Zweiter Teil: Regelungsbedürfnisse im Urheberrecht für die Verwendung von Werken als Trainings- und Inputdaten

Untersucht werden muss zunächst, wie das urheberrechtliche Regelungsregime im Kontext der Verwendung von Werken als Trainings- und Inputdaten ausgestaltet sein sollte. Fraglich ist damit zum einen, ob die Verwendung urheberrechtlich geschützter Werke im Rahmen der untersuchten Fallgruppen den betroffenen Urhebern vorbehalten bleiben sollte, oder ob eine Freistellung der Sachverhalte zweckmäßig ist. Der Freistellungsbegriff im Sinne der vorliegenden Untersuchung umfasst jede Regelung, die es Trainingsdatensammlern oder Verwendern von KI ermöglicht, Werke für den jeweiligen Zweck ohne individuelle Zustimmung der betroffenen Urheber zu verwenden, womit die Nutzungen im Regelfall nicht mehr durch die Urheber kontrolliert werden können.[49] Falls eine Freistellung zweckmäßig ist, muss darüber hinaus geklärt werden, durch welchen Freistellungsmechanismus die Erlaubnis umgesetzt sein sollte.

A. Herleitung der Beurteilungskriterien und geeigneter Freistellungsmechanismen

Unklar ist allerdings, nach welchen Kriterien die Zweckmäßigkeit einer Freistellung zu beurteilen ist, welche Regelungsmechanismen für die Umsetzung eines Freistellungsbedürfnisses grundsätzlich zur Verfügung stehen und anhand welcher Anhaltspunkte eine Wahl zwischen den verschiedenen Rechtsinstituten erfolgen kann.

49 Angelehnt an EuGH, Urt. v. 16.11.2016 - C-301/15, GRUR 2017, 62 Rn. 33 – Soulier; einen engeren Freistellungsbegriff verwendet hingegen Stieper, siehe *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 6, er differenziert beispielsweise noch einmal zwischen Freistellungen und gesetzlichen Lizenzen.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

I. Beurteilungskriterien für die Zweckmäßigkeit einer Freistellung

Ob eine Freistellung zweckmäßig ist, könnte von mehreren miteinander in Beziehung stehenden Faktoren abhängen.

1. Der Schutz der Urheberinteressen als Hauptzweck des Urheberrechts

Der Schutz der Schöpferinteressen ist Kernanliegen des Urheberrechts.[50] Das kommt bereits in § 1 UrhG zum Ausdruck. Umgekehrt ist der Schutz der Schöpferinteressen damit auch für regulatorische Überlegungen von wesentlicher Bedeutung.

a) Die Urheberinteressen

Abgesichert werden sollen zum einen die Vermögensinteressen der Urheber.[51] Das Urheberrecht zielt dabei darauf ab, dem Schöpfer ein Wirtschaftsgut zu verschaffen, dass nach dessen schöpferischer Tätigkeit zu seinem Unterhalt beiträgt.[52] Außerdem übt die Perspektive auf gerechte Entlohnung eine Anreizwirkung für schöpferische Leistung aus.[53] Aus diesen Gründen soll der Urheber vor einer unbefugten Auswertung seines Werkes geschützt werden.[54]

---

50 Siehe auch *Schack*, Urheber- und Urhebervertragsrecht, S. 2 Rn. 2.

51 BT-Drs. IV/270, S. 28.

52 BGH, Urt. v. 18.05.1955 - I ZR 8/54, GRUR 1955, 492 (496) – Grundig-Reporter; *Schack*, Urheber- und Urhebervertragsrecht, S. 6 Rn. 11; *Ulmer*, Urheber- und Verlagsrecht, S. 24 f. Im europäischen Recht auch Richtlinie 2006/115/EG des Europäischen Parlaments und des Rates vom 12. Dezember 2006 zum Vermietrecht und Verleihrecht sowie zu bestimmten dem Urheberrecht verwandten Schutzrechten im Bereich des geistigen Eigentums (Vermiet- und Verleih-RL) ErwGr. 5.

53 Zum folgenorientierten Begründungsansatz siehe *Loewenheim*, in: Schricker/Loewenheim, Urheberrecht, Einleitung zum UrhG Rn. 9; *Fisher*, Chicago-Kent Law Review 1998, 1203 (1213 f.); sehr kritisch dazu aber *Kleinemenke*, Fair Use im deutschen und europäischen Urheberrecht?, S. 50 Fn. 97 m. w. N. Tatsächlich scheint sich zu bestätigen, dass der aus dem Urheberrechtsschutz resultiernde Anreiz für Schöpfer geringer ist als angenommen, mit ersten vorsichtigen Ergebnissen hierzu *Handke/Girard/Mattes*, Fördert das Urheberrecht Innovation? Eine empirische Untersuchung, S. 105, 121.

54 Vgl. BT-Drs. IV/270, S. 27.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Darüber hinaus verkörpert jedes Werk aber auch die Persönlichkeit seines Urhebers.[55] Das Band, das zwischen Urheber und Werk geknüpft wird, muss neben den materiellen Interessen ebenfalls geschützt werden.[56] Daher sind auch ideelle Urheberinteressen zu berücksichtigen. Hierzu zählt insbesondere, sein Werk vor entstellenden Änderungen zu schützen.[57] Auch das Interesse eines Urhebers an der öffentlichen Anerkennung seiner Urheberschaft[58] sowie das Interesse daran, selbst darüber zu bestimmen, ob, wann und unter welchen Umständen sein Werk erstmalig öffentlich zugänglich gemacht wird,[59] ist zu berücksichtigen.

b) Grundrechtlicher Schutz der Urheberinteressen

Das Urheberrecht ist in seiner vermögensrechtlichen Ausprägung als geistiges Eigentum außerdem durch spezielle Grundrechtspositionen geschützt. Da die vermögensrechtliche Komponente des urheberrechtlichen Rechtsrahmens weitestgehend durch das Unionsrecht harmonisiert ist,[60] wird dieser Schutz vorrangig durch Art. 17 Abs. 2 EU-GRCh realisiert.[61] Im Übrigen fällt das Urheberrecht als geistiges Eigentum aber auch in den Schutzbe-

55 EuGH, Urt. v. 01.12.2011 - C-145/10, GRUR 2012, 166 Rn. 88 – Painer; *Schack*, Urheber- und Urhebervertragsrecht, S. 4 Rn. 7; außerdem Richtlinie 93/98/EWG des Rates vom 29. Oktober 1993 zur Harmonisierung der Schutzdauer des Urheberrechts und bestimmter verwandter Schutzrechte (Schutzdauer-RL) ErwGr. 17.

56 Vgl. schon BT-Drs. IV/270, S. 44, 45.

57 So schon BT-Drs. IV/270, S. 43; eines der drei Pfeiler des Urheberpersönlichkeitsrechtes, siehe *Peukert*, in: Schricker/Loewenheim, Urheberrecht, UrhG § 14 Rn. 1.

58 So schon RG, Urt. v. 08.04.1925 - I 251/24, RGZ 110, 393 (397); BGH, Urt. v. 16.06.1994 - I ZR 3/92, GRUR 1995, 671 (672) – Namensnennungsrecht des Architekten; BT-Drs. IV/270, S. 43.

59 „Grundnorm des Urheberschutzes", so *Peukert*, in: Schricker/Loewenheim, Urheberrecht, UrhG § 12 Rn. 1; außerdem bereits *v. Gierke*, Deutsches Privatrecht Band 1: Allgemeiner Teil u. Personenrecht, S. 790, 792.

60 Das gilt insbesondere für die Verwertungsrechte des Vervielfältigungs- und Verbreitungsrechts und des Rechts der öffentlichen Wiedergabe sowie der zugehörigen Schrankenbestimmungen, siehe *v. Ungern-Sternberg*, in: Schricker/Loewenheim, Urheberrecht, UrhG § 15 Rn. 1; *Stieper*, in: Schricker/Loewenheim, Urheberrecht, Vor UrhG §§ 44a ff. Rn. 25 ff. Siehe insgesamt zum Harmonisierungsstand des Urheberrechts *Nordemann*, in: Fromm/Nordemann, Urheberrecht, Einl. zum UrhG Rn. 43 ff.

61 Zum Urheberrecht als Teil des Schutzbereichs des Art. 17 EU-GRCh für alle EuGH, Urt. v. 29.07.2019 - C-469/17, GRUR 2019, 934 Rn. 57 – Afghanistan Papiere; *Jarass*, in: Jarass, Charta der Grundrechte der Europäischen Union, EU-GRCh Art. 17 Rn. 10. Selbst im mindestharmonisierten Bereich geben die Unionsgrundrechte den Mindeststandard vor, siehe EuGH, Urt. v. 26.02.2013 - C-399/11, NJW 2013, 1215 Rn. 60 –

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

reich des nationalen Art. 14 GG.[62] Soweit ein Anwendungsspielraum für nationale Grundrechte verbleibt,[63] muss damit auch Art. 14 GG berücksichtigt werden. Hieraus folgt zum einen, dass bei einer Freistellung der Verwendung von Werken der unantastbare Kernbereich beziehungsweise Wesensgehalt des Urheberrechts gewahrt bleiben muss. Er ist sowohl in der europäischen als auch in der nationalen Grundrechtsdogmatik verankert.[64] Weiterhin muss aber auch die Bedeutung des Art. 17 EU-GRCh sowie unter Umständen des Art. 14 GG im grundrechtlichen Gefüge beachtet werden.

Als besondere Ausprägung des allgemeinen Persönlichkeitsrechts genießen die im Urheberpersönlichkeitsrecht gebündelten ideellen Interessen des Urhebers darüber hinaus auch über Art. 2 Abs. 1 i. V. m. Art. 1 Abs. 1 GG besonderen grundrechtlichen Schutz.[65]

Wenn untersucht wird, ob und inwiefern eine Freistellung zweckmäßig ist, muss damit auch berücksichtigt werden, welche Bedeutung die in der jeweiligen Sachverhaltskonstellation betroffenen Grundrechte haben, wie sie eingeschränkt werden können und in welchem Verhältnis sie zu anderen geschützten Rechtspositionen stehen.

Melloni; sowie insgesamt zum Anwendungsbereich der Unionsgrundrechte weitergehend *Stieper*, in: Schricker/Loewenheim, Urheberrecht, Vor UrhG §§ 44a ff. Rn. 32a.

62 Zum verfassungsrechtlichen Eigentumsbegriff vgl. BVerfG, Beschl. v. 10.06.2009 - 1 BvR 706/08 u.a., NJW 2009, 2033 Rn. 215 m. w. N.; zum Urheberrecht als Eigentum im Sinne des Art. 14 GG erstmals BVerfG, Beschl. v. 07.07.1971 - 1 BvR 765/66, GRUR 1972, 481 (483) – Kirchen- und Schulgebrauch.

63 Zu einem umfassenden Überblick über die Abgrenzungskriterien zur Differenzierung zwischen den Anwendungsbereichen der nationalen und europäischen Grundrechte siehe insgesamt BGH, Urt. v. 30.04.2020 - I ZR 115/16, GRUR 2020, 843 Rn. 45 ff. – Metall auf Metall IV m. w. N.

64 Zum Kernbereich des Art. 14 GG: BVerfG, Beschl. v. 02.03.1999 - 1 BvL 7–91, NJW 1999, 2877 (2878); *Axer*, in: Epping/Hillgruber, Beck'scher Online-Kommentar Grundgesetz, GG Art. 14 Rn. 87. Zum Wesensgehalt des Unionsgrundrechts, der über Art. 52 Abs. 1 S. 1 EU-GRCh gewährt wird, siehe im Kontext der Eigentumsgarantie *Jarass*, in: Jarass, Charta der Grundrechte der Europäischen Union, EU-GRCh Art. 17 Rn. 32a m. w. N. aus der Rechtsprechung des EuGH.

65 *Peukert*, in: Schricker/Loewenheim, Urheberrecht, Vor UrhG §§ 12 ff. Rn. 6; *Specht-Riemenschneider*, in: Dreier/Schulze, Urheberrechtsgesetz, Vor UrhG §§ 12 ff. Rn. 5; *Krüger-Nieland*, in: FS Hauß, S. 215 (219 f., 223 f.). Der Fokus liegt dabei auf Art. 2 Abs. 1 GG, sodass im Hinblick auf die ideellen Urheberinteressen keine absoluten Einschränkungsvorbehalte bestehen, vgl. beispielsweise BVerfG, Beschl. v. 24.11.2009 - 1 BvR 213/08, ZUM 2010, 235 Rn. 73; anders aber *Schack*, GRUR 1985, 352 (353); *Schack*, Urheber- und Urhebervertragsrecht, S. 25 Rn. 48.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

c) Schlussfolgerung: Freistellung wegen Unerheblichkeit für die Urheber

Aus dem Schutzziel des Urheberrechts folgt im Umkehrschluss, dass Handlungen wie unter anderem das Web Scraping von Werken als Trainingsdaten vom Ausschließlichkeitsrecht der Urheber ausgeklammert werden können, wenn die ideellen oder materiellen Urheberinteressen durch sie nicht beeinträchtigt sind. Dieser regulatorische Ansatz wurde beispielsweise auch bei der Einführung der Schrankenbestimmungen aus §§ 55, 56 UrhG (Vervielfältigungen durch Sendeunternehmen und in Geschäftsbetrieben) verfolgt.[66]

2. Freistellungen zum Ausgleich von Nutzer- und Urheberinteressen

Zwar schützt das „Recht des Urhebers", wie es schon in § 1 UrhG zum Ausdruck kommt, im Grundsatz vorrangig die ideellen und materiellen Urheberinteressen. Neben den Interessen der betroffenen Urheber könnten bei der Evaluation von Freistellungsbedürfnissen aber auch verschiedene Interessen der betroffenen Nutzergruppen, also der Trainingsdatensammler, KI-Entwickler oder Anwender von KI-Systemen berücksichtigt werden müssen. Fraglich ist jedoch, wo der Anknüpfungspunkt zur Berücksichtigung dieser Nutzerinteressen liegt und wie eine Berücksichtigung methodisch abgebildet werden kann.

a) Anknüpfungspunkt des Interessenausgleichs

Um die betroffenen Nutzerinteressen einzubeziehen, könnten sich im Wesentlichen drei Anknüpfungspunkte sowohl aus der europäischen als auch aus der nationalen Rechtsordnung ergeben.

aa) Grundrechte der Nutzer

Möglich ist zum einen, dass auch die Interessen der Trainingsdatensammler, KI-Entwickler oder Anwender von KI-Systemen grundrechtlich geschützt sind. Wenn das Sammeln der urheberrechtlich geschützten Trai-

66 BT-Drs. IV/270, S. 32.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

ningsdaten, die Durchführung des Trainingsvorgangs selbst oder die Anwendung eines KI-Systems auf Grundlage urheberrechtlich geschützten Inputs Teil der Berufsausübung oder der unternehmerischen Tätigkeit ist, wird der Schutzbereich der Berufsfreiheit aus Art. 15 Abs. 1 EU-GRCh sowie aus Art. 12 GG oder der unternehmerischen Freiheit aus Art. 16 EU-GRCh eröffnet.[67] Je nach Anwendungsumfeld der KI könnten auch die Wissenschafts- und Pressefreiheit (Art. 11, 13 EU-GRCh sowie Art. 5 Abs. 1 S. 2, Abs. 3 GG) betroffen sein. Beim Einsatz von KI kollidieren demzufolge immer wieder nicht nur gegenläufige Interessen von Urhebern und Nutzern, sondern auch Grundrechtspositionen auf unionsrechtlicher oder nationaler Ebene miteinander. Diese Grundrechtspositionen müssen in der Folge zu einem angemessenen, gerechten Ausgleich gebracht werden.[68] Das folgt auf unionsrechtlicher Ebene bereits aus Art. 52 Abs. 1 S. 2 EU-GRCh.[69] Gerecht bedeutet, dass nicht nur die Urheberinteressen, sondern auch die Belange der Trainingsdatensammler, KI-Entwickler und Anwender von KI-Systemen gebührend berücksichtigt werden müssen. Ziel des Ausgleichs ist also, alle involvierten, grundrechtlich geschützten und gegenläufigen Interessen bestmöglich zur Geltung zu bringen.[70] Dies wird auf national-verfassungsrechtlicher Ebene als praktische Konkordanz bezeichnet.[71]

---

67 Zum Schutzbereich des Art. 12 GG vgl. *Jarass*, in: Jarass/Pieroth, Grundgesetz für die Bundesrepublik Deutschland: Kommentar, GG Art. 12 Rn. 10 m. w. N.

68 Für die Unionsgrundrechte: EuGH, Urt. v. 29.01.2008 - C-275/06, GRUR 2008, 241 Rn. 66 – Promusicae/Telefónica; EuGH, Urt. v. 22.01.2013 - C-283/11, ZUM 2013, 202 Rn. 60 – Sky Österreich; EuGH, Urt. v. 27.03.2014 - C-314/12, NJW 2014, 1577 Rn. 63 – UPC Telekabel; EuGH, Urt. v. 26.04.2022 - C-401/19, GRUR 2022, 820 Rn. 66 – Polen/Parlament und Rat. Aus dem nationalen Recht: BVerfG, Beschl. v. 24.02.1971 - 1 BvR 435/68, NJW 1971, 1645 (1647) – Mephisto; BVerfG, Beschl. v. 03.11.1987 - 1 BvR 1257/84 u. a., NJW 1988, 325 (326) – Herrnburger Bericht; vgl. auch BVerfG, Beschl. v. 26.05.1970 - 1 BvR 83, 244, 345/69, NJW 1970, 1729 (1730) – Dienstpflichtverweigerung; *Epping*, Grundrechte, S. 33 Rn. 79.

69 Vgl. dazu auch *Streinz/Michl*, in: Streinz, EUV/AEUV, EU-GRCh Art. 52 Rn. 17; *Schwerdtfeger*, in: Meyer/Hölscheidt, Charta der Grundrechte der Europäischen Union, EU-GRCh Art. 52 Rn. 39.

70 Im Unionsrecht folgt das schon daraus, dass der EuGH von einem „Gleichgewicht" der kollidierenden Grundrechte spricht, siehe EuGH, Urt. v. 29.01.2008 - C-275/06, GRUR 2008, 241 Rn. 66 – Promusicae/Telefónica; EuGH, Urt. v. 22.01.2013 - C-283/11, ZUM 2013, 202 Rn. 60 – Sky Österreich; EuGH, Urt. v. 27.03.2014 - C-314/12, NJW 2014, 1577 Rn. 63 – UPC Telekabel; EuGH, Urt. v. 26.04.2022 - C-401/19, GRUR 2022, 820 Rn. 66 – Polen/Parlament und Rat. Dies setzt eine bestmögliche Geltung aller kollidierender Grundrechtspositionen voraus. Für das nationale Recht siehe BVerfG, Beschl. v. 07.03.1990 - 1 BvR 266/86, 1 BvR 913/87, NJW 1990, 1982 (1983) – Bundesflagge.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Auch bei der Identifizierung von Freistellungsbedürfnissen kann dieser Ausgleichsmechanismus zur Anwendung gebracht werden. Mit seiner Hilfe lassen sich im konkreten Fall aufeinandertreffende, grundrechtlich geschützte Interessen von Urhebern sowie Trainingsdatensammlern, KI-Entwicklern oder Anwendern von KI-Systemen zum Ausgleich bringen.[72]

Folge der Grundrechtskollision ist auf nationaler Ebene nicht nur ein Abwägungsraum für die Judikative bei Auslegung des geltenden Rechts. Vielmehr folgt aus ihm schon ein Auftrag an den nationalen Gesetzgeber, die widerstreitenden Interessen zum Ausgleich zu bringen.[73] Auf unionsrechtlicher Ebene könnte eine entsprechende Förderungs- und damit auch Ausgleichspflicht innerhalb des Kompetenzbereichs des europäischen Normgebers mitunter über Art. 51 Abs. 1 S. 2 EU-GRCh begründet werden.[74] Aus beidem folgt, dass der Ausgleich grundrechtlich abgesicherter Urheber- und Nutzerinteressen ein Leitmotiv für die Ermittlung normativer Regelungsbedürfnisse ist. Grundrechtspositionen der Trainingsdatensammler, KI-Entwickler oder Anwender von KI-Systemen bilden damit einen ersten Anknüpfungspunkt, über den ihre Interessen bei der Identifizierung von Regelungsbedürfnissen berücksichtigt werden können.

bb) Berücksichtigung von Nutzerinteressen als sekundäres Regelungsziel des europäischen Urheberrechts

Darüber hinaus könnten Nutzerinteressen aber auch dann berücksichtigt werden, wenn kein angemessener Ausgleich spezieller, kollidierender Grundrechtspositionen erforderlich ist. Denn Regelungsziel der InfoSoc-RL[75] des Unionsgesetzgebers, die sowohl die Verwertungsrechte als auch

71 Beispielsweise in Begriff geprägt von *Hesse*, Grundzüge des Verfassungsrechts der Bundesrepublik Deutschland, S. 142 Rn. 317; verwendet beispielsweise in BVerfG, Beschl. v. 18.02.2019 - 1 BvR 2556/17, GRUR 2019, 606 Rn. 9 – Loud; *Epping*, Grundrechte, S. 93.

72 Dazu auch allgemein *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 45 ff.

73 BVerfG, Beschl. v. 27.01.2015 - 1 BvR 471/10, 1 BvR 1181/10, NJW 2015, 1359 Rn. 98 – Kopftuchverbot.

74 *Schwerdtfeger*, in: Meyer/Hölscheidt, Charta der Grundrechte der Europäischen Union, EU-GRCh Art. 51 Rn. 68.

75 Richtlinie 2001/29/EG des Europäischen Parlaments und des Rates vom 22. Mai 2001 zur Harmonisierung bestimmter Aspekte des Urheberrechts und der verwandten Schutzrechte in der Informationsgesellschaft (InfoSoc-RL).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

einen wesentlichen Teil der urheberrechtlichen Schrankenbestimmungen harmonisiert und damit einen wesentlichen Grundstein des europäischen Urheberrechts verkörpert, sowie der DSM-RL[76] ist nicht nur der Schutz der urheberrechtlichen Verwertungsinteressen. Vielmehr dienen die Richtlinien gemäß ErwGr. 31 S. 1 InfoSoc-RL sowie ErwGr. 6 S. 1 DSM-RL auch dem Rechts- und Interessenausgleich zwischen Rechtsinhabern und Nutzern. Beide Erwägungsgründe bringen einen allgemeinen Rechtsgedanken zum Ausdruck. Dieser wird auch in ErwGr. 1 verwaiste Werke-RL[77] anerkannt. Nach ihm bietet das europäische Urheberrecht einen Anknüpfungspunkt für die Berücksichtigung von Nutzerinteressen bei der Ausgestaltung des Rechtsrahmens unabhängig davon, ob diese Interessen durch spezielle Freiheitsrechte geschützt sind und damit in ein grundrechtlich geprägtes Spannungsverhältnis mit dem Urheberrecht eintreten.

cc) Sozialbindung des Urheberrechts

Ein dritter Bezugspunkt für die Berücksichtigung von Nutzerinteressen abseits des Ausgleichs spezieller, kollidierender Grundrechtspositionen könnte im nationalen Regelungszusammenhang zudem über Art. 14 GG bestehen, soweit für ihn im konkreten Fall Anwendungsspielräume verbleiben. Er sieht in Art. 14 Abs. 1 S. 2 GG einen Gesetzesvorbehalt vor. Dieser ermöglicht der Legislative, das Urheberrecht zur Verfolgung anderer Zwecke einzuschränken.[78] Hierbei muss aber stets die in Art. 14 Abs. 2 GG verankerte Sozialbindung des Eigentums berücksichtigt werden. Sie umfasst auch das Urheberrecht als geistiges Eigentum.[79]

76 Richtlinie 2019/790/EU des Europäischen Parlaments und des Rates vom 17. April 2019 über das Urheberrecht und die verwandten Schutzrechte im digitalen Binnenmarkt und zur Änderung der Richtlinien 96/9/EG und 2001/29/EG (DSM-RL).

77 Richtlinie 2012/28/EU des Europäischen Parlaments und des Rates vom 25. Oktober 2012 über bestimmte zulässige Formen der Nutzung verwaister Werke (verwaiste Werke-RL).

78 Vgl. BVerfG, Beschl. v. 07.07.1971 - 1 BvR 276/71, GRUR 1972, 487 – Schulfunksendungen.

79 So schon das RG, Urt. v. 26.04.1933 - ZR I 301/32, RGZ 140, 264 (270) – Verlagsrecht an sämtlichen Ausgaben; im Weiteren *Götting*, in: Loewenheim, Handbuch des Urheberrechts, § 30 Rn. 1; allgemein zur Sozialbindung und dem Gesetzesvorbehalt in Art. 14 GG BVerfG, Urt. v. 10.03.1981 - 1 BvR 92/71, 1 BvR 96/71, NJW 1981, 1257 – Private Gondelbahn.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Von der Sozialbindung des Eigentums geht eine legitimierende, gleichzeitig aber auch eine begrenzende Wirkung für Einschränkungen des Urheberrechts aus.[80] Das Urheberrecht darf nur zugunsten öffentlicher, also solcher Interessen eingeschränkt werden, die für die gesamte oder zumindest einen Teil der Gesellschaft von Bedeutung sind.[81] Die Sozialbindung ermöglich eine Einschränkung allerdings auch dann, wenn die öffentlichen Interessen im Speziellen nicht unmittelbar grundrechtlich geschützt sind. Allgemeininteressen können auch Interessen der Gesamtheit aller Trainingsdatensammler, KI-Entwickler oder Anwender von KI-Systemen als Nutzergruppen der Allgemeinheit sein. Möglich ist deswegen, dass die Interessen der Urheber über Art. 14 Abs. 1 S. 2, Abs. 2 GG auch gegenüber nicht in speziellen Grundrechten widergespiegelten Interessen dieser Nutzergruppen zurücktreten müssen.

Aus Art. 14 Abs. 1 S. 2 GG kann dabei ein konkreter Regelungsauftrag an den Gesetzgeber abgeleitet werden.[82] Er ist angehalten, proaktiv den Ausgleich zwischen kollidierenden Urheber- und öffentlichen Interessen vorzunehmen, um die Sozialbindung des geistigen Eigentums zu verwirklichen.[83] Bei der Analyse, wie die urheberrechtlichen Regelungen ausgestaltet sein sollen, bilden der Gesetzesvorbehalt des Art. 14 Abs. 1 S. 2 GG und die Sozialbindung des Urheberrechts damit einen dritten, national geprägten Anknüpfungspunkt, über den objektivierte Interessen der Trainingsdatensammler, KI-Entwickler oder Anwender von KI-Systemen berücksichtigt werden können.

b) Funktion des Ausgleichsmechanismus

Klärungsbedürftig bleibt, über welchen Mechanismus der Ausgleich von objektivierten Urheber- und Nutzerinteressen vorgenommen werden kann.

---

80 BVerfG, Beschl. v. 02.03.1999 - 1 BvL 7–91, NJW 1999, 2877 (2878); für das Urheberrecht: *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 43.

81 *Götting*, in: Loewenheim, Handbuch des Urheberrechts, § 30 Rn. 1; oft auch als „Allgemeininteressen" bezeichnet, vgl. beispielsweise BVerfG, Beschl. v. 25.10.1978 - 1 BvR 352/71, GRUR 1980, 44 (46) – Kirchenmusik; *Schack*, Urheber- und Urhebervertragsrecht, S. 48 Rn. 93; BT-Drs. IV/270, S. 30.

82 BVerfG, Urt. v. 06.12.2016 - 1 BvR 2821/11, 1 BvR 321/12, 1 BvR 1456/12, NJW 2017, 217 Rn. 218 – Atomausstieg; BVerfG, Beschl. v. 12.06.1979 - 1 BvL 19/76, NJW 1980, 985 (987).

