

§ 3. Inhalt des Urheberrechts

Das Urheberrecht schützt den Urheber nach § 11 UrhG in seinen geistigen und persönlichen Beziehungen zum Werk und in der Nutzung des Werkes. Dieser Schutz dient vor allem der Sicherung einer angemessenen Vergütung. Die wirtschaftliche Verwertung steht auch im Mittelpunkt der Auseinandersetzung über die Verwendung urheberrechtlich geschützter Werke beim Training generativer KI-Modelle. Die nachfolgende Analyse befasst sich deshalb vorrangig mit den Verwertungsrechten.⁸⁵

A. Überblick

Ein Ziel des Urheberrechts ist es, den Urheber an der wirtschaftlichen Verwertung seines Werkes zu beteiligen und damit wirtschaftlich abzusichern.⁸⁶ Das sogenannte Prinzip der angemessenen Vergütung spiegelt diese Ratio und ist als wesentlicher Grundgedanke des Urheberrechts zu achten.⁸⁷ Die in §§ 15 ff. UrhG nicht abschließend aufgezählten Verwertungsrechte stellen sicher, dass eine angemessene Vergütung erzielt werden kann. Für die wirtschaftliche Verwertung zentral sind die Rechte zur Vervielfältigung und zur körperlichen Verbreitung nach § 15 Abs. 1 Nr. 1 und Nr. 2 sowie §§ 16 und 17 UrhG. Daneben hat der Urheber das ausschließliche Recht, sein Werk in unkörperlicher Form öffentlich wiederzugeben. Bei digitalen Werken ist vor allem das Recht der öffentlichen Zugänglichmachung nach §§ 15 Abs. 2 Nr. 2, 19a UrhG von Bedeutung. Wie gezeigt werden kann,

85 Bei Training generativer KI-Modelle und bei der Erstellung KI-generierten Outputs überdies mögliche Eingriffe in das Urheberpersönlichkeitsrecht und die daraus resultierenden Ansprüche aus §§ 12 ff. UrhG werden nicht behandelt.

86 Vgl. nur RGZ 128, 102 (113) – Schlagerliederbuch; BGH GRUR 1955, 492 (497) – Grundig-Reporter („Leitgedanke des Urheberrechts, daß der Urheber tunlichst angemessen an den wirtschaftlichen Früchten zu beteiligen sei, die aus seinem Werk gezogen werden ...“); BGH NJW 1999, 1953 (1956) – Kopienversanddienst („... im gesamten Urheberrecht zu beachtenden Grundsatz, daß der Urheber tunlichst angemessen an dem wirtschaftlichen Nutzen seines Werkes zu beteiligen ist ...“); Schulze in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 11 Rn. 1; Schack, Urheber- und Urhebervertragsrecht, 10. Aufl. 2021, Rn. 429.

87 Bundestags-Drucksache 14/8058, 17 f.

greift das KI-Training mit urheberrechtlich geschützten Werken und Leistungsgegenständen⁸⁸ auf vielfache Art und Weise in diese Rechte ein.⁸⁹ Für die Begutachtung wird davon ausgegangen, dass die Schutzfrist für eine Vielzahl der zum Training generativer KI-Modelle genutzten Werke noch nicht abgelaufen ist.⁹⁰

B. Verletzungshandlungen

Die Verwertungsrechte der §§ 15 ff. UrhG und der Art. 2 ff. InfoSoc-Richtlinie⁹¹ erfassen auch und vor allem technische Vorgänge. Dies kann körperliche Vervielfältigungen aber auch Eingriffe in anderer Form betreffen. Die grundsätzlich uneingeschränkte Reichweite des Vervielfältigungsrechts wird in der Formulierung „auf jede Art und Weise und in jeder Form“ in Art. 2 InfoSoc-Richtlinie deutlich.⁹² Nach dem hiernach implizierten weiten

88 Die Begutachtung erfolgt vorrangig mit Blick auf die Verwertung von Werken im Sinne des § 2 Abs. 2 UrhG. Vergleichbare Zusammenhänge bestehen bei Verwertung anderer Inhalte und bei Eingriff in andere Schutzgegenstände, so insbesondere Computerprogramme (§ 69a UrhG), Leistungsschutzrechte der Presseverleger (§§ 87f ff. UrhG), der Datenbankhersteller (§§ 87a ff. UrhG) oder der Lichtbildner (§ 72 UrhG) und Filmhersteller (§§ 89, 94 UrhG).

89 Eine neue Nutzungsart des „Trainings generativer KI-Modelle“ wird man deshalb wohl noch nicht konstruieren müssen, wenngleich dies vereinzelt angedacht wird; vgl. z.B. zögerlich bei Hofmann WRP 2024, 11 (15). Auch neue Formen der Verwertung sind jedenfalls zunächst grundsätzlich exklusiv dem Urheber zugewiesen (vgl. Schulze in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 15 Rn. 9). Soweit Verwertungsrechte europarechtlich harmonisiert wurden, ist eine Erweiterung durch den nationalen Gesetzgeber zwar ausgeschlossen (vgl. EuGH GRUR 2014, 360 Rn. 33 ff. – Nils Svensson u.a./Retriever Sverige; EuGH GRUR 2014, 473 Rn. 42 ff. – OSA; BGH GRUR 2019, 813 Rn. 37 – Cordoba II). Dies gilt aber nicht außerhalb des durch InfoSoc- und DSM-Richtlinie harmonisierten Bereichs (vgl. EuGH GRUR 2012, 156 Rn. 93 ff. – Football Association Premier League; BGH GRUR 2013, 818 Rn. 12 – Die Realität I). Lässt sich eine neue Nutzungsart keinem der bereits harmonisierten Verwertungsrechte zuordnen, was für das Training generativer KI-Modelle gegebenenfalls zu untersuchen wäre, fällt diese ebenfalls unter die §§ 15 ff. UrhG (vgl. allgemein v. Ungern-Sternberg, Festschrift Bornkamm, 2014, 1007 (1014 f.); v. Ungern-Sternberg, Festschrift Loschelder, 2010, 415 (422)).

90 Vgl. für die Urheberrechte an Werken insbesondere §§ 64 ff. UrhG.

91 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, ABl. Nr. L 167 vom 22.6.2001, S. 10.

92 Vgl. grundlegend zur weiten Auslegung im Interesse eines hohen Schutzniveaus: v. Ungern-Sternberg, Festschrift Joachim Bornkamm, 2014, 1007 (1009 f.) m.w.N.

Verständnis ist als Vervielfältigung im Sinne des Gesetzes zudem nicht nur die vollständige Replikation anzusehen, sondern auch die Kopie von Teilen eines urheberrechtlich geschützten Werks sowie eine nur zeitweise Vervielfältigung.⁹³ Auch digitale Vervielfältigungen sind umfasst.⁹⁴ Vor diesem Hintergrund ist für generative KI-Modelle nach den Stadien des Trainingsprozesses sowie dem nachfolgenden Einsatz zu differenzieren. Wie gezeigt werden kann, kommt es nicht nur bei der Sammlung, Vorbereitung und Korpus-Speicherung der Trainingsdaten zu Vervielfältigungen (nachfolgend I.), sondern auch bei der Adaption der Parameter beim Training generativer KI-Modelle (nachfolgend II.). Zudem ist bei Einsatz der Modelle mit der Entstehung von Vervielfältigungen und Bearbeitungen der beim Training verwerteten Werke zu rechnen (nachfolgend III.). Im Hinblick auf das Recht der öffentlichen Zugänglichmachung in §§ 15 Abs. 2 Nr. 2, 19a UrhG und Art. 3 InfoSoc-Richtlinie ist schließlich an Handlungen anzuknüpfen, die einer Vervielfältigung und Bearbeitung von Trainingsdaten nachfolgen, insbesondere das Angebot der Nutzung von KI-Systemen im Internet (z.B. ChatGPT auf der OpenAI-Webseite) sowie die Ermöglichung des Downloads von KI-Modellen.

