



Mittelstand 4.0
Kompetenzzentrum
Chemnitz

Betrieb 4.0
machen!



**Recht
behalten!**



Nachgelesen

KI-Trainingsdaten schützen

Michael Rätze

Mittelstand-
Digital

Gefördert durch:



Bundesministerium
für Wirtschaft
und Energie

aufgrund eines Beschlusses
des Deutschen Bundestages

Für die Entwicklung und Optimierung eines KI-Systems werden eine Vielzahl an Daten benötigt. Oft genügen reine Rohdaten nicht, sondern es sind entsprechend aufbereitete KI-Trainingsdaten erforderlich. Deren Generierung bindet wirtschaftliche Kapazitäten, weshalb sich die Frage nach ihrem möglichen Schutz stellt.

In dieser Ausgabe unserer *Nachgelesen*-Reihe erfahren Sie:

- Was KI-Trainingsdaten sind,
- welcher Wert ihnen zukommt sowie
- welche rechtlichen Möglichkeiten existieren, diese Daten zu schützen.

Was sind KI-Trainingsdaten?

Der allgemeine Lebenszyklus eines KI-Systems verläuft regelmäßig dergestalt, dass zunächst Rohdaten zu Trainingsdaten veredelt und anschließend in die KI-Komponente eines Gesamtsystems eingepflegt werden.

Veredeln bedeutet, entsprechende Daten auszuwählen und zu ordnen oder zu sortieren und zur Verfügung zu stellen. Eine quantitativ und qualitativ hochwertige Grundlage an Trainingsdaten ist entscheidend für den

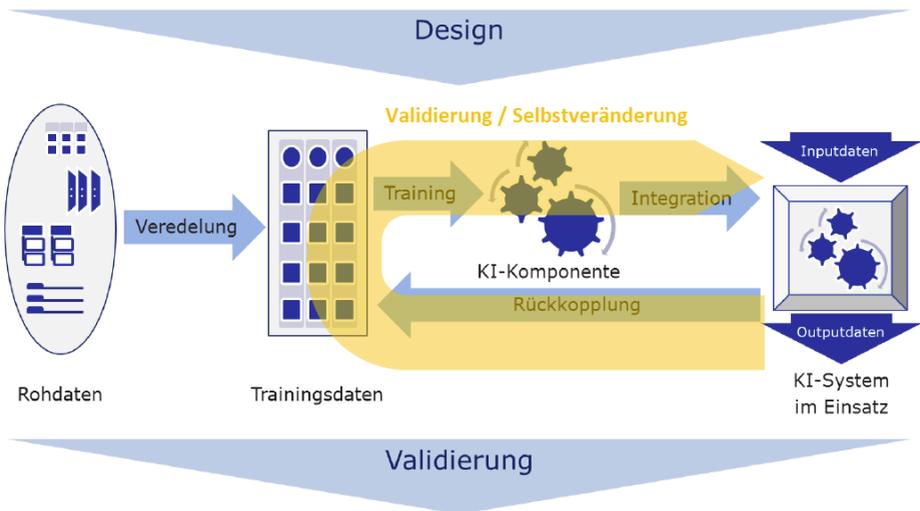


Abb. 1: Allgemeiner Lebenszyklus von KI-Systemen¹

Erfolg der KI-Komponente und damit des Gesamtsystems. Die Sammlung von Rohdaten und die darauffolgende Veredelung zu Trainingsdaten kann sehr aufwendig sein. Daher ergibt sich auch deren wirtschaftlicher Wert. Die Unternehmen investieren Zeit und Geld in die Veredelung der Rohdaten zu Trainingsdaten und

möchten häufig nicht, dass Konkurrenten sich dieser bedienen und so ihrerseits den wirtschaftlichen Aufwand für den Betrieb eines solches Systems minimieren können. Daher und aufgrund der Rolle von Trainingsdaten für KI-Systeme stellt sich die Frage, ob Trainingsdaten rechtlich geschützt werden können.