83 BVerfG, Beschl. v. 12.06.1979 - 1 BvL 19/76, NJW 1980, 985 (987).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Erster Anknüpfungspunkt zur Berücksichtigung der Interessen von Trainingsdatensammlern, KI-Entwicklern und Anwendern von KI-Systemen sind ihre Grundrechtspositionen, die mit der Grundrechtsposition der Urheber kollidieren können. Aus einer europarechtlichen Perspektive darf der Unionsgesetzgeber in diesen Fällen in jedes der betroffenen Grundrechte nur so weit eingreifen, wie es verhältnismäßig ist. Dies folgt unmittelbar aus Art. 52 Abs. 1 S. 2 EU-GRCh.[84] Auch beim Ausgleich kollidierender Grundrechtspositionen auf national-verfassungsrechtlicher Ebene kommt es im Rahmen der Anwendung des Grundsatzes der praktischen Konkordanz darauf an, inwieweit die Einschränkung des zurückstehenden Grundrechts zwingend erforderlich,[85] also verhältnismäßig ist.[86] Der auf beiden Stufen des europäischen Mehrebenensystems zum Tragen kommende Verhältnismäßigkeitsgrundsatz schließt dabei eine umfassende Abwägung der in Widerspruch zueinander stehenden Grundrechtspositionen ein.[87] Der Ausgleich der Urheberinteressen mit den grundrechtlich geschützten Interessen der bezeichneten Nutzergruppen käme infolgedessen vor allem durch eine umfassende Interessenabwägung zustande.

Eine Konkretisierung der Funktionsweise des Interessenausgleichsmechanismus, den ErwGr. 31 S. 1 InfoSoc-RL sowie ErwGr. 6 S. 1 DSM-RL in Bezug nehmen, findet in den jeweiligen Erwägungsgründen darüber hinaus nicht statt. Dort, wo die Durchführung des Interessenausgleichs von der Normsetzungs- auf die Anwendungsphase verlagert wurde, wie es insbesondere beim in Art. 5 Abs. 5 InfoSoc-RL verankerten Dreistufentest der Fall ist, wird jedoch deutlich, dass der bezweckte Interessenausgleich auch gemäß der Richtlinienkonzeption im Wesentlichen durch eine umfassende Interessenabwägung gewährleistet wird.[88] Im Rahmen des im Richtlinienrecht zum Ausdruck kommenden Interessenausgleichs kommt es demzufol-

84 Vgl. dazu auch EuGH, Urt. v. 22.01.2013 - C-283/11, ZUM 2013, 202 Rn. 47 ff. – Sky Österreich; EuGH, Urt. v. 27.03.2014 - C-314/12, NJW 2014, 1577 Rn. 63 – UPC Telekabel; *Streinz/Michl*, in: Streinz, EUV/AEUV, EU-GRCh Art. 52 Rn. 17; *Schwerdtfeger*, in: Meyer/Hölscheidt, Charta der Grundrechte der Europäischen Union, EU-GRCh Art. 52 Rn. 39.

85 BVerfG, Beschl. v. 26.05.1970 - 1 BvR 83, 244, 345/69, NJW 1970, 1729 (1730) – Dienstpflichtverweigerung; *Hesse*, Grundzüge des Verfassungsrechts der Bundesrepublik Deutschland, S. 142 Rn. 318.

86 *Hesse*, Grundzüge des Verfassungsrechts der Bundesrepublik Deutschland, S. 142 Rn. 318.

87 Für die unionsrechtliche Dogmatik *Kingreen*, in: Calliess/Ruffert, EUV/AEUV, EU-GRCh Art. 52 Rn. 70 ff. Für das nationale Recht siehe *Grabitz*, AöR 1973, 568 (575).

88 Zum Dreistufentest im Detail siehe 3. Teil A. III. 2. d).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

ge ebenfalls auf eine freie Gegenüberstellung der betroffenen Urheber- und Nutzerinteressen an.

Auch Einschränkungen des Urheberrechts, die aus der national angelegten Sozialbindung des Immaterialgüterrechts folgen, müssen sich am Verhältnismäßigkeitsgrundsatz messen lassen.[89] In diesen Fällen müssten die gegenläufigen Urheber- und Nutzerinteressen demzufolge ebenfalls in eine umfassende Abwägung zueinander gestellt werden.

Unabhängig davon, über welchen Anknüpfungspunkt die jeweilig betroffenen Nutzerinteressen berücksichtigt werden können, ist die umfassende Abwägung mit den Interessen betroffener Urheber damit zentraler Ausgleichsmechanismus. Hieraus folgt für die vorliegende Untersuchung, dass die umfassende Interessenabwägung als Ausgleichinstrument angewendet werden kann, um Freistellungsbedürfnisse zu identifizieren. Die urheberrechtliche Freistellung einer der im Folgenden zu untersuchenden Sachverhaltskonstellationen ist neben der Nichtbeeinträchtigung der Urheberinteressen demzufolge auch dann zweckmäßig, wenn eine umfassende Interessenabwägung dazu führt, dass den Belangen der betroffenen Nutzergruppe im konkreten Fall Vorrang vor den Urheberinteressen zu gewähren ist.[90]

3. Berücksichtigungsfähigkeit einer ökonomischen Analyse des Urheberrechts

Auch die ökonomische Analyse des Urheberrechts könnte darüber hinaus zur Identifizierung von Freistellungsbedürfnissen geeignet sein.

a) Überblick über die Funktionsweise

Nach der volkswirtschaftlich geprägten ökonomischen Analyse des Rechs müsse die Rechtsordnung das Ziel verfolgen, einen Zustand der Effizienz herbeizuführen.[91] Er ist erreicht, wenn alle wirtschaftlichen Mittel innerhalb der Gesellschaft so verteilt sind, dass die Bedürfnisse aller bestmöglich, also aus einem gesamtwirtschaftlichen Blickwinkel optimal, befriedigt

89 BVerfG, Beschl. v. 08.04.1987 - 1 BvR 564, 684, 877, 886, 1134, 1636, 1711/84; BVerfG, Beschl. v. 26.04.1995 - 1 BvL 19/94 u.a., DtZ 1995, 323 (324).

90 Vgl. insbesondere *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 96; *Peukert*, Urheberrecht, § 2 Rn. 5; BT-Drs. IV/270, S. 30.

91 *Schäfer/Ott*, Lehrbuch der ökonomischen Analyse des Zivilrechts, S. XI.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

werden.[92] In einem solchen Zustand bedeutete jede Besserstellung eines Beteiligten unweigerlich die Schlechterstellung einer anderen Person.[93]

Um zu bestimmen, ob eine rechtliche Regelung dieses Effizienzziel fördert, sind daher stets ihre Folgen in den Blick zu nehmen.[94] Wichtiger Maßstab zur Beurteilung der Folgen einer Regelung ist das Pareto-Kriterium. Hiernach ist ein Zustand einem anderen vorzuziehen, wenn er zumindest von einem Beteiligten präferiert, gleichzeitig aber von keinem anderen Beteiligten abgelehnt wird.[95] Ob der wirtschaftlich rational handelnde Mensch die Anpassung der Rechtsordnung gegenüber dem Verbleib im Ist-Zustand vorzieht, hängt maßgeblich davon ab, welche wirtschaftlichen Folgen die Rechtsänderung hätte. Die anzustellende Kosten-Nutzen-Abwägung muss stets positiv ausfallen.[96]

Neben Pareto-Superiorität können Anpassungen der Rechtsordnung das Effizient-Ziel auch fördern, wenn sie das Kaldor/Hicks-Kriterium erfüllen. Dazu muss es der wirtschaftlich bevorteilten Gruppe möglich sein, die Benachteiligten monetär zu entschädigen, ohne selbst ihren erlangten Vorteil vollständig aufzubrauchen.[97]

b) Eignung der ökonomischen Analyse des Rechts für die Beantwortung der aufgeworfenen Fragestellungen

Unklar ist aber, ob mit Hilfe der ökonomischen Analyse des Urheberrechts aussagekräftige Antworten auf die zu untersuchenden Fragestellungen gefunden werden können.

Hierfür spricht zunächst, dass die ökonomische Rechtsanalyse einen objektiven Maßstab verspricht, anhand derer Regelungsbedürfnisse zur Erreichung von Allokationseffizienz identifiziert werden können.[98] Darüber hinaus führt die ökonomische Analyse des Rechts zu einer Vereinfachung

92 *Schäfer/Ott*, Lehrbuch der ökonomischen Analyse des Zivilrechts, S. XI.
93 *Schäfer/Ott*, Lehrbuch der ökonomischen Analyse des Zivilrechts, S. 14.
94 Vgl. *Eidenmüller*, Effizienz als Rechtsprinzip, S. 3 ff.
95 *Eidenmüller*, Effizienz als Rechtsprinzip, S. 48; *Schäfer/Ott*, Lehrbuch der ökonomischen Analyse des Zivilrechts, S. 13.
96 *Kirchgässner*, Homo oeconomicus, S. 12.
97 *Eidenmüller*, Effizienz als Rechtsprinzip, S. 51.
98 Vgl. *Maamar*, Computer als Schöpfer, S. 92 („quasi-mathematische Argumentation").

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

komplexer Sachverhalte in eine Modellstruktur.[99] Sie hilft dabei, die vielschichtigen Sachverhalte Lösungen zuzuführen.[100]

Allerdings müssen bei der Ausgestaltung der Rechtsordnung auch der Grundrechtsschutz der Urheber, diverse Nutzergrundrechte und weitere Allgemeininteressen berücksichtigt werden.[101] Das Modelldenken verhindert jedoch, dass diese Bandbreite normativ abzubildender Interessen in die Beurteilung eines Freistellungsinteresses mit einbezogen werden können.[102] Die vielschichtigen Sachverhalte auf das rechtsökonomische Effizienzziel zu reduzieren, welches selbst weder in den Primärrechtsakten der Europäischen Union noch in der nationalen Verfassung[103] abgebildet ist, wird der tatsächlichen Komplexität der zu untersuchenden Sachverhaltskonstellationen nicht gerecht. Zwar könnten diese Aspekte durch eine spätere Korrektur des ökonomischen Analyseergebnisses zugunsten anderer, höherwertiger und insbesondere nicht ökonomischer Interessen berücksichtigt werden.[104] Gerade wenn das Ergebnis der ökonomischen Analyse des Urheberrechts aber nur insoweit für maßgeblich erachtet wird, als es mit anderen, höherwertigen Interessen nicht kollidiert, ist die zuvor vorgenommene Analyse der Regelungsbedürfnisse auf Basis ökonomischer Modelle nicht von Bedeutung.[105] Es käme vielmehr nur auf die sich anschließende Interessenabwägung an.

Darüber hinaus haben urheberrechtliche Regelungen, die auf konkrete, eng abgegrenzte Arbeitsschritte wie die hier in Rede stehenden bezogen sind, nur sektoral begrenzte Auswirkungen auf die zu erwartenden Kosten sowie die generierbaren Vorteile für Konsumenten, also Nutzer wie Trainingsdatensammler, KI-Entwickler oder Anwender von KI-Systemen. Gleiches gilt für die Kosten-Nutzen-Rechnung der Urheber. Aus solchen nanoökonomischen Beobachtungen lassen sich aber keine Auswirkungen

99 *Bisges*, ZUM 2014, 930 (930); *Maamar*, Computer als Schöpfer, S. 92.

100 *Bischoffshausen*, Die ökonomische Rechtfertigung der urheberrechtlichen Schutzfrist, S. 259 spricht von einer „Verringerung der Komplexität juristischer Problemstellungen"; Auch *Maamar*, Computer als Schöpfer, S. 92.

101 Dazu bereits oben unter 2. Teil A. I. 1. sowie 2.

102 So bereits auch *Fezer*, JZ 1988, 223 (224).

103 *Eidenmüller*, Effizienz als Rechtsprinzip, S. 444 f.; siehe außerdem BVerfG, Urt. v. 01.03.1979 - 1 BvR 532, 533/77, 419/78, 1 BvL 21/78, NJW 1979, 699 (702) – Mitbestimmung: Das BVerfG sieht im Grundgesetz keine Wirtschaftsordnung angelegt; vgl. außerdem *Fezer*, JZ 1988, 223 (224).

104 So *Schäfer/Ott*, Lehrbuch der ökonomischen Analyse des Zivilrechts, S. XVIII; auch *Maamar*, Computer als Schöpfer, S. 138.

105 Ähnlich auch *Fezer*, JZ 1988, 223 (228).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

auf die Innovations- und Verbreitungsanreize für Urheber, die Höhe der noch verbleibenden Urheberrente insgesamt oder ein sektoral hervorgerufenes Marktversagen ableiten. Es fehlt insbesondere auch an einer entsprechender Datenlage.[106] Eine volkswirtschaftliche Prognose, ob beispielsweise die Freistellung des Web Scrapings von Werken als Trainingsdaten spezifisch den Anreiz für Urheber verringert, neue Werke zu schaffen oder die Urheberrente erheblich mindert, lässt sich demzufolge nicht abgeben. Vorhersagen zu Auswirkungen auf Indikatoren, die die ökonomische Analyse der urheberrechtlichen Rechtsnormen maßgeblich mitprägen, sind in den hier betreffenden Fällen damit nicht möglich.

In diesem Zusammenhang muss auch berücksichtigt werden, dass leistungsfähige KI-Systeme erst seit wenigen Jahren auf dem Massenmarkt zur Verfügung stehen.[107] Es sind noch keine empirischen Untersuchungen betreffend die ökonomischen Auswirkungen von urheberrechtlichen Regelungen auf die Erreichung der Allokationseffizienz im Kontext des Einsatzes von KI durchgeführt worden.[108] Ohne eine empirische Datengrundlage ist

106 Allein eine im Jahr 2015 veröffentlichte Studie deutet an, dass die konkrete Ausgestaltung des urheberrechtlichen Schutzes keine Auswirkungen auf die Schaffenskraft der Urheber haben könnte und demzufolge überhaupt keinen Innovationsanreiz ausübe. Hierüber abschließende Aussagen zu treffen, sei mangels ausreichender Datenlage aber noch nicht möglich gewesen. Siehe *Handke/Girard/Mattes*, Fördert das Urheberrecht Innovation? Eine empirische Untersuchung, S. 105, 121.

107 Bildklassifikationsfähigkeiten von KI-Systemen haben sich in den letzten zehn Jahren um 27,7 % verbessert (*Maslej u. a.*, The AI Index 2023 Annual Report by Stanford University, S. 81.), Fähigkeiten der Objekterkennung um 26 % seit 2015 (*Maslej u. a.*, The AI Index 2023 Annual Report by Stanford University, S. 88.), der beste Q→AR-Score, ein Maßstab zur Bestimmung der Leistungsfähigkeit von Systemen für optisches Bildverständnis, hat sich seit 2018 um 66 % erhöht (*Maslej u. a.*, The AI Index 2023 Annual Report by Stanford University, S. 95.), der SuperGLUE-Score zur Bestimmung der Leistungsfähigkeit von KI-Systemen zum Verständnis englischer Sprache um 8 % seit 2019, außerdem wurde in diesem Bereich im Jahr 2020 die Human Baseline, die durchschnittliche menschliche Leistungsfähigkeit, überschritten (*Maslej u. a.*, The AI Index 2023 Annual Report by Stanford University, S. 100.), die Genauigkeit des Leseverständnisses von KI-Systemen nahm seit 2020 um rund 17 % zu (*Maslej u. a.*, The AI Index 2023 Annual Report by Stanford University, S. 101.), die Fehlerquote von Spracherkennungssystemen verbesserte sich seit 2017 von über 7 % auf 0,14 % (*Maslej u. a.*, The AI Index 2023 Annual Report by Stanford University, S. 109.), im Bereich der verwendeten Rechenhardware verdreifachte sich seit 2021 zugleich der Median der FLOP/s-Geschwindigkeit (*Maslej u. a.*, The AI Index 2023 Annual Report by Stanford University, S. 118.).

108 Das ergibt sich aus zwei systematischen Auswertungen der rechtsökonomischen Forschung, siehe *Peukert/Windisch*, Center for Law & Economics Working Paper Series ETH Zürich 2023; *Margoni*, 21 for 2021: Computational Uses – CREATe. Au-

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

die rechtsökonomische Analyse der aufgeworfenen Fragestellungen damit aber nicht belastbar.[109]

Zwar könnte dem entgegengehalten werden, eine spätere Änderung der zur Analyse genutzten Hypothesen und Prognosen beeinträchtige nicht die Tauglichkeit der ökonomischen Analyse des Rechts als Instrument für die Entscheidungsfindung als solche.[110] Vielmehr könnten die Analyseergebnisse bei einer Änderung der zugrundeliegenden empirischen Grundannahmen in der Zukunft angepasst werden.[111] Allerdings ist Ziel dieser Arbeit, bereits im Zeitpunkt der Untersuchung tragfähige Antworten auf die regulatorischen Fragestellung herauszuarbeiten. Eine Verweis auf die Reversibilität der regulatorischen Erwägungen ist in diesem Rahmen damit nicht ausreichend. Insgesamt ist die ökonomische Analyse des Urheberrechts mithin kein Instrument, welches im Rahmen dieser Untersuchung dafür geeignet ist, regulatorische Freistellungsbedürfnisse zu identifizieren.[112]

Weil das Effizienzziel auf die Maximierung der Wohlfahrt aller Gesellschaftsmitglieder gerichtet ist, erfüllt es aber dennoch die Voraussetzungen aus der Sozialbindung des geistigen Eigentums gemäß Art. 14 Abs. 1 S. 2, Abs. 2 GG. Damit kann es dennoch als Allgemeininteresse im Rahmen der freien, umfassenden Interessenabwägung berücksichtigt werden. Das gilt auch für das Coase Theorem, nach dem geringe beziehungsweise keine Transaktionskosten für die Übertragung von Gütern auf einem Markt dazu führen, dass dieser sich wohlfahrtsökonomisch autonom reguliert.[113]

ßerdem wurde zuletzt am 9.10.2024 eine ergebnislose Stichwortsuche zu „Artificial Intelligence" und „Künstliche Intelligenz" bei *copyright evidence* durchgeführt, einer enzyklopädischen Datenbank von empirischen Studien zum Urheberrecht, siehe https://www.copyrightevidence.org (zuletzt abgerufen am 9.10.2024).

109 Ähnlich auch *Hennemann*, Interaktion und Partizipation, S. 317; grundsätzlich auch *Maamar*, Computer als Schöpfer, S. 93, der im Anschluss aber dennoch eine rechtsökonomische Analyse vornimmt.

110 *Maamar*, Computer als Schöpfer, S. 93.

111 *Maamar*, Computer als Schöpfer, S. 93.

112 Anders *Maamar*, Computer als Schöpfer, S. 93.

113 *Coase*, The Journal of Law and Economics 1960, 1 (10, 15 f.); mit Beispielen *Schäfer/Ott*, Lehrbuch der ökonomischen Analyse des Zivilrechts, S. 80 ff.; zu einer kritischen Einordnung *Eidenmüller*, Effizienz als Rechtsprinzip, S. 91 ff.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

4. Techniksensibilität des Urheberrechts oder der „more technological approach"

Fraglich ist schließlich, inwiefern auch technische Notwendigkeiten beim Einsatz von KI im Rahmen der regulatorischen Analyse berücksichtigt werden können.

Zum einen könnte in diesem Zusammenhang ein *more technological approach*[114] genommen werden. Hiernach muss das Urheberrecht besonders für die systemischen Notwendigkeiten beim Einsatz von Technik aufgeschlossen sein.[115] Nutzungen urheberrechtlich geschützter Werke, die beim Einsatz von KI aus einer technischen Perspektive notwendig oder jedenfalls unvermeidlich sind, sollten demnach ermöglicht werden, insofern der Einsatz von KI als solches förderungswürdig ist. Neben der Freistellung aufgrund einer fehlenden Beeinträchtigung der Urheberinteressen oder einer umfassenden Abwägung der Urheber- und Nutzerinteressen wäre dies demzufolge ein dritter Ansatzpunkt, um Freistellungsbedürfnisse für die zu untersuchenden Fallgruppen zu identifizieren.

Gegen einen solchen *more technological approach* spricht allerdings, dass Aufgabe des Urheberrechts nicht primär die Steuerung des Einsatzes von Technik, sondern der Schutz der Urheber und ihrer Werke sowie der Ausgleich von Urheber- mit Nutzerinteressen ist.[116] Auch wenn die Reflexion technischer Möglichkeiten und Notwendigkeiten im urheberrechtlichen Kontext notwendig ist, können aus technischen Abläufen nicht unmittelbar normative Entscheidungen abgeleitet werden.[117]

Nichtsdestotrotz muss das Urheberrecht aber techniksensibel ausgestaltet und angewendet werden.[118] Das bedeutet, dass technische Besonderheiten bei der regulatorischen Ausgestaltung ebenfalls mitzuberücksichtigen sind.[119] In die umfassende Abwägung von Interessenabwägung können demzufolge auch technische Bedürfnisse mit einfließen.[120]

---

114 *Podszun/Grünberger*, ZGE 2014, 269.
115 *Podszun/Grünberger*, ZGE 2014, 269 (270).
116 *Hofmann*, ZGE 2016, 482 (501).
117 *Godt*, ZGE 2014, 279 (281); sich anschließend auch *Hofmann*, ZGE 2016, 482 (511).
118 *Hofmann*, ZGE 2016, 482 (511); *de la Durantaye/Hofmann*, ZUM 2021, 873 (877).
119 *Becker*, ZGE 2016, 239 (267); *Hofmann*, ZGE 2016, 482 (511); *de la Durantaye/Hofmann*, ZUM 2021, 873 (877); so schlussendlich auch *Grünberger*, ZUM 2015, 273 (275 f.).
120 *Becker*, ZGE 2016, 239 (267).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

5. Zusammenfassung: Maßstab für die regulatorische Analyse

Alles in allem sollten die zu untersuchenden Fallgruppen, also das Web Scraping von Werken zum Training von KI, die eigentliche Durchführung des Trainingsvorgangs auf Basis von Werken, die Speicherung urheberrechtlich geschützter Inputdaten für die Anwendung von KI-Systemen sowie die Vorverarbeitung der als Inputdaten erfassten Werke vom Ausschließlichkeitsrecht betroffener Urheber ausgenommen sein, wenn die Freistellung regulatorisch zweckmäßig ist. Das ist zum einen der Fall, wenn in den jeweiligen Sachverhaltskonstellationen keine Urheberinteressen berührt werden. Zum anderen ist eine Freistellung aber auch dann zweckmäßig, wenn die Nutzer- die betroffenen Urheberinteressen im Rahmen einer umfassenden Abwägung überwiegen. Hierbei können auch ökonomische, wohlfahrtsorientierte Aspekte berücksichtigt werden.

II. Bestimmung des Freistellungsmechanismus

Für die Umsetzung eines identifizierten Freistellungsbedürfnisses kommt es weiterhin darauf an, welche geeigneten Freistellungsmechanismen im urheberrechtlichen Regelungsregime überhaupt zur Verfügung stehen, welche Wirkungen und Unterschiede ihnen mit Blick auf die betroffenen Urheber jeweils zukommen und anhand welcher Beurteilungskriterien der interessengerechte Regelungsmechanismus identifiziert werden kann.

1. Zur Verfügung stehende Freistellungsmechanismen

Damit ein Regelungsmechanismus für eine Freistellung im Sinne der vorliegenden Untersuchung in Betracht kommt, muss er Trainingsdatensammlern, KI-Entwicklern oder Anwendern von KI-Systemen ermöglichen, die betroffenen Werke ohne individuelle Zustimmung der Urheber zu verwenden.

a) Unvergütete Schutzbereichsbegrenzung: Unmittelbar begrenzende Ausgestaltung der Verwertungsrechte und vergütungsfreie Schrankenbestimmung

Möglich ist dabei zum einen, die Freistellung ohne angemessene Vergütung betroffener Urheber auf Ebene des urheberrechtlichen Schutzbereichs

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

umzusetzen. Folge einer solchen Begrenzung wäre, dass die betreffende Sachverhaltskonstellation wie beispielsweise das Web Scraping von Werken zum Training von KI nicht vom Ausschließlichkeitsrecht der Urheber umfasst würde. Damit bestünde kein Verbotsrecht der Urheber, sodass die jeweilige Handlung erlaubnisfrei vorgenommen werden dürfte. Individuelle Nutzungsrechte einzuholen, wäre demzufolge nicht mehr nötig. Eine vergütungsfreie Begrenzung des urheberrechtlichen Schutzbereichs kann dabei einmal mit Hilfe einer vergütungsfreien Schrankenbestimmung (§§ 44a ff. UrhG) oder durch eine unmittelbar begrenzende Ausgestaltung der Verwertungsrechte in den §§ 15 ff. UrhG erfolgen.

b) Gesetzliche Lizenz

Eine Freistellung könnte darüber hinaus auch im Wege einer gesetzlichen Lizenz umgesetzt werden. Hierbei handelt es sich um eine Art von Schrankenbestimmung, durch die die betreffende Werknutzung zwar freigestellt, zugunsten der Urheber allerdings vergütungspflichtig ausgestaltet wird. Nutzer werden damit in die Lage versetzt, die gewünschten Werke beispielsweise zum Training ihrer KI zu speichern, ohne Individuallizenzen einholen zu müssen. Trotzdem müssen sie den Urhebern im Gegenzug eine dem wirtschaftlichen Wert der Nutzung entsprechende angemessene Vergütung gewähren.[121]

Die Bemessung dieser Vergütung kann pauschalisiert erfolgen, wenn eine genauere Abrechnung mit unzumutbarem Aufwand oder zu hohen Transaktionskosten verbunden oder sogar unmöglich wäre.[122] Insbesondere Trainingsdatensammler speichern mit Hilfe von Webcrawlern automatisiert große Datenmengen. Dabei haben sie zumeist keinen genauen Überblick über die gespeicherten Inhalte. Auch eine manuelle nachträgliche Durchsicht ist mit Rücksicht auf die Größe typischer Trainingsdatensets unzumutbar.[123] Eine gesetzliche Lizenz für das Web Scraping von Werken als Trainingsdaten könnte damit insbesondere mit einem pauschal veran-

121 Zur Vergütungshöhe: *Pflüger*, Gerechter Ausgleich und angemessene Vergütung, S. 145; insbesondere zu anderen bei der Bestimmung der Vergütungshöhe berücksichtigungsfähigen Aspekten: *Lauber-Rönsberg*, Urheberrecht und Privatgebrauch, S. 112 ff.

122 So z. B. bei der Vergütung zur Nutzung von Werken als Unterrichts- und Lehrmedien, § 60a UrhG, siehe *Hagemeier*, in: Ahlberg/Götting/Lauber-Rönsberg, Beck'scher Online-Kommentar Urheberrecht, UrhG § 60h Rn. 10.

123 Vgl. zur Größe der Trainingsdatensätze oben unter 1. Teil B. II. 3.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

schlagten gesetzlichen Vergütungsanspruch kombiniert werden, der über Verwertungsgesellschaften abgewickelt wird.

c) Erweiterte kollektive Lizenz

Eine Freistellung der zu untersuchenden Sachverhaltskonstellationen ist darüber hinaus auch über einen erweiterten kollektiven Lizenzmechanismus (ECL) möglich.[124] Hierbei handelt es sich um einen Lizenzmechanismus, der Verwertungsgesellschaften ermöglicht, Lizenzen auch zur Nutzung solcher Werke zu erteilen, die nicht Teil ihres Repertoires sind.[125] Bei Umsetzung der Freistellung im Wege eines ECL-Mechanismus müssten Nutzer wie Trainingsdatensammler zwar immer noch eine Lizenz für die Nutzung von Werken beispielsweise zum Training von KI abschließen. Da die Lizenzvereinbarung aber kollektiv mit einer Verwertungsgesellschaft getroffen und für eine unbestimmte Zahl von Werken einer Vielzahl von Urheber gelten würde, erfüllt der ECL-Mechanismus dennoch die Voraussetzungen, die an eine Freistellung im Sinne der vorliegenden Untersuchung gestellt werden.

