is widely used and was officially standardized in September 2022. All currently leading AI companies claim to respect opt-outs via robots.txt.¹⁷⁹ The REP is implemented via a robots.txt text file that even unexperienced rightsholders can install with low effort.¹⁸⁰

However, a missing robots.txt opt-out not necessarily indicates that rightsholders are not interested in preventing such uses. For one, this is due to the lack of granularity the REP offers. While it allows to express the rightsholder's preference down to a distinct resource location and on a bot-by-bot basis, the communicated preference is always a binary between allowing and disallowing and does not allow to distinguish between elements inside a resource.¹⁸¹ The standard also does not provide for a way yet to discriminate on the level of the purpose of a use.¹⁸² Further, users are free to give their bot any name – rightsholders in turn cannot ensure that if they only allow a certain bot of a company to do something, it really is this bot that is making the request and use.

The resulting difficulties for the attention economy of the internet are well illustrated by the use of copyright-protected content by Google: Following fears that opting out of the Google search bot would limit visibility for Google search in general, they introduced a separate bot, Google-Extended, that was to be used for the compilation of AI training data and represented therefore the right addressee for an AI opt-out.¹⁸³ However, as was revealed during Google's search antitrust trial, such opt-outs are not honored concerning search-specific AI products. To object to uses of works in the search context, rightsholders must opt out of being indexed for search – yet this will also remove them from a (while declining) still relevant source of traffic.¹⁸⁴

¹⁷⁹ See Overview of OpenAI Crawlers - OpenAI API, *supra* note 71 (concerning OpenAI); Google Can Train Search AI With Web Content Even After Opt-Out, *supra* note 175 (concerning Google DeepMind); Does Anthropic crawl data from the web, and how can site owners block the crawler?, *supra* note 71 (concerning Anthropic).

¹⁸⁰ See e.g. *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1113 (2006) (assuming that web site owners are generally able to communicate their preferences via robots.txt); Sag, *supra* note 26, at 1667, 1677 (attributing the success of robots.txt to its low cost, describing adding it as “trivial” and “virtually costless”); Michael W. Carroll, *A Realist Approach to Copyright Law's Formalities*, 28 BERKELEY TECH. L.J. 1511, 1522 (2013) (describing it as “trivially easy”).

¹⁸¹ It also does not allow to manage opt-outs involving overlapping rights, see European Union Intellectual Property Office, *The Development of Generative Artificial Intelligence from a Copyright Perspective*, 230-31, https://www.europarl.europa.eu/meetdocs/2024_2029/plmrep/COMMITTEES/JURI/DV/2025/05-12/2025.05.12_item6_Study_GenAIfromacopyrightperspective_EN.pdf.

¹⁸² *Id.* at 178-79 (pointing out that robots.txt does not allow for the disaggregation of different uses).

¹⁸³ Danielle Romain, *An Update on Web Publisher Controls*, GOOGLE BLOG (Sept. 28, 2023), <https://blog.google/technology/ai/an-update-on-web-publisher-controls/>.

¹⁸⁴ See Google Can Train Search AI With Web Content Even After Opt-Out, *supra* note 175 (stating that publishers can only decline having their data used in search AI if they opt out of being indexed for search and citing a Google spokesperson: “Google has a separate way for publishers to manage their content in Search via the well-established robots.txt web standard”).

For small users in particular, it comes down to a trade-off between visibility and protection of their works.¹⁸⁵

Additionally, the information-disclosing effect of the REP is limited. A robots.txt file by specification only communicates that certain uses may or may not occur without individual transactions. The protocol does not require information on the identity or contact details of rightsholders. This poses the risk that even after an opt-out, transactions are unavailable for willing parties which would represent the worst of both worlds: It recreates the same unfavorable sticky property rule that would be in place without an opt-out scheme yet places an additional burden on rightsholders to get there. Consequently, just because rightsholders did not opt-out,¹⁸⁶ it does not necessarily allow the inference that they saw no danger of substitution through the compilation of AI training data.