I. Sammlung, Vorbereitung und Speicherung der Trainingsdaten

Für das Training generativer KI-Modelle wird auf vielfache Art und Weise auf Datenbestände zugegriffen, die aus urheberrechtlich geschützten Werken bestehen.⁹⁵ Weitgehend unbestritten ist das Vervielfältigungsrecht in den folgenden Stadien und Phasen des Trainings berührt:

- Webscraping,
- Kopie der Daten in den Arbeitsspeicher,
- Vorbereitung der Daten für das Training sowie
- Erstellung und Speicherung des Korpus.

Die Digitalisierung und Herstellung von Kopien digitaler Werke in dauerhaft gespeicherter Form, etwa auf einer Festplatte, gilt als Vervielfältigung

93 Loewenheim in Schricker/Loewenheim, UrhR, 6. Aufl. 2020, § 16 Rn. 5 ff. und Rn. 11.

94 Vgl. nur BGH GRUR 1999, 325 (327) – Elektronische Pressearchive; zudem Loewenheim in Schricker/Loewenheim, UrhR, 6. Aufl. 2020, § 16 Rn. 16; Schack, Urheber- und Urhebervertragsrecht, 10. Aufl. 2021, Rn. 437 ff.

95 Siehe oben § 2.A. und B. (technischer Teil).

im Sinne des Gesetzes.⁹⁶ Überdies kann auch die Erstellung lediglich vorübergehender Kopien im Arbeitsspeicher eines Computers als Vervielfältigung eingeordnet werden. Dies umfasst insbesondere den Download aus dem Internet.⁹⁷ Entsprechend sind sowohl das Webscraping als auch die Speicherung der Trainingsdaten in einem Korpus als Vervielfältigungen einzuordnen.⁹⁸

Gleiches gilt grundsätzlich für Handlungen, mit denen die Datenbestände für das Training vorbereitet werden, z.B. der Normalisierung oder der Umwandlung in eine andere Repräsentation. Das Ergebnis kann dabei im weitesten Sinne als abgewandelte Codierung betrachtet werden. Beim Training von Diffusionsmodellen werden Daten (z.B. Bilddateien) überdies zum Teil noch weiterreichend verändert, etwa durch Hinzufügen von Rauschen (*noise*). Auch derartige Veränderungen können, ungeachtet einer Einordnung als Bearbeitung oder Umgestaltung im Sinne des § 23 UrhG, jedenfalls auch als Vervielfältigungen im Sinne des § 16 UrhG anzusehen sein.⁹⁹

96 Vgl. § 16 Abs. 2 UrhG; hierzu z.B. BGH GRUR 2014, 549 Rn. 18 – Meilensteine der Psychologie; BGH GRUR 2013, 503 Rn. 8 – Elektronische Leseplätze II; BGH GRUR 2010, 616 Rn. 36 – marions-kochbuch.de; Schack, Urheber- und Urhebervertragsrecht, 10. Aufl. 2021, Rn. 435.

97 BGH GRUR 2017, 266 Rn. 38 – World of Warcraft I; BGH GRUR 2011, 418 Rn. 12 – UsedSoft; BGH GRUR 2010, 616 Rn. 36 – marions-kochbuch.de; überdies z.B. Schack, Urheber- und Urhebervertragsrecht, 10. Aufl. 2021, Rn. 437; Konertz/Schönhof WRP 2024, 289 (290); Lux/Noll TLJ 2024, 111 (112).

98 Vgl. z.B. Steinrötter/Schauer in Barudi (Hrsg.), Das neue Urheberrecht, 2021, § 4 Rn. 9; Heerma in Wandtke/Bullinger, UrhG, 6. Aufl. 2022, § 16 Rn. 19; Bomhard InTeR 2023, 174 (175); Kögel InTeR 2023, 179 (180); Vesala IIC 54 (2023), 351 (353 f.); Pesch/Böhme GRUR 2023, 997 (1004); Baumann NJW 2023, 3673 (3674); Maamar ZUM 2023, 481 (483); Konertz/Schönhof WRP 2024, 289 (290); Kraetzig NJW 2024, 697 (698); Wulf/Löbeth GRUR 2024, 737 (739); für Großbritannien z.B. Guadamuz GRUR Int. 2024, 111 (115); für das US-amerikanische Recht z.B. Sobel Colum. J. L. & Arts. 41 (2017), 45 (61 ff.); Quang Berkeley Tech. L. J. 36 (2021), 1407 (1413 ff.).

99 Auch Bearbeitungen oder andere Umgestaltungen im Sinne des § 23 Abs. 1 Satz 1 UrhG sind Vervielfältigungen, soweit sie körperlich festgelegt sind. Siehe nur BGH GRUR 2023, 571 Rn. 28 – Vitrineneuchte; BGH GRUR 2022, 899 Rn. 56 – Porsche 911; BGH GRUR 2014, 65 – Rn. 36 – Beuys-Aktion. Ob rein technische Veränderungen als Bearbeitungen im Sinne des § 23 Abs. 1 UrhG anzusehen sind, ist zu bezweifeln (*arg. e* „persönliche geistige Schöpfungen“ in § 3 UrhG); in Betracht kommt aber eine Einordnung als andere Umgestaltung. Das Recht des Urhebers umfasst – mit Ausnahme der Fälle des Abs. 2 – nur die Veröffentlichung und Verwertung. Vgl. Spindler GRUR 2016, 1112 (1114); Obergfell, Festschrift Büscher, 2018, 223 (226); Pesch/Böhme GRUR 2023, 997 (1002 f.); Hofmann WRP 2024, 11 (12); Konertz/Schönhof WRP 2024, 289 (290). Die Rechte des § 23 Abs. 1 und Abs. 2 UrhG

II. Training eines generativen KI-Modells

Für das eigentliche Training generativer KI-Modelle muss technisch zwischen dem *Pre-Training* und dem *Fine-Tuning* unterschieden werden.¹⁰⁰ Während des *Pre-Trainings* wird aus der initialen Konzeption des Modells durch die ersten Trainingsläufe ein Basismodell geschaffen; teils wird dieses auch als *pre-trained model* oder *foundation model* bezeichnet.¹⁰¹ Im Anschluss daran erhält das Modell typischerweise durch weiteres Training mit speziellen Datenbeständen im Rahmen des *Fine-Tuning* besondere Fähigkeiten „antrainiert“. Ein basales *language model* kann so etwa durch *Fine-Tuning* in ein Modell mit der Fähigkeit zur Auswertung wissenschaftlicher Texte transformiert werden.¹⁰² In allen Stadien dieses Trainings sind Verwertungsrechte berührt: Das als möglichst umfassend definierte Vervielfältigungsrecht erfasst mit den digitalen Vervielfältigungen unbestritten auch Vorgänge, die bei der Verarbeitung von Daten durch Software und damit bei einer Verarbeitung in KNNs stattfinden.¹⁰³ Ebenfalls unbestritten scheint dabei, dass KI-Modelle die verwerteten Trainingsdaten in der Regel

gelten nach Abs. 3 dieser Vorschrift zudem nicht, wenn es sich um eine ausschließlich technisch bedingte Änderung eines Werkes im Rahmen der TDM-Nutzung handelt.

100 Siehe oben § 2.B.IV. (technischer Teil).

101 Vgl. z.B. Lee/Cooper/Grimmelmann, Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain, J. Copyright Soc'y of the U.S.A. (forthcoming 2024), S. 39 ff. (einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4523551 (zuletzt am 27. Juni 2024)).