ECL können beispielsweise durch Kulturerbe-Einrichtungen wie öffentlichen Bibliotheken zur Nutzung nicht verfügbarer Werke gemäß § 52 VGG abgeschlossen werden. Außenstehende Urheber, deren Werke von der erweiterten Wirkung der Lizenzvereinbarung betroffen sind, haben in Folge einer ECL die gleichen Rechte und Pflichten wie solche Urheber, die mit der Verwertungsgesellschaft Wahrnehmungsverträge geschlossen haben. Das betrifft insbesondere auch die Verteilung der Einnahmen aus der Verwertung der Werke (vgl. § 23 VGG). Damit eine Lizenzvereinbarung von der erweiterten Wirkung erfasst wird, müssen aber von der Verwertungsgesellschaft und Lizenznehmern einige Voraussetzungen erfüllt werden. Um ECL-Vereinbarungen abzuschließen, bedarf es insbesondere einer repräsentativen Verwertungsgesellschaft (vgl. § 51b VGG). Nur ihre Tarife haben sich als so angemessen erwiesen, dass es gerechtfertigt ist, sie auch auf Werke außenstehender Rechtsinhaber anzuwenden.[126]

Um ein Leerlaufen der Freistellungsregelung zu verhindern, wenn keine repräsentative Verwertungsgesellschaft besteht, muss der ECL-Mechanismus mit einer subsidiär eingreifenden Schrankenbestimmung verbunden

124 In diese Richtung auch *Raue*, ZUM 2021, 793 (796).

125 Zum Mechanismus genauer *Conrad*, ZUM 2022, 436; *Staats*, ZUM 2019, 703; *Staats*, ZUM 2020, 728; umfassend außerdem *Trumpke*, Exklusivität und Kollektivierung.

126 *Bergström*, GRUR Ausl 1962, 364 (372).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

werden. Ein entsprechender Mechanismus ist in den §§ 52 ff. VGG i. V. m. § 61d UrhG auch für die Nutzung vergriffener Werke vorgesehen. Nutzungen auf Basis der Schrankenbestimmung des § 61d UrhG erfolgen dabei vergütungsfrei.

d) Zusammenfassung

Im Ergebnis stehen zur Freistellung des Web Scrapings von urheberrechtlich geschützten Trainingsdaten, der eigentlichen Durchführung des KI-Trainings, der Speicherung von Werken als Inputdaten für die Anwendung von KI-Systemen sowie der Vorverarbeitung dieser urheberrechtlich geschützten Inputdaten vier Freistellungsmechanismen zur Verfügung. Neben einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte und einer unvergüteten Schrankenbestimmung wäre eine Freistellung auch im Wege einer vergütungspflichtigen Schrankenbestimmung (gesetzliche Lizenz) und mit über einen ECL-Mechanismus möglich.

2. Entscheidungskriterien bei der Wahl zwischen den Freistellungsmechanismen

Fraglich ist weiterhin, welche Entscheidungskriterien die Wahl zwischen den zur Verfügung stehenden Freistellungsmechanismen beeinflussen.

a) Angemessener Ausgleich für die Nutzungshandlung

Als Grundlage der Entscheidung dient zum einen, dass zwei der vier Freistellungsmechanismen (gesetzliche Lizenz, erweiterte kollektive Lizenz) eine (teilweise) Vergütung der betroffenen Urheber vorsehen. Die Frage danach, ob den betroffenen Urhebern für die in der jeweiligen Sachverhaltskonstellation maßgebliche Handlung wie beispielsweise der Speicherung ihrer Werke zum KI-Training eine Vergütung zustehen soll, kann damit als erstes Kriterium zur Wahl des geeigneten Freistellungsmechanismus herangezogen werden. Grundsätzlich sollte dem Urheber dabei eine Vergütung für die Freistellung zuerkannt werden, wenn nicht ein gesteigertes öffentli-

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

ches Interesse ausnahmsweise eine vergütungsfreie Nutzung rechtfertigt.[127] Ob das der Fall ist, muss anhand einer umfassenden Abwägung ermittelt werden.[128] Insbesondere, aber nicht nur ausgleichsbedürftig sind demnach Einschränkungen der für die Urheber besonders wirtschaftlich relevanten Kernverwertungsmöglichkeiten.[129] Insgesamt genügt es aber, wenn die Auslegung und Anwendung einer Freistellungsregelung genügend Spielräume lässt, über die die Vergütungsinteressen betroffener Urheber angemessen berücksichtigt werden können.[130] Ausreichend ist weiterhin, wenn „unter dem Strich"[131] eine angemessene Vergütung der Urheber für die Schöpfungsleistung gewährleistet bleibt.[132]

b) Darlegungs- und Beweislast bei vergütungsfreier Nutzung

Ist eine Vergütung der Urheber hingegen nicht nötig, stehen für die Freistellung eine vergütungsfreie Schrankenbestimmung sowie eine unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte zur Verfügung. Eine Differenzierung zwischen beiden Freistellungsmechanismen könnte zum einen anhand dar Darlegungs- und Beweislastverteilung vorgenommen werden, die jeweils aus den Freistellungsmechanismen resultiert. Hierfür ist notwendig, dass die Freistellungen im Wege einer Schrankenbestimmung einerseits und einer unmittelbar begrenzenden Ausgestaltung der Verwertungsrechte andererseits hinsichtlich der Darlegungs- und Beweislast unterschiedliche Wirkung entfalten.

Gegen eine unterschiedliche Wirkung von Schrankenbestimmung und unmittelbar begrenzender Ausgestaltung der Verwertungsrechte in Bezug auf die herbeigeführte Darlegungs- und Beweislastverteilung spricht, dass

127 BVerfG, Beschl. v. 25.10.1978 - 1 BvR 352/71, GRUR 1980, 44 (48) – Kirchenmusik; vgl. auch BVerfG, Urt. v. 11.10.1988 - 1 BvR 743/86 u. a., NJW 1992, 1307 (1308) – Justizvollzugsanstalten.

128 Vgl. *Lauber-Rönsberg*, Urheberrecht und Privatgebrauch, S. 112.

129 *Lauber-Rönsberg*, Urheberrecht und Privatgebrauch, S. 112.

130 BVerfG, Urt. v. 31.05.2016 - 1 BvR 1585/13, GRUR 2016, 690 Rn. 79 f. – Metall auf Metall.

131 BVerfG, Urt. v. 31.05.2016 - 1 BvR 1585/13, GRUR 2016, 690 Rn. 79 f. – Metall auf Metall.

132 BVerfG, Beschl. v. 07.07.1971 - 1 BvR 765/66, GRUR 1972, 481 (483) – Kirchen- und Schulgebrauch; BVerfG, Beschl. v. 25.10.1978 - 1 BvR 352/71, GRUR 1980, 44 (46)– Kirchenmusik; BVerfG, Urt. v. 31.05.2016 - 1 BvR 1585/13, GRUR 2016, 690 – Rn. 80 - Metall auf Metall.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

beide Freistellungsmechanismen das Ausschließlichkeitsrecht der Urheber gesetzlich ausgestalten und damit vom Zeitpunkt seiner Entstehung an in seiner Gestalt formen.[133] Sowohl die Schrankenbestimmungen als auch die begrenzende Ausgestaltung der urheberrechtlichen Verwertungsrechte wirken also tatbestandlich auf das Ausschließlichkeitsrecht der Urheber ein.[134] Demzufolge besteht auf einer materiell-rechtlichen Ebene grundsätzlich Wirkungsgleichheit zwischen beiden Rechtsinstituten. Das könnte eine weitere Differenzierung überflüssig machen.[135]

Neben der materiell-rechtlichen Ausgestaltung ist aber auch die Durchsetzung der Freistellungsregelungen zu berücksichtigen. Nach einem zivilprozessualen Grundsatz hat hierbei jede Partei die für sie günstigen Umstände zu beweisen.[136] In Bezug auf das Urheberrecht bedeutet das für Urheber, dass sie darlegen und beweisen müssen, dass ihre Rechtsposition überhaupt betroffen ist. Damit sind sie für Tatsachen betreffend die Werkqualität des streitgegenständlichen Inhalts[137] und darüber hinaus auch dafür beweispflichtig, dass ihre Verwertungsrechte berührt sind.[138] Hieraus folgt aber auch, dass Urheber bei einer auf Ebene der Verwertungsrechte durchgeführten Freistellung grundsätzlich beweispflichtig dafür wären, dass die streitige Handlung nicht in den Anwendungsbereich der Freistellungsregelung und damit in den Schutzbereich des Ausschließlichkeitsrechts fällt. Im Gegensatz hierzu trägt der Nutzer die Beweislast

133 Zur Rechtsnatur der urheberrechtlichen Schrankenbestimmungen *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 129 ff., insb. S. 131, 149 f.

134 *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 149 f.; aA. nur vereinzelt insbesondere in der Rechtsprechung, siehe beispielhaft LG Köln, Urt. v. 13.08.2020 - 14 O 77/19, GRUR-RR 2020, 482 Rn. 40 – Feindliche Übernahme.

135 Ähnlich das Bundesverfassungsgericht, welches von einer bloßen „Frage der Gesetzestechnik" spricht, siehe BVerfG, Beschl. v. 25.10.1978 - 1 BvR 352/71, GRUR 1980, 44 (45) – Kirchenmusik.

136 *Prütting*, in: Krüger/Rauscher, Münchener Kommentar zur ZPO, ZPO § 286 Rn. 114 f.

137 So beispielsweise BGH, Urt. v. 27.05.1981 - I ZR 102/79, GRUR 1981, 820 (822) – Stahlrohrstuhl II.

138 Beispielsweise zur öffentlichen Wiedergabe OLG Frankfurt a. M., Urt. v. 20.01.2015 - 11 U 95/14, ZUM-RD 2015, 296 (298); *v. Ungern-Sternberg*, in: Schricker/Loewenheim, Urheberrecht, UrhG § 15 Rn. 405 m. w. N.; zur Wiedererkennbarkeit eines Werkes innerhalb eines derivativen Erzeugnisses *Leistner*, GRUR 2019, 1008 (1010).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

für das Vorliegen einer für ihn günstigen Schrankenbestimmung.[139] Demzufolge hat die Ausgestaltung des Freistellungsmechanismus Auswirkungen darauf, welchem der Beteiligten die Darlegungs- und Beweislast für die Freistellungsvoraussetzungen obliegt.[140]

Dies gilt selbst dann, wenn einer nicht beweisbelasteten Partei eine sekundäre Darlegungslast auferlegt wird. Denn eine Pflicht zum vollständigen Gegenbeweis ergibt sich aus der sekundären Darlegungslast nicht.[141] Vielmehr sind substantiierte Ausführungen ausreichend, die darlegen, warum die vom Prozessgegner (hier: Urheber) vorgetragenen Tatsachen nicht vorliegen können.[142]

Alles in allem kann die Verteilung der Darlegungs- und Beweislast damit zur Wahl zwischen einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte und der Freistellung mittels einer unvergüteten Schrankenbestimmung herangezogen werden. Anzumerken bleibt allerdings, dass die gezielte Verteilung der Darlegungs- und Beweislast im Kontext konkreter Regelungsentwürfe nur über präzise, beweislastsensible Normausgestaltungen verwirklicht werden kann. Erweist es sich beispielsweise als interessengerecht, den Urheber für das Nicht-Vorliegen einer Freistellungsvoraussetzung darlegungs- und beweispflichtig zu machen, darf die Freistellungsregelung nicht in Form einer Regel-Ausnahme-Formulierung („dies gilt nicht, wenn"; „es sei denn") umgesetzt werden. Denn diese würde zu einer Umkehrung der Beweislast zurück auf den Werknutzer führen.

c) Schutz eingesetzter technischer Schutzmaßnahmen via § 95a UrhG

Neben der Darlegungs- und Beweislast könnte bei der Wahl zwischen einer Freistellung durch eine unmittelbar begrenzende Ausgestaltung der urheberrechtlichen Verwertungsrechte oder eine urheberrechtliche Schrankenbestimmung auch darauf abgestellt werden, ob technische Schutzmaßnahmen gegen die betreffende Verwendung der Werke von § 95a UrhG geschützt werden sollen. Denn ein solcher Schutz über § 95a UrhG könnte

139 St. Rspr., siehe bspw. BGH, Urt. v. 30.03.2017 - I ZR 19/16, GRUR 2017, 1233 Rn. 15 – Loud; *Leistner*, in: Schricker/Loewenheim, Urheberrecht, UrhG § 97 Rn. 28; *Flechsig*, GRUR 1993, 532 (536).

140 Zu dieser Verschiebung grundsätzlich schon *Dreier*, GRUR Int. 2015, 648 (650).

141 BGH, Urt. v. 30.03.2017 - I ZR 19/16, GRUR 2017, 1233 Rn. 15 – Loud.

142 Vgl. *Stadler*, in: Musielak/Voit, ZPO, ZPO § 138 Rn. 10a.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

voraussetzen, dass die technische Schutzmaßnahme auf die Abwehr von Handlungen abzielt, die als Teil der urheberrechtlichen Verwertungsrechte im Sinne der §§ 15 ff. UrhG dem Ausschließlichkeitsrecht der Urheber zugewiesen sind.

Hiergegen – und im Umkehrschluss für einen weiten, nicht an die §§ 15 ff. UrhG geknüpften Schutzbereich des § 95a UrhG – spricht zunächst, dass sich technische Maßnahmen nach dem Wortlaut des § 95a Abs. 2 S. 1 UrhG nicht gegen *Werknutzungen*, sondern nur gegen *Handlungen* richten müssen, die vom Rechtsinhaber nicht genehmigt worden sind. Zudem bezieht die Regelung in § 95a Abs. 2 S. 2 UrhG ausdrücklich auch zugangshindernde Maßnahmen ein.[143] Der Zugang zum Werk (und die anschließende Rezeption des Werkes) sind aber für sich genommen nicht Teil des urheberrechtlichen Ausschließlichkeitsrechts. Von einem über die urheberrechtlichen Nutzungsrechte hinausgehenden Verständnis des § 95a UrhG, welches auch den urheberrechtlich nicht relevanten Werkzugang erfasst, ging auch der historische Gesetzgeber bei der Einführung des § 95a UrhG aus.[144]

Allerdings ist die Verwendung des Rechtsbegriffs „Handlung" darauf zurückzuführen, dass der historische Gesetzgeber eine möglichst richtlinienahe Umsetzung anstrebte.[145] Im § 95a UrhG zugrundeliegenden Art. 6 InfoSoc-RL ist von Maßnahmen gegen *Handlungen* die Rede, die vom Rechtsinhaber nicht genehmigt worden sind. Beispielsweise in ErwGr. 16 InfoSoc-RL ist jedoch von einer „Haftung für Handlungen im Netzwerk-Umfeld" die Rede. Die *Haftung* kann aber nur entstehen, wenn ein Haftungstatbestand erfüllt ist. Hierfür bedarf es einer urheberrechtlich relevanten Nutzung. Dies spricht dafür, dass der Handlungsbegriff im hier verwendeten Kontext nur urheberrechtlich relevante Nutzungshandlungen umfasst. Nach Art. 3 Abs. 3 InfoSoc-RL gilt der Erschöpfungsgrundsatz zudem nicht für *Handlungen der öffentlichen Wiedergabe*, womit eine dem Urheber zugewiesene Werknutzung gemeint sein muss. Beides zeigt, dass der Handlungsbegriff innerhalb der InfoSoc-RL synonym zur urheberrechtlich relevanten Werknutzung zu verstehen ist. Daraus, dass das Gesetz den Schutz technischer Schutzmaßnahmen nach § 95a Abs. 2 S. 1 UrhG auf Maßnahmen erstreckt, die sich gegen *Handlungen* richten, die vom Urheber nicht genehmigt worden sind, kann demzufolge nicht geschlossen

143 *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 462.
144 BT-Drs. 15/38, S. 26.
145 BT-Drs. 15/38, S. 26.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

werden, dass § 95a UrhG einen über die urheberrechtlichen Verwertungsrechte hinausgehenden Schutzbereich eröffnet.

Im Gegenteil verfolgt die InfoSoc-RL das Ziel, ein hohes Schutzniveau für Urheber zu gewährleisten. Dieses Schutzziel ist auch in Art. 6 Abs. 3 S. 2 InfoSoc-RL verankert. Zweck des hohen Schutzniveaus ist, dass Urhebern eine angemessene Vergütung zuteilwird, auf Basis derer sie weiterhin schöpferisch tätig sein können.[146] Eine solche Vergütungsleistung ist aber nur für urheberrechtlich relevante Werknutzungen gerechtfertigt. Denn nur sie berühren die ausschließlichen Verwertungsrechte, welche den Urhebern insbesondere durch Art. 2 bis 4 InfoSoc-RL zur Verwirklichung des Schutzniveaus zugewiesen worden sind.

Darüber hinaus müssen technische Maßnahmen, die von § 95a UrhG geschützt werden sollen, nach dem Wortlaut der Regelung solche Handlungen verhindern oder erschweren, die vom Urheber nicht genehmigt worden sind (Art. 6 Abs. 3 S. 1 InfoSoc-RL respektive § 95a Abs. 2 S. 1 UrhG). Dies setzt denklogisch aber eine Zustimmungsbedürftigkeit durch den Urheber voraus. Diese kann nur dann vorliegen, wenn eine urheberrechtlich relevante Werknutzung stattfindet oder jedenfalls stattfinden soll.[147]

Alles in allem sprechen sowohl teleologische als auch systematische und grammatikalische Gründe dafür, dass der Umgehungsschutz aus § 95a UrhG nur gegen solche Nutzerhandlungen geltend gemacht werden kann, mittels derer ein Werk in einer Art und Weise genutzt wird, die dem Urheber über die Verwertungsrechte aus §§ 15 ff. UrhG zugewiesen worden ist.[148] Der Umgehungsschutz ist außerdem auf vorgelagerte[149] Handlungen

146 ErwGr. 4, 9, 10 InfoSoc-RL.

147 *Wandtke/Ohst*, in: Wandtke/Bullinger, Urheberrecht, UrhG § 95a Rn. 43; *Arlt*, in: Hoeren/Sieber/Holznagel, Handbuch Multimedia-Recht, Teil 7.7 (Schutz und Identifizierung durch technische Schutzmaßnahmen) Rn. 25; *Peukert/Leistner*, in: Loewenheim, Handbuch des Urheberrechts, § 40 (Schutz technischer Maßnahmen) Rn. 3; ähnlich wohl *Rieber-Mohn*, IIC 2006, 182 (186 f.), wobei der Autor ein vertragliches Zustimmungsrecht als ausreichend erachtet; a. A. *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 463, nach dem es für einen Schutz über § 95a UrhG ausreicht, wenn jedenfalls auch urheberrechtlich relevante Nutzungshandlungen erfasst werden.

148 EuGH, Urt. v. 23.01.2014 - C-355/12, GRUR 2014, 255 Rn. 25, 31 – Nintendo/PC Box; *Czychowski*, in: Fromm/Nordemann, Urheberrecht, UrhG § 95a Rn. 14; *Götting*, in: Schricker/Loewenheim, Urheberrecht, UrhG § 95a Rn. 45; *Wandtke/Ohst*, in: Wandtke/Bullinger, Urheberrecht, UrhG § 95a Rn. 43; *Arlt*, in: Hoeren/Sieber/Holznagel, Handbuch Multimedia-Recht, Teil 7.7 (Schutz und Identifizierung durch technische Schutzmaßnahmen) Rn. 25; *Peukert/Leistner*, in: Loewenheim, Handbuch des Urheberrechts, § 40 (Schutz technischer Maßnahmen) Rn. 3; *Specht-*

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

zu erstrecken, mit denen das Ziel verfolgt wird, das betreffende Werk in urheberrechtlich relevanter Weise zu nutzen.[150] Hierzu gehört auch der Zugang zum Werk, der in Art. 6 Abs. 3 S. 2 InfoSoc-RL respektive § 95a Abs. 2 S. 2 UrhG ausdrücklich mitumfasst wird.[151] Umgehungen zur Ermöglichung urheberrechtlich freier Verwendungen eines Werkes können über § 95a UrhG hingegen nicht untersagt werden. Besteht ein Bedürfnis danach, technische Schutzmaßnahmen über § 95a UrhG zu schützen, die gegen die jeweilige Verwendung von Werken in den zu untersuchenden Sachverhaltskonstellationen wie beispielsweise dem Web Scraping von Werken als KI-Trainingsdaten eingesetzt worden sind, spricht dies infolgedessen dafür, dass die Freistellung durch eine urheberrechtliche Schrankenbestimmung umgesetzt sein sollte.

d) Notwendigkeit eines Opt-out-Mechanismus

Sollen Urheber als Ausgleich für eine freigestellte Werknutzung eine Vergütung erhalten, kann die Freistellung im Wege einer gesetzlichen Lizenz oder mit Hilfe eines ECL-Mechanismus umgesetzt werden. Fraglich ist vor diesem Hintergrund aber, welche Unterschiede sich im Weiteren zwischen den beiden Rechtsinstituten ergeben, die eine Wahl des besser geeigneten Freistellungsmechanismus möglich machen.

Zum einen könnte die Notwendigkeit eines Opt-out-Mechanismus, den betroffener Urheber auslösen können, Einfluss darauf haben, welcher der beiden Regelungsmechanismen für eine Freistellung gewählt werden sollte. Unter einem Opt-out-Mechanismus wird dabei eine Regelung verstanden,

---

Riemenschneider, Diktat der Technik, S. 199, 203; ähnlich wohl *Rieber-Mohn*, IIC 2006, 182 (186 f.), wobei der Autor ein vertragliches Zustimmungsrecht als ausreichend erachtet; BT-Drs. 15/38, S. 26; a. A. *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 463, nach dem es für einen Schutz über § 95a UrhG ausreicht, wenn jedenfalls auch urheberrechtlich relevante Nutzungshandlungen erfasst werden; ebenso im norwegischen Urheberrecht, siehe dazu *Rieber-Mohn*, IIC 2006, 182 (191).

149 Für den Werkzugang *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 462 f.

150 Vgl. EuGH, Urt. v. 23.01.2014 - C-355/12, GRUR 2014, 255 Rn. 25, 31 – Nintendo/PC Box.

151 A. A. *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 462. Nach ihm spricht die Erfassung des Werkzugangs dafür, dass insgesamt auch urheberrechtlich freie Handlungen erfasst sind, solange die technische Schutzmaßnahme umgangen wird.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

die es Urhebern ermöglicht, die Freistellungswirkung durch einseitige Erklärung zu beenden. Ein Opt-out-Mechanismus kann dabei grundsätzlich sowohl im Rahmen gesetzlicher Lizenzen[152] als auch im Rahmen eines ECL-Mechanismus[153] vorgesehen werden.

Üblicherweise werden gesetzliche Vergütungsansprüche für Massennutzungen verwertungsgesellschaftenpflichtig ausgestaltet. Zu diesen Massennutzungen gehört auch das Web Scraping von Werken als Trainingsdaten für KI.[154] Sollten gesetzliche Lizenzen mit einem Opt-out-Mechanismus verknüpft werden, müsste die Opt-out-Erklärung allerdings unmittelbar am Werkstück, beispielsweise in den Metadaten einer Bilddatei, abgegeben werden. Denn eine Abwicklung der Nutzungshandlungen selbst über die Verwertungsgesellschaften, beispielsweise durch Abschluss eines Lizenzvertrags, findet nicht statt. Die Nutzung wäre vielmehr eo ipso erlaubnisfrei gestellt. Damit bestünde allerdings die Gefahr, dass Urheber von der Verwertungsgesellschaft eine aufgrund der Massennutzung pauschalisierte Ausschüttung erhielten, obwohl sie eine Vorbehaltserklärung zur Beendigung der Freistellungswirkung am Werkstück abgegeben haben.

Im Rahmen des erweiterten kollektiven Lizenzmechanismus wird der Opt-out hingegen direkt gegenüber der lizenzierenden Verwertungsgesellschaft erklärt, die die ECL-Vereinbarung abschließt (vgl. § 51 Abs. 2 VGG). Infolgedessen wäre die Verwertungsgesellschaft in diesem Fall in die Lage versetzt, eine ungerechtfertigte Vergütung vorbehaltserklärender Urheber zu verhindern. Wenn die Einrichtung eines Opt-out-Mechanismus zugunsten betroffener Urheber erforderlich ist, spricht dies im Ergebnis dafür, eine Freistellung im Wege eines ECL-Mechanismus umzusetzen.

e) Erhöhtes Bedürfnis nach Rechtssicherheit für Werknutzer

Ist kein Opt-out-Mechanismus zugunsten betroffener Urheber erforderlich, so spricht dies umgekehrt allerdings noch nicht ohne Weiteres dafür, eine Freistellung im Wege einer gesetzlichen Lizenz umzusetzen. Denn ein Freistellungsmechanismus ohne Opt-out-Mechanismus kann sowohl als ECL (ohne Widerspruchsmöglichkeit) als auch durch eine gesetzliche Lizenz

152 Beispielsweise der Rechtevorbehalt in § 49 Abs. 1 S. 1 UrhG.

153 Beispielhaft die Widerrufsmöglichkeiten des Rechtsinhabers in den erweiterten kollektiven Lizenzmechanismen, die durch Umsetzung der DSM-RL im VGG implementiert worden sind, siehe §§ 51 Abs. 2, 52 Abs. 2 VGG.

154 Dazu bereits siehe oben unter 2. Teil A. II. 2. a).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

(ohne Vorbehaltsmöglichkeit) umgesetzt werden. Bei der Wahl zwischen beiden Regelungsmechanismen könnte es in diesem Fall vielmehr darauf ankommen, ob die Trainingsdatensammler, KI-Entwickler oder Anwender von KI-Systemen ein besonders ausgeprägtes Bedürfnis nach Rechtssicherheit haben.

Gesetzliche Lizenzen sind als Rechtsnormen abstrakt-generell ausgestaltet. Dies geht mit einer Einschränkung der Rechtssicherheit in Bezug auf die konkrete Einzelfallanwendung einher. ECL beruhen hingegen auf Lizenzvereinbarungen, die im Rahmen der gesetzlichen Spielräume für den konkreten Einzelfall individualisiert werden können. Im direkten Vergleich beider Rechtsinstitute gewährleistet der ECL-Mechanismus damit mehr Rechtssicherheit.[155] Wenn ein besonderes ausgeprägtes Rechtssicherheitsbedürfnis besteht, ist daher eine Freistellung im Wege eines ECL-Mechanismus vorzuziehen. Das ist beispielsweise dann der Fall, wenn das Sammeln von Werken als Trainingsdaten oder andere Sachverhaltskonstellationen ein hohes wirtschaftliches Risiko für die Freistellungsbegünstigten mit sich bringen.

Bestehen keine überdurchschnittlichen Anforderungen an die von der Freistellungsregelung gewährleistete Rechtssicherheit, so ist die Werknutzung hingegen auf Basis einer gesetzlichen Lizenz niederschwelliger möglich. Denn ihre Freistellungswirkung tritt bereits von Gesetzes wegen ein. In diesen Fällen sollte die Freistellung daher im Wege einer gesetzlichen Lizenz umgesetzt werden.

