Hochschule Konstanz
Technik, Wirtschaft und Gestaltung

Autoren:
Lucas Zoller, LL.M.
Manuel Treiterer, LL.M.
Prof. Dr. Marc Strittmatter

Kurzgutachten

über die rechtliche Schutzfähigkeit von Machine-Learning-Modellen am Beispiel von dezentral trainierten neuronalen Netzen

erstattet dem
Karlsruher Institut für Technologie

Karlsruher Institut für Technologie

Konstanz/ Karlsruhe
im Februar 2022

Inhaltsverzeichnis

I. Arbeitsauftrag und Kurzbeschreibung des Projekts ... 1

II. Zusammenfassung der Ergebnisse ... 2

III. Einleitung ... 4

A. Darstellung eines neuronalen Netzes und dessen Bestandteile ... 5

B. Einsatz von Federated Learning Modellen im Fertigungsbereich anhand von neuronalen Netzen ... 7

IV. Die rechtliche Schutzfähigkeit von Maschinendaten ... 9

A. Rechtslage bezüglich Maschinendaten ... 9
1. Sachenrecht ... 9
2. Immaterialgüterrecht ... 11
a) Urhebergesetz ... 11
(1) Computerprogramm ... 11
(2) Datenbankwerk ... 12
(3) Datenbank ... 13
(a) Unabhängige Elemente ... 13
(b) Investitionsleistung ... 16
(i) Beschaffung ... 16
(ii) Überprüfung & Darstellung ... 19
b) Geschäftsgeheimnisgesetz ... 20
(1) Geheimsein ... 20
(2) Wirtschaftlicher Wert ... 21
(3) Berechtigtes Interesse ... 22
(4) Angemessene Geheimhaltungsmaßnahmen ... 22
3. Strafrecht ... 24
4. Deliktsrecht ... 26

B. Ergebnis ... 27

V. Die rechtliche Schutzfähigkeit von dezentral trainierten neuronalen Netzen ... 28

A. Schutz von untrainierten neuronalen Netzen ... 28
1. Netzarchitektur ... 28
a) Urhebergesetz ... 29
(1) Computerprogramm – § 69a UrhG ... 29
(2) Datenbankwerk – § 4 Abs. 2 UrhG ... 32
(3) Datenbank – § 87a UrhG ... 33
b) Geschäftsgeheimnisgesetz ... 34
2. Lern- und Optimierungsfunktion ... 36
a) Urhebergesetz ... 37
b) Geschäftsgeheimnisgesetz ... 38
3. Hyperparameter ... 39

B. Schutz von trainierten neuronalen Netzen ... 41
1. Test- und Trainingsdaten ... 41

a) Urhebergesetz ... 41
(1) Darstellungen wiss. oder techn. Art – § 2 Abs. 1 Nr. 7 UrhG ... 41
(2) Datenbankwerk – § 4 Abs. 2 UrhG ... 43
(3) Datenbank – § 87a UrhG ... 46
b) Geschäftsgeheimnisgesetz ... 48
2. Trainierte Gewichtungsdaten ... 50
a) Computerprogramm – § 69a UrhG ... 50
b) Datenbankwerk – § 4 Abs. 2 UrhG ... 52
c) Datenbank – § 87a UrhG ... 53
d) Geschäftsgeheimnisgesetz ... 54
C. Schutz von fusionierten neuronalen Netzen ... 56
1. Sammlungen von trainierten Gewichtungsdaten ... 56
a) Datenbankwerk – § 4 Abs. 2 UrhG ... 56
b) Datenbank – § 87a UrhG ... 57
c) Geschäftsgeheimnisgesetz ... 59
2. Fusionierte Gewichtungsdaten ... 60
a) Computerschutz – § 69a UrhG ... 60
b) Datenbankwerk – § 4 Abs. 2 UrhG ... 60
c) Datenbank – § 87a UrhG ... 61
d) Geschäftsgeheimnisschutz ... 61
VI. Regelungsansätze zur Lösung entstehender Zuordnungsfragen ... 64
A. Zentrale Stelle mit übergeordneter Position ... 64
B. Cross-Licensing/ Netzwerkverträge ... 66
C. Datentreuhand ... 67
VII. Weiterführende Rechtsfragen ... 70
Literaturverzeichnis ... V

Abbildungsverzeichnis

Abbildung 1: Skizzenhafte Darstellung des Zusammenspiels von Transferfunktion, Aktivierungsfunktion und Neuronen 6
Abbildung 2: Skizzenhafte Darstellung eines neuronalen Netzes 7
Abbildung 3: Schematische Darstellung eines Federated Learning-Prozesses 8
Abbildung 4: Federated Learning Konstellation (1/3)..... 64
Abbildung 5: Federated Learning Konstellation (2/3)..... 67
Abbildung 6: Federated Learning Konstellation (3/3)..... 69

Abkürzungsverzeichnis

BGB	Bürgerliches Gesetzbuch
DB-RL	EG-Richtlinie zum rechtlichen Schutz von Datenbanken
DGA-E	Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates über europäische Daten-Governance
EuGH	Europäischer Gerichtshof
GeschGehG	Geschäftsgeheimnisgesetz
KI	Künstliche Intelligenz
ML	Machine Learning
StGB	Strafgesetzbuch
TS-RL	EU-Richtlinie über den Schutz vertraulichen Know-hows und vertraulicher Geschäftsinformationen
UrhG	Urhebergesetz

I. Kurzbeschreibung des Service-Meister Projekts und Arbeitsauftrag

Das Service-Meister Projekt setzte sich 2019 im KI-Innovationswettbewerb des Bundesministeriums für Wirtschaft und Energie durch. Ziel des Projekts ist die Entwicklung einer Plattform, um Service-Prozesse mit künstlicher Intelligenz zu unterstützen. Konkret soll damit erreicht werden, Service-Wissen für Anlagen und Prozesse des Mittelstands verfügbar zu machen.
Diese Notwendigkeit besteht unter anderem aufgrund neuer Geschäftsmodelle und des ansteigenden Fachkräftemangels innerhalb des deutschen Mittelstands. Fachkräfte sollen beispielsweise mit der Unterstützung von digitalen Ratgebern befähigt werden, auch komplexe Dienstleistungen durchzuführen.
Das Projekt verfolgt mithin das Ziel, die Wettbewerbsfähigkeit des deutschen Mittelstands zu fördern.

Im Rahmen des Service-Meister Projekts stellte sich insbesondere die Frage, inwieweit Daten, Datensammlungen, künstliche Intelligenz (KI) und deren Umsetzungsformen sowie daraus entstehende Ergebnisse *de lege lata* rechtlich geschützt und zugeordnet werden können.
Das Karlsruher Institut für Technologie hat die Hochschule Konstanz - Technik, Wirtschaft und Gestaltung (HTWG) damit beauftragt, ein Kurzgutachten zu ausgewählten rechtlichen Fragestellungen zu erstellen.

Das vorliegende Kurzgutachten untersucht die Schutzfähigkeit von Maschinendaten und Machine-Learning-Modellen (ML-Modelle) am Beispiel des Federated Learning nach aktueller Rechtslage. Es ist insoweit als Kurzgutachten zu begreifen, als dass es eine konkrete Fragestellung der gegebenen Rechtssituation umfasst, jedoch keine weitergehenden Ansätze zur Rechtsfortbildung untersucht und entwickelt. Dies wäre einer weitergehenden Forschungsfrage vorbehalten. Die Schwerpunkt liegt darauf, eine Betrachtung der gegebenen Situation und der Möglichkeiten, mit dieser Situation auf Vertragsebene umzugehen.
Federated Learning ist eine kollaborative Form des maschinellen Lernens, bei welcher verschiedene dezentral organisierte Parteien mitwirken. Als konkreter Untersuchungsgegenstand werden ML-Modelle auf Basis neuronaler Netze herangezogen.
Des Weiteren werden am Ende des Gutachtens Regelungsansätze zur Lösung entstehender Zuordnungsfragen im Rahmen des Federated Learning aufgezeigt.

II. Zusammenfassung der Ergebnisse

Die Untersuchung der rechtlichen Schutzfähigkeit von dezentral trainierten neuronalen Netzen setzt eine vorangehende rechtliche Analyse der Rechtspositionen an maschinengenerierten Daten voraus. Maschinengenerierte Daten lassen sich allgemein definieren als maschinenlesbar codierte Information[1], die von einer Datenverarbeitungsanlage automatisch erzeugt und verarbeitet werden[2].

Weil Daten aufgrund ihrer fehlenden Verkörperung nicht als Sache gelten, können weder sachenrechtliche Eigentums- noch Besitzrechte an den Daten als solche begründet werden – sehr wohl hingegen an verkörperten Datenträgern.

Daten können durch strafrechtliche Normen geschützt sein, diese begründen jedoch, ähnlich wie deliktische Ansprüche, keine absoluten Rechte an Daten gegenüber jedermann.

Allgemein gibt es keine unmittelbar absolut wirkenden Rechtspositionen an Daten. Vielmehr können Daten, unter Erfüllung bestimmter Voraussetzungen anderer Schutzrechte, mittelbar mitgeschützt sein.

Der Datenbankschutz nach § 87a UrhG schützt beispielsweise die Datenbank und deren wesentlichen Teile als Ergebnis einer Investition. Vor einer Benutzung einzelner Daten durch Dritte schützt das Datenbankrecht des § 87a UrhG allerdings nicht.

Zudem schützt das GeschGehG Geschäftsgeheimnisse vor dem unbefugten Zugang durch Dritte. Daten, welche die Anforderungen eines Geschäftsgeheimnisses (insbesondere Informationseigenschaft, Geheimsein und angemessene Schutzmaßnahmen) erfüllen, dürfen demnach nicht durch einen unbefugten Zugang erlangt werden. Eine ausschließliche Rechtsposition an Daten, die nur dem Inhaber der Daten eine Nutzung oder Verwertung erlaubt, begründet das GeschGehG jedoch nicht. Vielmehr erhält der Geschäftsgeheimnisinhaber ein Abwehrrecht, um die faktische Ausschließlichkeitsposition an den Daten als Geschäftsgeheimnis aufrecht zu erhalten.

Aufbauend auf der rechtlichen Einordnung von maschinengenerierten Daten wurde die rechtliche Schutzfähigkeit von dezentral trainierten neuronalen Netzen untersucht. Dabei wurden neuronale Netze entlang eines Federated Learning-Prozesses in den Stadien untrainiert, trainiert und fusioniert betrachtet.

Ein *Federated* Learning-Prozess zeichnet sich dadurch aus, dass eine zentrale Stelle ein ML-Modell bei dezentral organisierten Parteien (auch „Clients“ genannt) trainiert, ohne dabei Einsicht auf die zugrundliegenden Daten der Clients zu erhalten. Die Clients übermitteln lediglich die trainierten Modelle zurück an die zentrale Stelle.