102 Das *Fine-Tuning* unterscheidet sich daher auch nicht grundsätzlich von der Erstellung des Basismodells. In beiden Fällen handelt es sich um KI-Trainingsvorgänge. Vgl. nochmals Lee/Cooper/Grimmelmann, Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain, J. Copyright Soc'y of the U.S.A. (forthcoming 2024), S. 42 f. (einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4523551 (zuletzt am 27. Juni 2024)).

103 Allgemeine Ansicht. Siehe etwa Käde, Kreative Maschinen und Urheberrecht – Die Machine Learning-Werkschöpfungskette vom Training über Modellschutz bis zu Computational Creativity, 2021, 68 ff.; Pesch/Böhme GRUR 2023, 997 (1004 ff.); v. Welser GRUR-Prax 2023, 516 (517); Heine GRUR-Prax 2024, 87 (87); Konertz/Schönhof WRP 2024, 289 (291 und 295); vgl. zudem ähnlich im US-Recht z.B. Lee/Cooper/Grimmelmann, Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain, J. Copyright Soc'y of the U.S.A. (forthcoming 2024), S. 68 (einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4523551 (zuletzt am 27. Juni 2024)) („Thus, the assembly of a dataset, the training of a model, the production of a generation, or a generative-AI system's use of a user-inputted prompt is a 'reproduction' within the meaning of copyright law.“).

nicht unmittelbar und dauerhaft in ihrem neuronalen Netz oder an anderer Stelle codieren und damit speichern. Die Verwendung des Dateninput beim Training des Systems führt vielmehr zur Adaption von Parameterwerten innerhalb des neuronalen Netzes. Dennoch wurde mehrfach nachgewiesen, dass etwa Sprachmodelle wie GPT-4 durch gezielte *prompts* dazu veranlasst werden können, Textpassagen aus dem Bestand ihrer Trainingsdaten unverändert oder weitgehend identisch auszuwerfen.¹⁰⁴ Auch bei Bildgeneratoren wie Stable Diffusion oder DALL-E können Trainingsdaten bei entsprechenden *prompts* wiedergegeben werden.¹⁰⁵ Wie gezeigt werden kann, kommt es tatsächlich zur Vervielfältigung der Trainingsdaten in den Parametern des KI-Modells. Die Bereitstellung zur Anwendung im Internet, so wie etwa bei ChatGPT auf der OpenAI-Webseite, ist daher auch als öffentliche Zugänglichmachung im Sinne der §§ 15 Abs. 2 Nr. 2, 19a UrhG einzuordnen.

1. Meinungsstand

Zum Teil wird das Phänomen des unveränderten, vollständigen oder teilweisen Auswurfs von Werken aus den Trainingsdatenbeständen zum Anlass genommen, von einer dauerhaften Speicherung der Trainingsdaten im KI-Modell auszugehen und daher auch eine urheberrechtlich relevante Vervielf-

104 Vgl. insoweit z.B. Henderson et al., J. Mach. Learning Research 23 (2023), 1; Carlini et al., Extracting Training Data from Large Language Models, in: Proceedings of the 30th USENIX Security Symposium (USENIX Security 21) 2633 (2021) (einsehbar unter: <https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting> (zuletzt am 21. August 2024)); Chang et al., Speak Memory: An Archaeology of Books Known to ChatGPT/GPT-4, in: Proceedings of the 2023 Conference in Empirical Methods in Natural Language Processing (einsehbar unter: <https://aclanthology.org/2023.emnlp-main.453/> (zuletzt am 15. August 2024)); aus dem juristischen Schrifttum z.B. Pesch/Böhme MMR 2023, 917 (919); Pesch/Böhme GRUR 2023, 997 (1001); Käde ZUM 2024, 174 (180); Sag Houston L. Rev. 61 (2023), 295 (326 ff.).

105 Carlini et al., Extracting Training Data from Diffusion Models, in: Proceedings of the 32nd USENIX Security Symposium (einsehbar unter: <https://doi.org/10.48550/arXiv.2301.13188> (zuletzt am 16. August 2024)); Somepalli et al., Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models, arXiv:2212.03860v3 (2022) (einsehbar unter: <https://doi.org/10.48550/arXiv.2212.03860> (zuletzt am 16. August 2024)); Somepalli et al., Understanding and Mitigating Copying in Diffusion Models, arXiv:2305.20086v1 (2023) (einsehbar unter: <https://doi.org/10.48550/arXiv.2305.20086> (zuletzt am 16. August 2024)).

fältigung anzunehmen.¹⁰⁶ Die überwiegende Gegenansicht verneint eine Vervielfältigung von Trainingsdaten in den Parametern allerdings mit dem Hinweis auf die typische Funktionalität generativer Modelle: Der Output solle durch die KI in jedem einzelnen Fall neu generiert werden. Es komme nicht zum determinierten Abruf statischer Informationszustände im Sinne eines „Auswendiglernens“ und der nachfolgenden Reproduktion des auswendig Gelernten. Dies sei gerade nicht Sinn und Zweck generativer Anwendungen.¹⁰⁷ Derartige Modelle dienen nicht der Speicherung von Daten, zielen daher auch nicht auf die Bevorratung ihrer Trainingsdaten, sondern ausschließlich auf deren Auswertung und die Schaffung neuer Inhalte.¹⁰⁸ Ergänzend wird vereinzelt auch noch darauf hingewiesen, dass die Annahme einer Vervielfältigung der Trainingsdaten jedenfalls daran scheitern müsse, dass eine wie auch immer geartete Speicherung nicht offen nach außen zu Tage trete.¹⁰⁹ Vor allem der letztgenannte Ansatz scheint sich auf die Begründung des Gesetzgebers und die höchstrichterliche Rechtsprechung stützen zu können, wonach für Vervielfältigungen gefordert wird, dass „das Werk den menschlichen Sinnen auf irgendeine Weise mittelbar oder unmittelbar wahrnehmbar zu machen“ sei.¹¹⁰ Jedenfalls auf den ersten Blick ist dies durchaus schlagkräftig: Die Adaption der Parameter in einem KNN ist tatsächlich nicht direkt mit den menschlichen Sinnen wahrnehmbar.

106 So wohl v. Welser GRUR-Prax 2023, 516 (517); Pesch/Böhme GRUR 2023, 997 (1005); Kögel InTeR 2023, 179 (180); K. Wagner MMR 2024, 298 (299); Sesing-Wagenpfeil ZGE 16 (2024), 212 (242 und passim); Mezei E.I.P.R. 46 (2024), 461 (463).

107 Vgl. nur Käde, Kreative Maschinen und Urheberrecht – Die Machine Learning-Werkschöpfungskette vom Training über Modellschutz bis zu Computational Creativity, 2021, 74 f.; Kögel in Bernzen u.a. (Hrsg.), Das IT-Recht vor der (europäischen) Zeitenwende?, 2023, 285 (287 f.); Söbbing/Schwarz RDt 2023, 415 (417); Baumann NJW 2023, 3673 (3674); Maamar ZUM 2023, 481 (482); Hofmann WRP 2024, II (12); Konertz/Schönhof WRP 2024, 289 (293); Konertz/Schönhof WRP 2024, 534 (536); Käde ZUM 2024, 174 (176 f.); zudem aus dem US-Schrifttum z.B. Sag Fordham L. Rev. 92 (2024), 1885 (1910, 1912 und passim).

108 Baumann NJW 2023, 3673 (3674); Hofmann WRP 2024, II (12); Heine GRUR-Prax 2024, 87 (87); Käde ZUM 2024, 174 (177); Peukert GRUR Int. 2024, 497 (506).