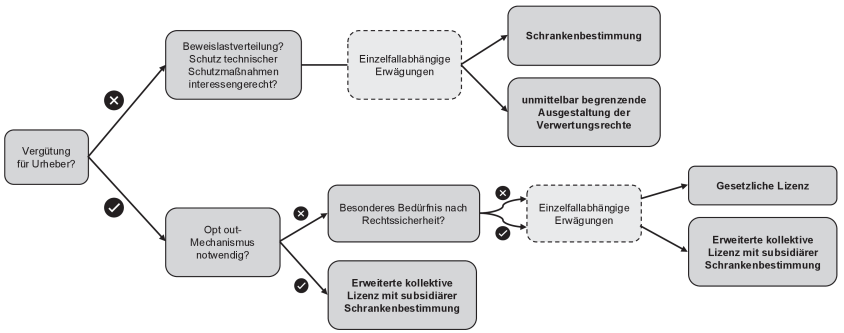
f) Weitere Aspekte und Zusammenfassung

Neben den herausgearbeiteten Entscheidungskriterien, die im Regelfall die Wahl zwischen den zur Verfügung stehenden Freistellungsmechanismen ermöglichen, können im Einzelfall aber auch andere Aspekte berücksichtigt werden.

155 *Trumpke*, Exklusivität und Kollektivierung, S. 578.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Alles in allem ist die Wahl des für eine Freistellung geeigneten Regelungsmechanismus anhand des folgenden Musters vorzunehmen:



B. Speicherung von urheberrechtlich geschützten Trainingsdaten aus Internetquellen mit Hilfe von Webcrawlern

Bevor ein KNN mit Werken aus Internetquellen trainiert werden kann, müssen die Werke zunächst abgespeichert werden. Dies geschieht insbesondere mit Hilfe automatisierter Webcrawler. Der Prozess wird auch als Web Scraping bezeichnet. Fraglich ist zunächst, ob das Web Scraping der Werke zur Verwendung als Trainingsdaten ohne individuelle Zustimmung betroffener Urheber möglich, also urheberrechtlich freigestellt sein sollte. Hiervon würden Entwickler von KI-Systemen und intermediäre Trainingsdatensammler profitieren, deren Tätigkeit darauf ausgerichtet ist, Trainingsdatensätze für Dritte zu erstellen und in Verkehr zu bringen. Damit ein regulatorisches Freistellungsbedürfnis besteht, muss es sich als zweckmäßig erweisen, das Web Scraping von Werken als Trainingsdaten vom Ausschließlichkeitsrecht betroffener Urheber auszuklammern.[156]

156 Zur Herleitung der regulatorischen Entscheidungskriterien insgesamt siehe unter 2. Teil A.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

I. Zweckmäßigkeit einer Freistellung

Die Freistellung des Web Scrapings von Werken als Trainingsdaten ist zweckmäßig, wenn die Interessen der Urheber, deren Werke für das Training einer KI abgespeichert werden, nicht berührt sind oder wenn eine umfassende Abwägung von Urheber- und Nutzerinteressen ergibt, dass den Nutzerinteressen beim Web Scraping von Werken als Trainingsdaten Vorrang zu gewähren ist.

1. Beeinträchtigung der Urheberinteressen

Die Speicherung von bereits im Internet verfügbar gemachten Werken geht nicht mit einer Veränderung der Inhalte einher. Mangels identifizierender Werkrezipienten beim rein-technischen KI-Training wird durch das sachverhaltstypische Weglassen der Urhebernennung im Trainingsprozess auch die Bekanntheit des Urhebers auf dem Kreativmarkt nicht geschmälert. Aus alldem folgt, dass vom Web Scraping weder das Erstveröffentlichungs- noch das Integritätsinteresse sowie das Interesse des Urhebers an der Anerkennung seiner Urheberschaft betroffen wird. Das Web Scraping beeinträchtigt daher keine geschützten ideellen Urheberinteressen.

Gegen eine Beeinträchtigung der urheberrechtlichen Verwertungsinteressen spricht weiterhin, dass für das Web Scraping notwendigerweise bereits Zugang zu den betroffenen Werken bestehen muss, womit die Werke bereits im Rahmen der Zugänglichmachung einmal verwertet worden sind.[157] Allerdings setzen insbesondere sämtliche Beschränkungen des Urheberrechts, die die Vervielfältigung von Werken durch Nutzer erlaubnisfrei stellen, denklogisch Zugang zu Repräsentationen dieser Werke voraus.[158] Hieraus folgt allerdings nicht zwingend, dass durch die Herstellung weiterer Werkstücke keine Verwertungsinteressen der betroffenen Urheber mehr beeinträchtigt sind. Vielmehr handelt es sich um zwei voneinander unabhängige Verwertungshandlungen: In einem ersten Schritt kann der Zugang zum Werk monetarisiert werden. Dies geschieht beispielsweise im Rahmen der Kinovorführung eines Filmwerks, für die Nutzer eine Vergütung entrichten. Im Anschluss hieran hat aber auch die Vervielfältigung des betreffenden Werkes weiterhin einen wirtschaftlichen Wert. Kinobesuchern

157 So *Raue*, CR 2017, 656 (661); *Raue*, GRUR 2017, 11 (14).

158 *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 58.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

wird nicht gestattet, im Anschluss an die Filmvorführung nun auch ein Vervielfältigungsstück des Filmwerks anzufertigen. Im Gegenteil müssten sie hierfür weitere Nutzungsrechte oder ein bereits angefertigtes Vervielfältigungsstück erwerben.

Eine Freistellung, nach der individuelle Nutzungsrechtseinräumungen durch jeden betroffenen Urheber für die Speicherung der Trainingsdaten nicht mehr notwendig wären, hätte vielmehr zur Folge, dass die Urheber mangels Nachfrage nicht mehr in der Lage wären, ihre Werke – unmittelbar oder mittelbar über Trainingsdatenanbieter – auf dem Markt für Trainingsdaten zu monetarisieren. Auch wenn es sich hierbei nicht um die übliche Kernverwertung urheberrechtlich geschützter Inhalte handelt, würde den Urhebern hiermit eine Möglichkeit zur Verwertung ihrer Werke genommen. Die Perspektive der Urheber auf eine den Lebensunterhalt und weitere schöpferische Tätigkeit sichernde Amortisation der Werkschöpfung würde infolgedessen geschmälert.

Hieraus folgt, dass das Web Scraping von Werken als Trainingsdaten die Verwertungsinteressen der Urheber beeinträchtigt.[159] Eine urheberrechtliche Freistellung kann demzufolge nicht bereits wegen einer Nichtbeeinträchtigung der Urheberinteressen erfolgen.

2. Abwägung der Urheberinteressen mit den Interessen der Trainingsdatensammler

Eine umfassende Abwägung von Urheber- und Nutzerinteressen könnte allerdings zur Folge haben, dass den Nutzerinteressen beim Web Scraping von Werken als Trainingsdaten Vorrang vor den Urheberinteressen zu gewähren ist.

a) Zu berücksichtigende Trainingsdatensammler- und Allgemeininteressen

Dem betroffenen Verwertungsinteresse der Urheber steht das Ziel der KI-Entwickler und intermediären Trainingsdatensammler gegenüber, mit Hilfe

159 So auch *Dregelies*, GRUR 2024, 1484 (1485); *Dornis/Stober*, Urheberrecht und Training generativer KI-Modelle, S. 150. Für das allgemeine Text- und Data-Mining anders hingegen *Hauk/Pflüger*, ZUM 2020, 383 (385); *Raue*, CR 2017, 656 (661); *Raue*, GRUR 2017, 11 (14); *Schack*, ZUM 2016, 266 (269); *Schack*, ZUM 2017, 802 (806); *Schack*, Urheber- und Urhebervertragsrecht, S. 311 Rn. 627; *Stieper*, GRUR 2020, 1 (4); *Stieper*, ZUM 2021, 776 (778); *Raue*, ZUM 2019, 684 (686), die davon ausgehen, dass die urheberrechtlichen Verwertungsrechte nicht berührt werden.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

von Web Scraping-Technologien urheberrechtlich geschützte Trainingsdaten zu generieren. In diesem Ziel könnten sich verschiedene Interessen bündeln, die zugunsten der Nutzer in die Abwägung eingestellt werden müssen. Gemein ist allen Fallkonstellationen aber, dass die für das Training genutzten Werke während des eigentlichen Lernvorgangs nur als Trägermedium für Muster und Zusammenhänge genutzt werden. Sie sollen aus den Werken extrahiert werden. Auch bei der menschlichen Informationserfassung werden Werke als Trägermedium, also nur um ihrer Informationsinhalte Willen, konsumiert. Infolgedessen ist zugunsten der Trainingsdatensammler zu berücksichtigen, dass das Ziel, Daten zum Training einer KI zu generieren, strukturelle Ähnlichkeiten zum menschlichen Informationsinteresse aufweist.[160]

Aufgrund des besonders hohen Daten- und Ressourcenaufwands[161] findet die Entwicklung von KI außerdem regelmäßig durch gewerblich handelnde Akteure statt. Damit ist das Web Scraping regelmäßig als Teil der Berufs- beziehungsweise unternehmerischen Freiheit der Datensammler geschützt, die in Art. 15 f. EU-GRCh sowie auf nationaler Ebene in Art. 12 Abs. 1, 2 GG verankert ist. Ebenso sind im Einzelfall auch Überschneidungen mit anderen Freiheitsrechten wie insbesondere der Wissenschaftsfreiheit (Art. 5 Abs. 3 GG) möglich.

Unabhängig davon, in welchem Zusammenhang das Web Scraping im konkreten Fall stattfindet, verfolgen sowohl KI-Entwickler als auch intermediäre Trainingsdatensammler direkt oder mittelbar das Ziel, durch das Training eines neuen KI-Systems den technologischen Fortschritt im Bereich des maschinellen Lernens zu fördern. Ziel der Trainingsdatensammler ist demzufolge auch die Innovationsförderung. Auch dieses Innovationsinteresse kann zugunsten der Trainingsdatensammler in die Abwägung aufgenommen werden.[162] Seine Berücksichtigung im urheberrechtlichen Regelungsregime zeigt sich bereits in einer anderen urheberrechtlichen Schrankenbestimmung, § 60c UrhG.[163]

160 Auch *Heesen/Jüngels*, RuZ 2021, 45 (53); *Raue*, CR 2017, 656 (661); *Hofmann*, WRP 2024, 11 Rn. 23.

161 Dazu bereits oben unter 1. Teil B II. 3.

162 So auch *Hofmann*, WRP 2024, 11 Rn. 22; *Trapova/Gervassis*, Kluwer Copyright Blog v. 16.3.2022; *Vesala*, IIC 2023, 351 (355); vgl. allgemein für das Text und Data-Mining auch *de la Durantaye*, GRUR 2017, 558 (561 f.); *Raue*, GRUR 2017, 11 (15); *Hilty/Moscon*, Modernisation of the EU Copyright Rules, S. 125.

163 Vgl. auch ErwGr. 5, 8 DSM-RL, auch der Richtliniengeber bezieht das Innovationsinteresse in seine Erwägungen ein.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Eine Freistellung des von Trainingsdatensammlern vorgenommenen Web Scrapings von Werken als Trainingsdaten kann also erfolgen, wenn die Abwägung des durch das Web Scraping betroffenen Verwertungsinteresses der Urheber mit dem von den KI-Entwicklern bzw. intermediären Trainingsdatensammlern verfolgten Innovationsinteresse und insbesondere ihrer Berufsfreiheit ergibt, dass den dargelegten Interessen der Datensammler im konkreten Fall Vorrang vor dem Verwertungsinteresse der Urheber zu gewähren ist.

b) Beeinträchtigungsintensitäten der gegenläufigen Interessen

Im Hinblick auf die betroffenen Urheberinteressen lässt sich hierfür zunächst feststellen, dass die Freistellung des Web Scrapings den Urhebern nicht die Möglichkeit nimmt, ihre Werke auszustellen, öffentlich wiederzugeben und Vervielfältigungsstücke der Werke in Verkehr zu bringen. Daher nimmt das Web Scraping den betroffenen Urhebern nicht die Möglichkeit, ihre Kernverwertung, also diejenigen Verwertungsvorgänge, bei denen die menschliche Rezeption der Werke in ihrer Eigenschaft als geistige Schöpfungen im Mittelpunkt steht, vorzunehmen.[164] Dies wäre aufgrund der gleichbleibenden Verwertungsform selbst dann der Fall, wenn auch kommerzielle Datensammler von einer Freistellungsregelung erfasst wären.[165] Dementsprechend sind die Verwertungsinteressen der Urheber vom Web Scraping zum Zwecke des Trainings einer KI nicht wesentlich betroffen.

Ähnliches könnte allerdings auch für die Ziele der KI-Entwickler oder intermediären Trainingsdatensammler gelten. Denn die mit dem Web Scraping im Regelfall ausgeübte Berufs- bzw. unternehmerische Freiheit im Sinne der Art. 15, 16 EU-GRCh ist nur in ihrer Ausformung als Ausübungsfreiheit berührt. Der Eingriff ist damit nicht intensiv.[166] Das ist auch bei der Kollision mit anderen Grundrechten wie dem über Art. 17 EU-GRCh geschützten Urheberrecht zu berücksichtigen.

164 Wie hier auch *Jager*, Artificial Creativity?, S. 347. So auch schon für das allgemein Text- und Data-Mining ohne spezifischen Bezug zu KI *de la Durantaye*, Allgemeine Bildungs- und Wissenschaftsschranke, S. 240; *European Copyright Society*, General Opinion on the EU Copyright Reform Package, S. 5.

165 So im Kontext des Dreistufentests auch *Hacker*, ZGE 2020, 239 (269); a. A. ebenfalls zum Dreistufentest allerdings *Jager*, Artificial Creativity?, S. 371; zum unspezifischen Text und Data-Mining außerdem *Triaille/Meeûs d'Argenteuil/Francquen*, Study on the legal framework of text and data mining (TDM), S. 99 f.

166 *Calliess*, in: Calliess/Ruffert, EUV/AEUV, EU-GRCh Art. 15 Rn. 16.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Das Innovationsinteresse ist darüber hinaus nicht ausdrücklich grundrechtlich abgebildet. Damit sind an seinen Schutz ebenfalls keine besonders hohen Anforderungen zu stellen. Zwar könnten im Einzelfall auch andere Grundrechte – insbesondere die Kommunikationsfreiheiten oder die Wissenschaftsfreiheit aus Art. 11, 13 EU-GRCh – betroffen sein. Da auch in diesen Fällen nur spezifische Formen der Freiheitsausübung, nicht aber die Freiheiten in ihrem Wesensgehalt berührt sind, wiegen auch die hier zu berücksichtigenden Nutzerinteressen im Verhältnis zum über Art. 17 EU-GRCh geschützten Verwertungsinteresse der Urheber nicht überproportional schwer. Bereits aus der Interessenlage als solcher ergibt sich demzufolge noch kein Überwiegen der Interessen der Urheber auf der einen oder der Trainingsdatensammler auf der anderen Seite.

c) Erwägungen gegen eine Freistellung

Dafür, das Web Scraping urheberrechtlich geschützter Trainingsdaten dem Ausschließlichkeitsrecht der Urheber zuzuweisen, spricht, dass das Sammeln der urheberrechtlich geschützten Trainingsdaten nicht mit dem Informationsvorgang selbst, also dem Rezipieren eines Werkes, vergleichbar ist.[167] Aufgrund etwaiger Parallelen hierzu kann daher zwar möglicherweise der eigentliche Trainingsvorgang der KI,[168] nicht aber das Web Scraping der Inhalte mit dem Werkgenuss gleichgesetzt und deswegen urheberrechtlich freigestellt werden.

Weiterhin werden beispielsweise Nachrichtenartikel als Trainingsdaten genutzt, um Anwendungen des Algorithmic Journalism – einer speziellen Art von Textgeneratoren, die journalistische Inhalte erzeugen – anzulernen. Gleiches gilt für Bildwerke, die genutzt werden, um bildgenerierende KI zu trainieren. Trainingsdatensammler verfolgen also mitunter das Ziel, ein KI-System herzustellen oder die Herstellung eines KI-Systems zu ermöglichen, das Konkurrenzprodukte zu den Werken produzieren kann, die zum Training des KI-Systems genutzt worden sind.[169] Mit dem Einsatz solch generativer KI-Systeme wird die Möglichkeit der Urheber, ihre Schöpfungs-

167 In dieser Hinsicht präzise *Hilty/Moscon*, Modernisation of the EU Copyright Rules, S. 20.

168 Dazu ausführlich unter 3. Teil B. II.

169 Dies stellt auch *de la Durantaye*, ZUM 2023, 645 (649) heraus. Ebenfalls im Kontext des Dreistufentests *Schack*, NJW 2024, 113 Rn. 8; *Dornis/Stober*, Urheberrecht und Training generativer KI-Modelle, S. 154 f., letztere an dieser Stelle aber etwas apodiktisch.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

leistungen insbesondere auch durch Kernverwertungsarten zu amortisieren, beeinträchtigt. Dies kann mittelbar auch bei der Freistellung des Web Scrapings von Werken für das KI-Training berücksichtigt werden.

In diesem Kontext könnten zum Training genutzte Werke im Einzelfall sogar im Output des trainierten KI-Systems identifizierbar sein. In der Folge bestünde die Gefahr, dass der Output der KI Urheberrechte der vom Training betroffenen Urheber verletzt.[170]

Gegen eine Freistellung spricht darüber hinaus, dass die Mehrzahl derjenigen Urheber, die das Web Scraping trotz einer urheberrechtlichen Freistellung verhindern wollen, nicht über die technischen Fähigkeiten verfügen, diesem Bedürfnis mit Hilfe gegen das Web Scraping gerichteter technischer Schutzmaßnahmen gerecht zu werden. Effektiver Schutz bestünde für betroffene Urheber damit nur, wenn sie ihre Werke überhaupt nicht im freien Internet verfügbar machen. Da gerade hier aber auch beispielsweise durch Lizenzeinnahmen für die öffentliche Zugänglichmachung der Werke oder Zugangsentgelte wesentliches Amortisationspotential liegt, könnten Urheber ihr Interesse nicht durchsetzen, ohne im Ergebnis doch mittelbar ihre Verwertungsmöglichkeiten zu beeinträchtigen, bei denen die Nutzung der Werke um ihren schöpferischen Gehalt Willen im Vordergrund steht. Diese ist Teil der Kernverwertung.

d) Erwägungen zugunsten einer Freistellung

Zugunsten einer Freistellung des Web Scrapings urheberrechtlich geschützter Trainingsdaten könnten allerdings ebenfalls eine Vielzahl verschiedener Argumente angeführt werden.

aa) Verhältnis zwischen der Speicherung zum Training und den Erzeugnissen des KI-Systems

Für eine Freistellung spricht zunächst, dass beim Training der KI selbst nach aktuellem Stand von Wissenschaft und Technik nicht vollständig ausgeschlossen werden kann, dass es in Einzelfällen zur Wiedererkennbarkeit von Trainingsdaten im Output kommt. Das Risiko für eine Identifizierbarkeit von Trainingsdaten im Output ist daher nicht restlos vermeidbar und

170 So im Kontext des Dreistufentests *Vesala*, IIC 2023, 351 (366 f.).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

gleichzeitig erheblich einzelfallabhängig. Das Sammeln von Werken zum Training von KI insgesamt zu verhindern, wird diesem Risikograd nicht gerecht. Es würde vielmehr zu einem überschießenden Schutz der Urheber führen. Zudem zeigt sich erst mit der Anwendung der KI, ob verwendete Trainingsdaten im Einzelfall im Output identifizierbar sind und ob es hierdurch zu Urheberrechtsverletzungen kommt. Die Generierung von Output durch KI geht aber auf einen zweiten Willensentschluss und demzufolge auf eine eigenständige natürliche Handlung zurück. Sie wird häufig auch von anderen Beteiligten durchgeführt als das Training. Sie ist deswegen vom Training einer KI abgrenzbar. Die Output generierenden Nutzer entrichten für die Generierung des Outputs zudem unter Umständen eine gesonderte Vergütung. Der Markt für die Anwendung eines KI-Systems und die Generierung von Output, an dem KI-Nutzer als Nachfrager teilnehmen, ist vom Markt für das Training von KI, dessen Teilnehmer KI-Entwickler sind, zu trennen. Die Generierung von Output ist deswegen wirtschaftlich klar vom Training abgrenzbar und selbstständig verkehrsfähig. Sie ist aus diesen Gründen ein eigenständiger Nutzungssachverhalt.[171] Ihre Rechtmäßigkeit ist daher auch eigenständig zu beurteilen. Gegen im Einzelfall auftretende Urheberrechtsverletzungen im Output von KI kann im jeweiligen Einzelfall eigenständig vorgegangen werden. Gleichzeitig kann es für die Freistellung der Datensammlung zum Training von KI aus einer zeitlich-logischen Perspektive nicht darauf ankommen, ob es zu einem späteren Zeitpunkt durch eine andere Nutzungshandlung, die Generierung des Outputs, zu Urheberrechtsverletzungen kommt.[172]

Inwieweit KI-Systeme auch unabhängig von Urheberrechtsverletzungen in konkretem Output in der Lage sind, Konkurrenzprodukte zu den betroffenen Werken zu generieren, hängt darüber hinaus von der Zielfunktion des trainierten KI-Systems sowie der Qualität des Outputs und der Nachfrage auf dem Kreativmarkt ab. Die mögliche Beeinträchtigung der Urheber ist dementsprechend einzelfallabhängig. Daraus folgt, dass die *Möglichkeit* zur Generierung von Konkurrenzprodukten nicht per se zum Ausbleiben

171 Außerdem eine eigenständige Nutzungsart im urhebervertragsrechtlichen Sinne (§ 31 Abs. 1 UrhG), vgl. zu den Kriterien für eine Abgrenzung eigenständiger Nutzungsarten *Wiebe*, in: Spindler/Schuster, Recht der elektronischen Medien, UrhG § 31 Rn. 9 m. w. N.

172 So im Kontext der Anwendung des Dreistufentests aber *Vesala*, IIC 2023, 351 (366 f.).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

einer Freistellungsregelung führen sollte.[173] Es bedarf vielmehr einer Betrachtung des jeweiligen Einzelfalls, insbesondere über den Dreistufentest bei der Anwendung einer Freistellungsregelung im konkreten Einzelfall.[174] Die Generierung von Output ist – wie bereits gezeigt werden konnte – ein eigenständiger, vom Training abgrenzbarer Nutzungssachverhalt, sodass er auch hinsichtlich der möglichen Konkurrenzsituation und dessen Auswirkungen vom Trainingsprozess losgelöst beurteilt werden sollte. Dies spricht im Ergebnis ebenfalls für eine einzelfallbezogene, am konkret generierten Output ansetzende, nicht aber eine systemische Betrachtung im Kontext des Trainingsprozesses.

bb) Kompensationswirkung einer Vergütungspflicht

Zudem könnte die Gewährung einer angemessenen Vergütung für die Nutzung der Werke zum Training von KI[175] die wirtschaftlichen Folgen für betroffene Urheber abmildern und dazu beitragen, den Substitutionsdruck zu senken. Dies gilt insbesondere deswegen, weil die wirtschaftliche Bedeutung der Werkverwendung beim Training von KI auch bei der Bestimmung der Vergütungshöhe berücksichtigt wird.[176]

cc) Art der Werkverwendung

Zu berücksichtigen ist darüber hinaus, dass mit dem Training von KI keine Nutzung der schöpferischen Elemente verwendeter Werke *als Werk* verbunden ist. Es findet vielmehr ein rein-technischer Vorgang statt, im Rahmen dessen die Werke als Informationsträger ausgewertet werden.[177]

Auch wenn das Sammeln urheberrechtlich geschützter Trainingsdaten zudem nicht selbst mit dem menschlichen Informationsvorgang, also dem

173 A. A., aber ohne eingehende Begründung *Dornis/Stober*, Urheberrecht und Training generativer KI-Modelle, S. 151, 154 f.

174 Für eine Einzelfallbetrachtung im Kontext des Dreistufentests auch *Leistner*, GRUR 2024, 1665 (1670); *Novelli u. a.*, Computer Law & Security Review 2024, (10).

175 Dazu siehe unter 2. Teil B. II. 1.

176 Zur Vergütungshöhe: *Pflüger*, Gerechter Ausgleich und angemessene Vergütung, S. 145; insbesondere zu anderen bei der Bestimmung der Vergütungshöhe berücksichtigungsfähigen Aspekten: *Lauber-Rönsberg*, Urheberrecht und Privatgebrauch, S. 112 ff.

177 *Hofmann*, WRP 2024, 11 Rn. 25; *Lennartz/Kraetzig*, GRUR 2024, 1258 (1260).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Werkgenuss, verglichen werden kann, besteht darüber hinaus ein untrennbarer Zusammenhang zwischen der beim KI-Training stattfindenden Musteranalyse und dem Sammeln der Werke mit Hilfe eines Webcrawlers. Besteht ein freihaltebedürftiges Verarbeitungsinteresse, beispielsweise das Interesse am Werkgenuss oder aber auch das Interesse an der Analyse der den Inhalten zugrundeliegenden Muster und Zusammenhänge, muss auch eine hierfür zunächst notwendige Vervielfältigung möglich sein. Nur so kann die rechtmäßige Benutzung des betreffenden Werkes ermöglicht werden. Diese Kausalität kommt bereits in der urheberrechtlichen Schrankenbestimmung für vorübergehende Vervielfältigungshandlungen insbesondere im Arbeitsspeicher eines Computers (§ 44a UrhG) zum Ausdruck.[178] In gleicher Weise muss auch eine Vervielfältigungshandlung, die für den „automatisierten Werkgenuss", also das Training einer KI,[179] erforderlich ist, zugunsten der Werknutzer freigestellt sein.[180]

dd) Innovationsbegünstigte und Bedeutung des Innovationsinteresses

Für eine Freistellung des Web Scrapings urheberrechtlich geschützter Trainingsdaten spricht weiterhin, dass von den technischen Innovationen, die durch umfangreich verfügbare Trainingsdatensätze ermöglicht werden, nicht nur die Innovatoren profitieren. Vielmehr kann das Innovationsinteresse als gesamtgesellschaftliches Anliegen verstanden werden. Insbesondere von generativen KI-Systemen, also solcher KI, die kreative Leistungsergebnisse produziert, profizieren auch wiederum Urheber selbst. Sie können die KI-Systeme als Werkzeug im Prozess des Werkschaffens einsetzen und Effizienzgewinne erzielen. Ungehinderter Zugang zu Trainingsdaten unabhängig von ihrer Werkqualität ist eine wichtige Voraussetzung dafür, die Entwicklung der KI insgesamt voranzubringen.[181] Andersherum droht durch einen mangelnden Zugang zu qualitativ hochwertigen Trainingsdaten, insbesondere verkörpert in urheberrechtlich geschützten Werken wie

178 *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 8, vgl. auch S. 56 ff.

179 Dazu siehe unter 2. Teil C. II. 1.

180 *Hilty/Moscon*, Modernisation of the EU Copyright Rules, S. 20; vgl. ebenfalls *Hofmann*, WRP 2024, 11 Rn. 23.