Mit Ausnahme der programmierten Netzarchitektur, kommt die rechtliche Prüfung zu dem Ergebnis, dass die betrachteten Bestandteile neuronaler Netze die Voraussetzungen der in

[1] Zech, Herbert: Daten als Wirtschaftsgut – Überlegungen zu einem „Recht des Datenerzeugers“ in: Computer und Recht, Heft 3, 2015, Seite 137; Zech, Information als Schutzgegenstand, Seite 55f

[2] Becker, Maximilian: „Schutzrechte an Maschinendaten und die Schnittstelle zum Personendatenschutz“ in: Festschrift für Karl-Heinz Fezer zum 70. Geburtstag, 2016, Seite 815f.

Frage kommenden urheberrechtlichen Schutzrechte in der Regel nicht erfüllen dürften. Die Besonderheiten eines Federated Learning-Prozesses haben dabei insbesondere keine Auswirkungen auf das Bestehen eines urheberrechtlichen Schutzes.

Ein Schutz als Geschäftsgeheimnis kommt hingegen grundsätzlich für alle Bestandteile eines neuronalen Netzes in Betracht, sofern die tatbestandlichen Voraussetzungen des GeschGehG erfüllt sind. Ausschlaggebend hierfür ist insbesondere das Vorliegen einer Information, deren Geheimsein und die Umsetzung angemessener Geheimhaltungsmaßnahmen durch den Geschäftsgeheimnisinhaber. Aufgrund der bestehenden Mehrpersonenkonstellation innerhalb eines Federated Learning-Prozesses, ergeben sich außerdem Schwierigkeiten hinsichtlich der Einstufung als Geschäftsgeheimnisinhaber.

Das GeschGehG selbst enthält keine Lösungsansätze bezogen auf die Rechtezuordnung in solchen Mehrpersonenkonstellationen, sodass bei fehlender vertraglicher Klarstellung sowohl die zentrale Stelle, als auch die teilnehmenden Clients als Geschäftsgeheimnisinhaber in Frage kommen können. Dies birgt wiederum Risiken, wie etwa eine fehlerhafte wirtschaftliche Verwertung und damit auch den Verlust des Geschäftsgeheimnisschutzes.

Mögliche Ansätze für eine Lösung der Rechtezuordnung in Federated Learning-Konstellationen skizziert das Gutachten im letzten Kapitel. Diese Ansätze sehen sowohl Grundzüge für vertragliche Lösungen zwischen den beteiligten Parteien vor, sowie den Einsatz von neutralen Dritten.

Durch die Betrachtung unterschiedlicher Lösungsansätze können die verschiedenen Interessen der beteiligten Parteien an der Rechtezuordnung und der Zugangseinräumung berücksichtigt werden.

III. Einleitung

ML-Modelle werden für die Entwicklung von Systemen mit künstlicher Intelligenz (KI) eingesetzt. Die Entwicklung derartiger Systeme erfordert in der Regel eine Vielzahl hochwertiger Test- und Trainingsdaten. Solche Daten können jedoch oftmals sensibel, datenschutzrechtlich relevant oder beides sein.

Federated Learning stellt eine alternative Herangehensweise dar, um ML-Modelle zu trainieren. Hierbei übermittelt eine zentrale Stelle untrainierte ML-Modelle an dezentral organisierte externe Parteien (auch „Clients“ genannt). Die zentrale Stelle kann beispielsweise ein Maschinenhersteller oder ein Hersteller eines Betriebssystems für Smartphones sein.
Die Clients wiederum wären in diesem Fall Maschinennutzer oder Nutzer von Smartphones. Bei diesen Clients werden die ML-Modelle lokal mit den Daten deren Maschinen oder Geräten trainiert. Die einzelnen Clients übermitteln die trainierten ML-Modelle wiederrum an die zentrale Stelle, welche die trainierten Modelle fusioniert. Unter der Fusionierung der Modelle versteht man die Kombination mehrerer trainierter Modelle zu einem aggregierten Modell. Die zentrale Stelle übermittelt das fusionierte Modell im Anschluss erneut an die Clients.

Dieser beschriebene Prozess zeichnet sich insbesondere dadurch aus, dass die zugrunde liegenden Test- und Trainingsdaten eines einzelnen Clients weder für die zentrale Stelle, noch für andere teilnehmende Clients, einsehbar sind. Diese Besonderheit erleichtert zugleich die Einhaltung regulatorischer Rahmenbedingungen wie beispielsweise datenschutzrechtliche Vorgaben oder auch Vertraulichkeitsanforderungen teilnehmender Clients.

Im Ergebnis ist die zentrale Stelle durch den Einsatz von Federated Learning in der Lage, Synergieeffekte durch das Netzwerk an teilnehmenden Clients zu generieren, ohne dass Letztere auf die Hoheit über ihre Daten verzichten müssen. Diese Synergieeffekte zeichnen sich im Regelfall durch eine bessere ML-Performance aus.

In der Praxis findet Federated Learning bereits in unterschiedlichen Branchen und Formen Anwendung. Neben Tech-Konzernen wie Google[3] oder Apple[4] nutzen auch mittlere und kleinere Unternehmen diesen Ansatz, um qualitativ hochwertige ML-Modelle zu erstellen. Nennenswert ist hierbei das ACR Data Science Institute im Gesundheitssektor[5] oder das deutsche Start-Up Unternehmen Prenode im klassischen Industriesektor.[6]

Nachfolgend wird die Schutzfähigkeit von ML-Modellen anhand von dezentral trainierten neuronalen Netzen untersucht.

Für ein einheitliches Verständnis des Untersuchungsgegenstands beschreibt das nachstehende Kapitel A zunächst überblicksartig den üblichen Aufbau und die Bestandteile

[3] vgl. Google: “Your chats stay private while Messages improves suggestions” in Google support online URL: https://support.google.com/messages/answer/9327902?hl=en, Abruf 14.02.2022.
[4] vgl. Apple: “Designing for privacy” in Apple WWDC online, URL: https://developer.apple.com/videos/play/wwdc2019/708, Abruf 14.02.2022.
[5] vgl. ACR Data Science Institute: "AI-LAB" in: dieselbe online, URL: https://ailab.acr.org, Abruf 14.02.2022.
[6] vgl. Prenode: "Decentralized Machine Learning für Industriemaschinen" in: dieselbe online, URL: https://prenode.de/de/industriemaschinen/, Abruf 14.02.2022.

neuronaler Netze. Das hieran anschließende Kapitel B skizziert ein Beispiel für den Einsatzbereich von Federated Learning in der Fertigungsindustrie.

Dies vorausgeschickt, widmet sich die rechtliche Prüfung zunächst der Frage nach der Schutzfähigkeit und den Rechten an Maschinendaten, um anschließend die rechtliche Schutzfähigkeit von neuronalen Netzen in einem Federated Learning-Prozess zu untersuchen. Hierbei werden neuronale Netze im untrainierten, trainierten und fusionierten Zustand betrachtet.

A. Darstellung eines neuronalen Netzes und typische Bestandteile

Im Folgenden werden als Verständnisvoraussetzung zunächst typische Bestandteile von neuronalen Netzen sowie Test- und Trainingsdaten erläutert, um einen Eindruck zu vermitteln, wie das Training neuronaler Netze in einem Federated Learning-Prozess erfolgt.

Test- und Trainingsdaten sind notwendig, um ein ML-Modell zielführend zu trainieren und dessen zu erwartende Performanz abschätzen zu können. Dabei wird der originale Datenbestand in Testdaten und Trainingsdaten getrennt. Während die Trainingsdaten zum Trainieren des neuronalen Netzes genutzt werden, werden die Testdaten verwendet, um die Fehlerquote und damit die Performanz nach Abschluss der Trainingsphase zu ermitteln.

Ein neuronales Netz besteht grundsätzlich aus den folgenden Elementen:[7]

- Netzarchitektur (Topologie): Die Netzarchitektur stellt das Fundament eines neuronalen Netzes dar und wird durch die Anzahl der Schichten, sowie die Anzahl der Neuronen pro Schicht definiert.

    - Eingabeschicht (engl. „input layer“): Die erste Schicht eines neuronalen Netzes bildet die Eingabeschicht. Hierbei wird der sogenannte Funktionsbereich (engl. „features“) der vorliegenden Daten verarbeitet.
    - Ausgabeschicht (engl. „output layer“): Die letzte Schicht eines neuronalen Netzes ist die Ausgabeschicht. Sie stellt das Ergebnis der Verarbeitung durch das neuronale Netz dar.
    - Verborgene Schicht (engl. „hidden layer“): Zwischen der Eingabeschicht und der Ausgabeschicht befinden sich verborgene Schichten. Hier findet die eigentliche Verarbeitung statt, indem Neuronen einen Satz gewichteter Eingaben aufnehmen und durch die Aktivierungsfunktion eine Ausgabe erzeugen.
    - Gewichte/ Gewichtungsdaten (engl. „weights“): Die Gewichte beschreiben die Intensität des Informationsflusses entlang der Verbindung zwischen zwei Neuronen. Neben den Gewichten ist die sogenannte Verzerrung (engl. „bias“) erwähnenswert. Die Hauptfunktion einer Verzerrung besteht darin, jedem Knoten einen trainierbaren konstanten Wert bereitzustellen. Dies unterstützt den erfolgreichen Lernprozess. Nachfolgend werden unter den Begriffen

[7] vgl. dazu Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 762-763.

„Gewichte“ und „Gewichtungsdaten“ sowohl die herkömmlichen Gewichte als auch die Verzerrung verstanden.

- Transferfunktion: Die Inputs der vorgelagerten Neuronen werden in die Transferfunktion gegeben und mit einem Faktor w (=Gewicht; engl. „weight“) multipliziert sowie mit einem Wert b, die sogenannte Verzerrung (engl. „bias“), addiert [Transferfunktion = f(x) = x*w + b].
- Aktivierungsfunktion: Die Aktivierungsfunktion erhält den Output der Transferfunktion, welcher mit einer linearen oder nicht-linearen Funktion verarbeitet wird. Diese bestimmt, ob und wie stark die nachfolgenden Neuronen aktiviert werden. In der Regel werden Aktivierungsfunktionen pro Schicht definiert. Die Aktivierungsfunktion wird nicht während des Trainings angepasst.