109 So im Ergebnis wohl Hofmann WRP 2024, II (12).

110 Vgl. nur Bundestags-Drucksache IV/270, S. 47 („Das Vervielfältigungsrecht umfasst nach dieser Definition die Herstellung jeder Art und Zahl von Vervielfältigungsstücken des Werkes, d. h. von körperlichen Festlegungen, die geeignet sind, das Werk den menschlichen Sinnen auf irgendeine Weise unmittelbar oder mittelbar wahrnehmbar zu machen (Bücher, Noten, Schallplatten und dergl.)“); weitgehend identisch in der Rechtsprechung z.B. BGH GRUR 2017, 793 Rn. 41 – Mart-Stam-Stuhl; BGH GRUR 2001, 51 (52) – Parfumflakon I; vgl. überdies K. Wagner MMR 2024, 298 (299); Konertz/Schönhof WRP 2024, 289 (290 und 293).

2. Korrekturen

Eine Vervielfältigung der urheberrechtlich geschützten Trainingsdaten „im Innern“ generativer KI-Modelle lässt sich allerdings weder unter Verweis auf die Funktionalität oder Zwecksetzung dieser Modelle, noch mit Blick auf die fehlende unmittelbare Wahrnehmbarkeit verneinen. Die gesetzlichen Voraussetzungen für das Vorliegen einer Vervielfältigung im Sinne des § 16 Abs. 1 UrhG sind vielmehr auch in der während und durch das Training bewirkten Adaption der Parameter eines generativen KI-Modells erfüllt.

a) Ausgangspunkt: Speicherung „im Innern“

Eine jedenfalls teilweise Memorisierung von Trainingsdaten „im Innern“ der KI-Modelle ist nicht bestritten. Verschiedene Untersuchungen haben gezeigt, dass zwischen 0,1 % und 10 % der Trainingsdaten *unmittelbar* in den Modellen gespeichert werden.¹¹¹ Auch eine Replikation von Trainingsdaten, die *nicht unmittelbar* gespeichert oder *memorisiert* werden und lediglich bei der Adaption der Parameter des Modells Verwendung finden, ist belegt.¹¹² Für die Frage des Vorliegens von Vervielfältigungen kann es nicht

111 Vgl. hierzu m.w.N. insbesondere Cooper/Grimmelmann, The Files are in the Computer: On Copyright, Memorization, and Generative AI, Chicago-Kent Law Review (forthcoming 2024), S. 49 (einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4803118 (zuletzt am 1. August 2024)).

112 Vgl. nochmals Henderson et al., J. Mach. Learning Research 23 (2023), 1; Carlini et al., Extracting Training Data from Large Language Models, in: Proceedings of the 30th USENIX Security Symposium (USENIX Security 21) 2633 (2021) (einsehbar unter: <https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting> (zuletzt am 21. August 2024)); Chang et al., Speak Memory: An Archaeology of Books Known to ChatGPT/GPT-4, in: Proceedings of the 2023 Conference in Empirical Methods in Natural Language Processing (einsehbar unter: <https://aclanthology.org/2023.emnlp-main.453/> (zuletzt am 15. August 2024)); Carlini et al., Extracting Training Data from Diffusion Models, in: Proceedings of the 32nd USENIX Security Symposium (einsehbar unter: <https://doi.org/10.48550/arXiv.2301.13188> (zuletzt am 16. August 2024)); Somepalli et al., Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models, arXiv:2212.03860v3 (2022) (einsehbar unter: <https://doi.org/10.48550/arXiv.2212.03860> (zuletzt am 16. August 2024)); Somepalli et al., Understanding and Mitigating Copying in Diffusion Models, arXiv:2305.20086v1 (2023) (einsehbar unter: <https://doi.org/10.48550/arXiv.2305.20086> (zuletzt am 16. August 2024)); aus dem juristischen Schrifttum vgl.

darauf ankommen, ob die unmittelbare Speicherung oder die Replikation in der Technologie oder der Funktionalität generativer KI-Modelle angelegt sind. Ebenso kann es keine Rolle spielen, ob die Entwickler oder Nutzer es auf eine Replikation angelegt haben und diese deshalb gezielt durch spezielle *prompts* evozieren. Der urheberrechtliche Vervielfältigungsbegriff ist als objektives Tatbestandsmerkmal unabhängig davon definiert, ob die Handelnden eine Vervielfältigung beabsichtigen. Am objektiven Vorliegen bestimmter Speicherungszustände für einen nicht unwesentlichen Teil der beim Training verwerteten Werke „im Innern“ der generativen KI-Modelle ist deshalb nicht zu zweifeln.¹¹³

b) Nicht gefordert: Unmittelbare Wahrnehmbarkeit

Auch das vereinzelt aufgestellte Erfordernis einer unmittelbaren Wahrnehmung zwingt nicht dazu, eine Vervielfältigung der Trainingsdaten in den Parametern zu verneinen. Der Gesetzgeber und die Gerichte verlangen für das Vorliegen einer Vervielfältigung im Sinne des § 16 UrhG lediglich, dass das Werk durch körperliche Festlegung „den menschlichen Sinnen auf irgendeine Weise mittelbar oder unmittelbar wahrnehmbar zu machen“ sei.¹¹⁴ Es spielt dabei keine Rolle, mit welchen Mitteln und auf welche Art und Weise eine Vervielfältigung körperlich festgelegt wird, oder ob zwischen Festlegung und sich anschließender menschlicher Wahrnehmung weitere Handlungen und Zwischenschritte erforderlich sind. Es genügt daher, dass ein Werk körperlich auf einem Werkträger fixiert wird, von dem es erst durch technische Hilfsmittel oder unter Nutzung bestimmter Fähigkeiten

Pesch/Böhme MMR 2023, 917 (919); Pesch/Böhme GRUR 2023, 997 (1001); Käde ZUM 2024, 174 (180); Sag Houston L. Rev. 61 (2023), 295 (326 ff.).

113 Vgl. insoweit auch instruktiv Cooper/Grimmelmann, *The Files are in the Computer: On Copyright, Memorization, and Generative AI*, Chicago-Kent Law Review (forthcoming 2024), S. 20 ff., insbesondere S. 28 f. (einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4803118 (zuletzt am 1. August 2024)).

114 Vgl. nochmals Bundestags-Drucksache IV/270, S. 47 („Das Vervielfältigungsrecht umfaßt nach dieser Definition die Herstellung jeder Art und Zahl von Vervielfältigungsstücken des Werkes, d. h. von körperlichen Festlegungen, die geeignet sind, das Werk den menschlichen Sinnen auf irgendeine Weise unmittelbar oder mittelbar wahrnehmbar zu machen (Bücher, Noten, Schallplatten und dergl.).“); zudem zu weitgehend identischen Formulierungen in der Rechtsprechung insbesondere BGH GRUR 2017, 793 Rn. 41 – Mart-Stam-Stuhl; BGH GRUR 2001, 51 (52) – Parfumflakon I.

wieder wahrnehmbar gemacht werden kann.¹¹⁵ Entsprechend besteht kein Zweifel daran, dass digitale Dateiversionen eines Werks, etwa in Form der Speicherung auf einer Festplatte, Vervielfältigungen im Sinne des Gesetzes sein können, obwohl stets technische Hilfsmittel eingesetzt werden müssen, um die digitalisierte Version des Werks sinnlich wahrnehmbar zu machen.¹¹⁶ Auch muss das Werk beim Vervielfältigungsvorgang keinesfalls unverändert bleiben: So ändern etwa die Verkleinerung oder Komprimierung nichts daran, dass es sich beim Ergebnis der Umwandlungsvorgänge um Vervielfältigungen handelt.¹¹⁷ Schließlich kann die Zerlegung eines Werks in einzelne Bestandteile, z.B. beim *Routing*, nichts am Vorliegen einer Vervielfältigung ändern, sofern das Werk am Ende wieder zusammengesetzt wird.¹¹⁸

115 Vgl. z.B. Heerma in Wandtke/Bullinger, UrhR, 6. Aufl. 2022, § 16 Rn. 4 f.; Dreier/Schulze/Schulze, UrhG, 7. Aufl. 2022, § 16 Rn. 6 (mit Blindenschrift als Beispiel); Sasing-Wagenpfeil ZGE 16 (2024), 212 (229); zudem zum US-Recht auch Lee/Cooper/Grimmelmann, Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain, J. Copyright Soc'y of the U.S.A. (forthcoming 2024), S. 74 (einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4523551 (zuletzt am 27. Juni 2024)).