Lastly, what also weighs heavily against a fairness assumption based on a lack of opt-outs is that for vast amounts of works that are compiled for AI training, it is not necessarily the rightsholder that uploaded the material to the open web. Rightsholders often do not control the availability of such works at all, not only if they were uploaded without their permission but also if the upload is authorized yet the rightsholders are technically unable to declare opt-outs, for example on large social media platforms.¹⁸⁷ A missing opt-out therefore cannot be attributed to an autonomous decision by the rightsholder as long as it is unclear whether the rightsholder decided on the circumstances of the use.¹⁸⁸ Due to the treatment of fair use as an affirmative defense,¹⁸⁹ it is ultimately up to the AI developer to establish this link between the work, missing opt-out and the rightsholder. In this respect, there are no reliance interests worthy of protection on the part of AI companies because it is obvious that not all content available on the open web is authorized by the rightsholder. A fair use case based on an opt-out scheme for the compilation of AI training data will therefore regularly fail at this stage.

¹⁸⁵ This invokes competition law problems through the tying of the search engine to the AI model, *see e.g.* European Commission, *Competition in Generative AI and Virtual Worlds*, Policy Brief No. 3/2024, 7, https://competition-policy.ec.europa.eu/document/download/c86d461f-062e-4dde-a662-15228d6ca385_en.

¹⁸⁶ Or even positively encouraged the use via robots.txt.

¹⁸⁷ *See also* U.S. Copyright Office, *supra* note 2, at 102 (reporting concerns regarding platform-level opt-outs).

¹⁸⁸ A similar problem arises for images that are uploaded without authorization by third parties and then used by search engines in thumbnails. The German Federal Court of Justice held that it is sufficient that the rightsholder has permitted the display online to any other person because such permission must have included uses by search engines of unauthorized third party uploads as it is generally known that search engines cannot distinguish whether an image has been uploaded by an authorised or unauthorised person, *see* German Federal Court of Justice, judgment of 19 October 2011 – I ZR 140/10 – Vorschaubilder II, GRUR 2012, 602 para. 28. Such a solution is, however, impossible where there are no explicit or implicit authorizations given.

¹⁸⁹ *See e.g.* Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 590 (1994).

bb. Good and Bad Faith

The inquiry into good and bad faith points out a problem for the use analyzed here. While the good faith of AI companies in employing, explaining and honoring an opt-out scheme generally supports a finding of fair use because they allow the autonomous decision of the rightsholder to take effect,¹⁹⁰ bad faith on its own may not carry any weight here. Anecdotal instances of AI companies effectively ignoring opt-outs are not rare.¹⁹¹ However, such ignorance can only weigh against a finding of fair use if the use would not be fair but for honoring an opt-out. It is unclear and beyond the scope of this article, though, whether the compilation of AI training data for the purposes of machine learning is not by itself a fair use.¹⁹² The crux lies in the fact that once an opt-out infrastructure is established and widely adopted, users will trust in that their opt-outs are honored and would therefore be particularly vulnerable if AI companies ignored them, irrespective of the fair use case.¹⁹³ Therefore, for bad faith to influence the analysis, AI companies must have at least given the impression that they will respect opt-outs, regardless of whether fair use requires it or not. Only then is it reasonable to put weight on the fact that they later acted in bad faith.

2. Three-Step-Test

A second important obstacle for opt-out schemes is the three-step test. Anchored in international copyright law, it functions as a limitation for exceptions and limitations of exclusive rights.¹⁹⁴ The test stipulates that limitations can apply (1) only in certain special cases that (2) do not conflict with the normal exploitation of a work and (3) do not unreasonably prejudice the legitimate interests of the rightsholders.

Especially for the exception or limitation for text and data mining that, according to the predominant view,¹⁹⁵ covers the compilation of AI training data in EU jurisdictions, the three-step test can be a significant hurdle. This is because for one there is a misguided¹⁹⁶ tendency for a

¹⁹⁰ See *supra* IV.1.a.bb.

¹⁹¹ See e.g. Elizabeth Lopatto, *Perplexity's Grand Theft AI*, THE VERGE (June 27, 2024), <https://www.theverge.com/2024/6/27/24187405/perplexity-ai-twitter-lie-plagiarism> (describing instances where a bot, presumably of the AI company Perplexity, aggressively aggregates content ignoring robots.txt instructions).