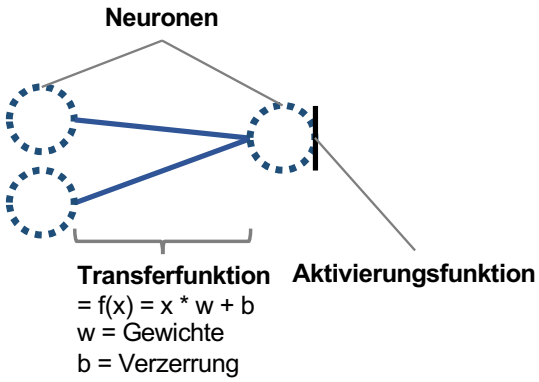

Abbildung *1*: Skizzenhafte Darstellung des Zusammenspiels von Transferfunktion, Aktivierungsfunktion und Neuronen

- Lern- und Optimierungsfunktion: Ziel der Optimierungsfunktion (engl. „loss function“) ist die Minimierung der Fehlerquote des Modells. Ein gängiges Verfahren hierzu ist die sogenannte Fehlerrückführung (engl. „backpropagation“), welche ausgehend von der Ausgabeschicht die Gewichte rekursiv anpasst, sodass die Fehlerquote minimiert wird.
- Hyperparameter: Die Hyperparameter sind Teil des Lern- und Optimierungs-Prozesses. Durch die Hyperparameter wird beispielsweise bestimmt, wie stark und wie schnell die Gewichte zwischen den Neuronen angepasst oder wie viele Optimierungsrunden durchgeführt werden sollen.
- Transformationsfunktion: Die Transformationsfunktion „übersetzt“ die Eingabedaten, welche das neuronale Netz verarbeiten soll. Beispielsweise wird eine schwarz-weiß Bilddatei mit 50x50 Pixeln passend durch die Transferfunktion auf 2500 Neuronen (=50*50) in der Eingabeschicht verteilt.
- Metadaten: Metadaten fallen automatisch beim Training an und liefern weitere Informationen über den Trainingsprozess. Beispiele für Metadaten in diesem Kontext sind die Fehlerquote/ -rate, weitere Performance-Daten des Modells oder auch der Zeitpunkt des Trainings.

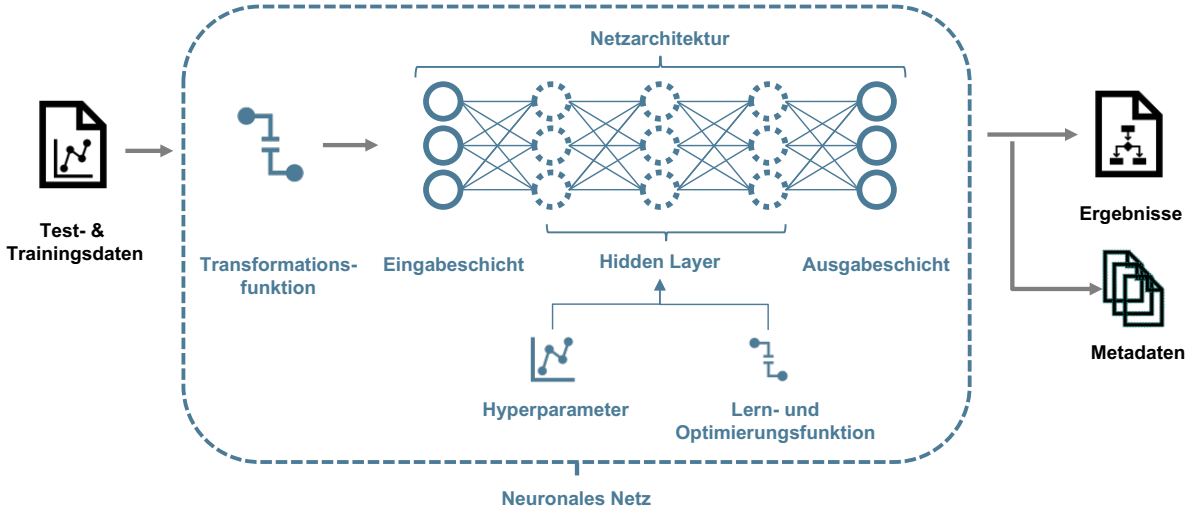

Abbildung 2: Skizzenhafte Darstellung eines neuronalen Netzes

B. Einsatz von Federated Learning Modellen im Fertigungsbereich anhand von neuronalen Netzen

Nachdem vorstehend neuronale Netze und deren Bestandteile kurz dargestellt wurden, sollen diese nachfolgend in den Kontext eines Federated Learning-Prozesses im Fertigungsbereich gesetzt werden.

Als Anwendungsbeispiel dient ein fiktiver Maschinenhersteller, welcher seinen Kunden mittels eines Federated Learning-Prozesses individuelle Predictive Maintenance Serviceleistungen anbietet. In diesem Fall wäre der Maschinenhersteller die zentrale Stelle und die Kunden die Clients innerhalb des Federated Learning-Prozesses. Als ML-Modell werden neuronale Netze verwendet, welche die Grundlage für die späteren Predictive Maintenance Serviceleistungen bilden.

Vor dem Trainingsprozess entscheidet der Maschinenhersteller über die konkrete Netzarchitektur, die Lern- und Optimierungsfunktion und die Hyperparameter des neuronalen Netzes. Der Maschinenhersteller entscheidet über die Hyperparameter, wie das Training der jeweiligen neuronalen Netze bei den Clients erfolgt. Nachdem der Maschinenhersteller das neuronale Netz erstellt hat, bildet dieses das untrainierte ML-Modell im Federated Learning-Prozess.

Der Maschinenhersteller übermittelt die untrainierten neuronalen Netze an seine Kunden und Maschinennutzer, wo das Training mit den lokalen Daten erfolgt. Dabei werden Daten verwendet, die im Zusammenhang mit der Prognose von Maschinenausfällen und Störungen stehen (beispielsweise Maschinentemperatur, Maschinenumgebungstemperatur, Maschinenauslastung oder Betriebsstunden). Konkret trainieren die neuronalen Netze anhand dieser Daten die Gewichtungsdaten der einzelnen Schichten innerhalb des neuronalen Netzes.

Da eine Vielzahl an Kunden des Maschinenherstellers mit diversen lokalen Daten am Federated Learning-Prozess teilnehmen, liegen am Ende unterschiedlich trainierte Modelle vor. Die Kunden übermitteln die trainierten Modelle zurück an den Maschinenhersteller, ohne dabei die zugrunde liegenden Daten zu transferieren.

Im Anschluss fusioniert der Maschinenhersteller die unterschiedlich trainierten Modelle mittels eines Fusionsalgorithmus mit dem Ziel, diese derartig zu aggregieren, dass das fusionierte Modell besser ist als die einzelnen Modelle, die bei den Kunden trainiert wurden.

Im hier betrachteten Anwendungsfall wäre die Erstellung von verlässlichen und sicheren Modellen für die Prognose von Maschinenausfällen und Störungen das Ziel des Federated Learning-Prozesses.

Durch den Einsatz von Federated Learning ist der Maschinenhersteller insbesondere in der Lage, verlässlichere Modelle zu erstellen und individuelle Besonderheiten (Maschinenkonfiguration, Maschinenauslastung, Maschinenumgebung) der jeweiligen Kunden zu berücksichtigen.

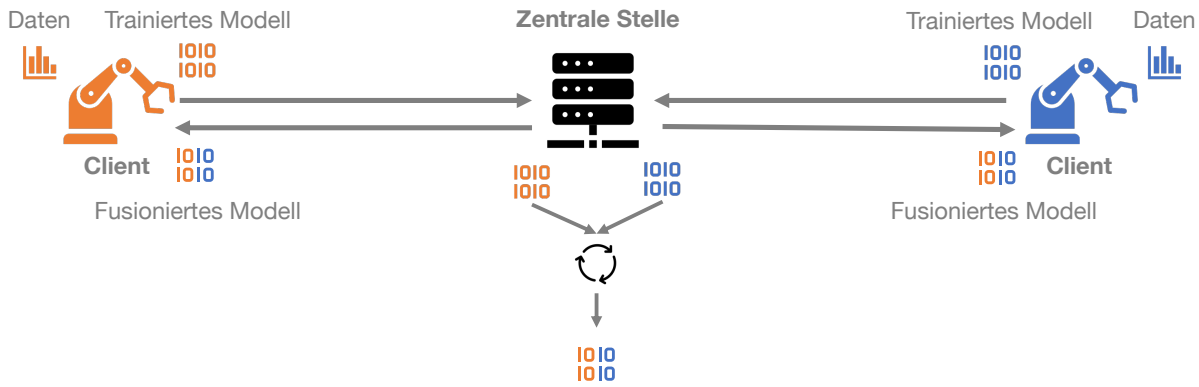

Abbildung 1: Schematische Darstellung eines Federated Learning-Prozesses

IV. Die rechtliche Schutzfähigkeit von Maschinendaten

Die Grundelemente des Federated Learning sind Computerprogramme und Daten. Besonders die Frage nach dem rechtlichen Schutz von Daten wirft viele Fragen auf, die für die rechtliche Einordnung von neuronalen Netzen in einem Federated Learning-Prozess entscheidend sind. Personenbezogene Daten, welche unter den Anwendungsbereich des Datenschutzrechts fallen, sind in den nachfolgenden Abschnitten explizit ausgeschlossen. Es sollen die Rechtspositionen an Maschinendaten, Industriedaten oder auch maschinengenerierte Daten aufgezeigt werden (nachfolgend "Maschinendaten“), welche keinen Personenbezug aufweisen. Maschinendaten lassen sich allgemein definieren als maschinenlesbar codierte Information[8], die von einer Datenverarbeitungsanlage automatisch erzeugt und verarbeitet werden[9].

A. Rechte an Maschinendaten

Ob an Maschinendaten Ausschließlichkeitsrechte bestehen können oder ob lediglich ein sogenannter Zugangsschutz besteht, wurde in den letzten Jahren in der Rechtswissenschaft intensiv diskutiert.[10] Neben der sachenrechtlichen Einordnung von Daten wurde der immaterial-güterrechtliche Schutz, der Schutz durch das Geschäftsgeheimnisgesetz, der strafrechtliche sowie der deliktische Schutz diskutiert. Im Folgenden wird der rechtswissenschaftliche Meinungsstand kurz dargestellt und auf den Untersuchungsgegenstand angewendet.

1. Sachenrecht

Der Gesetzgeber definiert Sachen in § 90 BGB als körperliche Gegenstände. Der Begriff eines Gegenstandes ist jedoch vom Gesetzgeber nicht legaldefiniert. Nach herrschender Meinung sind vom Gegenstandsbegriff alle individualisierbaren vermögenswerten Objekte und Güter umfasst, über die Rechtsmacht im Sinne von Herrschafts- oder Nutzungsrechten ausgeübt werden kann.[11] Gegenstände zeichnen sich also, anders als Sachen, nicht durch Körperlichkeit aus.[12]
Ohne Körperlichkeit kann an einem Gegenstand weder sachenrechtlicher Besitz noch Eigentum bestehen.[13] Körperlich ist ein Gegenstand, wenn er für den Menschen sinnlich wahrnehmbar, räumlich abgegrenzt und beherrschbar ist.[14] Diese Definition hilft nur bedingt,

---

[8] Zech, Herbert: Daten als Wirtschaftsgut – Überlegungen zu einem „Recht des Datenerzeugers“ in: Computer und Recht, Heft 3, 2015, Seite 137; Zech, Herbert: Information als Schutzgut, 2012, Seite 55ff.

[9] Becker, Maximilian: „Schutzrechte an Maschinendaten und die Schnittstelle zum Personendatenschutz“ in: Festschrift für Karl-Heinz Fezer zum 70. Geburtstag, 2016, Seite 815f.