181 COM(2018) 237 final, S. 12; vgl. außerdem *European Commission u. a.*, Study on copyright and new technologies: copyright data management and artificial intelligence (SMART 2019/0038), S. 208 f.; *Hofmann*, WRP 2024, 11 Rn. 22 m. w. N.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

zum Beispiel wissenschaftlicher Fachliteratur, ein Qualitätsdefizit bei der Entwicklung von KI-Systemen. Denn die Güte des Systems verhält sich proportional zur Güte der verwendeten Trainingsdaten.[182] Defizitäre KI-Systeme haben hingegen Effizienzverluste zur Folge, sodass ihre Entwicklung gesellschaftlich nicht erwünscht ist.[183]

Aus einer ökonomischen Perspektive muss auch berücksichtigt werden, dass es Anwendungsszenarien gibt, für die der Markt keine geeigneten Trainingsdaten zur Verfügung stellt. Dies kann einmal auf die Art der benötigten Trainingsdaten, zum Beispiel umgangssprachliche Formulierungen, zum anderen aber auch auf die Menge der benötigten Trainingsdaten zurückzuführen sein.[184] Ohne eine Freistellung der Speicherung der benötigten, urheberrechtlich geschützten Trainingsdaten wären Innovationen in den jeweiligen Anwendungsbereichen erschwert. Das sollte vor allem generative KI-Systeme betreffen. Dies widerspräche aber dem Interesse an einer Förderung technischer Innovationen, das sowohl im UrhG[185] als auch durch den Richtliniengeber[186] anerkannt worden ist.

ee) Praktische Erwägungen

Aus einer praktischen Perspektive für eine Freistellung des Web Scrapings von Werken zum Training von KI spricht zudem, dass beim Web Scraping von Inhalten aus dem Grenzbereich der urheberrechtlichen Schutzfähigkeit, beispielsweise beim Sammeln von Zeitungsartikeln zum Training einer Algorithmic Journalism-Anwendung, weder von KI-Entwicklern oder anderen Trainingsdatensammlern noch durch den Webcrawler in jedem Fall beurteilt werden kann, ob der abgespeicherte Inhalt als Werk Schutz genießt und daher die urheberrechtlichen Rahmenbedingungen beachtet werden müssen.[187] Die mit dieser Rechtsunsicherheit einhergehenden Ab-

182 *de la Durantaye*, ZUM 2023, 645 (647); *Margoni/Kretschmer*, GRUR Int. 2022, 685 (687); aus einer etwas universelleren Perspektive *Hacker*, ZGE 2020, 239 (243); im Anschluss an die vorgenannten *Hofmann*, WRP 2024, 11 Rn. 22.

183 *Hofmann*, WRP 2024, 11 Rn. 22.

184 Für den Textgenerator „GPT-3" waren Trainingsdaten mit einem Umfang von 300 Milliarden Wörtern notwendig, siehe *Brown u. a.*, arXiv:2005.14165 2020, S. 8.

185 Beispielsweise in §§ 44b, 60c, 60d UrhG.

186 ErwGr. 5 DSM-RL.

187 Vgl. *European Commission u. a.*, Study on copyright and new technologies: copyright data management and artificial intelligence (SMART 2019/0038), S. 273;

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

schreckungseffekte können nur durch eine Freistellung der Nutzung auch urheberrechtlich geschützter Inhalte beseitigt werden.

ff) Ökonomischer Vorteil aus einer Freistellung

Selbst wenn quantitativ und qualitativ hinreichende Trainingsdaten auf dem Trainingsdatenmarkt verfügbar wären, müssten KI-Entwickler oder intermediäre Trainingsdatensammler schließlich einen nicht überschaubaren Aufwand betreiben, um Lizenzvereinbarungen mit jedem betroffenen Urheber abzuschließen. Die für Trainingsdatensammler hierdurch entstehenden Transaktionskosten wären sehr hoch und könnten von vielen Trainingsdatensammlern auch nicht erbracht werden.[188] Schließlich sind bei der Wiedergabe von Werken im Internet auch oft die betreffenden Rechtsinhaber nicht bekannt, womit auch ein individueller Lizenzabschluss nicht möglich wäre.[189] Das gilt vor allem für Alltagswerke der sogenannten kleinen Münze. Auch das hätte eine Innovationshemmung zur Folge.[190] Insbesondere wäre aber eine die ökonomische Effizienz verwirklichende, selbstgesteuerte Güterallokation durch Transaktionen zwischen den Urhebern und Trainingsdatensammlern bei einer fehlenden Kenntnis über die betroffenen Urheber nicht möglich. Gerade aufgrund der im Zusammenhang zu KI bestehenden Prognoseunsicherheiten, die eine manuelle Güterzuordnung ausschließen,[191] sollte dieser Allokationsmechanismus aus einer ökonomischen Perspektive aber möglich sein.

Prohibitiv hohe Transaktionskosten und Probleme bei der Identifizierung von Rechtsinhabern führen darüber hinaus dazu, dass eine unterbleibende Freistellung des Web Scrapings von Werken zum Training zu einer

---

Ebenso im Zusammenhang zur im Entwurf der KI-VO vorgeschlagenen Transparenzpflicht *de la Durantaye*, ZUM 2023, 645 (656).

188 *Bomhard*, InTeR 2023, 174 (175); *Hofmann*, WRP 2024, 11 Rn.22; *Novelli u. a.*, Computer Law & Security Review 2024, (10); *Raue*, ZUM 2021, 793 (796); *Steinbrecher*, MMR 2019, 639 (640); *European Commission u. a.*, Study on copyright and new technologies: copyright data management and artificial intelligence (SMART 2019/0038), S. 273; *Senftleben u. a.*, JIPITEC 2022, 67 Rn. 12; *Senftleben*, IIC 2023, 1535 (1346 f.).

189 *de la Durantaye*, ZUM 2023, 645 (656).

190 So auch *Hacker*, ZGE 2020, 239 (261); *de la Durantaye*, Allgemeine Bildungs- und Wissenschaftsschranke, S. 238; *Vesala*, IIC 2023, 351 (355); vgl. außerdem *The Intellectual Property Office (GB)*, Modernising Copyright: A modern, robust and flexible framework, S. 37.

191 Dazu siehe oben unter 2. Teil A. I. 3. b).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

lose-lose-Situation für Urheber und Trainingsdatensammler führt: Mangels Freistellung können Trainingsdatensammler kein KI-Training auf Basis der Werke durchführen, gleichzeitig erzielen Urheber mit ihren Werken aber auch keine Vergütung durch das Training von KI auf Basis ihrer Werke, weil prohibitiv hohe Transaktionskosten individuelle Lizenzabschlüsse verhindern.[192] Damit wären die Amortisationsinteressen der Urheber durch eine vergütungspflichtige Freistellung sogar besser geschützt als durch eine individuallizenzliche Lösung.

e) Schlussfolgerungen

Insgesamt sprechen bessere Gründe für eine Freistellung des Web Scrapings von Werken zum KI-Training. Vereinzelte Argumente gegen eine Freistellung konnten hingegen entkräftet werden. Dementsprechend ist den Interessen der Trainingsdatensammler vor den Verwertungsinteressen betroffenen Urheber Vorrang zu gewähren. Infolgedessen ist eine Freistellung des Web Scrapings von Werken als Trainingsdaten insgesamt zweckmäßig und sollte aus einer regulatorischen Perspektive daher erfolgen.[193]

II. Freistellungsmechanismus

Die Zweckmäßigkeit der Freistellung des Web Scrapings von Werken als Trainingsdaten wirft im Weiteren allerdings die Frage auf, durch welchen Regelungsmechanismus eine solche Freistellung konkret umzusetzen ist.

---

192 *Senftleben*, IIC 2023, 1535 (1548).

193 Zustimmend *Hacker*, ZGE 2020, 239 (261, 269); *Hofmann*, WRP 2024, 11 Rn. 18; *Lennartz/Kraetzig*, GRUR 2024, 1258 (1260); *Trapova/Gervassis*, Kluwer Copyright Blog v. 16.3.2022; grundsätzlich auch *Jager*, Artificial Creativity?, S. 371, wenn auch mit Bedenken hinsichtlich einer Erfassung kommerzieller Handlungen; in Bezug auf allgemeine Text- und Data-Mining-Aktivitäten auch *Schack*, ZUM 2016, 266 (269); *Schack*, ZUM 2017, 802 (806, 807); *Schack*, Urheber- und Urhebervertragsrecht S. 311 Rn. 627; *de la Durantaye*, Allgemeine Bildungs- und Wissenschaftsschranke, S. 237 ff.; *Cocoru/Boehm*, An analytical review of text and data mining practices and approaches in Europe - Policy recommendations in view of the upcoming copyright legislative proposal, S. 7; *European Copyright Society*, General Opinion on the EU Copyright Reform Package, S. 5; *Hilty/Moscon*, Modernisation of the EU Copyright Rules, S. 125. Ebenfalls ErwGr. 8-11 DSM-RL.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

1. Angemessener Ausgleich für die Nutzungshandlung

Zur Beantwortung dieser Fragestellung muss zunächst herausgearbeitet werden, ob den Urhebern ein angemessener Ausgleich für das Web Scraping ihrer Werke gezahlt werden sollte.

a) Rechtfertigungsansätze für eine vergütungsfreie Nutzung

Gegen einen finanziellen Ausgleich spricht, dass die Kernverwertungsmöglichkeiten der Urheber durch das Web Scraping der Werke, um sie als Trainingsdaten zu verwenden, nicht unmittelbar beeinträchtigt werden.[194] Die Werke werden zudem nicht *als Werk*, also ihrer geistig-schöpferischen Gehalte wegen, sondern lediglich als Informationsträger ausgewertet.[195] Darüber hinaus profitieren anwendungsbezogen nicht nur die KI-Entwickler, sondern auch andere Beteiligtengruppen und mitunter sogar Urheber selbst von KI-Neuentwicklungen. Mit umfangreichen Trainingsdatensätzen wird zudem auch die Entwicklung der KI als solche vorangebracht.[196] Nicht zuletzt wird verhindert, dass qualitativ minderwertige KI-Systeme Schäden für die Allgemeinheit verursachen.[197] Das mit dem Sammeln der urheberrechtlich geschützten Trainingsdaten verfolgte Innovationsinteresse ist damit auch von gesamtgesellschaftlicher Relevanz.[198] Es ist demzufolge ein öffentliches Interesse, dass eine vergütungsfreie Nutzung grundsätzlich rechtfertigen kann.

Weiterhin gegen einen angemessenen Ausgleich der betroffenen Urheber spricht, dass beim Web Scraping vorrangig solche Quellen untersucht werden, die im *Surface Web*, also dem Teil des Internets verfügbar sind, der ohne weiteres für Webcrawler und damit auch für die Öffentlichkeit verfügbar ist.[199] Damit sind auch die für das KI-Training durch Webcrawler vervielfältigten Werke überwiegend bereits öffentlich verfügbar gewesen.

194 Dazu bereits siehe oben unter 2. Teil B. I. 1.

195 *Hofmann*, WRP 2024, 11 Rn. 25; *Lennartz/Kraetzig*, GRUR 2024, 1258 (1260).

196 COM(2018) 237 final, S. 12; vgl. außerdem *European Commission u. a.*, Study on copyright and new technologies: copyright data management and artificial intelligence (SMART 2019/0038), S. 208 f.

197 *Dregelies*, GRUR 2024, 1484 (1484).

198 In diese Richtung auch *Dregelies*, GRUR 2024, 1484 (1484).

199 Vgl. zu Web Crawling und dem Hidden Web *Gupta/Bhatia*, IJCTT 2014, 111 (111); zum Begriff des Surface Web siehe *Lewandowski*, in: Kuhlen/Lewandowski/Semar/Womser-Hacker, Grundlagen der Informationswissenschaft, S. 391 (395).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Da demzufolge zumeist bereits Zugang zum Werk bestand, könnte hieraus geschlussfolgert werden, dass durch das Web Scraping für die betroffenen Urheber im Regelfall kein ausgleichsbedürftiger Schaden entsteht.[200] In diesem Kontext spricht ebenfalls gegen eine angemessene Vergütung, dass Urheber von im Internet abrufbaren Werken auch die Möglichkeit haben, ihre Werke mit Hilfe digitaler Rechtemanagementsysteme gegen den Zugriff Dritter zu schützen. Hierzu gehört auch, den Zugriff und in der Folge auch das KI-Training auf Basis der Werke von einer angemessenen Vergütung abhängig zu machen.[201] Mit Blick auf den vom Urheberrecht zur Verfügung gestellten Rechtsrahmen könnten für Trainingsdatensammler, die Werke nur kurzzeitig zum Training von KI abspeichern, Abgrenzungsschwierigkeiten zwischen der vergütungsfreien Schrankenbestimmung des § 44a UrhG für vorübergehende Vervielfältigungshandlungen und einer vergütungspflichtigen Freistellung des Web Scrapings von Werken als Trainingsdaten entstehen, die zu Rechtsunsicherheit führen würden.[202] Weiterhin besteht die Gefahr, dass die jeweils anteilige Anrechnung des mit der KI erzielten Gewinnes auf die große Zahl betroffener Urheber mitunter dazu führen könnte, dass die Vergütungshöhe für einzelne Urheber unerheblich und zudem praktisch schwer zu bestimmen ist.[203]

Schließlich könnte einer Vergütungspflicht für das Web Scraping urheberrechtlich geschützter Werke zum Training von KI zuletzt entgegengehalten werden, dass viele der betroffenen Urheber aufgrund der niedrigen urheberrechtlichen Schutzvoraussetzungen[204] und seiner Formfreiheit[205] ihre Werke wie beispielsweise Beiträge in sozialen Netzwerken oder kurze Blogposts gar nicht als Werke verwerten. Das führt dazu, dass sie weder in einer Verwertungsgesellschaft organisiert sind, noch ein Interesse daran haben, ihre urheberrechtlichen Befugnisse aktiv wahrzunehmen. Infolge-

200 So für das allgemeine, nicht Trainingsdaten-spezifische Text und Data-Mining *de la Durantaye*, Allgemeine Bildungs- und Wissenschaftsschranke, S. 243; *de la Durantaye*, GRUR 2017, 558 (562), mit Verweis auf BT-Drs. 18/12329, S. 41. Ähnlich *Stieper*, ZUM 2021, 776 (778), nachdem der Schaden für Rechtsinhaber auch in nicht wissenschaftsbezogenen Fällen minimal sei.

201 *Hofmann*, WRP 2024, 11.

202 *Dreier*, in: Dreier/Schulze, Urheberrechtsgesetz, UrhG § 44b Rn. 14; BT-Drs. 19/27426, S. 88.

203 *Jager*, Artificial Creativity?, S. 371.

204 Selbst ein Satz oder einzelne Satzteile können urheberrechtlichen Schutz begründen, siehe EuGH, Urt. v. 16.07.2009 - C-5/08, GRUR 2009, 1041 Rn. 47 – Infopaq.

205 Art. 5 Abs. 2 S. 1 RBÜ, dazu ferner *de la Durantaye/Hofmann*, ZUM 2021, 873 (876 f.).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

dessen besteht ein Risiko, dass die von Verwertungsgesellschaften eingezogene Vergütung nicht zu einer Entschädigung der tatsächlich betroffenen Urheber führt.[206] Insgesamt sprechen damit einige Gründe dafür, vom Web Scraping betroffenen Urhebern keinen finanziellen Ausgleich für das Web Scraping ihrer Werke zum Training von KI zu gewähren.[207]

b) Argumente für eine Vergütungspflicht

Allerdings müssen für das KI-Training gesammelte Inhalte vor dem eigentlichen Trainingsverfahren zumindest annotiert, häufig aber auch in qualitativer Hinsicht aufbereitet werden. Dieser Prozess ist ressourcen- und insbesondere zeitintensiv.[208] Darüber hinaus müssen Datensätze unter Umständen auch nach Abschluss des Trainingsverfahrens aufbewahrt werden, um Anforderungen an die Nachvollziehbarkeit der durchgeführten Forschung und Entwicklung zu erfüllen sowie ein Weitertraining des KI-Modells im Anwendungsprozess zu ermöglichen.[209] Intermediäre Trainingsdatensammler speichern die Werke außerdem, um Trainingsdatensätze zu erstellen, die nicht nur vorübergehend im Sinne des § 44a UrhG auf dem Trainingsdatenmarkt angeboten werden. Die für das KI-Training gesammelten Daten werden damit regelmäßig nicht nach einer unerheblichen Zeit automatisch wieder gelöscht. Das betrifft auch Werke, die für das Training von KI gespeichert worden sind. Die Voraussetzungen der urheber-

206 *Triaille/Meeûs d'Argenteuil/Francquen*, Study on the legal framework of text and data mining (TDM), S. 103 f.

207 So für das allgemeine Text und Data-Mining *de la Durantaye*, GRUR 2017, 558 (562); *Heesen/Jüngels*, RuZ 2021, 45 (53); *Raue*, CR 2017, 656 (661); *Schack*, ZUM 2016, 266 (282); *Schack*, GRUR 2021, 904 (907); *Stieper*, ZUM 2021, 776 (778); *Stieper*, GRUR 2020, 1 (4); *de la Durantaye*, Allgemeine Bildungs- und Wissenschaftsschranke, S. 243.

208 Mit einem anschaulichen Beispiel *Paaß/Hecker*, Künstliche Intelligenz, S. 15 f.; *Tenyks*, The Unseen Cost of "Low Quality" Large Datasets.

209 Teilweise ist in Konzepten für das sogenannte Data Lifecycle Management in Bezug auf die KI-Entwicklung und den Umgang mit Trainingsdaten überhaupt keine Löschung der gesammelten Trainingsdaten vorgesehen, siehe beispielsweise *Mishra*, Medium/towards data science v. 12.2.2022. In der Norm ISO/IEC 8183 zum Data Lifecycle Management bei der Entwicklung von KI wird die Löschung beispielsweise neben der Archivierung und Umwidmung mit anschließender Weiterverwendung als eine der Möglichkeiten für das sogenannte Data Decommissioning genannt, siehe dort unter Punkt 6.10. Die Norm empfiehlt weiterhin, bestimmte, selbst festzulegende Datenkategorien für ein späteres Systemaudit aufzubewahren.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

rechtlichen Schrankenbestimmung für vorübergehende Vervielfältigungen (§ 44a UrhG) sind praktisch demzufolge im Regelfall nicht erfüllt.[210] Die befürchteten Abgrenzungsprobleme bestehen demzufolge zwar beim Web Scraping von sonstigen Inhalten zum Zwecke der Informationsextraktion, nicht aber beim Web Scraping von Werken, die als Trainingsdaten für KI verwendet werden sollen.

Auch daraus, dass bereits Zugang zum Werkstück besteht, kann nicht geschlussfolgert werden, dass Urheber durch das Web Scraping keinen ausgleichsbedürftigen Schaden erleiden würden. Denn zum einen führt die bereits bestehende Zugänglichkeit eines Werkes nicht dazu, dass Nutzer auf dieser Basis nicht weitere Verwertungshandlungen vornehmen können, die im Folgenden zu einer Intensivierung der Beeinträchtigung der Urheberinteressen führen. Aus diesem Grund unterliegt zum Beispiel auch die Privatkopiefreiheit Begrenzungen (§ 53 UrhG), auch wenn hierfür denklogisch bereits Zugang zu einem – nicht offensichtlich rechtswidrig hergestellten – Werkstück bestehen muss.[211]

In diesem Zusammenhang kann zugunsten der Urheber außerdem berücksichtigt werden, dass das Innovationsinteresse der Allgemeinheit im Gegensatz zum von Art. 17 EU-GRCh beziehungsweise subsidiär Art. 14 GG umfassten Urheberrecht nicht von einem eigenen, speziellen Freiheitsgrundrecht geschützt ist. Nicht die Vergütungsfreiheit, sondern der Werkzugang ist außerdem relevant für das Vorantreiben von Innovationen. Zugang kann aber im Zusammenspiel mit einer angemessenen Vergütung der Urheber erreicht werden.

KI-Entwicklung wird wegen des hohen Datenaufwandes im Übrigen insbesondere von gewerblich handelnden Akteuren durchgeführt. Wirtschaftliche Individualinteressen bei der Entwicklung von KI können jedoch aus sich heraus keine vergütungsfreie Nutzung rechtfertigen. Dass eine Verquickung von Allgemeininteressen wie dem Interesse an Innovationen durch KI-Entwicklung mit starken wirtschaftlichen Individualinteressen im Regelfall zu einer Vergütung derjenigen Urheber führt, die von der Freistellungsregelung betroffen sind, zeigen in systematischer Hinsicht auch die Schrankenbestimmungen zur Herstellung von Sammlungen für den

210 Zu den Voraussetzungen der Schrankenbestimmung aus § 44a UrhG vgl. EuGH, Urt. v. 16.07.2009 - C-5/08, GRUR 2009, 1041 Rn. 62 ff. – Infopaq; zu § 44a UrhG und dem Training von KI vgl. ebenfalls *Dreier*, in: Dreier/Schulze, Urheberrechtsgesetz, UrhG § 44a Rn. 4. Im Detail zu § 44a UrhG siehe auch 3. Teil A. III. 1.

211 *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 58.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

religiösen Gebrauch (§ 46 UrhG) sowie zur Herstellung von Schulbüchern (§ 60b UrhG).

Insbesondere bei der Entwicklung generativer KI-Systeme werden urheberrechtlich geschützte Werke außerdem als Trainingsdaten verwendet, damit nach Abschluss des Trainingsprozesses mit den verwendeten Werken ähnliche oder identische Erzeugnisse generiert werden können. Die später generierten Erzeugnisse sind damit grundsätzlich geeignet, die zum Training verwendeten Werke zu ersetzen. Ziel der Generierung des Outputs generativer KI ist typischerweise außerdem eine menschliche Rezeption der erzeugten Inhalte. Damit beeinträchtigen KI-Entwickler oder -Anwender, die mit solchen Systemen am Markt auftreten, die Kernverwertungsmöglichkeiten betroffener Urheber. Gerade in diesen Fällen sollten Urheber nicht die vergütungsfreie Nutzung ihrer Werke dulden müssen.[212] Weil Trainingsdatensammler mit den Anbietern solcher KI-Systeme auf vertraglicher Ebene unmittelbar oder mittelbar verbunden sind und sie durch die Datenaggregation einen konkreten Einfluss darauf haben, welche Urheber von den Auswirkungen des generativen KI-Systems betroffen sind, kann diese Beeinträchtigung mittelbar auch den Trainingsdatensammlern zugerechnet werden.

Zugunsten der Urheber muss in diesem Zusammenhang auch berücksichtigt werden, dass sich der wirtschaftliche Trainingsvorteil im später folgenden Vorteil fortsetzt, der aus der Anwendung eines KI-Systems generierbar ist. Allein durch die Anwendung generativer KI-Systeme, die nicht unwesentlich auf Grundlage urheberrechtlich geschützter Werke trainiert werden, ergibt sich eine Produktivitätssteigerung von 2,4 bis 4,4 Billionen US-Dollar jährlich.[213] Diese Wertschöpfung geht auch auf die Nutzung urheberrechtlich geschützter Werke zurück. Urheber sollten an ihr daher auch beteiligt werden. Dass eine solche Beteiligung entlang der Wertschöpfungskette dem urheberrechtlichen Regelungskomplex systemimmanent ist, zeigen nicht zuletzt auch die Regelungen zum Urhebervertragsrecht und insbesondere § 32a UrhG.

Unabhängig vom Output des KI-Systems können gewerblich tätige Trainingsdatensammler durch die Vermarktung ihres Produktes, beispielsweise einer KI-Anwendung, außerdem Einkünfte erzielen. Gleiches gilt für

212 *European Commission u. a.*, Study on copyright and new technologies: copyright data management and artificial intelligence (SMART 2019/0038), S. 207; ebenso *Schack*, NJW 2024, 113 Rn. 24, 26; sich anschließend außerdem *Senftleben*, IIC 2023, 1535 (1538); vgl. zudem *Wandtke*, MMR 2017, 367 (368).

213 *McKinsey*, The economic potential of generative AI: The next productivity frontier.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

intermediäre Trainingsdatensammler, die Trainingsdatensätze an KI-Entwickler veräußern. Der Trainingsdatenmarkt hatte im Jahr 2024 ein Volumen von knapp unter drei Milliarden US-Dollar.[214] Teil hiervon sind auch urheberrechtlich geschützte Werke. Lediglich die Urheber, deren Werke die Entwicklung der KI als Trainingsgrundlage ermöglicht haben, würden bei der vergütungslosen Freistellung des Web Scrapings nicht an der Generierung von Einnahmen auf Basis ihrer eigenen Werke teilhaben. Dieser Value-Gap[215] könnte durch eine angemessene Ausgleichszahlung der wirtschaftlichen Profiteure, also der KI-Entwickler und intermediären Trainingsdatensammler, geschlossen werden. Das Werke auch Teil des Trainingsdatenmarktes sind, spricht dafür, sie an dieser Wertschöpfung auch partizipieren zu lassen.

Selbst wenn darüber hinaus ein gewisses Risiko bestünde, dass eine von Verwertungsgesellschaften eingezogene Vergütung nicht zu einer Entschädigung der tatsächlich betroffenen Urheber führte, ändert das aber nichts daran, dass betroffenen Urhebern aus den oben genannten Gründen ein Ausgleich für die Werknutzungen zustehen muss. Keine Vergütungspflicht vorzusehen, weil mitunter praktische Schwierigkeiten bei der Verteilung der Einnahmen bestehen, bringt betroffene Urheber in eine umso schlechtere Stellung, wobei das Problem selbst, die Wahrnehmung der Urheberrechte und insbesondere der Vergütungsansprüche, nicht adressiert wird. Gleiches gilt für den Fall, in dem Urheber wegen des geringen Einfluss ihres einzelnen Werkes auf das Training einer KI insgesamt nur eine unwesentliche Vergütung erhalten können.

c) Schlussfolgerung

Alles in allem sprechen damit durchgreifende Gründe für eine vergütungspflichtige Freistellung des Web Scrapings von Werken zum Training von KI. Das Innovationsinteresse der Trainingsdatensammler ist infolgedessen kein gesteigertes öffentliches Interesse, das eine vergütungsfreie Nutzung der Werke zum Training von KI ausnahmsweise rechtfertigt.[216] Demzufol-

214 *Fortune Business Insights*, AI training data set market.

215 Im hier verwendeten Sinne eine Diskrepanz zwischen den aus dem Umgang mit dem Werk erzielten Einnahmen der KI-Wirtschaft und der Vergütung der Urheber, deren Werke für das Training einer KI verwendet wurden, siehe zum Begriff m. w. N. *Wimmers/Barudi*, GRUR 2017, 327 (328).

216 Vgl. zu einer Vergütungspflicht auch *de la Durantaye*, ZUM 2023, 645 (659).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

ge ist die regulatorisch zweckmäßige Freistellung des Web Scrapings von urheberrechtlich geschützten Trainingsdaten nur gegen eine angemessene Vergütung der betroffenen Urheber vorzusehen.[217]

2. Notwendigkeit eines Opt-out-Mechanismus

Fraglich ist weiterhin, ob die Freistellung des Web Scrapings im Wege einer gesetzlichen Lizenz oder mit Hilfe eines ECL-Mechanismus vorgenommen werden sollte. Dafür sind Aspekte der Rechtssicherheit und die Notwendigkeit eines Opt-out-Mechanismus entscheidend.

a) Vorteile eines Opt-out-Mechanismus

Für einen Opt-out-Mechanismus, mit dessen Hilfe betroffene Urheber das Web Scraping ihrer Werke verhindern können, spricht zunächst, dass Urheber hierdurch in die Lage versetzt werden, ihre Sekundärinteressen vor der Beeinträchtigung durch Trainingsdatensammler zu schützen. Enthält beispielsweise ein Schriftwerk autobiografische Züge, so möchte der Urheber unter Umständen vermeiden, dass sein Werk zum Training einer KI verwendet wird. So kann er verhindern, dass das KI-System persönlichkeitsrelevante Inhalte aus seinem Werk nachbildet. Dieses Ziel wird erreicht, indem die Freistellung des Web Scrapings der betreffenden Werke durch Erklärung des Urhebers (rückwirkend) beendet wird.