KI-Trainingsdaten rechtlich schützen

Der gegenwärtigen Rechtslage entsprechend kommen hierfür der urheberrechtliche Schutz der Darstellungen wissenschaftlicher oder technischer Art, der Datenbank- und Computerprogrammenschutz, der wettbewerbsrechtliche Leistungsschutz und das Geschäftsgeheimnis in Betracht.

1. Schutz als Darstellung wissenschaftlicher oder technischer Art

§ 2 Abs. 1 Nr. 7 UrhG (Urheberrechtsgesetz) schützt die Darstellungen wissenschaftlicher oder technischer Art, wie Zeichnungen, Pläne, Karten, Tabellen und plastische Darstellungen. Für einen Schutz genügt bereits die Darstellung einfachster wissenschaftlicher Erkenntnisse.² Eine Aufzählung bloßer Fakten genügt regelmäßig nicht.³ Denn Werke im Sinne

des Urheberrechts müssen Merkmale einer persönlich geistigen Schöpfung aufweisen, § 2 Abs. 2 UrhG. Hier muss der Einzelfall und speziell der Prozess der Veredelung betrachtet werden. Erschöpft sich die Veredelung in der bloßen Zusammenstellung einzelner Daten, selbst wenn dies große Mühe und erheblichen Zeitaufwand bedeutet, ist ein Schutz wohl zu verneinen.⁴ Werden die Rohdaten durch die Veredelung jedoch dergestalt aufbereitet, dass sie eine innovative Einteilung, Anordnung oder Art der Sammlung darstellen, kann die notwendige Schöpfungshöhe erreicht sein.⁵ Dann wäre der Schutz durch das Urheberrecht gegeben.

2. Schutz als Datenbank

Ein Schutz der Trainingsdaten als Datenbank kann als sog. Datenbankwerk gem. § 4 Abs. 2 UrhG oder als sog. sui-generis-Schutz („(nach

eigener Art“) gem. §§ 87 a ff. UrhG erfolgen. Beide Schutzmöglichkeiten stellen auf den gleichen Datenbankbegriff ab. Sie unterscheiden sich lediglich darin, dass § 4 Abs. 2 UrhG eine persönlich geistige Schöpfung erfordert und §§ 87 a ff. UrhG wesentliche Investitionen für die Beschaffung, Überprüfung oder Darstellung der unabhängigen Elemente verlangt. Datenbanken sind demnach jede Sammlung von Werken, Daten oder anderen unabhängigen Elementen, die systematisch oder methodisch angeordnet und einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sind.⁶

Die persönlich geistige Schöpfung für den Schutz nach § 4 Abs. 2 UrhG, vergleichbar dem Schutz nach § 2 Abs. 2 UrhG, ließe sich mit hinreichender Innovation der Einteilung, Anordnung oder Art der Sammlung bejahen. Die für den Schutz nach § 87a ff. UrhG erforderlichen wesentlichen Investitionen sind grundsätzlich weit zu verstehen. Erfasst sind nicht lediglich finanzielle Investitionen, sondern auch der Einsatz von Zeit, Arbeit und Energie.⁷ Dies trifft auf diejenigen zu, die die Rohdaten zu Trainingsdaten veredelt und verspricht einen entsprechenden Schutz.

3. Schutz als Computerprogramm

Als dritte Möglichkeit kommt der Schutz der Trainingsdaten als Computerprogramm gem. § 69 a ff. UrhG

in Frage. Danach sind Computerprogramme im Sinne des Gesetzes Programme in jeder Gestalt, einschließlich des Entwurfsmaterials. Den Begriff greift auch die DIN 44300 auf und beschreibt ihn als „eine zur Lösung einer Aufgabe vollständige Anweisung zusammen mit allen erforderlichen Vereinbarungen“.⁸ Computerprogramme werden demnach sehr weit gefasst. Geschützt sind etwa Betriebssysteme, Anwendungsprogramme, Hilfsprogramme, Suchmaschinen, E-Mail-Software, Software-Agenten oder Route-Software.⁹ Dennoch fallen Trainingsdaten als bloße Datensammlung, da sie selbst keine Befehls- oder Steueranweisungen an eine andere Komponente enthalten, wohl nicht darunter.⁹