[10] vgl. u.a. von Oelffen, Sabine: „Eigentum an Daten“ in: Ballestrem, Johannes; et al. (Hrsg.): Künstliche Intelligenz - Rechtsgrundlagen und Strategien in der Praxis, 2020, Seite 77ff.; vgl. Stender, Vorwachs Jutta; Steege, Hans: "Wem gehören unsere Daten?", in: NJOZ, Heft 36, 2018, Seite 1361ff.; Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 336ff.

[11] vgl. Stressemann, in: MüKo, BGB, § 90 Rn. 1; Fritzsche, in: BeckOK BGB, § 90 Rn. 4.

[12] vgl. Fritzsche, in: BeckOK BGB, § 90 Rn. 4.

[13] vgl. Fritzsche, in: BeckOK BGB, § 90 Rn. 5.

[14] vgl. Fritzsche, in: BeckOK BGB, § 90 Rn. 5.

da die Wahrnehmbarkeit die Beherrschbarkeit voraussetzt, die wiederum nur vorliegt, sofern eine Sache im Raum abgegrenzt werden kann.[15]
Die räumliche Abgrenzung setzt dabei eine eigene körperliche Begrenztheit oder eine Begrenztheit aus einer Fassung in einem Behälter oder durch eine künstliche Kennzeichnung voraus.[16]

Daten sind kein exklusives Gut und nicht-rival, da sich Daten beliebig vervielfältigen lassen und sich ihre Verbreitung faktisch grundsätzlich kaum kontrollieren lässt.[17] Somit fehlt eine räumliche Abgrenzung bei Daten. Ebenso sind Daten nicht sinnlich wahrnehmbar. Selbst wenn man Daten auf einem Datenträger oder mittels eines Computers als wahrnehmbar ansehen könnte, bleiben sie als solche, ähnliche wie Energie in einem Akku, nicht sinnlich wahrnehmbar.[18]

Aufgrund fehlender Körperlichkeit scheiden Daten als Sachen im Sinne des § 90 BGB aus.[19] Folglich kann an Daten weder Besitz noch Eigentum begründet werden. Davon zu unterscheiden ist ein vorliegender körperlicher Datenträger auf dem nicht-körperliche Daten gespeichert sind. Der Datenträger unterliegt sachenrechtlichen Vorschriften aufgrund seiner körperlichen Sacheigenschaft. Die darauf enthaltenen Daten bleiben hingegen ein nicht-körperlicher Gegenstand und somit ein immaterielles Gut.[20]

Daten stellen jedoch ein individualisierbares, vermögenswertes Objekt dar. Über Daten kann Rechtsmacht im Sinne von Herrschafts- und Nutzungsrechten ausgeübt werden, sodass sich Daten als – wenn auch nicht körperlicher – Gegenstand im rechtlichen Sinne einordnen lassen.[21]

Zusammenfassung – Sachenrecht
Daten stellen keine Sache dar, sondern immaterielle Güter, die als Gegenstände im rechtlichen Sinne einzustufen sind.[22] Der Datenträger als solches stellt eine Sache im Sinne des § 90 BGB dar, jedoch bleiben die Daten auf dem Datenträger ein immaterielles Gut bzw. ein Gegenstand im zivilrechtlichen Sinne.[23]
An Daten kann weder sachenrechtliches Eigentum noch Besitz begründet werden, sondern nur an einem entsprechenden Datenträger.

[15] vgl. Fritzsche, in: BeckOK BGB, § 90 Rn. 5.
[16] vgl. Fritzsche, in: BeckOK BGB, § 90 Rn. 7; vgl. Stressemann, in: MüKo, BGB, § 90 Rn. 8.
[17] vgl. Zech, Herbert: Information als Schutzgut, 2012, Seite 327.
[18] vgl. Fritzsche, in: BeckOK BGB, § 90 Rn. 6.
[19] vgl. Stender, Vorwachs Jutta; Steege, Hans: "Wem gehören unsere Daten?", in: NJOZ, Heft 36, 2018, Seite 1362.
[20] vgl. Fritzsche, in: BeckOK BGB, § 90 Rn. 25, 27; BGH, Urt. v. 13.10.2015 – VI ZR 271/14, NJW 2016, 1094, Rn. 20.
[21] vgl. BGH, Beschl. v. 21.9.2017 – I ZB 8/17 – GRUR 2018, 222, Rn. 15.
[22] vgl. Stressemann, in: MüKo, BGB, § 90 Rn. 25; vgl. Fritzsche, in: BeckOK BGB, § 90 Rn. 25; OLG Brandenburg, Urt. v. 6.11.2019 – 4 U 123/19 – NZI 2020, Rn. 42; vgl. BGH, Beschl. v. 21.9.2017 – I ZB 8/17, GRUR 2018, 222, Rn. 15; BGH, Urt. v. 04.11.1987 – VIII ZR 314/86, NJW 1988, 407.
[23] vgl. Fritzsche, in: BeckOK BGB, § 90 Rn. 25, 27; BGH, Urt. v. 13.10.2015 – VI ZR 271/14, NJW 2016, 1094, Rn. 20; vgl. Bartsch, Michael: "Software als Rechtsgut" in: Computer und Recht, Heft 9, 2010, Seite 558; vgl. Heydn, Truiken J.: "Identitätskrise eines Wirtschaftsguts: Software im Spannungsfeld zwischen Schuldrecht und Urheberrecht" in: Computer & Recht, Heft 12, 2010, Seite 772; vgl. Schur, Nico: Die Lizenzierung von Daten, 2020, Seite 42.

2. Immaterialgüterrecht

Da Daten immaterielle Güter darstellen, liegt es nahe, Rechte an Daten im Lichte des Immaterialgüterrechts zu prüfen. Dabei benennenswert sind die Schutzrechtspositionen im Urhebergesetz (UrhG) in Form von Datenbankwerken (§ 4 II UrhG), Computerprogrammen (§ 69a UrhG) und Datenbanken (§ 87a UrhG). Darüber hinaus soll der Schutz als Geschäftsgeheimnis im Sinne des Geschäftsgeheimnisgesetz (GeschGehG)[24] untersucht werden.
Nachfolgend wird dargestellt, ob und inwiefern Daten die Voraussetzungen dieser genannten Schutzrechte erfüllen können.

a) Urhebergesetz

Die aus dem Urhebergesetz infrage kommenden Schutzrechte an Daten lassen sich grundsätzlich in Urheberrechte (§§ 4, 69a UrhG) und verwandte Leistungsschutzrechte (§ 87a UrhG) trennen.
Sowohl Urheberrechte als auch verwandte Leistungsschutzrechte räumen dem jeweiligen Inhaber Ausschließlichkeitsrechte gegenüber jedermann ein, sodass ein Schutz durch diese beiden Rechtspositionen von hoher wirtschaftlicher Bedeutung ist.
Urheberrechte erfordern insbesondere eine persönliche geistige Schöpfung des Urhebers, wohingegen Leistungsschutzrechte entstehen, indem entweder eine Leistung als ausübender Künstler oder eine kaufmännisch-organisatorischen Tätigkeit zum Kulturschaffen beiträgt.[25]

(1) Computerprogramm

Computerprogramme sind gemäß § 69a Abs. 1 UrhG Programme in jeder Gestalt, einschließlich des Entwurfsmaterials. Von einer genaueren Legaldefinition des Computerprogramms sah sowohl der europäische als auch der deutsche Gesetzgeber ab, da ein gesetzlich definierter Begriff veralten könnte.[26]
Aufgrund der fehlenden Legaldefinition zieht die Rechtsprechung regelmäßig die Definitionen der WIPO und der DIN 44300 heran.[27] Gemäß § 1 (i) der Mustervorschriften der WIPO ist ein Computerprogramm „eine Folge von Befehlen, die nach Aufnahme in einen maschinenlesbaren Träger fähig sind zu bewirken, dass eine Maschine mit informationsverarbeitenden Fähigkeiten eine bestimmte Funktion oder Aufgabe oder ein bestimmtes Ergebnis anzeigt, ausführt oder erzielt."

---

[24] Gesetz zum Schutz von Geschäftsgeheimnissen vom 18. April 2019 (BGBl. I S. 466).
[25] vgl. Dreier, in: Dreier/Schulze, UrhG, Einleitung, Rn. 1.
[26] vgl. EG-Kommission: „Vorschlag für eine Richtlinie des Rates über den Rechtsschutz von Computerprogrammen", KOM (88) 89/C 91/95 vom 12.04.1989, Seite 6; vgl. Bundestag: Entwurf eines Zweiten Gesetzes zur Änderung des Urheberrechtsgesetzes, in: Bundestag online, Drucksache 12/4022, 1992, URL: https://dserver.bundestag.de/btd/12/040/1204022.pdf, Abruf 14.02.2022, Seite 9.
[27] vgl. BGH, Urt. v. 09.05.1985 – I ZR 52/83, NJW 1986, 192, 196; OLG Hamburg, Urt. v. 12.3.1998 – 3 U 226/97, MMR 1999, 230; Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 761-769; EuGH, Urt. v. 22. 12. 2010 – C-393/09, GRUR 2011, 220, 220.

Die Rechtsprechung wie auch weite Teile der Literatur orientieren sich an der Definition der WIPO, wonach ein Computerprogramm grundsätzlich eine Folge von Steuerungsbefehlen in einem konkreten Programmcode enthalten muss.[28]

Daten und Programmcode unterscheiden sich jedoch in genau diesem Merkmal. Daten zeichnen sich in der Regel dadurch aus, dass diese keine Folge von Steuerungsbefehlen verkörpern. Daten gelten ohne Weiteres auch nicht als geschütztes Entwurfsmaterial im Sinne des § 69a Abs. 1 UrhG eines Computerprogramms, da ein geschütztes Entwurfsmaterial nur vorliegt, sofern es zur Entwicklung eines Computerprogramms führen kann.[29]
Betrachtet man die Systematik der §§ 69a, 4 Abs. 2 und 87a UrhG erkennt man, dass auch der Gesetzgeber klar zwischen Daten ohne Steuerungsbefehl und Programmcode unterscheidet, da für Daten nur der Schutz durch das Datenbankwerk in § 4 Abs. 2 UrhG und der Datenbankschutz in § 87a UrhG in Frage kommen soll.[30]

Im Ergebnis sind einfache Daten kein geschütztes Computerprogramm im Sinne des § 69a UrhG.