---

217 Auch befürwortet von *Dreier*, in: Dreier/Schulze, Urheberrechtsgesetz, UrhG § 44b Rn. 14; *Dregelies*, GRUR 2024, 1484 (1486); *Geiger*, IIC 2024, 1129 (1152); *Geiger*, in: FS Hilty, S. 67 (79 f.); *Kowala*, IIC 2024, 1604 (1618); *Papastefanou*, CR 2023, 1 Rn. 49; *Schack*, NJW 2024, 113 Rn. 24, 26 (vor dem Hintergrund des Dreistufentests); für kommerzielles KI-Training auch *Wulf/Löbeth*, GRUR 2024, 737 (740 f.); im Grundsatz für eine Vergütungspflicht, aber mit anderer Ausgestaltung (Anknüpfung an den Betrieb des KI-Systems) auch *Senftleben*, IIC 2023, 1535 (1549 ff.); *Jacques/Flynn*, GRUR Int 2024, 1137 (1141, 1148), dazu weitergehend siehe unten unter 4. Teil A. II. 4.; weiterhin im Hinblick auf allgemeines Text und Data-Mining auch *Spindler*, ZGE 2018, 273 (294); *Wandtke*, MMR 2017, 367 (368); *Hilty/Moscon*, Modernisation of the EU Copyright Rules, S. 142. anders aber *Becker*, GRUR 2024, 505 (507); *Hofmann*, WRP 2024, 11 Rn. 20; *Konertz/Schönhof*, WRP 2024, 289 Rn. 35; *Raue*, ZUM 2021, 793 (797); *Jager*, Artificial Creativity?, S. 371: nicht kalkulierbar, für einzelne Urheber zu vernachlässigen; in Bezug auf allgemeines Text und Data Mining, aber zur neuen Schrankenregelung auch *Stieper*, GRUR 2020, 1 (4).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Darüber hinaus eröffnet ein Opt-out-Mechanismus den vom Web Scraping betroffenen Urhebern aber auch die Möglichkeit, statt einer gesetzlichen Freistellung Individuallizenzen für die Nutzung ihrer Werke zum Training von KI zu erteilen.[218] Dies ist insbesondere für Urheber von Bedeutung, die nur bereit sind, Lizenzverträge zur Nutzung ihrer Werke für KI-Training zu einem höheren als dem marküblichen Preis abzuschließen. Letztlich schützt ein Opt-out-Mechanismus damit auch das Selbstbestimmungsrecht der Urheber.[219]

Ein Opt-out-Mechanismus würde schließlich den Vorteil mit sich bringen, dass Urhebern als Ultima Ratio ein Werkzeug zur Verfügung gestellt würde, mit dessen Hilfe sie in die Lage versetzt wären, die Nutzung ihrer Werke als Trainingsdaten für solche KI-Systeme zu verhindern, die ihre Kernverwertung besonders intensiv beeinträchtigen.

b) Praktische Schwierigkeiten und weitere Nachteile eines Opt-out-Mechanismus

Allerdings muss auch berücksichtigt werden, dass der Abschluss individueller Lizenzvereinbarungen für das automatisierte Web Scraping von urheberrechtlich geschützten Trainingsdaten häufig de facto nicht möglich oder aufgrund der Zahl der betroffenen Rechtsinhaber jedenfalls unzumutbar ist.[220] Da es praktisch demzufolge kaum zu Individuallizenzen kommen wird, besteht für Urheber infolgedessen kein Spielraum, im Rahmen dessen von der erweiterten Selbstbestimmung Gebrauch gemacht werden könnte. Im Gegenteil führen Schwierigkeiten bei der Identifizierung von Trainingsdatensammlern, die Werke ohne Erlaubnis der Urheber abspeichern, dazu, dass die trotz einer Opt-out-Erklärung stattfindende Speicherung von Werken zum Training von KI nur mit erheblichem Aufwand verhindert und durch Schadensersatzforderungen monetär ausgeglichen werden kann. Eine gesetzlich ausgestaltete und pauschalisierte Vergütung erreicht, unabhängig von konkreten Nutzungsnachweisen, alle Urheber.[221] Sie sind

218 ErwGr. 18 S. 2 DSM-RL.

219 *Raue*, ZUM 2021, 793 (796).

220 Dazu bereits ausführlicher siehe oben unter 2. Teil B. I. 2.

221 Alle bereits entstandenen und zukünftigen Ansprüche aus der Hersteller- und Betreibervergütung werden bei Meldung eines Werkes durch Urheber pauschal und einmalig abgegolten, siehe beispielsweise § 3 Nr. 12 Verteilungsplan *VG Wort*.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

mit ihr demzufolge bessergestellt als durch einen Opt-out-Mechanismus.[222] Weiterhin ist der Schutz sonstiger, beispielsweise allgemein-persönlichkeitsbezogener Interessen abseits der geistigen und persönlichen Beziehung des Urhebers zu seinem Werk nicht Zweck des Urheberrechts.

Abgesehen davon spricht gegen einen Opt-out-Mechanismus außerdem, dass eine Vorbehaltserklärung insbesondere bei digitalen Werkwiedergaben (im Internet) mit geringem Aufwand abgegeben werden kann.[223] Neben den faktischen Hürden, die die praktische Wirksamkeit der Vorbehaltserklärung betreffen, droht der Freistellungsmechanismus damit de jure leer zu laufen.[224]

Urheber erklären einen Nutzungsvorbehalt zudem ex ante. Mit Beginn einer öffentlichen Wiedergabe der Werke haben sie also noch keine Kenntnis darüber, für welche Arten von KI-Systemen die Werke als Trainingsdaten später einmal genutzt werden sollen. Dies erschwert ihnen die effektive und zielgerichtete Anwendung des Opt-out-Mechanismus als Ultima Ratio zur Verhinderung von Beeinträchtigungen ihrer Kernverwertungsarten. Hinzu kommt, dass gerade urheberrechtlich geschützte Inhalte „kleiner Münze" wie Beiträge in sozialen Netzwerken oft über Plattformen von Drittanbietern und nicht im Herrschaftsbereich der Urheber selbst öffentlich zugänglich gemacht werden. Insofern für die Erklärung des Opt-outs in diesen Fällen die Mitwirkung der Plattformbetreiber und Infrastrukturanbieter notwendig ist, käme es durch den Opt-out-Mechanismus infolgedessen nicht zur Stärkung der Selbstbestimmung der Urheber. Zudem ist ein Ausgleich zwischen den Interessen betroffener Urheber und Nutzer bereits im Rahmen der Identifizierung des Freistellungsinteresses vorgenommen worden. Daher ist ein Opt-out-Mechanismus zur Wahrung der Urheberinteressen nicht erforderlich.[225]

c) Schlussfolgerungen

Alles in allem sprechen demzufolge bessere Gründe dafür, bei der Freistellung des Web Scrapings von Werken zum Training von KI keinen Opt-out-Mechanismus für Urheber vorzusehen.[226]

---

222 Auch dazu bereits siehe oben unter 2. Teil B. II. 1. So auch *Schack*, NJW 2024, 113 Rn. 25.

223 *Hacker*, ZGE 2020, 239 (259).

224 Ähnlich auch *Hacker*, ZGE 2020, 239 (259); *Hofmann*, WRP 2024, 11 Rn. 18.

225 A. A. hingegen *Raue*, in: Dreier/Schulze, Urheberrechtsgesetz, VGG § 52 Rn. 1.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

3. Erhöhtes Rechtssicherheitsbedürfnis

Trainingsdatensammler könnten beim Web Scraping der Werke weiterhin ein erhöhtes Rechtssicherheitsbedürfnis haben. Dann wäre für sie eine Lizenzlösung, also eine Freistellung im Wege eines ECL-Mechanismus, vorzugswürdig. Für die Beurteilung des Rechtssicherheitsbedürfnisses sind unter anderem die Folgen eines möglichen Rechtsverstoßes von Bedeutung. Sie beeinflussen maßgeblich das mit dem Web Scraping Werken zum Training von KI einhergehende wirtschaftliche Risiko.

Für ein erhöhtes Rechtssicherheitsbedürfnis der Trainingsdatensammler spricht zunächst, dass das Web Scraping von Werken als Trainingsdaten typischerweise eine Vielzahl von Urhebern betrifft. Bestehen hinsichtlich der zu erfüllenden Freistellungsvoraussetzungen strukturelle Defizite bei der Speicherung der Werke, droht Trainingsdatensammlern, sich Ansprüchen von einer Mehrzahl von Rechtsinhabern ausgesetzt zu sehen. Das führt zu einer relativen Erhöhung des prozessualen Kostenrisikos für Trainingsdatensammler.

Allerdings ist das Prozessrisiko, dem sich Trainingsdatensammler aussetzen, wegen der erschwerten Aufdeckungsmöglichkeiten für Urheber ohnehin begrenzt. Darüber hinaus folgen auf das Web Scraping der Werke stets die Vorverarbeitung und ggf. Annotation des Trainingsmaterials. Sobald die vorbereitenden Maßnahmen abgeschlossen sind, kann der eigentliche Trainingsvorgang beginnen. Die Nutzung von Trainingsdaten folgt strukturell infolgedessen stets dem gleichen Muster. Der Sachverhalt kann wegen seiner konzeptuellen Uniformität damit auch in einer abstrakt-generellen

226 So auch *Geiger*, in: FS Hilty, S. 67 (79 f.); *Hofmann*, WRP 2024, 11 Rn. 20; *Chiou*, JIPITEC 2019, 398 Rn. 23; *Margoni/Kretschmer*, GRUR Int. 2022, 685 (686); *Schack*, NJW 2024, 113 Rn. 25; *Trapova/Gervassis*, Kluwer Copyright Blog v. 16.3.2022; im Ergebnis auch *Brockmeyer*, Text und Data Mining, S. 166 ff.; in diese Richtung, jedoch lediglich bezogen auf allgemeines Text und Data-Mining im nicht-kommerziellen, wissenschaftlichen Kontext außerdem *Triaille/Meeûs d'Argenteuil/Francquen*, Study on the legal framework of text and data mining (TDM), S. 107 ff.; anders *Hacker*, ZGE 2020, 239 (259): „sachgerechter Ausgleich von Verwertungs- und Informationsinteressen".
Auch praktisch kann in einem ersten Fall die Einführung eines Opt-outs beobachtet werden: Nach Protesten einiger der die Plattform *Artstation* nutzenden Künstler, die herausfanden, dass die dort hochgeladenen, digitalen Vervielfältigungsstücke ihrer Werke ohne ihr Wissen zum Training eines KI-Bildgenerators genutzt wurden, führte der Plattformbetreiber *Epic Games* eine Funktion ein, mit Hilfe derer die Künstler der Nutzung ihrer Werke zum KI-Training beim Hochladen manuell widersprechen können, siehe *Bünte*, heise online v. 16.12.2022.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Rechtsnorm abgebildet werden. Es besteht daher kein Konkretisierungsbedarf durch einzelfallspezifische Lizenzverträge.

Weiterhin können die urheberrechtlichen Beseitigungs-, Unterlassungs- und Löschungsansprüche (§§ 97 Abs. 1, 98 UrhG) nur gegen konkrete Urheberrechtsverletzungen gerichtet werden. Die rechtsverletzende Vervielfältigung eines urheberrechtlich geschützten Trainingsdatums führt daher zu Ansprüchen gegen den Trainingsdatensammler in Bezug auf das konkrete Werk als Trainingsdatum. Die trainierte KI[227] oder ein vollständiger Trainingsdatensatz sind von den spezifischen Beseitigungs- oder Löschungsansprüchen damit nicht betroffen.[228] Im Falle einer Schadensersatzhaftung (§ 97 Abs. 2 UrhG) wäre möglicherweise nur der Verkehrswert des betreffenden Werkes als Trainingsdatum zu ersetzen (Lizenzanalogie).[229] Soll hingegen der Verletzergewinn abgeschöpft werden, kommt es maßgeblich darauf an, welcher Gewinn tatsächlich auf der streitigen Urheberrechtsverletzung, also der Verletzung der Rechtsposition am jeweiligen Werk, beruht.[230] Insbesondere Gewinn, der aufgrund anderer Umstände erzielt worden ist, muss dem Trainingsdatensammler verbleiben.[231] Da eine KI mit einer großen Menge von Trainingsdaten angelernt wurde und einzelne, auf rechtsverletzende Weise aggregierte Werke dementsprechend nur geringen Einfluss auf den Trainingsdatensatz insgesamt und damit auch das trainierte KI-System haben, kann der betroffene Urheber nur einen kleinen Teil des mit dem KI-System erzielten Gewinns herausverlangen. Gleiches gilt für das Verhältnis zwischen einzelnen, urheberrechtsverletzend gespeicherten Werken und den Trainingsdatensätzen, in denen Werke in urheberrechtsverletzender Weise enthalten sind. Selbst wenn ein Trainingsdatensammler solchen Ansprüchen ausgesetzt ist, sind seine wesentlichen Investitionen daher nicht bedroht.

Für Trainingsdatensammler ist das wirtschaftliche Risiko bei der Verletzung von Urheberrechten an als Trainingsdaten genutzten Werken alles in allem daher nicht außergewöhnlich hoch. Auch aufgrund der schematisierbaren Nutzungshandlung besteht beim Web Scraping von Werken zum Training von KI infolgedessen kein besonderes Rechtssicherheitsbedürfnis für Trainingsdatensammler, welches eine Freistellung im Wege eines ECL-

227 Dazu im Detail unter 3. Teil B. IV.
228 Wie hier auch *Linke*, in: Kuschel/Asmussen/Golla, Intelligente Systeme - intelligentes Recht: GRUR Junge Wissenschaft Hamburg 2020/2021, S. 179 (197).
229 Vgl. BGH, Urt. v. 22.03.1990 - I ZR 59/88, GRUR 1990, 1008 (1010) – Lizenzanalogie.
230 BGH, Urt. v. 30.01.1959 - I ZR 82/57, GRUR 1959, 379 (380) – Gasparone; *Wolff/Bullinger*, in: Wandtke/Bullinger, Urheberrecht, UrhG § 97 Rn. 76 m. w. N.
231 BGH, Urt. v. 24.07.2014 - I ZR 27/13, GRUR 2015, 269 Rn. 21 – K-Theory.

https://doi.org/10.5771/9783748963462-85 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Mechanismus nahelegt. Stattdessen kann die Freistellung als abstrakt-generelle Rechtsnorm umgesetzt werden.

4. Schlussfolgerung für den Freistellungsmechanismus

Die Freistellung des Web Scrapings von Werken für das Training von KI sollte als abstrakt-generelle Freistellungsregelung realisiert sein. Trotz der Zweckmäßigkeit der Freistellung ist jedoch ein angemessener Ausgleich für die Speicherung der Werke zugunsten betroffener Urheber zu gewährleisten. Ein Opt-out-Mechanismus zur Beendigung der Freistellungwirkung sollte den Urhebern allerdings nicht zur Verfügung gestellt werden. Insgesamt ist die Freistellung des Web Scrapings von Werken zum Training von KI im Urheberrecht daher als gesetzliche Lizenz, also vergütungspflichtige Schrankenbestimmung, vorzusehen.[232]

III. Unterscheidung zwischen kommerzieller und nicht-kommerzieller Tätigkeit

Die Freistellung des Web Scrapings von urheberrechtlich geschützten Trainingsdaten könnte weiterhin unterschiedslos sowohl kommerzielle als auch nicht-kommerzielle Verarbeitungszwecke erfassen. Dagegen spricht zwar, dass das Web Scraping im nicht-kommerziellen Umfeld wegen der einge-

232 So auch *Geiger*, IIC 2024, 1129 (1152 f.); anders aber *Margoni/Kretschmer*, GRUR Int. 2022, 685 (689), die eine unmittelbar begrenzende Ausgestaltung der urheberrechtlichen Verwertungsrechte fordern. Ebenso wohl auch *Lennartz/Kraetzig*, GRUR 2024, 1258 (1260). So zum allgemeinen Text und Data-Mining mit einigen Einschränkungen grundsätzlich auch *de la Durantaye*, Allgemeine Bildungs- und Wissenschaftsschranke, S. 214, 240 f.; *Triaille/Meeûs d'Argenteuil/Francquen*, Study on the legal framework of text and data mining (TDM), S. 99 ff.; außerdem der Unionsgesetzgeber, siehe ErwGr. 8, 18 DSM-RL; eine unmittelbar begrenzende Ausgestaltung der Verwertungsrechte zugunsten des allgemeinen Text und Data-Mining fordern hingegen hingegen *Schack*, ZUM 2016, 266 (269, 282); *Schack*, ZUM 2017, 802 (806); *Schack*, GRUR 2021, 904 (907); *Schack*, Urheber- und Urhebervertragsrecht, S. 311 Rn. 627; *European Copyright Society*, General Opinion on the EU Copyright Reform Package, S. 5; sogar noch weitergehend (Ausschluss jeder digitalen Vervielfältigung aus dem urheberrechtlichen Schutzbereich und Fokus auf die nachfolgenden Verwertungshandlungen) *Cocoru/Boehm*, An analytical review of text and data mining practices and approaches in Europe - Policy recommendations in view of the upcoming copyright legislative proposal, S. 7.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

schränkten Verwertung der gesammelten Trainingsdaten oder entwickelten KI-Systeme eine geringere Belastung für betroffene Urheber bedeutet.[233] Andererseits besteht die Befürchtung, dass eine auch auf kommerzielle Handlungen ausgeweitete Freistellung nicht mit dem Dreistufentest vereinbar ist, insbesondere weil kein bestimmter Sonderfall mehr freigestellt wäre, so wie es Art. 5 Abs. 5 InfoSoc-RL voraussetzt.[234] Allerdings würden sich für kommerziell tätige Trainingsdatensammler und KI-Entwickler Wettbewerbsnachteile im Vergleich zu anderen, für sie günstigeren Rechtsordnungen ergeben. Diese würden zu einer Abwanderung der KI-Industrie führen.[235] Abgrenzungsschwierigkeiten zwischen kommerzieller und nicht-kommerzieller Tätigkeit würden weiterhin Rechtsunsicherheiten hervorrufen, die Abschreckungseffekte hervorruft.[236] Insbesondere ergab die Abwägung von Nutzer- und Urheberinteressen ein umfassendes Freistellungsinteresse aber auch ohne Einschränkungen hinsichtlich der Kommerzialität des Web Scrapings. Das zeigt, dass das Freistellungsbedürfnis auch für kommerzielle Nutzungszwecke uneingeschränkt besteht. Auch aus dieser Perspektive bedarf es daher keiner Einschränkung auf nicht-kommerzielle Tätigkeiten.[237] Neben alldem verdeutlicht das U.S.-Unternehmen *OpenAI*, dass es keinen Garanten dafür gibt, dass ein im nicht-kommerziellen, wissenschaftlich orientierten Kontext trainiertes KI-System nach Abschluss der urheberrechtlich relevanten Nutzungshandlungen und einer Neuausrichtung des Unternehmensziels nicht doch zu kommerziellen Zwecken eingesetzt wird. Gleiches gilt für Ausgründungen aus Hochschulen, die aus innovationspolitischer Sicht auch wünschenswert sind. Dann ist eine Beschränkung der Freistellungswirkung auf die Tätigkeit nicht-kommerzieller Trainingsdatensammler ohnehin obsolet.

---

233 Vgl. auch *Triaille/Meeûs d'Argenteuil/Francquen*, Study on the legal framework of text and data mining (TDM), S. 102 f.

234 *Jager*, Artificial Creativity?, S. 371; vorläufiges Ergebnis zum allgemeinen Text und Data-Mining bei *Triaille/Meeûs d'Argenteuil/Francquen*, Study on the legal framework of text and data mining (TDM), S. 102 f.

235 Vgl. *Raue*, GRUR 2017, 11 (15 f.); *Hargreaves u. a.*, Standardisation in the area of innovation and technological development, notably in the field of text and data mining, S. 64; *Senftleben*, IIC 2023, 1535 (1348).

236 Im Kontext von Text und Data-Mining zu sontigen Datenanalysezwecken *Cocoru/Boehm*, An analytical review of text and data mining practices and approaches in Europe - Policy recommendations in view of the upcoming copyright legislative proposal, S. 8; *Hilty/Moscon*, Modernisation of the EU Copyright Rules, S. 129.

237 Für das allgemeine Text und Data-Mining ähnlich auch *European Copyright Society*, General Opinion on the EU Copyright Reform Package, S. 5.

https://doi.org/10.5771/9783748963462-85 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Transaktionskosten müssen zudem auch für kommerzielle Trainingsdatensammler gesenkt werden.[238] Auch im Rahmen ihrer Tätigkeit besteht ein erhebliches Innovationspotential.[239] Auch wenn sich die Freistellungswirkung auf kommerziell tätige Trainingsdatensammler erstreckt, ist die Freistellung darüber hinaus dennoch inhaltlich auf das Web Scraping von Werken zum KI-Training beschränkt. Das ist eine abgegrenzte Nutzungsart. Sie kann nicht deswegen kein Sonderfall im Sinne des Dreistufentests mehr sein, weil mehr Werknutzer im Sonderfall von der Freistellungswirkung profitieren. Ob ein Sonderfall vorliegt, richtet sich nutzungsbezogen nach der Reichweite der Freistellung und nicht nach dem von der Regelung angesprochenen Adressatenkreis.[240] Demzufolge wäre eine Ausweitung der Freistellungswirkung auf kommerziell tätige Werknutzer auch mit dem Dreistufentest vereinbar.

Insgesamt überwiegen damit die Gründe dafür, das Web Scraping von Werken zum Training von KI gleichermaßen für kommerziell und nichtkommerziell tätige Trainingsdatensammler freizustellen.[241]

IV. Zwischenergebnis

Das Web Scraping von Werken, um sie nach einer Aufbereitung als Trainingsdaten für die Entwicklung von KI zu verwenden, kann sowohl von KI-Entwicklern selbst als auch von intermediären Trainingsdatensammlern durchgeführt werden. Letztere stellen Trainingsdatensätze zusammen, die auf dem Markt für Trainingsdaten erworben werden können.

Das Web Scraping beeinträchtigt die Verwertungsinteressen betroffener Urheber nur geringfügig. Gleichzeitig verfolgen Trainingsdatensammler schutzwürdige Innovationsinteressen. Abschlüsse von Lizenzverträgen zur Nutzung der Werke sind aufgrund praktischer Schwierigkeiten allerdings nicht möglich. Demzufolge sollte das Web Scraping von Werken zum Trai-

238 Zum allgemeinen Text und Data-Mining *Raue*, CR 2017, 656 (656).

239 So auch für das allgemeine Text und Data-Mining *Hilty/Moscon*, Modernisation of the EU Copyright Rules, S. 129; ebenfalls der Richtliniengeber, siehe ErwGr. 18 UAbs. 1 DSM-RL.

240 Vgl. ähnlich *Loewenheim*, in: Schricker/Loewenheim, Urheberrecht, UrhG § 44a Rn. 14, der von „eng umrissenen Tatbeständen" spricht.

241 So auch *Hacker*, ZGE 2020, 239 (269); a. A. allerdings *Jager*, Artificial Creativity?, S. 371; anders zum unspezifischen Text und Data-Mining außerdem *Triaille/Meeûs d'Argenteuil/Francquen*, Study on the legal framework of text and data mining (TDM), S. 99 f.

https://doi.org/10.5771/9783748963462-85 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

ning von KI urheberrechtlich freigestellt sein. Die Freistellung ist im Wege einer vergüteten Schrankenbestimmung umzusetzen. Von der Schrankenbestimmung sollten sowohl kommerzielle als auch nicht-kommerzielle Trainingsdatensammler profitieren. Ein Opt-out-Mechanismus, mit dem Urheber die Freistellungswirkung beenden können, ist hingegen nicht vorzusehen.

C. Durchführung des Trainings auf Basis urheberrechtlich geschützter Trainingsdaten

Nach der Speicherung der für das Training bestimmten Werke mittels Webcrawlern wird auf Basis der Werke der eigentliche Trainingsvorgang durchgeführt. Dieser beginnt im Regelfall mit einer Vorverarbeitung der zu verwendenden, urheberrechtlich geschützten Trainingsdaten. Einzelfallabhängig werden hierbei erhebliche Änderungen an den für das Training bestimmten digitalen Werkexemplaren vorgenommen. An die Vorverarbeitung der Werke schließt sich der eigentliche Trainingsvorgang im engeren Sinne, also die Analyse der Werke durch das KI-System, an. Ziel des Trainingsvorgangs ist eine Konfiguration der im KNN enthaltenen numerischen Parameter. Diese erfolgt in einer Art und Weise, die im Anschluss eine bestmögliche Qualität des vom KI-System generierten Outputs gewährleistet.

Klärungsbedürftig ist auch für die Durchführung dieses Trainingsprozesses auf Basis von Werken, ob und inwiefern regulatorische Freistellungsbedürfnisse bestehen. Entscheidend ist damit die Zweckmäßigkeit der Freistellung. Insofern eine Freistellung des Trainings erfolgen sollte, stehen hierfür verschiedene Freistellungsmechanismen bereit.[242]

242 Zur Herleitung der regulatorischen Entscheidungskriterien insgesamt siehe unter 2. Teil A.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

I. Zweckmäßigkeit der Freistellung

Eine Freistellung des eigentlichen KI-Trainingsprozesses ist zweckmäßig, wenn hierdurch keine Urheberinteressen beeinträchtigt werden oder die betroffenen Urheberinteressen die gegenläufigen Nutzerinteressen im Rahmen einer Abwägung überwiegen. Für die an dieser Stelle vorzunehmende Analyse und Gegenüberstellung der Urheber- und Nutzerinteressen kann dabei insgesamt auf das zur Speicherung urheberrechtlich geschützter Trainingsdaten Ausgeführte verwiesen werden.[243] Die Interessenlage ist insoweit identisch. Daraus folgt, dass die Durchführung des Trainings auf Basis urheberrechtlich geschützter Werke zwar die Verwertungsinteressen betroffener Urheber beeinträchtigt. Dennoch überwiegen die verfolgten Nutzerinteressen in einer Gesamtabwägung. Damit erweist sich eine Freistellung der Durchführung des Trainings auf Basis urheberrechtlich geschützter Werke als zweckmäßig.