116 Vgl. nur BGH GRUR 2010, 616 (619) – marions-kochbuch.de; OGH GRUR Int. 2007, 626 (627) – St. Stephan; zudem z.B. Schack, Urheber- und Urhebervertragsrecht, 10. Aufl. 2021, Rn. 435; Dreier/Schulze/Schulze, UrhG, 7. Aufl. 2022, § 16 Rn. 7; K. Wagner MMR 2024, 298 (299); Käde ZUM 2024, 174 (181); Sasing-Wagenpfeil ZGE 16 (2024), 212 (236); überdies nochmals Lee/Cooper/Grimmelmann, Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain, J. Copyright Soc'y of the U.S.A. (forthcoming 2024), S. 74 (einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4523551 (zuletzt am 27. Juni 2024)); Cooper/Grimmelmann, The Files are in the Computer: On Copyright, Memorization, and Generative AI, Chicago-Kent Law Review (forthcoming 2024), S. 30 f. (einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4803118 (zuletzt am 1. August 2024)); instruktiv auch Sobel, Elements of Style: Copyright, Similarity, and Generative AI, Harv. J. L. & Tech. 38 (forthcoming 2024), 1 (23 ff.) (einsehbar auf SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4832872 (zuletzt am 28. Juni 2024)).

117 BGHZ 44, 288 (293); BGH GRUR 1990, 669 (673) – Bibelreproduktion; BGH GRUR 2010, 628 (629) – Vorschaubilder; zudem z.B. Schack, Urheber- und Urhebervertragsrecht, 10. Aufl. 2021, Rn. 435; Pesch/Böhme GRUR 2023, 997 (1002).

118 Während die Vervielfältigung für die Zwischenstadien, insbesondere die Zerlegung in Bestandteile, umstritten ist, besteht Einigkeit darüber, dass jedenfalls im endgültigen Zusammensetzen wieder eine Vervielfältigung zu sehen ist. Vgl. nur Loewenheim in Schricker/Loewenheim, UrhR, 6. Aufl. 2020, § 16 Rn. 23; Dreier/Schulze/Schulze, UrhG, 7. Aufl. 2022, § 16 Rn. 12; Heerma in Wandtke/Bullinger, UrhR, 6. Aufl. 2022, § 16 Rn. 21; Pesch/Böhme GRUR 2023, 997 (1002).

Auch die Tatsache, dass die für das Training verwerteten Werke lediglich „parametrisiert“ abgebildet werden, zwingt nicht zu einer abweichenden Beurteilung. Insoweit besteht kein Unterschied zu anderen Varianten der digitalen Speicherung, z.B. auf Computerfestplatten¹¹⁹: Auch für diese Fälle wird nicht gefordert, dass die Speicherung einem bestimmten, etwa formwahrenden Muster oder System folgt. Entsprechend spielt auch die konkrete Gestalt der „Abbildung“ von Vervielfältigungen im Vektorraum des Modells keine Rolle. Jedenfalls am Ende kann aus dem Modell heraus – einen passenden *prompt* vorausgesetzt – wieder ein sinnvoll kombiniertes und sinnlich wahrnehmbares Erzeugnis entstehen, welches das gesamte Werk oder einen substantiellen Teil davon wiedergibt. Das technische Verfahren für die Aufspaltung, die Veränderung und die anschließende Reproduktion ist auch insoweit nicht relevant.¹²⁰

c) Method in flux: „Wahrnehmbarmachung“

Zweifel bestehen lediglich im Hinblick darauf, dass die Replikation der Trainingsdaten bei Einsatz generativer KI-Modelle nicht als standardisierte Funktion angelegt ist. Es gibt – jedenfalls im Moment – keine zuverlässige Routine oder Methode, die es erlaubt, umfassend, kontrolliert und effizient auf die Speicherungszustände im Innern der KI-Modelle zuzugreifen und diese *ad hoc* „abzurufen“. Die beobachteten Vervielfältigungen, vor allem die Reproduktionsvorgänge beim KI-Einsatz, werden vielmehr als technischer Defekt (*overfitting*) und als nach dem Nutzungszweck der KI-Modelle unerwünscht beschrieben.¹²¹

Eine transparente und zuverlässige Routine zur „Wahrnehmbarmachung“ wird vom Gesetzgeber und den Gerichten allerdings gerade nicht gefordert. Zur Erinnerung: Es bedarf für das Vorliegen einer Vervielfälti-

119 So auch Pesch/Böhme GRUR 2023, 997 (1005); instruktiv zudem Cooper/Grimmelmann, The Files are in the Computer: On Copyright, Memorization, and Generative AI, Chicago-Kent Law Review (forthcoming 2024), S. 30 ff. (einschbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4803118 (zuletzt am 1. August 2024)).

120 So im Ergebnis auch K. Wagner MMR 2024, 298 (299); Sesing-Wagenpfeil ZGE 16 (2024), 212 (236 f.).

121 Auf dieser Grundlage verneint die überwiegende Ansicht eine Vervielfältigung durch eine Adaption der Parameter in einem KNN. Siehe nochmals oben § 3.B.II.1 und die dortigen Nachweise.

gung lediglich einer körperlichen Festlegung des Werkes, die „geeignet ist, das Werk den menschlichen Sinnen auf irgendeine Weise mittelbar oder unmittelbar wahrnehmbar zu machen“.¹²² Von einer zuverlässigen oder unfehlbaren Methode ist keine Rede. Dies zu fordern wäre auch absurd, kann die Existenz einer verkörperten Vervielfältigung doch keinesfalls vom uneingeschränkten Zugang zu dieser Verkörperung abhängen. Im Schrifttum wird insoweit zutreffend darauf verwiesen, dass auch in anderem Zusammenhang, etwa bei digitalen Medien, nicht nach einem reibungslosen Funktionieren der technischen Hilfsmittel oder Abläufe gefragt wird; auch der Aufwand darf keine Rolle spielen. Mit einer bestimmten Wahrscheinlichkeit ist überdies auch bei der bewährten Technik mit Fehlfunktionen beim Abruf zu rechnen; dies ändert aber nichts daran, dass die gespeicherten digitalen Informationszustände als Vervielfältigungen im Sinne des Urheberrechts einzuordnen sind.¹²³ Dementsprechend wird eine Vervielfältigung von Daten etwa auch bei Verlust des Passworts für den Zugang zu den verschlüsselten Datenbeständen nicht verneint.¹²⁴

Erkenntnisse der KI-Forschung belegen überdies, dass die gezielte Reproduktion von Trainingsdatenbeständen nicht als marginales und vernachlässigbares Phänomen einzuordnen ist. Es ist auch keinesfalls ausgeschlossen, dass mit der Zeit – ein besseres Verständnis der Vorgänge „im Inneren“ generativer KI-Modelle vorausgesetzt – weitere Möglichkeiten eines transparenten und effizienten Abrufs der Speicherungszustände bekannt werden.¹²⁵ Allein an der Tatsache, dass die Optionen der standardisierten Abrufbarkeit im Moment noch begrenzt erscheinen, kann sich die Frage der Vervielfältigung deshalb nicht entscheiden.¹²⁶

122 Vgl. nochmals Bundestags-Drucksache IV/270, S. 47; zudem BGH GRUR 2017, 793 Rn. 41 – Mart-Stam-Stuhl („Jede körperliche Festlegung eines Werkes, die geeignet ist, das Werk den menschlichen Sinnen auf irgendeine Art mittelbar oder unmittelbar wahrnehmbar zu machen, stellt eine Vervielfältigung iSd § 16 I UrhG dar ...“); überdies BGH GRUR 2001, 51 (52) – Parfumflakon I.