(2) Datenbankwerk

Um die Schutzvoraussetzungen eines Datenbankwerks im Sinne des § 4 Abs. 2 UrhG zu erfüllen, muss zunächst ein Sammelwerk vorliegen. Gemäß der Legaldefinition in § 4 Abs. 1 UrhG ist ein Sammelwerk eine Sammlung von Werken, Daten oder anderen unabhängigen Elementen, die aufgrund der Auswahl oder Anordnung der Elemente eine persönliche geistige Schöpfung darstellen. Entscheidend ist also das Vorliegen einer persönlichen geistigen Schöpfung im Hinblick auf die Auswahl oder Anordnung,[31] die sich aus dem Ergebnis eines unmittelbaren und zielgerichteten geistigen Schaffens- bzw. Gestaltungsprozesses[32] durch einen Menschen als Schöpfer (§ 7 UrhG) ergibt.[33]

In den Industriekontext gesetzt, werden Sammlungen von Maschinendaten regelmäßig durch Industrieanlagen automatisch erzeugt und verarbeitet, ohne die unmittelbare Einwirkung durch einen Menschen. Eine persönliche geistige Schöpfung liegt allerdings nicht vor, sofern eine Maschine nicht nur als Hilfs- bzw. Ausführungsmittel eingesetzt wird.[34] Unabhängig davon, ob ein Mensch die auf einer Anlage ausgeführte Software programmiert hat oder eine Anlage durch einen Menschen entweder zur Verarbeitung angeregt oder direkt gesteuert wird, kann deshalb bei der Erzeugung und Verarbeitung von Maschinendaten in der Regel nicht von einer persönlichen geistigen Schöpfung ausgegangen werden. Ebenso scheidet eine persönliche geistige Schöpfung hinsichtlich der Auswahl und Anordnung aus, sofern diese auf technischen

[28] vgl. BGH, Urt. v. 09.05.1985 – I ZR 52/83, NJW 1986, 192, 196; OLG Hamburg, Urt. v. 12.3.1998 – 3 U 226/97, MMR 1999, 230; Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 761-769; EuGH, Urt. v. 22. 12. 2010 – C-393/09, GRUR 2011, 220, 220.
[29] vgl. EuGH, Urt. v. 22. 12. 2010 – C-393/09, GRUR 2011, 220, 220.
[30] vgl. Dreier, in: Dreier/Schulze, UrhG, § 69a, Rn. 12; vgl. Grützmacher, in: Wandtke/Bullinger, UrhG, § 69a, Rn. 17.
[31] vgl. EuGH, Urt. v. 1. 3. 2012 – C-604/10, GRUR 2012, 386, Rn. 33, 42.
[32] vgl. Ahlberg, in: BeckOK, UrhG, § 2, Rn. 54.
[33] vgl. Ahlberg, in: BeckOK, UrhG, § 2, Rn. 52.
[34] vgl. Ahlberg, in: BeckOK, UrhG, § 2, Rn. 55.

Erwägungen, Regeln oder Zwängen beruhen.[35] In solchen Konstellationen steht nicht mehr ein Mensch und dessen persönliche geistige Schaffungsleistung im Mittelpunkt, sondern technische Gegebenheiten und Zwecke.

Automatisierte Sammlungen von Maschinendaten erfüllen nicht das Kriterium einer persönlichen geistigen Schöpfung und somit auch nicht die Schutzvoraussetzungen eines Datenbankwerks.

(3) Datenbank

Eine fehlende persönliche geistige Schöpfung ist für die Erfüllung der Voraussetzungen des Datenbankherstellerrechts im Sinne des § 87a UrhG unschädlich. Dieser sui generis Schutz von Datenbanken setzt kein Werk im Sinne des § 2 Abs. 2 UrhG voraus, da gemäß § 87a Abs. 1 UrhG nicht von einem „Sammelwerk“ wie in § 4 Abs. 2 UrhG die Rede ist, sondern nur von einer „Sammlung“.[36]
Das Datenbankherstellerrecht schützt keine persönliche geistige Schöpfung, sondern eine Investitionsleistung in die Herstellung einer Datenbank durch die Gewährung von Ausschließlichkeitsrechten an der jeweiligen Datenbank.[37]

Gemäß § 87a Abs. 1 UrhG ist eine Datenbank eine 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 und deren Beschaffung, Überprüfung oder Darstellung eine nach Art oder Umfang wesentliche Investition erfordert.

Der Datenbankschutz sui generis stellt dabei einige Voraussetzungen, die sich jedoch auf zwei Hauptmerkmale konzentrieren. So spielt die Unabhängigkeit der Elemente und eine wesentliche Investitionsleistung in die Datenbank eine entscheidende Rolle. Die in der Datenbank enthaltenen Elemente müssen unabhängig voneinander und systematisch oder methodisch angeordnet und einzeln zugänglich sein.

Der § 87a UrhG spricht zwar von einer Sammlung von Werken, Daten oder anderen unabhängigen Elementen, jedoch umfasst der Begriff des Elements sowohl Werke als auch Daten.[38] Unter einem Element versteht man alles, was für den Menschen wahrnehmbar ist und einen Informationsgehalt besitzt.[39] Hierunter fallen unter anderem Töne, Texte, Bilder, Zahlen oder Fakten sowie Daten und auch urheberrechtlich geschützte Werke.[40]

Die vom Gesetzesgeber geforderte wesentliche Investitionsleistung muss nach Art oder Umfang für die Beschaffung, Überprüfung oder Darstellung der Datenbank erforderlich sein. Diese beiden Merkmale (Unabhängigkeit der Elemente und wesentliche Investitionsleistung) werden im Nachfolgenden einzeln betrachtet.

(a) Unabhängige Elemente

[35] vgl. EuGH, EuGH, Urt. v. 1. 3. 2012 – C-604/10, GRUR 2012, 386, Rn. 33, 42.
[36] vgl. OLG Köln, Urt. v. 15.12.2006 – 6 U 229/05, MMR 2007, 443 – 446.
[37] vgl. Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 3, 4.
[38] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 9.
[39] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 9.
[40] vgl. Erwägungsgrund 17 der DB-RL.

§ 87a UrhG setzt eine Sammlung von unabhängigen Elementen voraus. Eine Sammlung kann nicht auf eine bestimmte Anzahl an Elementen festgelegt werden. So gilt, dass ein sogenannter Sammlungscharakter durch die Menge der enthaltenen Elemente erkennbar sein muss. Eine geringe Anzahl von Elementen erfüllt daher regelmäßig nicht die Voraussetzung einer Sammlung.[41]
Die entsprechenden Elemente müssen in der Sammlung systematisch oder methodisch angeordnet und einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sein. Dabei bezieht man sich im Hinblick auf die Anordnung auf die Zugriffsebene und nicht auf die tatsächliche Speicherung der Elemente.[42] Entscheidend ist, dass die Person, welche auf die Sammlung zugreift, die Möglichkeit hat, jedes Element in der Sammlung zu finden.[43]

Ausschlaggebend ist also die Verknüpfung des Abfragemittels mit der in der Sammlung enthaltenen Elemente auf der Zugriffsebene.[44] Für diesen Schritt müssen die Elemente strukturiert angeordnet sein. Dabei genügt eine alphabetische, numerische oder chronologische Anordnung.[45]
Als Abfragemittel kommen Systeme in Frage, welche die Lokalisierung jedes Elements in der Sammlung ermöglichen. Ein Index, ein Inhaltsverzeichnis, eine Gliederung oder eine andere Art der Einteilung können daher als Abfragemittel dienen.[46] Es muss gewährleistet sein, dass durch das Abfragemittel die Elemente einzeln zugänglich sind. Dies ist erfüllt, wenn jedes Element aus der Datenbank einzeln abrufbar ist.[47]

Maschinendaten stellen unzweifelhaft Elemente einer Datenbank dar. Inwiefern diese systematisch oder methodisch in Form einer Sammlung angeordnet sind und einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sind, ist stark vom Einzelfall abhängig. In der Regel liegen Maschinendaten mit einer inneren Struktur (Beispiel: Maschine 1, Sensor A: 120°C) vor, sodass sich diese wiederum systematisch oder methodisch anordnen und mithilfe eines Abfragesystems auch einzeln abrufen lassen.
Ungeordneten Datensammlungen, welche gegebenenfalls noch aufbereitet werden müssen, um sie für ein Abfragesystem kompatibel zu machen, fehlt in der Regel die geforderte innere Struktur. Diese unstrukturierten Datensammlungen können ohne Weiteres keine Datenbank im Sinne des § 87a UrhG bilden.[48]

Weitaus kritischer ist die Frage nach der Unabhängigkeit von Maschinendaten. Der Datenbankschutz in § 87a UrhG verlangt eine Sammlung von unabhängigen Elementen. Bei der Beurteilung der Unabhängigkeit ist die folgende Herangehensweise zu beachten.

Ein Element ist unabhängig, wenn es von anderen Elementen getrennt werden kann, ohne dass die Trennung den selbstständigen Informationswert des Elements beeinträchtigt.[49]

---

[41] vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 4; Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 3, 4; Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 11.
[42] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 21.
[43] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 19.
[44] vgl. OLG Köln, Urt. v. 15.12.2006 - 6 U 229/05 – MMR 2007, 443, 444.
[45] vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 7.
[46] vgl. EuGH, Urt. v. 9. 11. 2004 – C-444/02 - GRUR 2005, 254, Rn. 30, 31.
[47] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 14.
[48] vgl. EuGH, Urt. v. 9. 11. 2004 – C-444/02 - GRUR 2005, 254, Rn. 29; vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 24; vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 7.
[49] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 12; vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 6; vgl. Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 21.

Der Informationswert eines einzelnen Elements ist dabei aus der Perspektive eines interessierten Dritten zu beurteilen und nicht aus Sicht eines typischen Nutzers.[50] Für einen interessierten Dritten ergibt sich ein Informationswert aus dem einzelnen Element, wenn dieses eine sachdienliche Information liefert.[51]

Ausgehend von diesem Verständnis bleibt lediglich zu klären, inwiefern ein Element sich von einem anderen Element abgrenzt. Unter den Elementbegriff fällt grundsätzlich alles, was für den Menschen wahrnehmbar ist und einen Informationsgehalt besitzt.[52] Dieses weite Verständnis lässt somit neben Zahlen, Daten oder Fakten auch Töne, Texte und Bilder als Elemente zu.[53] Die Kombination dieser verschiedenen Werte oder Daten kann wiederum ein einzelnes Element bilden. Der EuGH fasste beispielsweise die Daten, Uhrzeiten und Namen von Fußballmannschaften pro Begegnung an den einzelnen Tagen einer Fußballmeisterschaft zu einem unabhängigen Element zusammen.[54] Er hielt in einer späteren Entscheidung ebenso fest, dass bei der Beurteilung der Unabhängigkeit von Elementen auch Datenkombinationen unabhängige Elemente verkörpern können.[55]
Der EuGH macht die Unabhängigkeit eines herausgelösten Elements davon abhängig, ob sich ein Dritter für das Element interessiert.[56] Dies sei dann erfüllt, wenn die Elemente eine sachdienliche Information liefern.[57] Im konkreten Fall bejahte der EuGH das Vorliegen einer sachdienlichen Information, da das Unternehmen die Elemente der Datenbank entnahm, um sie weiter zu verwerten.[58]

Auch die nationale Rechtsprechung hat die Unabhängigkeit von Elementen bei der Kombination verschiedener Einzelinformationen bejaht.[59]