II. Freistellungsmechanismus

Die Wahl eines geeigneten Freistellungsmechanismus richtet sich zum einen danach, ob den betroffenen Urhebern als Kompensation für die Freistellung des KI-Trainings auf Basis ihrer Werke eine angemessene Vergütung gewährt werden sollte. Sofern dies der Fall ist, ist eine Freistellung im Wege einer gesetzlichen Lizenz oder eines ECL-Mechanismus umzusetzen. Falls eine Vergütung nicht notwendig ist, ist die Freistellung demgegenüber als unvergütete Schrankenbestimmung oder durch unmittelbar begrenzende Ausgestaltung der urheberrechtlichen Verwertungsrechte interessengerecht. Die Wahl zwischen gesetzlicher Lizenz und ECL-Mechanismus richtet sich insbesondere danach, ob den Urhebern ein Opt-out-Mechanismus zur Verfügung zu stellen ist, mittels dessen sie die Freistellungswirkung beenden können. Von Bedeutung ist außerdem, ob seitens der KI-Entwickler ein besonderes Rechtssicherheitsbedürfnis besteht. Die Wahl zwischen unvergüteter Schrankenbestimmung und unmittelbar begrenzender Ausgestaltung der urheberrechtlichen Verwertungsrechte richtet sich insbesondere danach, inwiefern ein Schutz eingesetzter technischer Schutzmaßnahmen gegen das Training von KI über § 95a UrhG interessengerecht erscheint. Auch die Beweislastverteilung zwischen Urheber und Nutzer ist maßgeblich.[244]

243 Dazu siehe oben unter 2. Teil B. I.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

I. Angemessene Vergütung der Urheber

Den Urhebern ist für eine Beschränkung ihrer absoluten Rechtsposition am Werk grundsätzlich eine angemessene Vergütung zu gewähren, wenn nicht ein gesteigertes öffentliches Interesse ausnahmsweise eine vergütungsfreie Nutzung rechtfertigt.[245] Ob das der Fall ist, muss anhand einer Abwägung ermittelt werden.[246] Dabei ist ausreichend, dass „unter dem Strich" eine angemessene Vergütung der Urheber gewährleistet bleibt.[247]

Mit dem KI-Training wird ein öffentliches Innovationsinteresse verfolgt, sodass eine vergütungsfreie Verwendung der Werke nach einer umfassenden Interessenabwägung grundsätzlich möglich ist.[248] Für eine Vergütung der Urheber bei der Verwendung ihrer Werke zum KI-Training spricht, dass hinsichtlich der Interessenlage kaum Unterschiede zur Generierung, also Speicherung der Werke mit Hilfe von Webcrawlern bestehen, die von Urhebern ebenfalls nur gegen Zahlung einer angemessenen Vergütung zu dulden ist.[249]

Allerdings werden die als Trainingsdaten genutzten Werke beim KI-Training mit dem Ziel analysiert, die numerischen Parameter des KI-Systems zu konfigurieren. Dabei wird die Gesamtheit aller den Trainingsdaten zugrundeliegenden Muster und Zusammenhänge, also Informationen, auf die Parameter des KI-Systems übertragen. Kern des KI-Trainings ist also eine inhalts- und informationsbezogene Analyse der Trainingsdaten. Die Analyse der in einem Werk enthaltenen Informationen ist aber auch Wesensmerkmal des Werkgenusses.[250] Dass die Informationsaufnahme dabei nicht unmittelbar durch einen Menschen stattfinden muss, sondern auch durch Maschinen erfolgen kann, zeigt der urheberrechtliche Schutz von Computerprogrammen, deren „Rezeption" im ersten Schritt durch Compu-

244 Zur Herleitung der Differenzierungskriterien siehe oben unter 2. Teil A. II. 2.
245 BVerfG, Beschl. v. 25.10.1978 - 1 BvR 352/71, GRUR 1980, 44 (48) – Kirchenmusik; vgl. auch BVerfG, Urt. v. 11.10.1988 - 1 BvR 743/86 u. a., NJW 1992, 1307 (1308) – Justizvollzugsanstalten.
246 Vgl. *Lauber-Rönsberg*, Urheberrecht und Privatgebrauch, S. 112.
247 BVerfG, Beschl. v. 07.07.1971 - 1 BvR 765/66, GRUR 1972, 481 (483) – Kirchen- und Schulgebrauch; BVerfG, Beschl. v. 25.10.1978 - 1 BvR 352/71, GRUR 1980, 44 (46)– Kirchenmusik; BVerfG, Urt. v. 31.05.2016 - 1 BvR 1585/13, GRUR 2016, 690 – Rn. 80 - Metall auf Metall.
248 Vgl. zur Begründung des öffentlichen Innovationsinteresses oben unter 2. Teil B. II. 1.
249 Dazu 2. Teil B. II. 1.
250 So auch *Sucker*, Der digitale Werkgenuss im Urheberrecht, S. 39 f., 42 (Werkgenuss im engen Sinne).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

ter erfolgt.[251] Der (automatisierte) Werkgenuss und das KI-Training sind als inhaltsbezogene Analysen der Werke demzufolge miteinander vergleichbar.[252] Hiergegen spricht auch nicht, dass das KI Training den Werkgenuss sowohl quantitativ als auch qualitativ übersteigt.[253] Denn dass die *Art und Weise* zweier Werkverwendungen zu strukturellen Ähnlichkeiten zwischen beiden Verwendungsarten führt, kann nicht dadurch in Frage gestellt werden, in welchem *Maß* die Verwendung der Werke stattfindet. Ein überdurchschnittlich aufmerksamer Vielleser wird ebenso wenig gegenüber einem durchschnittlichen Rezipienten schlechter gestellt.

Die Rezeption urheberrechtlich geschützter Werke ist jedoch vergütungsfrei möglich. Dies spricht dafür, auch die eigentliche Durchführung des KI-Trainings auf Basis von Werken vergütungsfrei zu stellen.

Darüber hinaus konnte gezeigt werden, dass bereits die vorgelagerte Generierung, also Speicherung der urheberrechtlich geschützten Trainingsdaten mit Hilfe von Webcrawlern, nur gegen einer angemessene Vergütung der betroffenen Urheber freigestellt sein sollte.[254] Wenn im Folgenden auch die eigentliche Durchführung des KI-Trainings nur gegen eine Vergütung der Urheber erlaubnisfrei gestellt würde, drohte eine Doppelvergütung des wirtschaftlich einheitlichen Sachverhalts zulasten der KI-Entwickler, die sowohl die Werke mit Hilfe von Web Scraping in Trainingsdatensätzen speichern als auch das KI-Training durchführen.

Außerdem ist durch die Entschädigung für die Speicherung der Werke bereits „unter dem Strich" eine angemessene Vergütung betroffener Urheber für die Verwendung der Werke zum Training von KI gewährleistet. Insgesamt sprechen in einer Gesamtabwägung damit durchgreifende Gründe dafür, dass das gesteigerte öffentliche Interesse der KI-Entwickler ausnahmsweise eine vergütungsfreie Durchführung des KI-Trainings rechtfertigt. Demzufolge sollte die Freistellung des eigentlichen Trainingsvorgangs

251 *Kleinkopf*, Text- und Data-Mining, S. 177 f.

252 Ebenso *Kögel*, DSRITB 2023, 285 (287); *Lennartz/Kraetzig*, GRUR 2024, 1258 (1259); vgl. zum allgemeinen Text und Data-Mining und dem Werkgenuss weiterhin *Bomhard*, in: Ahlberg/Götting/Lauber-Rönsberg, Beck'scher Online-Kommentar Urheberrecht, UrhG § 44b Rn. 3; *Raue*, ZUM 2019, 684 (686); *Cocoru/Boehm*, An analytical review of text and data mining practices and approaches in Europe - Policy recommendations in view of the upcoming copyright legislative proposal, S. 8; im Hinblick auf allgemeines Text und Data-Mining kritisch, letztendlich aber zustimmend allerdings *Kleinkopf*, Text- und Data-Mining, S. 172 ff., 181; a. A. hingegen *Dornis/Stober*, Urheberrecht und Training generativer KI-Modelle, S. 103 ff.

253 *Schack*, NJW 2024, 113 Rn. 9; in diese Richtung wohl auch *Dornis*, GRUR 2021, 784 (103 ff.).

254 Dazu siehe oben unter 2. Teil B. II. 1.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

entweder im Wege einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte oder einer unvergüteten Schrankenbestimmung umgesetzt werden.

2. Schutz technischer Schutzmaßnahmen über § 95a UrhG

Die Wahl zwischen der Freistellung im Wege einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte und einer vergütungsfreien Schrankenbestimmung richtet sich weiterhin insbesondere nach dem Bedürfnis zum Schutz technischer Schutzmaßnahmen über § 95a UrhG, die gegen das KI-Training eingesetzt worden sind. Gegen einen derartigen Schutz spricht insbesondere, dass das KI-Training neue technische Anwendungen hervorbringt und eng mit dem öffentlichen Innovationsinteresse verbunden ist. Der Schutz technischer Schutzmaßnahmen sollte allerdings nur für solche Maßnahmen gelten, die nicht die Entwicklung neuer technischer Geräte behindern.[255] Außerdem sichert der Schutzmechanismus umfassend die Verwertung von Werken im digitalen Umfeld ab.[256] Das KI-Training berührt aber nicht die Kernverwertungsmöglichkeiten betroffener Urheber, weil das Werk nicht als Werk, also um seines schöpferischen Gehalts Willen, verarbeitet wird.

Im Ergebnis sind technische Schutzmaßnahmen zur Verhinderung des KI-Trainings dementsprechend nicht in den Schutzbereich des § 95a UrhG einzubeziehen. Das spricht für eine Umsetzung des regulatorischen Freistellungsinteresses im Wege einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte.

3. Vergleich zum Werkgenuss

Bereits gezeigt werden konnte, dass die eigentliche Durchführung des KI-Trainings mit dem Werkgenuss vergleichbar ist.[257] Er wird nicht von den urheberrechtlichen Verwertungsrechten erfasst.[258] Diese teleologische

255 ErwGr. 48 InfoSoc-RL.

256 Zur Differenz zwischen urheberrechtlichem Schutzbereich und Schutzumfang des § 95a UrhG siehe *Specht-Riemenschneider*, GRUR 2019, 253 (258).

257 Dazu m. w. N. siehe unmittelbar oben unter 2. Teil C. II. 1.

258 *Heerma*, in: Wandtke/Bullinger, Urheberrecht, UrhG § 15 Rn. 10 m. w. N. außerdem bereits BT-Drs. IV/270, S. 28. Nach *Sucker* gilt das jedenfalls für den analogen Werk-

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Parallele spricht ebenfalls dafür, die Durchführung des KI-Trainings auf Basis urheberrechtlich geschützter Trainingsdaten durch eine unmittelbar begrenzende Ausgestaltung der urheberrechtlichen Verwertungsrechte (in den §§ 15 ff. UrhG) freizustellen.

III. Zwischenergebnis

Alles in allem bestehen regulatorische Freistellungsinteressen, aufgrund derer die Durchführung des KI-Trainings auf Basis von Werken urheberrechtlich freigestellt sein sollte. Die vergütungsfreie Freistellung ist im Wege einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte umzusetzen. Die eigentliche Durchführung des KI-Trainings muss jedoch von der Datenaggregation abgegrenzt werden, für die vorangehend bereits das Interesse an einer Freistellung im Wege einer vergüteten Schrankenbestimmung identifiziert werden konnte.

D. Speicherung von urheberrechtlich geschützten Inputdaten

Freistellungsbedürftig könnte darüber hinaus auch die Speicherung von urheberrechtlich geschützten Inputdaten sein.[259] Diese können beispielsweise mittels Sensortechnik wie Kameras oder Mikrofonen oder durch Webcrawler erfasst, aber auch manuell in ein KI-System eingegeben werden. Hierzu gehört auch die Verwendung von geschützten Werken als Prompt eines generativen KI-Systems. Von der Speicherung urheberrechtlich geschützter Trainingsdaten unterscheidet sie sich dadurch, dass in der Anwendungsphase des KI-Systems, in der die Inputdaten verarbeitet werden, nur vorübergehende Speicherungen der Werke notwendig sind.

---

genuss, während der digitale Werkgenuss mittelbar durch die zwangsläufig stattfindenden Vervielfältigungshandlungen vom urheberrechtlichen Ausschließlichkeitsrecht erfasst wird, siehe *Sucker*, Der digitale Werkgenuss im Urheberrecht, S. 47 ff., 89. Diese – zutreffende – folgenorientierte Betrachtungsweise hat aber keinen Einfluss darauf, dass der eigentliche menschliche Analysevorgang, also der Werkgenuss im engeren Sinne, auch im digitalen Kontext nicht dem Ausschließlichkeitsrecht der Urheber zugewiesen ist.

259 Zur Herleitung der regulatorischen Entscheidungskriterien insgesamt siehe unter 2. Teil A.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

I. Freistellungsbedürfnis

Die Freistellung der Speicherung von urheberrechtlich geschützten Inputdaten ist dann vorzunehmen, wenn sie sich insgesamt als zweckmäßig erweist.

1. Beeinträchtigung der Urheberinteressen

Eine Zweckmäßigkeit der Freistellung kann sich zum einen daraus ergeben, dass die Interessen der Urheber, deren Werke als Inputdaten erfasst und verarbeitet werden, im vorliegenden Fall nicht berührt sind. In Betracht kommen dabei neben dem materiellen Verwertungsinteresse auch die immateriellen Anerkennungs-, Integritäts- und Erstveröffentlichungsinteressen betroffener Urheber.

Weil allein durch die Speicherung eines Werkes zur Verwendung als Inputdatum keine Umgestaltung stattfindet, scheidet eine Beeinträchtigung des Integritätsinteresses der Urheber aus. Mangels menschlicher Wahrnehmung der vervielfältigten Inputdaten im KI-System ist auch nicht das mit dem Amortisationsbedürfnis verknüpfte Interesse der Urheber betroffen, öffentlich als Schöpfer des Werkes anerkannt zu werden. Da die Speicherung von als Inputdaten zu verwendenden Werken nur vorübergehend ist und keine Verfügbarmachung der Inhalte gegenüber Dritten erfolgt, ist auch das Erstveröffentlichungsinteresse der Urheber nicht berührt. Insgesamt sind die immateriellen Urheberinteressen von der Speicherung urheberrechtlich geschützter Inputdaten demzufolge nicht beeinträchtigt.

Fraglich bleibt aber, ob die Speicherung der Werke zur Verwendung als Inputdaten die Vermögensinteressen der Urheber beeinträchtigt. Das ist der Fall, wenn sie die Perspektive der Urheber auf eine Amortisation ihrer Schöpfungsleistung schmälert.

Inputdaten verwirklichen ihren Zweck in der Anwendungsphase der KI, ermöglichen also den operativen Einsatz des KI-Systems. Je nach Aufgabenstellung und den hierfür benötigten Inputdaten müssen sie sofort oder zeitnah verarbeitet werden. Das ist beispielsweise bei KI-Systemen der Fall, die audiovisuelle Daten für die Steuerung autonomer Fahrzeuge nutzen. Gleiches gilt für Anwendungen im nachrichtenbezogenen Tätigkeitsbereich, die auf eine zeitnahe Verarbeitung aktueller journalistischer Veröffentlichungen zur automatisierten Erstellung von Pressespiegeln angewiesen sind. Hierbei ist im Vorhinein auch nicht bekannt, welche Werke für die Anwendung

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

des KI-Systems benötigt werden. Dies hängt insbesondere vom konkreten Anwendungsszenario und der Umgebung des KI-Systems ab. Somit ist auch ein vorbeugender individueller Lizenzerwerb nicht möglich. Eine individuelle Lizenzierung der Nutzung erfasster Werke kann in diesen Anwendungsfällen daher regelmäßig nicht erfolgen. Mangels Nachfrage wird es in solchen Anwendungsszenarien daher nicht zu Lizenzabschlüssen für die Nutzung von Werken als Inputdaten kommen. Bei der Verwendung von Werken als Inputdaten für solche KI-Systeme wird die Amortisationsmöglichkeit der Urheber und damit ihr Verwertungsinteresse demzufolge nicht unmittelbar beeinträchtigt.

Allerdings können Inputdaten selbst in den beschriebenen Fallkonstellationen und auch darüber hinaus jedenfalls nach der Anwendung des KI-Systems zu anderen Zwecken weiterverwendet werden. Denkbar ist insbesondere eine Verwendung der Daten zum Training eines KI-Systems.[260] Da den Inhalten im Rahmen solcher Sekundärverwertungen eine wirtschaftliche Bedeutung zukommen kann, ist bereits bei der Speicherung der urheberrechtlich geschützten Inputdaten das Verwertungspotenzial der erfassten Werke mit in den Blick zu nehmen. Aus diesem Grund beeinträchtigt die Speicherung urheberrechtlich geschützter Inputdaten in den skizzierten Fallkonstellationen die Verwertungsinteressen der Urheber mittelbar.

Darüber hinaus ist nicht auszuschließen, dass in anders gelagerten, voraussehbareren und nicht zeitkritischen Fallkonstellationen wie der gezielten Anwendung generativer KI-Systeme Spielräume für individuellen Lizenzerwerb zur Nutzung spezifisch ausgewählter Werke als Inputdaten bestehen. Daraus würde eine unmittelbare Beeinträchtigung der Verwertungsinteressen betroffener Urheber folgen. Demzufolge ergibt sich die Zweckmäßigkeit einer Freistellung der Speicherung urheberrechtlich geschützter Inputdaten nicht schon aus einer fehlendenden Beeinträchtigung der Urheberinteressen.

2. Abwägung der Urheber- und Nutzerinteressen

Eine Freistellung der Speicherung von urheberrechtlich geschützten Inputdaten bei der Anwendung von KI-Systemen könnte aber trotzdem zweckmäßig sein, wenn die Nutzerinteressen die betroffenen Verwertungsinteressen der Urheber im Rahmen einer umfassenden Abwägung überwiegen.

260 Dazu siehe unter I. Teil B. II. 3.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Zugunsten der KI-Nutzer ist das Interesse an einer reibungslosen Anwendung des KI-Systems in die Abwägung einzustellen. Dieses Interesse kann, je nach spezifischem Anwendungsfall, auch durch die Berufsfreiheit (Art. 12 GG) oder andere Freiheitsrechte geschützt sein.

Für die Urheber spricht, dass innerhalb der Sphäre der KI-Anbieter oder -Nutzer eine Kontrolle über die Verwendung der als Inputdaten erfassten Werke nicht mehr möglich ist. Es besteht damit die Gefahr, dass ein Schutz vor Sekundärverwertungen nach der Speicherung der Werke als Inputdaten nicht mehr effektiv sichergestellt werden kann. Gerade diese Sekundärverwertungen können allerdings die Verwertungsinteressen der Urheber beeinträchtigen.

Jedoch dürfte die praktisch häufigste Beeinträchtigung der wirtschaftlichen Interessen der Urheber aus einer sich an die Verwendung der Werke als Inputdaten anschließenden Nutzung zum KI-Training folgen. Gerade für diesen Zweck bestehen aber regulatorische Freistellungsinteressen,[261] sodass ein vorgelagerter Schutz der Urheber in Bezug auf diesen Fall nicht als notwendig erscheint.

Darüber hinaus ist insbesondere bei der unkontrollierten Erfassung von Inputdaten aus Bereichen, in denen nicht ausschließlich urheberrechtlich geschützte Werke verarbeitet werden, weder vorherseh- noch kontrollierbar, ob, zu welchem Umfang und welche Werke konkret verarbeitet werden. Wäre die Speicherung von Werken als Inputdaten in diesen Fällen nicht erlaubnisfrei möglich, hätte das zur Folge, dass der Einsatz der KI-Systeme nur unter Inkaufnahme von Rechtsverletzungen oder insgesamt nicht möglich wäre. Eine solche prohibitive Wirkung kann vom Urheberrecht allerdings nicht ausgehen. Andersherum verhält sich das Urheberrecht, wie beispielsweise ErwGr. 31 InfoSoc-RL oder Art. 3 f. DSM-RL für neue Medien oder automatisierte Text und Data Mining-Prozesse zum Ausdruck bringen, offen gegenüber dem Einsatz neuer Technologien.[262]

Im Übrigen spricht zugunsten der Nutzer, dass die Verwertungsinteressen der Urheber in einigen Fallkonstellationen nur mittelbar betroffen sind. Selbst wenn eine unmittelbare Interessenbeeinträchtigung stattfindet,

261 Siehe oben unter 2. Teil B. I. 2.

262 Auch der EuGH betont regelmäßig die Funktion gewisser urheberrechtlicher Bestimmungen, sowohl die Entwicklung als auch den Einsatz neuer Technologien zu ermöglichen und zu gewährleisten, siehe u.a. EuGH, Urt. v. 04.10.2011 - C-403/08, 429/08, GRUR 2012, 156 Rn. 164 - Football Association Premier League u. Murphy; EuGH, Urt. v. 05.06.2014 - C-360/13, GRUR 2014, 654 Rn. 24 – PRCA/NLA. In diesem Sinne schreibt auch Hofmann: „Die moderne Infrstruktur hängt am Urheberrecht", siehe *Hofmann*, GPR 2017, 176 (176).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

ist die Verwendung von Werken als Inputdaten darüber hinaus nur reintechnischer Natur. Die Kernverwertungsarten, bei denen die menschliche Rezeption der Werke in ihrer Eigenschaft als geistige Schöpfungen im Mittelpunkt steht, werden nicht berührt. Insofern die Kernverwertungsarten nachfolgend durch den Output des KI-Systems beeinträchtigt werden, ist dies auf Ebene des Outputs und für den jeweiligen Einzelfall zu untersuchen. Die Verwendung des generierten Outputs ist von der Verwendung der Werke als Inputdaten zu trennen. Alles in allem wiegen die Interessen der KI-Anbieter beziehungsweise -Nutzer bei der Verwendung von Werken als Inputdaten damit schwerer als die Beeinträchtigung der Verwertungsinteressen der Urheber.

3. Schlussfolgerung

Aus der umfassenden Abwägung von Urheber- und Nutzerinteressen ergibt sich, dass den Nutzerinteressen vor dem Verwertungsinteresse betroffener Urheber Vorrang zu gewähren ist. Damit ist eine urheberrechtliche Freistellung der Speicherung von Werken als Inputdaten insgesamt zweckmäßig.

II. Freistellungsmechanismus

Geklärt werden muss weiterhin, mit Hilfe welches Mechanismus die Freistellung der Speicherung urheberrechtlich geschützter Inputdaten umgesetzt sein sollte. Neben einer vergütungsfreien Schrankenbestimmung und einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte kann hierfür auch auf eine gesetzliche Lizenz oder einen ECL-Mechanismus zurückgegriffen werden.

1. Notwendigkeit eines angemessenen Ausgleichs

Fraglich ist zunächst, ob den betroffenen Urhebern ein angemessener Ausgleich für die Verwendung ihrer Werke als Inputdaten gezahlt werden sollte. Die Notwendigkeit eines angemessenen Ausgleichs für betroffene Urheber spricht dabei für die Freistellung mittels einer gesetzlichen Lizenz oder eines ECL-Mechanismus. Soll ein angemessener Ausgleich hingegen nicht gewährt werden, spricht dies weiterhin für die Freistellung im Wege einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Verwertungsrechte oder einer vergütungsfreien Schrankenbestimmung. Grundsätzlich ist dabei jede Einschränkung des Verfügungsrechts der Urheber ausgleichspflichtig, sofern nicht ausnahmsweise gesteigerte öffentliche Interessen eine vergütungsfreie Nutzung der Werke und damit auch die Einschränkung des urheberrechtlichen Verwertungsrechts rechtfertigen.[263]

Ohne Inputdaten kann eine KI nicht in der Anwendungsphase eingesetzt werden. Sie werden zur Erzeugung eines Outputs benötigt. Inputdaten sind damit stets kausal für die Funktionsfähigkeit von KI. Das gilt demzufolge auch für die vorgelagerte Speicherung der als Inputdaten verwendeten Inhalte. Die Förderung von KI-Innovationen ist darüber hinaus ein öffentliches Anliegen.[264] Innovationen zu fördern, welche anschließend nicht angewendet werden können, würde die Innovationsförderung allerdings ad absurdum führen. Die Anwendung von KI auf Basis von urheberrechtlich geschützten Inputdaten ist demzufolge in das öffentliche Innovationsinteresse mit einzubeziehen. Es kann grundsätzlich eine vergütungsfreie Nutzung rechtfertigen.

Klärungsbedürftig ist im Weiteren aber, ob das öffentliche Innovationsinteresse eine vergütungsfreie Speicherung der urheberrechtlich geschützten Inputdaten tatsächlich rechtfertigt.[265] Das ist im Rahmen einer umfassenden Abwägung zu ermitteln.[266] Dabei ist maßgeblich, ob „unter dem Strich"[267] eine angemessene Vergütung der betroffenen Urheber für ihre Schöpfungsleistung verbleibt.[268] In systematischer Hinsicht gegen eine Vergütung der Urheber spricht, dass das Urheberrecht auch an anderer Stelle (§§ 44a, 69d UrhG) dazu dient, die normale Benutzung eines Werkes sicherzustellen.[269] Weil der Werkzugang in diesen Fällen rechtmäßig erfolgt,

263 BVerfG, Beschl. v. 25.10.1978 - 1 BvR 352/71, GRUR 1980, 44 (48) – Kirchenmusik; vgl. auch BVerfG, Urt. v. 11.10.1988 - 1 BvR 743/86 u. a., NJW 1992, 1307 (1308) – Justizvollzugsanstalten.

264 Dazu siehe oben unter 2. Teil B. II. 1.

265 Vgl. zum Kriterium BVerfG, Beschl. v. 25.10.1978 - 1 BvR 352/71, GRUR 1980, 44 (48) – Kirchenmusik; vgl. auch BVerfG, Urt. v. 11.10.1988 - 1 BvR 743/86 u. a., NJW 1992, 1307 (1308) – Justizvollzugsanstalten.

266 Vgl. *Lauber-Rönsberg*, Urheberrecht und Privatgebrauch, S. 112.

267 BVerfG, Urt. v. 11.10.1988 - 1 BvR 743/86 u. a., NJW 1992, 1307 (176) – Justizvollzugsanstalten; BVerfG, Urt. v. 31.05.2016 - 1 BvR 1585/13, GRUR 2016, 690 Rn. 74 – Metall auf Metall.

268 BVerfG, Beschl. v. 07.07.1971 - 1 BvR 765/66, GRUR 1972, 481 (483) – Kirchen- und Schulgebrauch; BVerfG, Beschl. v. 25.10.1978 - 1 BvR 352/71, GRUR 1980, 44 (46) – Kirchenmusik.

269 Vgl. zu § 44a: *Stieper*, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 8 f.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

kann die rein akzessorische Nutzungssicherung vergütungsfrei erfolgen. Ähnlich könnten die Dinge bei einer Freistellung der Speicherung urheberrechtlich geschützter Inputdaten liegen: Denn auch mit ihr wird der Zweck verfolgt, die urheberrechtlich grundsätzlich nicht beschränkte Funktionalität der KI, also ihre normale Benutzung, sicherzustellen. Die Vervielfältigung der urheberrechtlich geschützten Werke ist nur technisch bedingt.[270]

In einigen Anwendungsfällen ist die Speicherung urheberrechtlich geschützter Inputdaten darüber hinaus nur zufällig und beiläufig. Da es in diesen Fällen mitunter nicht auf die schöpferischen Eigenschaften des Werkes ankommt, schmälert dies den wirtschaftlichen Vorteil, der aus der Verwendung des Werkes als solches gezogen wird. Er bildet die Basis für die Bestimmung einer angemessenen Vergütung.[271]

Insgesamt sprechen damit wesentliche Gründe für eine vergütungsfreie Nutzung der betroffenen Werke. Das öffentliche Innovationsinteresse bei der Speicherung urheberrechtlich geschützter Inputdaten rechtfertigt damit ausnahmsweise eine vergütungsfreie Nutzung. Demzufolge ist für die Speicherung urheberrechtlich geschützter Inputdaten kein angemessener Ausgleich für Urheber vorzusehen.

2. Schutz eingesetzter technischer Schutzmaßnahmen

Mangels Notwendigkeit einer angemessenen Vergütung könnte die Freistellung der Speicherung urheberrechtlich geschützter Inputdaten im Wege einer nicht vergüteten Schrankenbestimmung oder durch eine unmittelbar begrenzende Ausgestaltung der urheberrechtlichen Verwertungsrechte in den §§ 15 ff. UrhG umgesetzt werden.