123 Vgl. nochmals instruktiv Cooper/Grimmelmann, *The Files are in the Computer: On Copyright, Memorization, and Generative AI*, Chicago-Kent Law Review (forthcoming 2024), S. 45 f. (einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4803118 (zuletzt am 1. August 2024)); ähnlich wohl auch Sesing-Wagenpfeil ZGE 16 (2024), 212 (239 ff.).

124 Vgl. hierzu Pesch/Böhme GRUR 2023, 997 (1005); ebenso Sesing-Wagenpfeil ZGE 16 (2024), 212 (246 mit dortiger Fn. 175).

125 Siehe bereits oben § 2.D.III. (technischer Teil).

126 Cooper/Grimmelmann, *The Files are in the Computer: On Copyright, Memorization, and Generative AI*, Chicago-Kent Law Review (forthcoming 2024), S. 28 f.

Schließlich ist aus praktischer Perspektive zu beachten, dass in Fällen der tatsächlichen Replikation bei Einsatz eines generativen KI-Modells – z.B. dem Auswurf eines literarischen Werkes durch ChatGPT – kaum Zweifel am Vorliegen einer Vervielfältigung „im Innern“ des KI-Modells besteht. Insbesondere ChatGPT lässt sich sogar das ausdrückliche Geständnis entlocken, dass bestimmte Werke (etwa Liedtexte) zwar repliziert werden *könnten*, dies aber zur Vermeidung von Urheberrechtsverstößen unterbleibe.¹²⁷ Die spontane, identische oder weitgehend identische Erschaffung *ex nihilo* ist bei umfangreichen und anspruchsvollen Inhalten jedenfalls sehr unwahrscheinlich.¹²⁸

3. Zwischenergebnis

Auch in der Ausbildung der Parameter eines generativen KI-Modells ist eine Vervielfältigung der zum Training eingesetzten, urheberrechtlich geschützten Werke zu sehen. Daher sind das Verbreitungsrecht nach § 17 Abs. 1 UrhG, vor allem aber das Recht der öffentlichen Zugänglichmachung nach §§ 15 Abs. 2 Nr. 2, 19a UrhG betroffen, wenn der Hersteller oder Anbieter eines KI-Modells dieses für Anwender zum Download bereithält. Am Vorliegen der für den Tatbestand erforderlichen Kriterien – insbesondere der Kenntnis des Handelnden, seinem Erwerbszweck sowie der Öffentlichkeit¹²⁹ – besteht kein Zweifel. Ebenso ist die Bereitstellung eines generativen KI-Modells im Rahmen eines Systems für die interaktive Anwendung durch beliebige Nutzer (z.B. ChatGPT auf der OpenAI-Webseite) eine öffentli-

(einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4803118 (zuletzt am 1. August 2024)).

127 Vgl. die Beispiele bei Sesing-Wagenpfeil ZGE 16 (2024), 212 (248).

128 Vgl. nochmals Cooper/Grimmelmann, The Files are in the Computer: On Copyright, Memorization, and Generative AI, Chicago-Kent Law Review (forthcoming 2024), S. 24 f. (einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4803118 (zuletzt am 1. August 2024)); überdies auch Sesing-Wagenpfeil ZGE 16 (2024), 212 (238 f.).

129 Vgl. zu diesen Vorgängen allgemein EuGH GRUR 2016, 1152 Rn. 32 ff. – GS Media/Sanoma ua; EuGH GRUR 2017, 610 Rn. 29 ff. – Stichting Brein/Wullems; EuGH GRUR 2017, 790 Rn. 24 ff. – Stichting Brein/Ziggo ua [The Pirate Bay]; EuGH GRUR 2021, 1054 Rn. 61 ff. – Peterson/Google ua u. Elsevier/Cyando [YouTube und uploaded]; zudem BGH GRUR 2018, 178 Rn. 19 – Vorschaubilder III; BGH GRUR 2019, 950 Rn. 19 ff. – Testversion; BGH GRUR 2022, 1308 Rn. 69 ff. – Youtube II; überdies z.B. Götting in BeckOK/UrhR, 42. Edition, Stand: 1.5.2024, § 19a UrhG Rn. 3 ff.; Eichelberger ZUM 2023, 660 (661 f.).

che Zugänglichmachung. Schließlich eröffnet der Anbieter auch damit – einen passenden *prompt* vorausgesetzt – den Zugriff auf urheberrechtlich geschützte Werke.¹³⁰

III. Genese von KI-Output

Werden die für das Training verwerteten Werke vollständig oder teilweise, jedenfalls im Wesentlichen unverändert, als Ergebnis des Einsatzes eines generativen KI-Systems ausgeworfen, kann auch der KI-Output in die Rechte des Urhebers eingreifen.¹³¹ Je nachdem wie nah die ausgeworfenen Ergebnisse an den beim Training verwerteten Werken liegen, kann es sich um Vervielfältigungen im Sinne des § 16 UrhG oder um Bearbeitungen oder andere Umgestaltungen im Sinne des § 23 UrhG handeln.¹³² Wird der Output im Anschluss weiterverwendet, etwa durch Download, Upload oder Vertrieb, kommen weitere Verletzungstatbestände in Betracht (z.B. §§ 17, 19a UrhG).

1. Vervielfältigung und Bearbeitung: Entscheidung nach Einzelfall

Die Beurteilung des Eingriffs in das Vervielfältigungs- oder das Bearbeitungsrecht hängt von den Umständen des Einzelfalls ab. Unter das Tatbestandsmerkmal der Vervielfältigung fällt nicht nur die identische Wiedergabe der zum Training verwerteten Werke. Der künstlerische oder literarische Stil ist zwar nach herrschender, wenngleich zweifelhafter Doktrin nicht

130 Vgl. insoweit nochmals allgemein EuGH GRUR 2016, 1152 Rn. 32 ff. – GS Media/Sanoma ua; EuGH GRUR 2017, 610 Rn. 29 ff. – Stichting Brein/Wullems; EuGH GRUR 2017, 790 Rn. 24 ff. – Stichting Brein/Ziggo ua [The Pirate Bay]; EuGH GRUR 2021, 1054 Rn. 61 ff. – Peterson/Google ua u. Elsevier/Cyando [YouTube und uploaded]; zudem z.B. BGH GRUR 2010, 628 Rn. 19 – Vorschaubilder I; überdies z.B. v. Ungern-Sternberg in Schricker/Loewenheim, Urheberrecht, 6. Aufl. 2020, § 19a UrhG Rn. 57.

131 Nicht analysiert werden soll der Fall, dass ein schutzfähiges Werk vom Nutzer bei Einsatz des KI-Systems – und damit nach dem eigentlichen Training – eingegeben wird und es in der Folge zum Output einer Vervielfältigung, Bearbeitung oder Umgestaltung dieses Werkes kommt.