Wendet man diese Erkenntnisse auf Maschinendaten an, lassen sich mehrere Maschinendaten zu einem Element zusammenfassen. Inwiefern sich Datenkombinationen als ein Element charakterisieren, ließ der EuGH unbeantwortet. Ein Ansatz wäre auch hier die Perspektive eines interessierten Dritten einzunehmen und aus dieser zu beurteilen, welche Datenkombinationen eine sachdienliche Information liefern könnte.[60] Solange diese Frage nicht eindeutig durch die Rechtsprechung beantwortet ist, bleibt die Rechtunsicherheit hinsichtlich dieser Beurteilung jedoch bestehen.[61]

---

[50] vgl. EuGH, Urt. v. 9. 11. 2004 – C-444/02 - GRUR 2005, 254, Rn. 34.
[51] vgl. EuGH, Urt. v. 9. 11. 2004 – C-444/02 - GRUR 2005, 254, Rn. 34.
[52] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 9.
[53] vgl. Erwägungsgrund 17 der DB-RL.
[54] vgl. EuGH, Urt. v. 9. 11. 2004 – C-444/02 - GRUR 2005, 254, Rn. 32, 33; EuGH, Urt. v. 1. 3. 2012 – C-604/10, GRUR 2012, 386, Rn. 26.
[55] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187, Rn. 20.
[56] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187, Rn. 28, 29.
[57] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187, Rn. 28, 29.
[58] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187, Rn. 28, 29.
[59] vgl. OLG München, Urt. v. 9.5.2019 – 29 U 1048/18, GRUR-RR 2020, 1; Österreichischer Oberster Gerichtshof, Beschluss vom 10. 7. 2001 – 4 Ob 155/01, GRUR Int. 2002, 452; BGH, Urt. v. 6.5.1999 – I ZR 199/96, MMR 1999, 470.
[60] vgl. Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 339.
[61] vgl. hierzu auch Wiebe, Andreas: "Landkarten als Datenbanken: Der Informationswert von Daten" in: GRUR-Prax, Heft 3, 2016, Seite 50: spricht sich sogar aus, dass der Hersteller selbst definieren sollte, was ein Element in seiner Datenbank darstellt; Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 339.

Von vorangegangenen Erkenntnissen ausgehend, lässt sich für Maschinendaten festhalten, dass diese grundsätzlich unabhängige Elemente bilden können. Ob eine Kombination von Maschinendaten ein Element mit Unabhängigkeitscharakter darstellt, ist davon abhängig, ob diese Kombination für Dritte eine sachdienliche Information liefert.
Welchen Umfang Maschinendaten haben müssen, damit diese eine sachdienliche Information verkörpern, ist stark vom Einzelfall abhängig. Vorstellbar wäre dies für Maschinendaten, die für die Nutzung von Services wie „Predictive Maintenance“ relevant sind oder über Optimierungsmöglichkeiten Auskunft geben.

(b) Investitionsleistung

Wie bereits aufgezeigt, ist für die Erfüllung des Datenbankschutzes nach § 87a UrhG eine nach Art oder Umfang wesentliche Investitionsleistung in die jeweilige Datenbank notwendig. Die wesentliche Investitionsleistung muss laut Gesetzestext für die Beschaffung, Überprüfung oder Darstellung der Datenbank erforderlich sein.

Beim vorangestellten Kriterium einer wesentlichen Investitionsleistung handelt es sich um einen unbestimmten Rechtsbegriff, der von der Rechtsprechung und der herrschenden Meinung als Minimalanforderung interpretiert wird.[62] Sogenannte “Allerweltsinvestitionen” sollen vom Schutz ausgeschlossen sein,[63] sodass Investitionsleistungen erforderlich sind, die keine nach objektiver Betrachtung unbedeutende, von jedermann leicht zu erbringende Aufwendung darstellen.[64] Die Beschaffung öffentlich leicht zugänglicher Daten wäre ein Beispiel für eine nicht wesentliche Investitionsleistung.[65]

Zusätzlich müssen Investitionsleistungen auf deren Berücksichtigungsfähigkeit untersucht werden. Grundsätzlich sind menschliche, finanzielle oder auch technische Ressourcen, die für die Beschaffung, Überprüfung oder Darstellung einer Datenbank erforderlich sind, berücksichtigungsfähig.[66]

Nachfolgend werden die drei Investitionsleistungen – Beschaffung, Überprüfung & Darstellung – nach deren Berücksichtigungsfähigkeit dargestellt.

(i) Beschaffung

Nach einer Grundsatzentscheidung des EuGH müssten Beschaffungsleistungen in die Sammlung und die Erzeugung von Daten unterschieden werden.[67] Die Erzeugung von Daten sei dabei als eine nicht berücksichtigungsfähige Beschaffungsleistung einzustufen, sodass solche Aufwendungen auch nicht in der Wesentlichkeitsbeurteilung berücksichtigt werden

[62] vgl. BGH, Urt. v. 1. 12. 2010 – I ZR 196/08, GRUR 2011, 724, Rn. 23; vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 14, 15; vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 54; Wiebe, in: Spindler/Schuster, UrhG, § 87a, Rn. 16.
[63] vgl. BGH, Urt. v. 1. 12. 2010 – I ZR 196/08, GRUR 2011, 724, Rn. 23; vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 54; Schack, Haimo: „Urheberrechtliche Gestaltung von Webseiten unter Einsatz von Links und Frames“ in: MMR, 2001, Heft 1, Seite 12.
[64] vgl. BGH, Urt. v. 1. 12. 2010 – I ZR 196/08, GRUR 2011, 724, Rn. 23.
[65] vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 15.
[66] vgl. EuGH, Urt. v. 9.11.2004 – C-338/02, GRUR 2005, 244, Rn. 28.
[67] vgl. EuGH, Urt. v. 9.11.2004 – C-338/02, GRUR 2005, 244, Rn. 31.

dürfen.[68] Grundsätzlich zeichne sich das berücksichtigungsfähige Sammeln von Daten dadurch aus, dass bereits vorhandene Elemente, die als solche in der „Natur“ vorkommen, beobachtet und gemessen werden.[69] Meteorologische[70] oder geologische Daten[71] gelten als vorhandene Daten, da diese allgemein verfügbar sind und somit auch von jedermann gesammelt werden können. Dies ist jedoch kritisch zu betrachten, da meteorologische oder geologische Daten auch nicht per se vorhanden und allgemein verfügbar sind. Beispielsweise könnte der Zugang zu einem Grundstück für das Beobachten und Messen benötigt werden und dieser Zugang jedoch durch den Eigentümer verwehrt werden.[72]

Eine Investition in die Erzeugung von Daten lässt sich von einer Sammlungsleistung dahingehend unterscheiden, dass bei einer Erzeugung die Daten erst durch die Messung generiert werden und nur der Datenerzeuger die erzeugten Daten kennt.[73] Das Erzeugen von Daten kann in diesem Kontext auch als „Erfinden“ neuer Daten verstanden werden.[74]
Dabei liegt es nahe, maschinengenerierte Daten unter dem Mantel der nicht berücksichtigungsfähigen Datenerzeugung zu fassen. Diese Annahme scheint jedoch in ihrer Absolutheit abwegig, zumal die Grenze zwischen Sammeln und Erzeugen von Daten in der Praxis nicht eindeutig ist.[75]

In einem Fall scheint die Rechtslage jedoch weitestgehend sicher zu sein: Sogenannte „Sole-Source“-Datenbanken, bei denen ausschließlich eine Person die enthaltenen Daten kennt, da diese die Daten selbst generiert hat, fallen unter den Begriff der nicht berücksichtigten Datenerzeugung.[76] Bei dieser Art von Datenbank besteht eine Monopolgefahr, weil die darin enthaltenen Daten nicht frei verfügbar sind und nicht durch eigene vergleichbare

[68] vgl. EuGH, Urt. v. 9.11.2004 – C-338/02, GRUR 2005, 244, Rn. 31.
[69] vgl. EuGH, Urt. v. 9.11.2004 – C-338/02, GRUR 2005, 244, Rn. 31; vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 56; vgl. Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 341; Wiebe, in: Spindler/Schuster, UrhG, § 87a, Rn. 12; vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 13.
[70] vgl. OLG Köln, Urt. v. 15.12.2006 – 6 U 229/05, MMR 2007, 443.
[71] vgl. LG München, Urt. v. 9. 11. 2005 – 21 O 7402/02, GRUR 2006, 225, 226; vgl. hierzu auch Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 341.
[72] vgl. Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 341.
[73] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 36; vgl. Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 341.
[74] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 271; vgl. Leistner, Matthias: "Datenbankschutz - Abgrenzung zwischen Datensammlung und Datengenerierung" in: Computer und Recht, Heft 1, 2018, Seite 20.
[75] vgl. EU-Kommission: „Evaluation of Directive 96/9/EC on the legal protection of databases“, SWD (2018) 146 final vom 25.4.2018, Seite 36; vgl. Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 341; vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 273.
[76] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 271; vgl. Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 341.

Anstrengungen erstellt werden können.[77] Nach Ansicht des EuGH sind solche Datenbanken nicht schutzfähig.[78]

Um die vorangehenden Erkenntnisse richtig auf Maschinendaten anwenden zu können, ist zu beachten, dass eine Vielzahl an Maschinendaten durch ähnliche Aufwendungen von jedermann erzeugt bzw. generiert werden können.[79] Dies ist entscheidend, da das Datenbankherstellerrecht nicht die Freiheit der Datennutzung einschränkt, sondern einen Investitionsanreiz für die Erstellung von Datenbanken gewährt.[80]
Maschinengenerierte Daten müssen daher immer darauf untersucht werden, inwiefern diese eine „Sole-Source"-Datenbank bilden. Viele maschinengenerierte Daten können auch von Dritten unter ähnlichen Aufwendungen beobachtet und gesammelt werden.

Ebenso sind sogenannte „Spin-Off"-Datenbanken, also Datenbanken, die als reines Nebenprodukt aus „einer anders ausgerichteten geschäftlichen Haupttätigkeit entstehen", schutzwürdig.[81]

Entgegen der Meinung mancher Stimmen in der Literatur[82] und der EU-Kommission,[83] sprach sich der EuGH jedoch für den Schutz von „Spin-Off"-Datenbanken aus: „In diesem Zusammenhang schließt der Umstand, dass die Erstellung einer Datenbank mit Ausübung einer Haupttätigkeit verbunden ist, […] als solcher nicht aus, dass diese Person den Schutz durch das Schutzrecht sui generis beanspruchen kann […]."[84]

Für diese Ansicht spricht zudem, dass der Richtlinientext keine Beschränkung für solche Datenbanken vorsieht.[85] Zumal ein zu enges Verständnis hinsichtlich der Abgrenzung von „Spin-Off"-Datenbanken zu erheblichen Rechtsunsicherheiten in einer angehenden Datenwirtschaft führen würde.[86] Ebenso schließe solch ein Verständnis einen wesentlichen Teil moderner Datensammlungen aus dem Schutzbereich aus.[87]

---

[77] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 271.
[78] vgl. EuGH, Urt. v. 9.11.2004 – C-338/02, GRUR 2005, 252, Rn. 31.
[79] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 271-272; Leistner, Matthias: "Datenbankschutz - Abgrenzung zwischen Datensammlung und Datengenerierung" in: Computer und Recht, Heft 1, 2018, Seite 20; vgl. Schmidt, Kirsten Johanna; Zech Herbert: „Datenbankherstellerschutz für Rohdaten?", in: Computer und Recht, Heft 7, 2017, Seite 422.
[80] vgl. Schmidt, Kirsten Johanna; Zech Herbert: „Datenbankherstellerschutz für Rohdaten?", in: Computer und Recht, Heft 7, 2017, Seite 426.
[81] Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 275.
[82] vgl. Schur, Nico: Die Lizenzierung von Daten, 2020, Seite 51, 58.
[83] vgl. EU-Kommission: „Evaluation of Directive 96/9/EC on the legal protection of databases", SWD (2018) 146 final vom 25.4.2018, Seite 36.
[84] EuGH, Urt. v. 9. 11. 2004 – C-338/02, GRUR 2005, 252, Rn. 29.
[85] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 276.
[86] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 275.
[87] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 275.