4. Schutz als wettbewerblcher Leistungsschutz

Eine weitere Schutzmöglichkeit kann das Wettbewerbsrecht bieten, speziell der ergänzende wettbewerbsrechtliche Leistungsschutz gem. § 4 Nr. 3 UWG (Gesetz gegen den unlauteren Wettbewerb). Zweck dieser Norm ist es, Unternehmen vor der Ausbeutung von ihnen geschaffener Leistungen mittels unlauterer Mittel oder Methoden durch einen Mitbewerber zu schützen.¹⁰ Es kann die Gefahr bestehen, dass sich ein Mitbewerber der Trainingsdaten bedient, um sein eigenes KI-System zu trainie-

ren.¹¹ Voraussetzung ist u. a. aber, dass die Trainingsdaten eine wettbewerbliche Eigenart ausweisen.¹² Eine solche könnte angenommen werden, wenn die konkrete Ausgestaltung oder bestimmte Merkmale der Trainingsdaten geeignet sind, die interessierten Verkehrskreise auf ihre betriebliche Herkunft oder ihre Besonderheiten hinzuweisen.¹³ Für den bloßen Trainingsdatensatz dürfte dies nur schwerlich möglich sein.¹⁴ Denkbar wären solche Ausgestaltungen und Merkmale eher für die KI-Komponente oder das Gesamtsystem, welche sich von anderen Lösungen durch ihre Funktionalität abheben können.

5. Schutz als Geschäftsgeheimnis

Trainingsdaten können ein Geschäftsgeheimnis gem. § 2 Nr. 1 GeschGeh (Geschäftsgeheimnisgesetz) sein. Dafür sind drei Voraussetzungen zu erfüllen. Es muss sich um eine geheime Information handeln, die von wirtschaftlichem Wert ist und

Gegenstand geeigneter Geheimhaltungsmaßnahmen ist. Der Begriff der Information wird vom Gesetzgeber und der juristischen Literatur sehr weit verstanden. Erfasst sein sollen etwa Herstellungsverfahren, Kunden- und Lieferantenlisten, Geschäftsstrategien, Unternehmensdaten, Marktanalysen, Prototypen, Formeln und Rezepte.¹⁵ Man könnte Informationen auch schlicht als Wissen um eine Tatsache definieren.¹⁶ Trainingsdaten bzw. deren Zusammenstellung lassen sich als solche Information begreifen. Ihnen kommt auch ein wirtschaftlicher Wert gestalt zu, dass lediglich das entwickelnde Unternehmen seine KI-Komponente damit trainieren kann und – wie bereits erwähnt – deren Erstellung teils mit einem erheblichen Arbeitsaufwand einhergeht. Letztlich sind angemessene Geheimhaltungsmaßnahmen zu ergreifen, also ein Tätigwerden auf Seiten des Unternehmens. Diese Maßnahmen können faktischer oder rechtlicher Natur sein.¹⁷ Fehlen entsprechende Maßnahmen, kann der Schutz durch das GeschGeh entfallen.¹⁸



Fazit

Trainingsdaten lassen sich über unterschiedliche Wege schützen. Namentlich durch das Urheberrechtsgesetz als Darstellung wissenschaftlicher oder technischer Art gem. § 2 Abs. 1 Nr. 7 UrhG, als Datenbank gem. § 4 Abs. 2 UrhG und dane-

ben gem. §§ 87a ff. UrhG und letztlich über das Geschäftsgeheimnisgesetz. Als nicht zielführend stellte sich der Schutz als Computerprogramm gem. §§ 69a ff. UrhG und der wettbewerbliche Leistungsschutz gem. § 4 Nr. UWG heraus.