¹⁹² See e.g. Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 TEX. L. REV. 743 (2020); Sag, *supra* note 67.

¹⁹³ See *supra* IV.1.a.bb.

¹⁹⁴ See Art. 9(2) Berne Convention, Art. 13 TRIPS, Art. 10 WCT, Art. 5(4) InfoSoc Directive. In this regard, it may also limit the application of fair use.

¹⁹⁵ See e.g. Regional Court of Hamburg, judgment of 27 September 2024 – 310 O 227/23 – LAION, LTZ 2025, 65; JONATHAN PUKAS, WERKNUTZUNG UND KÜNSTLICHE INTELLIGENZ 137 f. (2025).

¹⁹⁶ Cf. Dan Wielsch, *Relationales Urheberrecht. Die Vielen Ummelten Des Urheberrechts*, 5 IPJ 274, 284 (2018) (pointing out the structural preference of the three-step test towards economic interests, manufactured through its interpretation).

rather strict interpretation of the test requirements limited to their wording. Additionally, the test – at least under current practice – not only binds legislators but also courts in their interpretation and application of limitations.¹⁹⁷ This means that the requirements of the steps must be complied with on a case-by-case basis.

a. The General Influence of Opt-Outs

aa. Step 1: Certain Special Cases

On the first test step, as a basic rule limitations must constitute certain special cases, meaning that the scope must be known and particularized.¹⁹⁸ The specialty criterion has both a qualitative and quantitative dimension: Quantitatively, the number of rights and works affected and the privileged group of users must remain exceptional; qualitatively, the limitation must be justifiable by a rational reason of public policy.¹⁹⁹

Limitations or exceptions with opt-out schemes therefore must be special cases in this sense. This is doubtful: Under an opt-out scheme the “normal” state of exclusivity is set aside in favor of an initial access to works. In other words, a basic freedom to use becomes the new “normal” for a certain use so that the relationship between exclusivity and access is reversed. However, this does not conflict with the purpose of the first test stage. The first step narrows limitations down to certain special cases to preserve the rightsholders’ ability to profit from their works.²⁰⁰ This stands under the assumption that if a limitation is in place there is no possibility for rightsholders to enjoy exclusivity. Under an opt-out scheme, in contrast, rightsholders still have the possibility to opt-out and switch to exclusivity where they may exploit their work – there is no either-or relationship between a limitation or exception and exclusivity. Therefore, limitations with opt-outs generally are compliant with the first test step yet do not influence the analysis in a certain way.

bb. Step 2: No Conflict with Normal Exploitation

¹⁹⁷ See e.g. Eleonora Rosati, No Step-Free Copyright Exceptions: The Role of the Three-Step in Defining Permitted Uses of Protected Content (Including TDM for AI-Training Purposes) (Nov. 10, 2023), <https://papers.ssrn.com/abstract=4629528> (arguing under reference to CJEU judgments that not only national legislators but also member state courts are addressees of the test).

¹⁹⁸ See MARTIN SENFTLEBEN, COPYRIGHT, LIMITATIONS AND THE THREE-STEP TEST: AN ANALYSIS OF THE THREE-STEP TEST IN INTERNATIONAL AND EC COPYRIGHT LAW 137 (2004).

¹⁹⁹ See Panel Report, United States—Section 110(5) of the U.S. Copyright Act, ¶ 6.108, WTO Doc. WT/DS160/R (adopted July 27, 2000); *id.* at 140 ff. (rejecting, however, a quantitative approach).

²⁰⁰ Cf. Sprigman, *supra* note 10, at 567 (assuming that the rightsholder’s ability to exploit his work is the relevant inquiry under the provision). See also Panel Report, *supra* note 199, ¶ 6.109 (making a connection to the normal exploitation of the work under the second test step).