132 v. Welser GRUR-Prax 2023, 57 f.; Pesch/Böhme GRUR 2023, 997 (1005); Vesala IIC 54 (2023), 351 (354); v. Welser GRUR-Prax 2023, 516 (517); Konertz WRP 2023, 796 (802); Heine GRUR-Prax 2024, 87 (87); K. Wagner MMR 2024, 298 (300).

geschützt.¹³³ Eine Vervielfältigung kann allerdings bereits bei Übernahme von Teilen eines Werks in den KI-Output vorliegen.¹³⁴ Auch Bearbeitungen und andere Umgestaltungen fallen zudem nach § 23 Abs. 1 Satz 1 UrhG grundsätzlich in den Bereich des Ausschließlichkeitsrechts des Urhebers. Für die Beurteilung, ob in diesen Fällen die geistig-schöpferische Gestaltung und Ausdrucksform als relevanter Schutzgegenstand auf unzulässige Art und Weise übernommen wurde, ist zu fragen, ob die „Eigenart in der Nachbildung erhalten bleibt und ein übereinstimmender Gesamteindruck besteht.“¹³⁵ Vor diesem Hintergrund besteht rechtspraktisch grundsätzlich kein Unterschied zu klassischen Szenarien der Bearbeitung und Umgestaltung durch einen *Menschen*, also ohne KI-Einsatz. Für die Beurteilung von KI-Output gilt es entsprechend, das betroffene Original mit dem künstlichen Erzeugnis zu vergleichen. Jedenfalls wenn der schöpferische Ausdruck und die Eigenart des Originals, und dazu kann und muss eben auch der individuelle und prägnante Stil eines Künstlers oder Autors zählen,¹³⁶ im KI-Output übernommen wurde, etwa bei Auswurf eines Bildes „Llama by Banksy“ oder bei Nachbildung von Charakteren und Figuren aus Comics,

133 Vgl. z.B. BGH GRUR 1970, 250 (250 f.) – Hummelfiguren III; BGH GRUR 1977, 547 (550) – Kettenkerze; Schulze in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 2 Rn. 45; überdies aber auch Ulmer, Urheber- und Verlagsrecht, 3. Aufl. 1980, § 19 V.2., S. 124 („Die Individualität der Werke kann von verschiedener Stärke sein. Am stärksten ist die Individualität, wenn das Werk den Stempel der Persönlichkeit des Urhebers trägt. Nicht nur bei künstlerischen, sondern auch bei wissenschaftlichen Werken kann durch Stil, Sprachgestaltung oder Gedankenführung der Schöpfer im Werk erkennbar sein.“); für eine instruktive Kritik (zur vergleichbaren Doktrin im US-Recht) vgl. Sobel, Elements of Style: Copyright, Similarity, and Generative AI, Harv. J. L. & Tech. 38 (forthcoming 2024), 1 (27 ff.) (einsehbar auf SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4832872 (zuletzt am 28. Juni 2024)).

134 Vgl. nur Bomhard/Gajec RD 2021, 472 (474); v. Welser GRUR-Prax 2023, 57 (58); Maamar ZUM 2023, 481 (489).

135 BGH GRUR 2023, 571 Rn. 28 – Vitrineneuchte; BGH GRUR 2022, 899 Rn. 56 ff. – Porsche 911; BGH GRUR 2014, 65 Rn. 36 – Beuys-Aktion; zudem Pesch/Böhme GRUR 2023, 997 (1005 f.); K. Wagner MMR 2024, 298 (300). Vgl. zur „Wiedererkennbarkeit“ bei Gegenständen mit Leistungsschutz auch BGH GRUR 2020, 843 Rn. 25 – Metall auf Metall IV; EuGH GRUR 2019, 919 Rn. 31 und Rn. 38 – Metall auf Metall III.

136 Vgl. nochmals Ulmer, Urheber- und Verlagsrecht, 3. Aufl. 1980, § 19 V.2., S. 124 („Die Individualität der Werke kann von verschiedener Stärke sein. Am stärksten ist die Individualität, wenn das Werk den Stempel der Persönlichkeit des Urhebers trägt. Nicht nur bei künstlerischen, sondern auch bei wissenschaftlichen Werken kann durch Stil [sic!], Sprachgestaltung oder Gedankenführung der Schöpfer im Werk erkennbar sein.“).

Romanen oder Filmen (z.B. „Snoopy“), dürfte der Schutzbereich des Urheberrechts bereits ohne eine umfassende Nachbildung oder sklavische Kopie des Originalwerks berührt sein.¹³⁷

Der Schutzbereich wird nach § 23 Abs. 1 Satz 2 UrhG allerdings verlassen, wenn das „neu geschaffene Werk einen hinreichenden Abstand zum benutzten Werk“ aufweist. Nach der Rechtsprechung des Bundesgerichtshofs erfordert dies, dass „die Nachbildung über eine eigene schöpferische Ausdruckskraft verfügt und die entlehnten eigenpersönlichen Züge des Originals angesichts der Eigenart der Nachbildung verblassen.“¹³⁸ Für Fälle eines entsprechend transformativen KI-Output ist die Anwendung der Norm allerdings nur dann anzunehmen, wenn kein Erfordernis an die schöpferische Qualität der neuen Gestaltung gestellt wird.¹³⁹ Wollte man für § 23 Abs. 1 Satz 2 UrhG das Vorliegen eines „Werks“ im Sinne des § 2 Abs. 2 UrhG verlangen, wäre die Anwendbarkeit auf Fälle der KI-Kreativität hingegen zu bezweifeln.¹⁴⁰

137 Vgl. hierzu und zu den Beispielen: Guadamuz GRUR Int. 2024, III (124 f.); zudem Pesch/Böhme GRUR 2023, 997 (1006); kritisch und instruktiv zudem Sobel, *Elements of Style: Copyright, Similarity, and Generative AI*, Harv. J. L. & Tech. 38 (forthcoming 2024), 1 (27 ff.) (einsehbar auf SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4832872 (zuletzt am 28. Juni 2024)); zum sogenannten *Snoopy problem* überdies Sag Houston L. Rev. 61 (2023), 295 (327 ff.); Samuelson, *Fair Use Defenses in Disruptive Technology Cases*, forthcoming U.C.L.A. L. Rev. 2024, S. 69 (einsehbar auf SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4631726 (zuletzt am 22. Juni 2024)); Lee/Cooper/Grimmelmann, *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain*, J. Copyright Soc'y of the U.S.A. (forthcoming 2024), S. 78 ff. (einsehbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4523551 (zuletzt am 27. Juni 2024)).

138 BGH GRUR 2023, 571 Rn. 28 – Vitrineneuchte; BGH GRUR 2022, 899 Rn. 56 – Porsche 911; BGH GRUR 2010, 628 Rn. 17 – Vorschaubilder; zudem K. Wagner MMR 2024, 298 (300); ähnlich Maamar ZUM 2023, 481 (490).

139 So etwa Pesch/Böhme GRUR 2023, 997 (1003 und 1005); wohl auch Maamar ZUM 2023, 481 (490); vgl. zudem BGH GRUR 2022, 899 Rn. 58 – Porsche 911 („Es kommt ... nicht mehr darauf an, ob die neue Gestaltung die Anforderungen an ein urheberrechtlich geschütztes Werk erfüllt.“).

140 Vgl. insoweit allg. Schulze in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 23 Rn. 38; zudem Konertz/Schönhof WRP 2024, 534 (537). Da KI keine eigenständigen „Werke“ erschaffen kann, dürfte die Ausnahme des § 23 Abs. 1 Satz 2 UrhG jedenfalls dann keine Anwendung finden, wenn es an hinreichend eigenschöpferischem Input vonseiten des Nutzers fehlt. Vgl. zum Schöpferprinzip und der Einordnung technisch autonom generierter KI-Erzeugnisse grundsätzlich Dornis GRUR 2019, 1252 (1255); Dornis GRUR 2021, 784 (785 ff.); Dornis E.I.P.R. 43 (2021), 570 (571 ff.); Dornis Yale J. Law & Tech. 22 (2020), 1 ff. Die denkbaren Szenarien eines eigenschöpferischen Inputs des Nutzers bei Einsatz des generativen KI-Systems (insbesondere in Form

2. Caveat: Keine „künstliche Doppelschöpfung“

Schließlich ist zu bezweifeln, ob die Berufung auf eine „künstliche Doppelschöpfung“ als Verteidigung gegen die vollständige, keinen hinreichenden Abstand wahrende oder gegen eine zumindest wiedererkennbare Vervielfältigung angeführt werden kann.