Wendet man diese Erkenntnisse auf Maschinendaten an, so sind Investitionen in die Beschaffung dieser Daten berücksichtigungsfähig, sofern sie von jedem Dritten mit ähnlichem Aufwand gesammelt werden können.

Die Sammlung von Produktionsdaten (Stückzahlen, Geschwindigkeit, Verzögerung, usw.) aus einer Wertschöpfungskette oder von Fahrzeugdaten kann beispielsweise als berücksichtigungsfähige Beschaffungsleistung angesehen werden.[88]

(ii) Überprüfung & Darstellung

Neben Beschaffungsleistungen können auch Investitionsleistungen in die Überprüfung und Darstellung der Datenbankelemente zu einem Datenbankschutz führen.

Unter Überprüfung von Datenbankelementen ist die Kontrolle der Zuverlässigkeit, der Richtigkeit oder das Korrigieren von unzutreffenden oder veralteten Daten zu verstehen.[89] Jedoch muss die Überprüfungsleistung von der Beschaffungsleistung trennbar sein, da ansonsten eine Investition in die Überprüfung nicht berücksichtigt werden kann, falls die damit verbundene Beschaffungsleistung als nicht-berücksichtigungsfähiges Erzeugen von Daten einzustufen ist.[90]

Unter Investitionen, die sich auf die Darstellung von Datenbanken beziehen, versteht man Mittel, welche einer Datenbank ihre Funktion verleihen.[91] Konkret meint dies Investitionen in Mittel für die systematische oder methodische Anordnung und Zugänglichkeit der enthaltenen Elemente.[92] Somit sind Kosten für ein Abfragesystem, eine Indexierung[93] oder auch für Nutzungsrechte an einem Computerprogramm berücksichtigungsfähig.[94] Ebenso inbegriffen sind Aufwendungen, welche in die Bereitstellung der technischen Infrastruktur der Datenbank getätigt werden, insbesondere deren Erhaltung, Pflege und Wartung.[95]

Wer Datenbankhersteller und somit Inhaber der Ausschließlichkeitsrechte an einer Datenbank ist, hängt davon ab, wer die Investition in die Datenbankherstellung tätig (vgl. § 87a Abs. 2 UrhG). Dem Datenbankhersteller steht gemäß § 87b UrhG das ausschließliche Recht zu, die Datenbank insgesamt oder einen nach Art und Umfang wesentlichen Teil der Datenbank zu vervielfältigen, zu verbreiten und öffentlichen wiederzugeben.

---

[88] vgl. Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 341-342; Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 276.

[89] vgl. Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 47, 48; Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 37.

[90] vgl. EuGH, Urt. v. 9. 11. 2004 – C-203/02, GRUR 2005, 244, 247; Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 37; vgl. Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 44; Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 277.

[91] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 38.

[92] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 38.

[93] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 38.

[94] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 39.

[95] vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 13.

Es ist zu beachten, dass insbesondere die Vervielfältigung nicht wesentlicher Teile einer geschützten Datenbank nicht zu einer Verletzung des Datenbankherstellerrechts führt, sofern die Schrankenbestimmung in § 87b Abs.1 S.2 UrhG greift.[96]

b) Geschäftsgeheimnisgesetz

Damit Maschinendaten unter den Schutzbereich des GeschGehG fallen, müssen diese die Voraussetzungen des § 2 Nr. 1 GeschGehG erfüllen. Danach ist ein Geschäftsgeheimnis eine Information, die geheim ist, einen wirtschaftlichen Wert besitzt und durch angemessene Geheimhaltungsmaßnahmen geschützt ist. Darüber hinaus verlangt der deutsche Gesetzgeber ein berechtigtes Interesse an der Geheimhaltung.

(1) Geheimsein

Damit Maschinendaten als Geschäftsgeheimnis im Sinne des GeschGehG gelten, müssen diese geheim, d.h. nicht allgemein bekannt sein. Ob eine Information geheim ist, wird anhand von drei Kriterien geprüft: den Bezugspunkt der Geheimhaltung, den Grad der Geheimhaltung und die Bestimmung des maßgeblichen Verkehrskreises.[97]

Eine Information ist nach § 2 Nr. 1 lit. a GeschGehG geheim, wenn sie weder insgesamt noch in der genauen Anordnung und Zusammensetzung ihrer Bestandteile bekannt oder ohne Weiteres zugänglich ist. Folglich sind Informationen, deren Bestanteile alle geheim sind, unkritisch.[98] Regelmäßig werden jedoch nicht alle Bestandteile geheim sein. Von Bedeutung dürfte hingegen meist sein, ob die genaue Anordnung und Zusammensetzung der Bestandteile allgemein bekannt ist. Danach können geheime Bestandteile einer Information mit allgemein bekannten Bestandteilen einer anderen Information verknüpfen werden, die im Ergebnis als solche unbekannt sind.[99] Auch Zusammensetzungen von rein bekannten Bestandteilen in einer besonderen Weise, können eine Information darstellen, die nicht allgemein bekannt ist.[100]

Für die Geheimniseigenschaft nach § 2 Nr. 1 lit. a GeschGehG dürfen Informationen nicht „ohne Weiteres zugänglich“ sein. Da für die Generierung eines Geschäftsgeheimnisses überwiegend Zeit und Kosten aufgewendet wurden[101], sind diese Größen für die Beantwortung dieses Kriteriums besonders relevant: je mehr Zeit und Kosten ein Dritter investieren muss, um die gleiche Information (auch in ihrer Zusammensetzung, s.o.) zu erhalten, desto eher ist davon auszugehen, dass diese nicht ohne Weiteres zugänglich ist.[102] Umgekehrt gilt, dass eine Information nicht geheim ist, wenn sie durch keinen größeren Zeit- und Kostenaufwand zugänglich ist.[103]

---

[96] vgl. Gaster, Jens: „Sui-generis Recht der Datenbankrichtlinie“ in: Hoeren, Thomas; Sieber, Ulrich; Holznagel, Bernd (Hrsg.): Handbuch Multimedia-Recht, 2021, Teil 7.6, Rn. 119a.
[97] Vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 7.
[98] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 8.
[99] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 8.
[100] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 8.
[101] Hauck, in: Heermann/Schlingloff, MüKoUWG, GeschGehG § 2 Rn. 9.
[102] vgl. Hauck, in: Heermann/Schlingloff, MüKoUWG, GeschGehG § 2 Rn. 9.
[103] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 9; vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 276.

Bei der Beurteilung der Bekanntheit einer Information ist gemäß § 2 Nr. 1 lit. a GeschGehG der maßgebliche Personenkreis heranzuziehen, welcher üblicherweise mit dieser Art von Information umgeht. Somit steht ein Fachpublikum und nicht die Gesamtbevölkerung bei der Beurteilung der Bekanntheit im Zentrum der Betrachtung.[104] Inwiefern sich dieses Fachpublikum abgrenzt, ist stark vom Einzelfall abhängig und dahingehend durch die betreffende Information bestimmt.[105] Dass eine Information einem größeren Personenkreis bekannt ist, schadet der Geheimniseigenschaft dabei nicht per se, solange dieser begrenzt ist.[106]

Für Maschinendaten bedeutet dies, dass bei der Beurteilung des Geheimnischarakters jeweils auf den Einzelfall abzustellen ist. Maschinendaten sind in der Regel nicht in der Anordnung und Zusammensetzung dem abzustellenden Personenkreis allgemein bekannt oder ohne weiteres zugänglich.[107] Vor allem sind regelmäßig Daten, die sich aus Maschinendaten ableiten, einem entsprechendem Fachpublikum nicht allgemein bekannt oder ohne Weiteres zugänglich.[108]

(2) Wirtschaftlicher Wert

Darüber hinaus verlangt der Gesetzgeber, dass ein Geschäftsgeheimnis einen wirtschaftlichen Wert besitzt. Dieses Merkmal kann als *de-minimis*-Schwelle verstanden werden. Es genügt bereits zur Bejahung eines wirtschaftlichen Wertes das Vorliegen eines kommerziellen Potentials.[109] Unschädlich ist dabei, ob es tatsächlich bereits potenzielle Abnehmer gibt.[110] Sofern die Beurteilung eines wirtschaftlichen Wertes Schwierigkeiten verursacht, kann mithilfe folgender Kontrollfrage ein wirtschaftlicher Wert festgestellt werden: Könnte die unbefugte Nutzung oder Offenlegung der Information die Interessen des Inhabers dadurch schädigen, „dass das wissenschaftliche oder technische Potenzial, die geschäftlichen oder finanziellen Interessen, die strategische Position oder die Wettbewerbsfähigkeit dieser Person untergraben“ werden?[111]

Diese Frage verdeutlich noch einmal, dass die Voraussetzung eines wirtschaftlichen Wertes dementsprechend niedrig anzusetzen ist. Davon abzugrenzen ist dennoch belanglose Information (bspw. kontextlose Datenpunkte außerhalb einer Datensammlung[112]), die als solche keinen Wert darstellt und somit nicht unter den Geschäftsgeheimnisschutz fällt.[113]

---

[104] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 9.
[105] vgl. Dörner, in: Schuster/Grützmacher, GeschGehG, § 2, Rn. 27.
[106] vgl. Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG, § 2, Rn. 34.
[107] vgl. Schur, Nico: Die Lizenzierung von Daten, 2020, Seite 67.
[108] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 276; vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 443.
[109] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 443; Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 16.
[110] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 16.
[111] vgl. Erwägungsgrund 14 der TS-RL; vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 443.
[112] Krüger/Wiencke/Koch: "Der Datenpool als Geschäftsgeheimnis" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 6, 2020, Seite 580.
[113] vgl. Erwägungsgrund 14 der TS-RL; Schur, Nico: Die Lizenzierung von Daten, 2020, Seite 68.