Bei der Wahl zwischen beiden Regelungsmechanismen ist zum einen entscheidend, ob Urhebern, die ihre Werke trotz der Freistellungsregelung mit Hilfe technischer Maßnahmen gegen die Verwendung als Inputdaten schützen wollen, der Schutz des § 95a UrhG zukommen soll. Der Schutzbereich der Regelung geht über den originären Schutzbereich des urheberrechtlichen Ausschließlichkeitsrechts hinaus und sichert Urhebern umfas-

270 Das war auch schon bei § 44a UrhG maßgeblich, siehe BT-Drs. 15/38, S. 18.

271 Zur Vergütungshöhe bei vergütungspflichtigen Schrankenbestimmungen: *Pflüger*, Gerechter Ausgleich und angemessene Vergütung, S. 145. Für die Tarife der Verwertungsgesellschaften ergibt sich das aus § 39 Abs. 1 S. 1 VGG.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

send die Verwertung ihrer Werke im digitalen Umfeld.[272] Dennoch muss auch beim Schutz von technischen Schutzmaßnahmen der Grundsatz der Verhältnismäßigkeit beachtet werden.[273] Deswegen sind nur solche Maßnahmen schutzfähig, die nicht den normalen Betrieb elektronischer Geräte und deren technische Entwicklung beeinträchtigen.[274] Andersherum dürfen keine Handlungen untersagt werden, deren wirtschaftlicher Zweck nicht in der Umgehung technischer Schutzvorkehrungen besteht.[275] Von Bedeutung ist außerdem, dass ein Nutzer zunächst den Aufwand erbringen muss, eine Nutzung durch die Zurverfügungstellung geeigneter Umgehungsmittel vom Urheber ermöglichen zu lassen (§ 95b UrhG), wenn er technisch geschützte Werke insbesondere auf Basis urheberrechtlicher Schrankentatbestände rechtmäßig nutzen möchte. Der Einsatz technischer Schutzmaßnahmen verlagert demzufolge die Aktionslast auf die Werknutzer.[276]

Gegen einen Schutz technischer Schutzmaßnahmen, die die Nutzung von Werken als Inputdaten für KI verhindern, spricht, dass die Verwertungsinteressen der Urheber in einigen Fallkonstellationen nur mittelbar betroffen sind. Selbst wenn es zu einer unmittelbaren Beeinträchtigung kommt, wird diese stets nur durch eine rein-technische Nutzung der Werke hervorgerufen. Da hierbei nicht die Rezeption der Inhalte bezweckt wird, sind von der Speicherung der Werke als Inputdaten nicht die Kernverwertungsarten der Urheber betroffen. Vor dem Hintergrund des Grundsatzes der Verhältnismäßigkeit erscheint es ungerechtfertigt, ihnen dafür den weitreichenden Schutz des § 95a UrhG zuzubilligen.

Darüber hinaus kann eine KI nicht ohne Inputdaten angewendet werden.[277] Sind KI-Systeme auf eine Verarbeitung von Werken auf Inputebene angewiesen, beeinträchtigen die technischen Schutzmaßnahmen demzufolge den normalen Betrieb dieser elektronischen KI-Systeme. Dass der Betrieb solcher KI-Systeme nicht grundsätzlich rechtlich missbilligt wird, zeigt jedoch bereits die Existenz der KI-VO.

Mit dem Einsatz von KI-Systemen, welche (unter anderem) Werke als Inputdaten verarbeiten, werden darüber hinaus eigene wirtschaftliche Zwecke verfolgt. Diese sind in der jeweiligen Zielfunktion und dem Anwendungsumfeld der KI-Systeme konkretisiert. Der wirtschaftliche Zweck

272 Zur Differenz zwischen urheberrechtlichem Schutzbereich und Schutzumfang des § 95a UrhG siehe *Specht-Riemenschneider*, GRUR 2019, 253 (258).
273 ErwGr. 48 S. 3 InfoSoc-RL.
274 ErwGr. 48 S. 1 InfoSoc-RL.
275 ErwGr. 48 S. 3 InfoSoc-RL.
276 So ebenfalls *Specht-Riemenschneider*, GRUR 2019, 253 (259).
277 Zur technischen Funktionalität siehe oben unter 1. Teil B. II. 2.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

der Verarbeitungshandlungen besteht demzufolge nicht in der Umgehung eingerichteter technischer Schutzvorrichtungen, sondern richtet sich nach dem jeweiligen Zweck des KI-Systems.

Außerdem verlagern technische Schutzmaßnahmen die Aktionslast auf den Werknutzer. Sie müssten gegebenenfalls den Anspruch aus § 95b UrhG durchsetzen. Da die Speicherung und Verarbeitung urheberrechtlich geschützter Inputdaten in diversen Anwendungskonstellationen wie auch das KI-Training automatisch ablaufen, besteht für eine Vielzahl von KI-Anwendern aber keine Möglichkeit, dieser Aktionslast nachzukommen. Allgemein werden KI-Systeme oftmals eingesetzt, um Effizienz- und Temposteigerungen bei der Bewältigung von Anwendungsproblemen zu erzielen. Aus beiden Gründen wäre in einer Vielzahl der Fallkonstellationen nicht mit einer Geltendmachung des Anspruchs aus § 95b UrhG zu rechnen. Diese Verlagerung der Aktionslast auf die KI-Anwender und dessen Legitimation durch § 95a UrhG sind daher nicht interessengerecht.

Aus alldem folgt, dass ein Rechtsschutz für technische Schutzmaßnahmen, deren Zweck es ist, die Speicherung von Werken als Inputdaten für KI zu verhindern, nicht gewährt werden sollte. Das spricht im Ergebnis für die Freistellung des Sachverhalts durch eine unmittelbar begrenzende Ausgestaltung der urheberrechtlichen Verwertungsrechte.

III. Zwischenergebnis

Eine Freistellung der Speicherung urheberrechtlich geschützter Inputdaten ist aufgrund eines Überwiegens der Nutzerinteressen zweckmäßig. Da für diese Freistellung keine Ausgleichsleistung für Urheber vorzusehen ist und von Urhebern eingesetzte technische Schutzmaßnahmen nicht vom Schutz des § 95a UrhG profitieren sollen, ist die Freistellung im Wege einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte umzusetzen. Hierfür kommt es de lege lata insbesondere auf die Reichweite des Vervielfältigungsrechts aus § 16 UrhG an.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

E. Vorverarbeitung urheberrechtlich geschützter Inputdaten durch Normalisierung und Feature Encoding

Freistellungsbedürfnisse könnten sich weiterhin auch für die Vorverarbeitung urheberrechtlich geschützter Inputdaten durch Normalisierung und Feature Encoding ergeben.[278] Dabei werden die Daten auf eine vergleichbare Skala transponiert (normalisiert) und anschließend in eine numerische Darstellungsform überführt („Feature Encoding").

I. Normalisierung der Inputdaten

Analysiert werden muss dabei zum einen die Normalisierung der als Inputdaten erfassten Werke. Ob ein regulatorisches Freistellungsbedürfnis besteht, richtet sich auch hierbei nach der Zweckmäßigkeit der Freistellung.[279] Für eine Umsetzung der Freistellung stehen im Folgenden eine vergütete oder vergütungslose Schrankenbestimmung, der ECL-Mechanismus oder die unmittelbar begrenzende Ausgestaltung der urheberrechtlichen Verwertungsrechte zur Verfügung.[280]

1. Zweckmäßigkeit der Freistellung

Eine Freistellung der Normalisierung von bei der Anwendung eines KI-Systems als Inputdaten erfassten Werken ist zweckmäßig, wenn durch die Normalisierung keine Urheberinteressen beeinträchtigt werden oder die bei der Normalisierung verfolgten Nutzerinteressen die betroffenen Urheberinteressen überwiegen.

Zwar müssen die Inputdaten oft aus dem situationsspezifischen Anwendungsumfeld generiert werden. Damit besteht für Urheber in vielen Fällen kein Markt, um ihre Werke als Inputdaten für KI zu verwerten.[281] Nichtdes-

278 Zur Bedeutung beider technischen Begriffe siehe oben unter 1. Teil B. II. 2.
279 Zur Herleitung der regulatorischen Entscheidungskriterien insgesamt siehe unter 2. Teil A.
280 Zur detaillierten Herleitung und Beschreibung der Entscheidungskriterien siehe oben unter 2. Teil A. II. 2.
281 Dazu bereits oben unter 2. Teil D. I. 1.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

totrotz können auch normalisierte Inputdaten sekundär verwertet werden. Dies umfasst insbesondere auch das Training eines KI-Systems auf Basis der durch die Anwendung der KI gewonnenen Datensätze. Hierfür macht es keinen Unterschied, ob die Daten normalisiert worden sind oder nicht. Ohnehin muss auch für das Training im Regelfall eine Normalisierung der Daten stattfinden. Normalisierte Trainingsdatenkorpora werden mitunter auch bereits öffentlich angeboten.[282] Auch die Herstellung normalisierter Kopien der Inputdaten greift demzufolge mittelbar in die Verwertungsinteressen der Urheber ein.

Darüber hinaus könnte die Normalisierung auch die Integritätsinteressen der als Inputdaten erfassten Urheber beeinträchtigen. Denn bei der Normalisierung der Werke werden beispielsweise Abkürzungen in Textinhalten aufgelöst, Begriffe auf ihren Wortstamm zurückgeführt und die Textinhalte tokeniziert. Bildinhalte können je nach Bedarf skaliert, zugeschnitten und insbesondere hinsichtlich der wiedergegebenen Farbschemata angepasst werden. Die Normalisierung der urheberrechtlich geschützten Trainingsdaten geht also mit einer Veränderung der Inhalte einher. Allerdings wird dem Urheber mit dem Schutz des Integritätsinteresses zum einen das Fortbestehen seines Werkes zugesichert.[283] Dieses ist durch die Normalisierung der Werke aber nicht gefährdet. Darüber hinaus soll verhindert werden, dass der Urheber mit Inhalten identifiziert wird, die in der jeweiligen Form nicht auf ihn zurückgehen.[284] Neben der Veränderung wird dafür allerdings vorausgesetzt, dass das veränderte Werkstück auch in der Öffentlichkeit wahrgenommen wird oder zumindest wahrgenommen werden kann.[285] Zwischen der Normalisierung der urheberrechtlich geschützten Inputdaten und dem Ablauf des Algorithmus auf Basis der vorverarbeiteten Werke gibt es jedoch kein Publikum, welches die durch die Normalisierung umgearbeiteten Darstellungen der Werke rezipieren könnte. Infolgedessen besteht auch keine Gefahr, dass Urheber in diesem Rahmen veränderte Inhalte zugerechnet werden. Demzufolge berührt die die Normalisierung urheberrechtlich geschützter Inputdaten nicht die Integritätsinteressen der Urheber.

282 Beispielsweise der MINT-Datensatz, siehe http://yann.lecun.com/exdb/mnist/ (zuletzt abgerufen am 18.10.2023).

283 Vgl. BGH, Urt. v. 21.02.2019 - I ZR 98/17, ZUM 2019, 508 Rn. 33 – HHole (for Mannheim).

284 BGH, Urt. v. 01.10.1998 - I ZR 104/96, ZUM 1999, 146 (148) – Treppenhausgestaltung.

285 Vgl. auch *Peukert*, in: Schricker/Loewenheim, Urheberrecht, UrhG § 14 Rn. 1, 13.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Insgesamt beeinträchtigt die Normalisierung urheberrechtlich geschützter Inputdaten die Verwertungs-, nicht aber die Integritätsinteressen betroffener Urheber. Um die Zweckmäßigkeit der Freistellung der Datennormalisierung zu bestimmen, bedarf es dennoch einer Abwägung der betroffenen Urheber- und Nutzerinteressen.

Die Abwägung unterscheidet sich dabei inhaltlich nicht von der Abwägung, die für die originäre Speicherung der urheberrechtlich geschützten Inputdaten bei der Anwendung von KI-Systemen durchgeführt wurde.[286] Denn aus den gleichen Gründen, wie auch die Erfassung und Speicherung von Werken als Inputdaten freigestellt sein sollte, muss auch die Normalisierung freigestellt sein, die einer Verwendung der Werke als Input für KI vorausgeht. Wäre das nicht der Fall, würde eine Freistellung der Speicherung von urheberrechtlich geschützten Inputdaten vielmehr ins Leere laufen. Im Ergebnis führt ein Überwiegen der Nutzerinteressen daher dazu, dass auch die Normalisierung der als Inputdaten generierten Werke urheberrechtlich freigestellt werden sollte.[287]

2. Freistellungsmechanismus

Auch bei der Wahl eines geeigneten Freistellungsmechanismus für die Normalisierung urheberrechtlich geschützter Inputdaten kommt es darauf an, ob den Urhebern für die Verwendung ihrer Werke als KI-Input eine angemessene Vergütung gezahlt werden muss. Ist das nicht der Fall, kann im Weiteren danach differenziert werden, welcher der Beteiligten im Verletzungsprozess die Beweislast tragen sollte. Entscheidungserheblich ist außerdem, ob ein Schutz technischer Schutzmaßnahmen über § 95a UrhG interessengerecht ist.[288]

Auch die Normalisierung ist technisch bedingt und erzeugt für KI-Anwender unmittelbar keinen eigenen wirtschaftlichen Vorteil. Die Verwertungsinteressen der Urheber sind nicht in ihrem Kernbereich betroffen. Im Wesentlichen kann betreffend des Freistellungsmechanismus daher wieder auf die Erwägungen zur Speicherung der Inputdaten zurückgegriffen werden.[289] Das öffentliche Innovationsinteresse rechtfertigt daher ausnahmsweise eine vergütungsfreie Nutzung auch bei der Normalisierung von als

286 Dazu siehe unmittelbar oben unter 2. Teil D. I.
287 Zu einem ähnlich gelagerten Problem vgl. auch *Raue*, GRUR 2017, 11 (15, 17).
288 Im Detail zur Herleitung der Entscheidungskriterien siehe oben unter 2. Teil A. II.
289 Dazu oben 2. Teil D. II.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Inputdaten erfassten Werken für die Anwendung von KI.[290] Die Normalisierung der Inputdaten geht darüber hinaus zumeist mit einer Veränderung der Werke einher. Technische Schutzmaßnahmen gegen diese Veränderungen beeinträchtigen demzufolge direkt die Funktionsfähigkeit von KI. Wegen der Automatisierung der Datenverarbeitung bei der Anwendung von KI ist auch keine Verlagerung der Aktionslast auf die Werknutzer interessengerecht. Daher sind technische Schutzmaßnahmen gegen die Normalisierung von als Inputdaten erfassten Werken auch nicht in den Schutzbereich des § 95a UrhG einzubeziehen.[291] Insgesamt sollte die Normalisierung von Inputdaten daher durch eine unmittelbar begrenzende Ausgestaltung der urheberrechtlichen Verwertungsrechte freigestellt sein.

3. Zwischenergebnis

Eine Freistellung der Normalisierung urheberrechtlich geschützter Inputdaten ist, wie auch im Rahmen der zuvor stattfindenden Speicherung der Inhalte, aufgrund überwiegender Nutzerinteressen zweckmäßig. Da diese Freistellung ohne Ausgleichsleistung für Urheber vorzusehen ist und von Urhebern eingesetzte technische Schutzmaßnahmen nicht vom Schutz des § 95a UrhG profitieren sollen, ist die Freistellung im Wege einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte umzusetzen.

II. Feature Encoding der normalisierten Inhalte

An die Normalisierung der urheberrechtlich geschützten Inputdaten schließt sich das Feature Encoding der Inhalte an. Dabei werden sie in ein numerisches Format, im Regelfall einen Tensor, transformiert. Auch das Feature Encoding könnte aus einer urheberrechtsregulatorischen Perspektive freigestellt werden müssen, wenn sich die Freistellung nach den bereits hergeleiteten Kriterien als zweckmäßig erweist.[292]

290 Vgl. oben 2. Teil D. II. 1.
291 Vgl. oben 2. Teil D. II. 2.
292 Zur Herleitung der regulatorischen Entscheidungskriterien insgesamt siehe unter 2. Teil A.

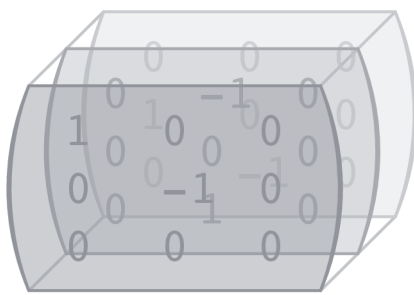
https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

1. Zweckmäßigkeit einer Freistellung

Eine urheberrechtliche Freistellung des Feature Encodings ist zweckmäßig, wenn durch den Verarbeitungsschritt von keine Urheberinteressen berührt werden oder die Nutzer- die gegenläufigen Urheberinteressen im Rahmen einer Gesamtabwägung überwiegen.

Fraglich ist dabei allem voran, ob durch das Feature Encoding der erfassten Werke die Verwertungsinteressen der Urheber beeinträchtigt werden. Hierfür spricht zunächst, dass sowohl bei der Speicherung als auch bei der anschließenden Normalisierung der Inhalte eine Beeinträchtigung der Verwertungsinteressen der Urheber stattfindet. Auch das Feature Encoding ist Teil dieses Prozesses, sodass die bereits angestellten Erwägungen auch auf das Feature Encoding übertragen werden könnten.

Allerdings werden die Werke beim Feature Encoding in ein numerisches Repräsentationsformat überführt. Die gebildeten Tensoren können unterschiedliche Komplexität aufweisen und insbesondere auch mehrdimensional sein. Beispielsweise werden dreidimensionale Tensoren benutzt, um Bildinhalte numerisch darzustellen. Jede Ebene des Tensors repräsentiert dabei einen Kanal des RGB-Farbmodells, mit dem sich die Farbwerte der einzelnen Pixel beschreiben lassen. Jedem Pixel des Bildes ist übereinanderliegend auf jeder Ebene des Tensors ein Zahlenwert zugeordnet, der den Pixelwert des Farbkanals im RGB-Spektrum repräsentiert. Ein solcher Tensor könnte wie folgt aussehen:



Für einen Rezipienten besteht bei diesen rein-numerischen Darstellungen keine Möglichkeit mehr, den Inhalt oder die Formgebung eines Werkes, also die jeweiligen schutzbegründenden Elemente, zu identifizieren. Selbst wenn Tensor und Ursprungswerk nebeneinander abgebildet sein sollten, sind die durchschnittlichen Rezipienten nicht mehr in der Lage, die nu-

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

merischen Darstellungen den Ursprungswerken zuzuordnen. In Folge des Feature Encodings können die Werke demzufolge nicht mehr *als Werk* verwertet werden. Eine Rezeption der schöpferischen Merkmale scheidet vielmehr aus. Damit berührt das Feature Encoding nicht die Kernverwertungsmöglichkeiten der Urheber.

Im Gegensatz zur Speicherung und Normalisierung der Inputdaten könnte darüber hinaus auch keine Sekundärverwertung der Inhalte mehr realisierbar sein. Denn die Sekundärverwertung von Inputdaten wird maßgeblich davon geprägt, dass die als Inputdaten generierten Werke später als Trainingsdaten weitergenutzt werden können. Als potenzielle Trainingsdaten kommt ihnen auch ein wirtschaftlicher Wert zu. Bereits in Tensoren umgewandelte, urheberrechtlich geschützte Inputdaten können aufgrund mangelnder menschlicher Wahrnehmbarkeit jedoch nicht mehr manuell annotiert werden. Eine solche Annotation durch Menschen ist aktuell aber zumeist noch notwendig.[293] Ohne menschliche Wahrnehmbarkeit der verkörperten Informationen ist die Vorverarbeitung der Daten insgesamt nur eingeschränkt möglich. Das betrifft insbesondere auch eine Klassifizierung und Vorsortierung der erfassten Inhalte. Aus alledem folgt, dass nach dem Feature Encoding der generierten Inputdaten keine Sekundärverwertung der umgewandelten Werke mehr möglich ist. Damit sind durch das Feature Encoding insgesamt keine Verwertungsinteressen der Urheber betroffen. Wie bei der Normalisierung der erfassten Werke fehlt es außerdem an einem Rezipientenkreis, der für eine Beeinträchtigung des Integritätsinteresses des Urhebers notwendig ist.[294]

Insgesamt werden durch das Feature Encoding der Werke keine Urheberinteressen beeinträchtigt, sodass eine Freistellung bereits danach zweckmäßig ist.

2. Freistellungsmechanismus

Mithilfe welches Regelungsmechanismus die Freistellung erfolgten sollte, richtet sich auch hier im Grundsatz nach den bereits herausgearbeiteten Kriterien.[295]

293 Vgl. *Paaß/Hecker*, Künstliche Intelligenz, S. 45, 143, 463, die von einer manuellen Annotation der Trainingsdaten ausgehen; *Rasmussen/Kirk/Moeslund*, Sensors 2022, 1596 S. 1 („traditional time-consuming manual annotation approach").

294 Dazu siehe oben unter 2. Teil E. I. 1.

295 Dazu also insgesamt siehe oben unter 2. Teil A. II.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

Fraglich ist also zunächst, ob Urhebern für das Feature Encoding der Werke eine angemessene Vergütung zu gewähren ist. Dann ist die Freistellung im Wege einer gesetzlichen Lizenz oder eines ECL-Mechanismus umzusetzen, wobei es bei der Wahl zwischen beiden Rechtsinstituten auf die Notwendigkeit eines Opt-out-Mechanismus sowie das Rechtssicherheitsbedürfnis der KI-Anwender ankommt. Ist hingegen keine Vergütung notwendig, kommt es insbesondere auf die Beweislastverteilung zwischen Urheber und KI-Anwender und das Schutzbedürfnis für technische Schutzmaßnahmen gegen das Feature Encoding der als Inputdaten gespeicherten Werke an. Mit Hilfe dieser Kriterien kann ermittelt werden, ob die Freistellung im Wege einer unvergüteten Schrankenbestimmung oder einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte erfolgen soll.

Wie auch schon die Speicherung und Normalisierung urheberrechtlich geschützter Inputdaten ist auch das Feature Encoding obligatorischer Teil des Funktionsablaufs eines KI-Systems. Um die Anwendung von KI insgesamt zu ermöglichen und dem hiermit einhergehenden Innovationsinteresse Rechnung zu tragen, muss folglich nicht nur die Speicherung der als Inputdaten generierten Werke und ihre Normalisierung, sondern auch die Vorverarbeitung ebendieser Daten gleichermaßen ermöglicht werden. Da der Markt für Urheber durch das Feature Encoding der Werke nicht beeinträchtigt wird, würde jede Ausgleichzahlung zudem eine Überkompensation bedeuten. Demzufolge sollte auch das Feature Encoding vergütungsfrei vorgenommen werden können.[296]

Klärungsbedürftig ist damit nur noch, ob eine Freistellung des Feature Encodings durch eine unmittelbar begrenzende Ausgestaltung der urheberrechtlichen Verwertungsrechte oder eine vergütungsfreie Schrankenbestimmung umgesetzt werden sollte. Das Rechtsschutzregime aus § 95a UrhG, welches dabei in den Blick zu nehmen ist, dient dem Schutz der urheberrechtlichen Verwertungsinteressen.[297] Im Umkehrschluss bedeutet das, dass ein Schutz technischer Schutzmaßnahmen vor Umgehung über § 95a UrhG nur dann gerechtfertigt ist, wenn die Verwertungsinteressen der von der Freistellung betroffenen Urheber berührt sind. Das ist beim Feature Encoding der Eingabedaten allerdings nicht der Fall. Ein Schutz von gegen das Feature Encoding der Werke eingesetzten technischen Maßnahmen über § 95a UrhG ist infolgedessen nicht interessengerecht. Das bedeutet, dass auch die Freistellung des Feature Encodings, also der Umwandlung urhe-

296 Im Detail dazu siehe oben unter 2. Teil D. II. 1.
297 Dazu ausführlich siehe oben unter 2. Teil A. II. 2. c).

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

berrechtlich geschützter Inputdaten in numerische Repräsentationen, aus einer regulatorischen Perspektive im Wege einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte umgesetzt werden sollte.[298]

III. Zwischenergebnis

Aus der regulatorischen Bedarfsanalyse folgt, dass sowohl die Normalisierung als auch das Feature Encoding urheberrechtlich geschützter Inputdaten, also deren Vorverarbeitung, im Wege einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte freigestellt sein sollten. De lege lata ist hierfür insbesondere die Reichweite des Vervielfältigungsrechts aus § 16 UrhG entscheidend.

F. Zusammenfassung der regulatorischen Bedarfsanalyse

Alles in allem muss bei der Identifizierung von Regelungsbedürfnissen für die Verwendung von Werken als Trainings- sowie Inputdaten für KI-Systeme untersucht werden, ob die genannten Fallgruppen urheberrechtlich freigestellt, also dem Ausschließlichkeitsrecht betroffener Urheber entzogen, oder ob Urhebern weiterhin ermöglicht werden soll, ihr Verbietungsrecht in den betreffenden Fällen auszuüben. Zentral für diese Beurteilung ist eine umfassende Abwägung von Urheber- und Nutzerinteressen. Die rechtökonomische Analyse bietet hingegen keinen geeigneten Ansatzpunkt für eine regulatorische Bedarfsanalyse. Trotzdem können ökonomische Argumente auch im Rahmen der freien Abwägung berücksichtigt werden.

Ist eine Freistellung zweckmäßig, stellt das Urheberrecht vier Mechanismen zur Verfügung, mit Hilfe derer die identifizierten Regelungsbedürfnisse für den Einsatz von KI umgesetzt werden können. Zum einen ist die Freistellung im Wege einer unmittelbar begrenzenden Ausgestaltung der

298 Zu einem ähnlich gelagerten Problem vgl. auch *Raue*, GRUR 2017, 11 (17). Er befürwortet im Kontext des Text und Data-Mining die Freistellung der Transformation der gesammelten, urheberrechtlich geschützten Inhalte in ein analysefähiges Format.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -

urheberrechtlichen Verwertungsrechte oder einer vergütungslosen Schrankenbestimmung möglich. Darüber hinaus könnte sie aber auch im Wege einer gesetzlichen Lizenz oder mit Hilfe eines ECL-Mechanismus umgesetzt werden. Neben Fragen der Beweislastverteilung und der Schutzbedürftigkeit eingesetzter technischer Schutzmaßnahmen über § 95a UrhG kommt es bei der Wahl des geeigneten Freistellungsmechanismus insbesondere darauf an, ob den Urhebern, die in der jeweiligen Fallgruppe vom KI-Einsatz betroffen sind, eine angemessene Vergütung für die Verwendung ihrer Werke gewährt werden sollte. Berücksichtigungsfähig ist darüber hinaus, ob ein Opt-out-Mechanismus zugunsten der Urheber interessengerecht ist, mit Hilfe dessen die Freistellungswirkung einer Regelung beendet werden kann. Ein besonderes Bedürfnis nach Rechtssicherheit fällt ebenfalls ins Gewicht. Im Einzelfall können aber auch andere Aspekte berücksichtigt werden.

Mit Hilfe des skizzierten, regulatorischen Untersuchungsmechanismus kann ermittelt werden, dass das Web Scraping von Werken zur Verwendung als Trainingsdaten durch eine vergütete Schrankenbestimmung ermöglicht sein sollte. Hierbei sind kommerzielle und nicht-kommerzielle Trainingsdatensammler gleichermaßen zu privilegieren. Ein Opt-out-Mechanismus für Urheber ist hingegen nicht vorzusehen. Auch die auf die Datenaggregation folgende, eigentliche Durchführung des KI-Trainings auf Basis der gesammelten Werke sollte aus einer regulatorischen Perspektive urheberrechtlich freigestellt sein. Zweckmäßig ist hierbei eine Freistellung im Wege einer unmittelbar begrenzenden Ausgestaltung der urheberrechtlichen Verwertungsrechte, die den Trainingsvorgang selbst aus dem urheberrechtlichen Ausschließlichkeitsrecht ausklammert.

Die Analyse der urheberrechtlichen Regelungsbedürfnisse ergibt darüber hinaus, dass sowohl die Speicherung urheberrechtlich geschützter Inputdaten als auch die Vorverarbeitung im Sinne der Normalisierung und des Feature Encodings der erfassten Inhalte durch eine unmittelbar begrenzende Ausgestaltung der Verwertungsrechte urheberechtlich freigestellt sein sollten. Infolgedessen kommt weder eine Vergütung der Urheber für die betreffenden Werkverwendungen noch ein Schutz von Verhinderungsmaßnahmen über § 95a UrhG in Betracht. De lege lata wird es für eine unmittelbar begrenzende Ausgestaltung der urheberrechtlichen Verwertungsrechte insbesondere auf den Schutzbereich des Vervielfältigungsrechts aus § 16 UrhG ankommen.

https://doi.org/10.5771/9783748963462-35 - am 12.01.2026, 05:30:13. https://www.inlibra.com/de/agb - Open Access -