Für Rechtsverletzungen durch einen unmittelbar handelnden Menschen gilt: Handelt der scheinbare Rechtsverletzer bei der Erschaffung des mit einem älteren Werk übereinstimmenden Erzeugnisses ohne das ältere Werk zu kennen, schafft er ein schutzfähiges neues Werk. Diese sogenannte Doppelschöpfung greift weder in das Bearbeitungsrecht des § 23 Abs. 1 S. 1, noch in die Verwertungsrechte der §§ 15 ff. UrhG des Urhebers des älteren Werkes ein.¹⁴¹ Teleologische Grundlage dieser Lücke ist ein Spezifikum der urheberrechtlichen Schutzanforderung: Ein Werk muss nicht objektiv neu sein, es genügt die subjektive Neuheit des Geschaffenen als persönliche geistige Schöpfung des Urhebers. Erschafft der nachfolgende Schöpfer subjektiv neu, handelt es sich entsprechend um sein *eigenes* Werk, nicht um das Werk des zeitlich früheren Schöpfers. Umgekehrt liegt keine Doppelschöpfung vor, wenn beim Schaffensprozess das Werk eines anderen übernommen wurde.¹⁴²

Für KI-generierte Ergebnisse stellt sich die vorrangige Frage, ob KI überhaupt zu einer „Schöpfung“ durch Schaffung subjektiv neuer Inhalte in der Lage ist. Dieser Aspekt wird für die „künstliche Doppelschöpfung“ bislang kaum diskutiert,¹⁴³ ist allerdings vor dem aufgezeigten teleologischen Hintergrund entscheidend: KI kann keine „Werke“ im Sinne des § 2 Abs. 2 UrhG erschaffen, kommen nach dem anthropozentrischen Schöpferprinzip

kreativer *prompts*) sollen hier nicht vertieft analysiert werden. Vgl. insoweit z.B. Finke ZGE 15 (2023), 414 (424 f.)

141 Vgl. nur BGH GRUR 1988, 810 (811) – Fantasy; BGH GRUR 1971, 266 (268) – Magdalenenarie; Bullinger in Wandtke/Bullinger, UrhR, 6. Aufl. 2022, § 23 Rn. 48.

142 Schulze GRUR 1984, 400 (410); Loewenheim in Schricker/Loewenheim, Urheberrecht, 6. Aufl. 2020, § 23 Rn. 36; Leistner in Schricker/Loewenheim, Urheberrecht, 6. Aufl. 2020, § 2 Rn. 64 f.; Finke ZGE 15 (2023), 414 (424 f.); kritisch insoweit allerdings Sesing-Wagenpfeil ZGE 16 (2024), 212 (226 und 389).

143 Überwiegend wird die Doppelschöpfung deshalb auch bei KI-Output grundsätzlich für möglich gehalten, so z.B. bei Baumann NJW 2023, 3673 (3677); Nordemann GRUR 2024, 1 (1 f.); Hofmann WRP 2024, 11 (17); Konertz/Schönhof WRP 2024, 534 (537); wohl auch bei Konertz WRP 2023, 796 (803); Käde ZUM 2024, 174 (182 Fn. 49); K. Wagner MMR 2024, 298 (301); zum britischen Recht vgl. Guadamuz GRUR Int. 2024, 111 (123).

doch nur Menschen als Schöpfer in Betracht.¹⁴⁴ Der Einwand der Doppelschöpfung muss deshalb bereits an dieser Stelle scheitern. Auch die vollkommen unabhängige Reproduktion eines geschützten Werkes durch eine KI, welche nachweisbar nicht mit diesem Werk trainiert wurde und auch sonst keinen Zugriff auf das Werk hatte, fällt daher nicht unter den Tatbestand der Doppelschöpfung.¹⁴⁵

Wollte man den Einwand der Doppelschöpfung dennoch mit der wohl überwiegenden Ansicht zulassen,¹⁴⁶ wäre zu fragen, ob der Output unabhängig vom beeinträchtigten Werk – also ohne Kenntnis – erstellt wurde.¹⁴⁷ Da zufällige Identität vor allem in Fällen der weitreichenden Übereinstimmung unwahrscheinlich ist, muss, wer sich auf eine Doppelschöpfung beruft, den Anscheinsbeweis gegen sich gelten lassen, dass er durch das ältere Werk inspiriert wurde und deshalb keine Doppelschöpfung vorliegt.¹⁴⁸ Bei KI-Systemen muss der Doppelschöpfungseinwand, falls das Ausgangswerk Bestandteil der Trainingsdaten war, ohne Erfolg bleiben. Die für menschliche Doppelschöpfungen bestehende Vermutung muss entsprechend gelten: Das Original war jedenfalls im „künstlichen Unterbewusstsein“ des Systems vorhanden und wurde bei der Produktion des KI-Outputs berücksichtigt.¹⁴⁹ Der Einwand wäre allenfalls dann denkbar, wenn Nutzer und Entwickler des KI-Systems belegen könnten, dass das Original nicht in den

144 Dies ist allgemeine Ansicht, auch in anderen Rechtsordnungen. Vgl. umfassend Dornis GRUR 2019, 1252 (1255); Dornis GRUR 2021, 784 (785 ff.); Dornis Yale J. Law & Tech. 22 (2020), 1 ff.

145 So zutreffend Finke ZGE 15 (2023), 414 (424 f.); a.A. Konertz WRP 2023, 796 (803); Hofmann WRP 2024, II (17); Konertz/Schönhof WRP 2024, 534 (537).

146 Vgl. nochmals Baumann NJW 2023, 3673 (3677); Nordemann GRUR 2024, I (1 f.); Hofmann WRP 2024, II (17); Konertz/Schönhof WRP 2024, 534 (537); wohl auch bei Konertz WRP 2023, 796 (803); Käde ZUM 2024, 174 (182 Fn. 49); K. Wagner MMR 2024, 298 (301); Sesing-Wagenpfeil ZGE 16 (2024), 212 (238).

147 Vgl. hierzu allgemein BGH GRUR 1971, 266 (268) – Magdalenenarie; KG GRUR-RR 2002, 49 (50) – Vaterland; Bullinger in Wandtke/Bullinger, UrhR, 6. Aufl. 2022, § 23 Rn. 47.

148 BGH GRUR 1971, 266 (268) – Magdalenenarie; noch strenger (Vollbeweis): OLG Köln GRUR 2002, 43 (44) – Klammerpose; zudem Bullinger in Wandtke/Bullinger, UrhR, 6. Aufl. 2022, § 23 Rn. 49.

149 Vgl. nochmals BGH GRUR 1971, 266 (268) – Magdalenenarie; zudem Konertz WRP 2023, 796 (803); Nordemann GRUR 2024, I (1 f.); Apel RD 2024, 109 (114 f.); K. Wagner MMR 2024, 298 (301).

Trainingsdaten enthalten war und überdies auch im weiteren Verlauf des KI-Einsatzes keine Eingabe erfolgte.¹⁵⁰

3. Zwischenergebnis

Auch bei einem Einsatz von generativen KI-Modellen durch die Nutzer kann es zur Vervielfältigung und Bearbeitung der Trainingsdaten kommen. Die Beurteilung der Rechtsverletzung ist stets eine Frage des Einzelfalls, insbesondere der Nähe des Gesamteindrucks von Original und KI-Output. Der Einwand der „künstlichen Doppelschöpfung“ ist in diesen Fällen bereits dogmatisch, jedenfalls aber auch rechtspraktisch ausgeschlossen.

150 Ähnlich Käde ZUM 2024, 174 (182) (Werk darf nicht Bestandteil der Trainingsdaten sein); zudem auch nochmals Guadamuz GRUR Int. 2024, 111 (123 f.).