Under the second test step, limitations must not conflict with a normal exploitation of the work, meaning the ways through which rightsholders typically extract economic value.²⁰¹ Normal is understood in both an empirical as well as normative sense, encompassing the conventional as well as potential and future forms of exploitation.²⁰² A conflict occurs if a use enters into economic competition with these forms of exploitation so that rightsholders are deprived of significant commercial gains.²⁰³

Uses under opt-out schemes are very likely not to amount to such a conflict. This is because, for one the schemes usually apply to scenarios where rightsholders generally have a low interest in economic exploitation.²⁰⁴ Furthermore, regularly there will be no conventional economic exploitation to conflict with. As pointed out above, opt-out use cases surface in situations where individual transactions are impractical due to disproportionate or even prohibitively high transaction costs.²⁰⁵ Under these circumstances, it is not the limitation with an opt-out scheme that deprives rightsholders of their normal exploitation, but the market segment itself which is not intact. On the contrary, the opt-out scheme serves to preserve the exploitation opportunities of such rightsholders and facilitates their autonomous choice to become effective.²⁰⁶

Concerning a potential or future normal exploitation, the ability to opt-out directly influences the analysis in a way that is comparable to the fair use inquiry. The opt-out right enables rightsholders to suspend the effect of a limitation at will. Thereby, it prevents a conflict with the normal exploitation to arise in the first place: If such a conflict is looming, rightsholders will make use of the opt-out. If they refrain from doing so, either there is no relevant economic competition because of the limitation,²⁰⁷ or the rightsholders have no interest in exclusivity – in both cases, the requirements of the second test step are fulfilled.²⁰⁸ However, to make this

²⁰¹ See Panel Report, *supra* note 199, ¶ 6.183.

²⁰² See SENFLEBEN, *supra* note 198, at 178 f. See also Panel Report, *supra* note 199, ¶ 6.166.

²⁰³ See also Panel Report, *supra* note 199, ¶ 6.183.

²⁰⁴ See *supra* III.3.a.

²⁰⁵ *Id.*

²⁰⁶ See *supra* III.3.b.

²⁰⁷ Cf. Sprigman, *supra* note 10, at 562 (stating generally that a "failure to comply is a signal that the net present value of expected future revenues from a work is lower than the cost of compliance").

²⁰⁸ See VAN GOMPEL, *supra* note 8, at 191 (pointing out that authors can escape a limitation through an opt-out mechanism and retain exclusivity); Martin Senftleben, *How to Overcome the Normal Exploitation Obstacle: Opt-out Formalities, Embargo Periods, and the International Three-Step Test*, 2014 BTLJ COMMENT. 1, 14 (2014) (arguing that rightsholders can regulate the impact on their exploitation strategy and thereby safeguard the normal exploitation); Alexander Peukert, *A Bipolar Copyright System for the Digital Network Environment*, 28 HASTINGS COMM. & ENT. L.J. 1, 65 (2005) (stating that an opt-out mechanism helps to avoid a conflict with the normal exploitation of the work); but see Jane C. Ginsburg, *"With Untired Spirits and Formal Constancy": Berne Compatibility of Formal Declaratory Measures to Enhance Copyright Title-Searching*, 28 BERKELEY TECH. L.J. 1583 (2013) (criticizing the burden national unharmonized opt-out exceptions put on foreign authors).

inference the conditions laid out above again must be given: Rightsholders must be aware of both the use as well as the opt-out right, and the opt-out must be feasible with reasonable effort for interested rightsholders.²⁰⁹

cc. Step 3: No Unreasonable Prejudice to Legitimate Interests

The third test step addresses limitations that harm legitimate interests of rightsholders. Such harm must not be unreasonable, meaning disproportionate with regards to the relevant objective of the limitation.²¹⁰ Therefore, any prejudice must be balanced with the interest that justifies the limitation. Additionally, a possible remuneration granted under a limitation may mitigate the harm.²¹¹

Like the second test step, the opt-out right again equips rightsholders with means to prevent and mitigate harm.²¹² Rightsholders may simply opt-out which prevents any harm that a limitation could pose. If they refrain from doing so, either there is no harm, or the harm brought about by the limitation is not unreasonable. This is because if a rightsholder prefers the situation under a limitation over an exclusive exploitation, it means that she is better off compared to an exclusive exploitation. And if the rightsholder is in a better position compared to the “normal” exploitation (which is always deemed reasonable under the test), there cannot be unreasonable harm. Just like under the second test step, though, this requires both awareness of the rightsholder concerning both use and opt-out option, and opt-out means that are available with reasonable effort.²¹³

b. Application to Art. 4 DSM Directive

When applying the findings to the limitation for text and data mining in Art. 4 DSM Directive,²¹⁴ under the first test step, the opt-out has no influence on compliance with the test. The limitation certainly involves a use that is both known and sufficiently limited in its scope. It is also special in a qualitative sense, as it serves to promote innovation by lowering the level of legal uncertainty,²¹⁵ as well as in a quantitative sense because it only involves the reproduction right and addresses only an exceptional use case.