Anmerkungen

- ¹ Positionspapier der Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder-06.11.2019.
- ² BGH GRUR 2011, 803 Rn. 43; Bullinger, in: Wandtke/Bullinger, UrhR, 3. Aufl., § 2 UrhG Rdnr. 132; Loewenheim, in: Schrickler/Loewenheim, § 2 UrhG Rdnr. 197.
- ³ OLG Hamburg UFITA 51/1968, 383.
- ⁴ BGH GRUR 1980 227, 231.
- ⁵ Hacker, GRUR 2020, 1025, 1028.
- ⁶ Dreier, Dreier/Schulze, § 87a UrhG, Rn. 3.
- ⁷ Dreier, Dreier/Schulze, § 87a UrhG, Rn. 12.
- ⁸ Kaboth/Spies, BeckOK UrhG 30.Ed., § 69a, Rn. 2.
- ⁹ Dreier, Dreier/Schulze, § 69a UrhG, Rn. 12.
- ¹⁰ Micklitz/Schirmbacher, Spindler/Schuster, 4. Aufl. 2019, UWG § 4 Rn. 40.
- ¹¹ Köhler in Köhler/Bornkamm/Feddersen, UWG, 38. Aufl. 2020, § 4 Rn. 3.3.
- ¹² Hacker, GRUR 2020, 1025, 1032.
- ¹³ BGH GRUR 2010, 80 Rn.23 – LIKEaBIKE; GRUR 2012, 58 Rn.43 – Seilzirkus.
- ¹⁴ Hacker, GRUR 2020, 1025, 1032.
- ¹⁵ BT-Drs. 19/4724, 24.
- ¹⁶ BeckOK UWG/Kalbfus, 4. Ed. 29.6.2017, UWG § 17 Rn. 11.
- ¹⁷ Alexander, Köhler/Bornkamm/Feddersen, UWG 21. Aufl., § 2 GeschGeh, Rn. 49.
- ¹⁸ Fuhlrott/Hiéramente DB 2019, 967, 968.

Autor

Michael Rätze ist wissenschaftlicher Mitarbeiter an der Professur für Privatrecht und Recht des geistigen Eigentums von Prof. Dr. Dagmar Gesmann-Nuissl an der Technischen Universität Chemnitz. Im Mittelstand 4.0-Kompetenzzentrum Chemnitz ist er als Fachkoordinator Recht 4.0 tätig.

michael.raetze@betrieb-machen.de

Weitere Informationen

Das Mittelstand 4.0-Kompetenzzentrum Chemnitz gehört zu Mittelstand-Digital. Mit Mittelstand-Digital unterstützt das Bundesministerium für Wirtschaft und Energie die Digitalisierung in kleinen und mittleren Unternehmen und dem Handwerk.

Was ist Mittelstand-Digital?

Mittelstand-Digital informiert kleine und mittlere Unternehmen über die Chancen und Herausforderungen der Digitalisierung. Die geförderten Kompetenzzentren helfen mit Expertenwissen, Demonstrationszentren, Best-Practice-Beispielen sowie Netzwerken, die dem Erfahrungsaustausch dienen. Das Bundesministerium für Wirtschaft und Energie (BMWi) ermöglicht die kostenfreie Nutzung aller Angebote von Mittelstand-Digital.

Der DLR Projektträger begleitet im Auftrag des BMWi die Kompetenzzentren fachlich und sorgt für eine bedarfs- und mittelstandsgerechte Umsetzung der Angebote. Das Wissenschaftliche Institut für Infrastruktur und Kommunikationsdienste (WIK) unterstützt mit wissenschaftlicher Begleitung, Vernetzung und Öffentlichkeitsarbeit.

Weitere Informationen finden Sie unter www.mittelstand-digital.de

IMPRESSUM:

Herausgeber:

Mittelstand 4.0-Kompetenzzentrum Chemnitz
Geschäftsstelle
c/o Technische Universität Chemnitz
Prof. Dr.-Ing. habil. Ralph Riedel
DE - 09107 Chemnitz
Tel: 0371 531 19935
Fax: 0371 531 819935
E-Mail: info@betrieb-machen.de
Web: www.betrieb-machen.de
www.kompetenzzentrum-chemnitz.digital

Redaktion & Gestaltung

Michael Rätze & Anikó Lessi

Bildnachweis Titel:

Bild @ Sebdeck - Freepik.com

Veröffentlichung:

August 2021