²⁰⁹ See *supra* III.3.b.cc, IV.1.a.aa.

²¹⁰ See SENFTLEBEN, *supra* note 198, at 253 f.

²¹¹ *Id.* at 237-38. See also Panel Report, *supra* note 199, ¶ 6.229.

²¹² Cf. Sprigman, *supra* note 10, at 562 (arguing, for the similar setting of a “new-style formality” in place, that insofar as “legitimate interests” relate to a rightsholder’s ability to capture rents on the basis of the exclusive rights, there is no interference under the second test step at least concerning economic interests).

²¹³ See *supra* III.3.b.cc, IV.1.a.aa.

²¹⁴ See *supra* II.

²¹⁵ See Recs. 8, 18 DSM Directive.

For both the second and the third test step, the decisive question is whether the limitation provides for the conditions to assume that rightsholders intentionally did not make use of the opt-out. Only this allows the inference that under step two no relevant conflict with a potential or future normal exploitation occurs, and under step three legitimate interests of rightsholders are not harmed or such harm is not unreasonable.²¹⁶

The requirements of Art. 4 DSM Directive for the EU member states, however, fall short of this. Concerning the awareness requirement, rightsholders must at least be put in a position that makes it likely that they find out about the use and their ability to opt-out. Yet, Art. 4 DSM Directive does not address this point directly. There is, for example, no affirmative information obligation of users. Rather, the provision seems to assume abstract knowledge of rightsholders because it presupposes that they will make use of an opt-out on their own. Of course, this is true for some rightsholders and use cases but the findings concerning the general awareness of uses for AI training from above show that for some rightsholders and especially concerning past uses the requirement will not necessarily be fulfilled.²¹⁷ The opt-out structure of the limitation therefore is unable to ensure conformity with the second and third test stage on its own, instead, the circumstances of each individual case – whether a rightsholder was aware abstractly – must be taken into account.

With regards to the difficulty of the opt-out mechanism, for the opt-out to be functional it must be effortless enough to allow unwilling rightsholders to switch to exclusivity with proportionate effort.²¹⁸ Art. 4(3) DSM Directive has two requirements: The reservation of rights must be made expressly and in an appropriate manner. The requirement of an express reservation does not amount to an undue burden for rightsholders. As the normative case for opt-outs rests on a positive decision of the rightsholder in favor of opting out,²¹⁹ it does not cause additional effort to do so explicitly. This also works in favor of preventing reliance interests of users to be harmed.

The second requirement to declare the reservation in an appropriate manner is, in principle, also not objectionable: It includes, for example, opt-outs in form of contractual agreements or via unilateral declaration.²²⁰ For the important use case of content made publicly available online, though, only machine-readable means are considered appropriate.²²¹ This is intended to ensure

²¹⁶ See *supra* IV.2.a.bb., cc.

²¹⁷ See *supra* III.3.cc.

²¹⁸ See *id.*

²¹⁹ See *supra* III.2.b.

²²⁰ Rec. 18 DSM Directive.

²²¹ *Id.*

that uses can actually be carried out automatically – without such a requirement the search cost problem that originally plagued users concerning the rightsholder’s identity would be merely shifted to the question whether a reservation of rights was declared.

However, as the legislator refrained from a further definition, the problem of the different standards also influences Art. 4(3) DSM Directive. The probably most widely used REP does not put an unreasonable burden on rightsholders, especially considering the benefit it has for users that are conducting automated uses. Yet, just like under the fair use analysis, the informative value of a rightsholder’s inaction is limited due to the low granularity of permission control. Also robots.txt does not solve the information problem that justifies the opt-out solution in the first place because the identity of the rightsholder is not given to the user. As the ability of using AI to recognize opt-outs improves, though, the effort will tend to decrease further in the future.²²²

Similar to the fair use analysis, the lack of control over third party uploads will often prevent an inference to be drawn from a failure to opt-out.²²³ Only where a work clearly was made available online under the control of the rightsholder, the opt-out may ensure compliance with the second and third test step. It is therefore particularly important for users to record the exact location of compiled content to be able to demonstrate a connection to the rightsholder. However, where such a connection is not clear, an unused opt-out can at best have an indicative value for the stages of the three step test.

V. Conclusion

Opt-out schemes are anything but irregular in copyright law. Enabling rightsholders to autonomously decide on uses of their works will sometimes be served best by requiring prior authorization, but not always. Where both rightsholders and user are generally interested in a use but disproportionate transaction costs prevent them from transacting, opt-out schemes provide a way out. They employ a flexible form of entitlement, a property rule, that combines the advantages of both static (zero-price) liability and property rules while offsetting their disadvantages. From the perspective of rightsholders, this allows for lawful uses between willing parties that could otherwise not occur while unwilling rightsholders remain capable of opting out. For users, this means they only must burden an infringement risk if they have the means to control it because

²²² See e.g. Regional Court of Hamburg, judgment of 27 September 2024 – 310 O 227/23 – LAION, LTZ 2025, 65 (stating obiter that an opt-out in natural language located in the terms of use is appropriate due to the general ability of the user to compare images with attached natural language image descriptions and check for matches via AI when compiling a dataset).

²²³ See *supra* IV.1.b.aa.

the opt-out forces rightsholders to disclose the missing information about their ownership. Individual transactions are therefore only necessary when they are feasible.

However, two requirements need to be present for the scheme to work: Rightsholders must be aware of both use and opt-out possibility, and the opt-out must be straightforward enough so that rightsholders may make use of it. Under these conditions, opt-out schemes integrate into both fair use and the three-step test analysis in the form of an inference: If rightsholders refrain from opting out, the use likely does not cause relevant harm.

Concerning the compilation of AI training data though, at least under current conditions, this inference is unreliable: It is uncertain whether rightsholders had and have the necessary awareness; also the prevalent opt-out standards neither provide for enough granularity nor ensure that rightsholders had the opportunity to decide whether to opt-out. And even if an opt-out is present, some standards do not lead to a disclosure of information to users that allows both parties to transact. Opt-out schemes are therefore at risk of resulting in the worst of both worlds: Higher burdens for rightsholders without inducing more uses in the interest of both parties.

So, although opt-out schemes have their rightful place in copyright law, for the compilation of AI training data they are in danger of becoming dysfunctional. This is a consequence of the legislator's choice to offload the decision-making power to rightsholders instead of resolving the underlying conflict between exclusivity and access themselves. Even though opt-out schemes prove to be an interesting and viable option for entitlement design under copyright law under the circumstances laid out above, the copyright conflict around AI training data is served better with a substantive decision on the lawfulness of uses – an outcome, which would also allow for a detailed balancing of the interests involved.

If policymakers nonetheless continue to rely on opt-out frameworks for AI training, this analysis suggests several reforms necessary to make such schemes functional. First, AI developers should be required to actively inform rightsholders about training uses, either individually or through centralized registries, thereby addressing the awareness deficit. Second, opt-out standards should be harmonized and provide sufficient granularity to allow rightsholders to distinguish between different types of AI uses without forfeiting visibility in search engines or other services. Third, opt-out mechanisms should mandate disclosure of contact information, enabling willing parties to transact where desired. Fourth, AI developers should bear the burden of establishing that works were made available under the control of the rightsholder before relying on a missing opt-out. Finally, grace periods and notice requirements modeled on the copyright restoration framework could protect reasonable reliance interests of AI developers while preserving rightsholder autonomy. Without such reforms, the current opt-out landscape will continue to

produce outcomes that serve neither the interests of rightsholders nor the broader public interest in innovation that copyright law is designed to promote.