Ein wirtschaftlicher Wert oder ein wirtschaftliches Potential von Maschinendaten ist in Anbetracht dessen derzeitiger Rolle nur schwer abzustreiten. Zumal durch die Verknüpfung bestehender Daten mit weiteren Daten eine erneute Auswertung möglich ist und sich dadurch neue Daten und Erkenntnisse ableiten lassen.[114]

Maschinendaten besitzen somit in der Regel einen wirtschaftlichen Wert.[115]

(3) Berechtigtes Interesse

Das sogenannte berechtigte Interesse ist ein Tatbestandsmerkmal, welches in der zugrundeliegenden EU-Richtlinie über den Schutz vertraulichen Know-hows und vertraulicher Geschäftsinformationen (TS-RL)[116] nicht enthalten ist. Diese nationale Abweichung ist jedoch irrelevant, da ebenfalls der Erwägungsgrund 14 TS-RL von einem legitimen Interesse an der Geheimhaltung eines entsprechenden Geschäftsgeheimnisses ausgeht.[117]
Das berechtigte Interesse ermöglicht es den Gerichten, eine eigenständige Einstufung und eine Willkürkontrolle durchzuführen.[118] Ein Geschäftsgeheimnisinhaber, der böswillig gegen eine unbefugte Nutzung vorgeht, die im Zweifelsfall auch strafrechtliche Sanktionen nach sich ziehen, können durch das Tatbestandmerkmal unterbunden werden.[119] So können sich beispielsweise illegale Handlungen nicht auf den Geschäftsgeheimnisschutz berufen, da hierfür ein berechtigtes Interesse an der Geheimhaltung fehlt.[120]

Somit wird nur in Ausnahmefällen das Tatbestandmerkmal nicht erfüllt.

(4) Angemessene Geheimhaltungsmaßnahmen

Gemäß dem Wortlaut des § 2 Nr. 1 lit. B GeschGehG muss der rechtmäßige Inhaber eines Geschäftsgeheimnisses den Umständen nach angemessene Geheimhaltungsmaßnahmen treffen. Es handelt sich somit um eine Obliegenheit seitens des Inhabers. Diese Obliegenheit sollte nicht unterschätzt werden, da die mangelnde und fehlerhafte Durchführung zum Verlust des Geschäftsgeheimnisschutzes führt.[121]

---

[114] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 443; vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 261.

[115] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 443; vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 261.

[116] Richtlinie 2016/943/EU des Europäischen Parlaments und des Rates vom 08.06.2016 über den Schutz vertraulichen Know-hows und vertraulicher Geschäftsinformationen (Geschäftsgeheimnisse) vor rechtswidrigem Erwerb sowie rechtswidriger Nutzung und Offenlegung, Abl.EU Nr. L 157 v. 15.6.2016, Seite 1 ff.

[117] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 70.

[118] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 71.

[119] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 71.

[120] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 444.

[121] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 443; vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 20.

Zu beachten ist dabei, dass durch das Tatbestandsmerkmal der Angemessenheit immer auf die konkrete Information und den jeweiligen Inhaber abzustellen ist.[122] So sind beispielsweise der Wert des Geheimnisses, die Größe des Unternehmens, die Kosten und Üblichkeit vergleichbarer Inhaber Orientierungspunkte für die Beurteilung.[123]

Geheimhaltungsmaßnahmen können dabei technische, organisatorische und auch rechtliche Maßnahmen sein. Verschlüsselungen durch Passwörter, Zugriffsberechtigungen oder auch vertragliche Geheimhaltungsmaßnahmen werden bei der Beurteilung der Angemessenheit berücksichtigt. Inwiefern für Maschinendaten bereits Geheimhaltungsmaßnahmen aufgrund von anderweitigen IT-Sicherheitsmechanismen vorliegen, ist im Einzelfall zu prüfen.

Maschinendaten können somit Geschäftsgeheimnisse im Sinne des GeschGehG darstellen.[124] Bestätigt wird dieses Ergebnis durch den Erwägungsgrund 14 der TS-RL.[125] Dieser beschreibt, dass unter den Schutzumfang der TS-RL neben Know-How und Geschäftsinformationen auch technologische Information fallen sollen.[126]

Inhaber eines Geschäftsgeheimnisses können sowohl natürliche als auch juristische Personen sein, welche die rechtmäßige Kontrolle über das jeweilige Geschäftsgeheimnis besitzen. Rechtmäßige Kontrolle liegt in der Regel bei den Personen vor, welche das Geschäftsgeheimnis selbst erstellten.[127] Ebenso kann eine rechtmäßige Kenntniserlangung über vertragliche Regelungen erfolgen.[128]

Zu beachten ist, dass das GeschGehG keine Ausschließlichkeitsrechte an geschützten Geschäftsgeheimnissen einräumt.[129] Das Gesetz schützt den Inhaber lediglich vor dem rechtswidrigen Zugang zu seinem Geschäftsgeheimnis. Dabei wird auch dieser Zugang durch explizit erlaubte Handlungen (§ 3 GeschGehG) explizit eingeschränkt, sodass beispielsweise Geschäftsgeheimnisse durch das sogenannte „Reverse Engineering“ grundsätzlich erlangt werden dürfen. Das GeschGehG untersagt die Überwindung des faktischen Zugangs durch die Gewährung von Abwehrrechten.[130] Geschäftsgeheimnisschutz ist also in erster Linie Zugangsschutz.[131]

[122] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 444; vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 20, 26.
[123] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 444.
[124] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 267-268; Schur, Nico: Die Lizenzierung von Daten, 2020, Seite 86-87.
[125] Richtlinie 2016/943/EU des Europäischen Parlaments und des Rates vom 08.06.2016 über den Schutz vertraulichen Know-hows und vertraulicher Geschäftsinformationen (Geschäftsgeheimnisse) vor rechtswidrigem Erwerb sowie rechtswidriger Nutzung und Offenlegung, Abl.EU Nr. L 157 v. 15.6.2016, Seite 1 ff.
[126] vgl. Erwägungsgrund 14 der TS-RL.
[127] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 81.
[128] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 82.
[129] vgl. Schur, Nico: Die Lizenzierung von Daten, 2020, Seite 86.
[130] vgl. Schur, Nico: Die Lizenzierung von Daten, 2020, Seite 72; 86.
[131] vgl. Schur, Nico: Die Lizenzierung von Daten, 2020, Seite 86.

Zusammenfassung – Immaterialgüterrecht:
Fasst man die Erkenntnisse aus der immaterialgüterrechtlichen Analyse zusammen, lassen sich maschinengenierte Daten regelmäßig nicht als geschütztes Computerprogramm (§ 69a UrhG) und nicht als Datenbankwerk (§ 4 Abs. 2 UrhG) schützen.
Der Schutz einer Datensammlung als Datenbank (§ 87a UrhG) und als Geschäftsgeheimnis gemäß § 2 Nr. 1 GeschGehG kommt jedoch insbesondere für strukturierte maschinengenerierte Datensammlungen in Betracht.

3. Strafrecht

Die für den Schutz von Daten relevanten Straftatbestände in §§ 202a, 202b, 202c, 202d und 303a StGB orientieren sich bei der Beurteilung von Daten auf die gesetzliche Legaldefinition in § 202a Abs. 2 StGB. Danach fallen unter den Datenbegriff nur Daten, „die elektronisch, magnetisch oder sonst nicht unmittelbar wahrnehmbar gespeichert sind oder übermittelt werden“. Der strafrechtliche Schutz von Daten besteht unabhängig von ihrem Informationsgehalt und unabhängig davon, ob die enthaltenen Informationen einen bestimmten Verarbeitungszweck haben.[132]
Diese weite Legaldefinition bedeutet wiederum, dass auch einfachste Maschinendaten ohne nennenswerten Inhalt unter die Legaldefinition fallen.[133] Die Straftatbestände in §§ 202a, 202b, 202c, 202d und 303a StGB sind somit auf eine Vielzahl von Daten anwendbar.

Im strafrechtlichen Sinne erfolgt eine Zuordnung von Daten über den sogenannten Skripturakt. Diese Ansicht wird nach aktueller Kenntnis sowohl von der Literatur[134], als auch von der Rechtsprechung[135] anerkannt. Der Skribent ist derjenige, der die gespeicherten Daten durch die Eingabe oder Ausführung eines Programms selbst erstellt hat.[136] Danach ist er der Berechtigte hinsichtlich der Daten.[137] Als Berechtigter kann auch eine Person in Frage kommen, wenn diese Zugang durch den Skribenten erhält oder der Zugang zu Daten durch den Skribenten geduldet wird.[138]

Eine Zuordnung zivilrechtlicher Rechte an Maschinendaten lassen die genannten Vorschriften dennoch nicht zu. Vor allem bei Mehrpersonenkonstellationen ist eine Zuordnung auf den Skribenten schwierig zu beurteilen. Als Skribent von Daten könnte beispielsweise der

[132] vgl. Graf, in: MüKo, StGB, §202a, Rn. 12.
[133] vgl. Schur, Nico: Die Lizenzierung von Daten, 2020, Seite 87.
[134] vgl. Zech, Herbert: Daten als Wirtschaftsgut – Überlegungen zu einem „Recht des Datenerzeugers“ in: Computer und Recht, Heft 3, 2015, Seite 143; Zech, Herbert: Information als Schutzgut, 2012, Seite 399; Hecker, in: Schönke/Schröder, StGB, § 303a, Rn. 3.
[135] vgl. OLG Naumburg, Urt. v. 27.8.2014 – 6 U 3/14, DAR 2015, 27, Rn. 21.
[136] vgl. Hoeren, Thomas: "Dateneigentum - Versuch einer Anwendung von § 303a StGB im Zivilrecht" in: MMR, Heft 8, 2013, Seite 487; vgl. auch Zech, Herbert: Daten als Wirtschaftsgut – Überlegungen zu einem „Recht des Datenerzeugers“ in: Computer und Recht, Heft 3, 2015, Seite 143.
[137] vgl. Hoeren, Thomas: "Dateneigentum - Versuch einer Anwendung von § 303a StGB im Zivilrecht" in: MMR, Heft 8, 2013, Seite 487; vgl. auch Zech, Herbert: Daten als Wirtschaftsgut – Überlegungen zu einem „Recht des Datenerzeugers“ in: Computer und Recht, Heft 3, 2015, Seite 143.
[138] vgl. Weidemann, in: BeckOK, StGB, §202a, Rn. 20.

„Verursacher der technischen Herstellung, der wirtschaftlich Verantwortliche oder derjenige, der den Herstellungsvorgang planerisch-technisch beherrscht“, in Frage kommen.[139]

Darüber hinaus räumen die Straftatbestände in §§ 202a, 202b, 202c, 202d und 303a StGB keiner Person Ausschließlichkeitsrechte an Daten ein. Vielmehr werden die Daten durch strafbewerte Sanktionen durch den Gesetzgeber geschützt.

Zusammenfassung – Strafrecht:
Das StGB verbietet die §§ 202a, 202b, 202c, 202d und 303a StGB festgelegten Handlungen durch strafbewerte Maßnahmen, ohne dabei eine Zuordnung von Rechten zu verfolgen. Daten werden durch die Strafvorschriften in ihrer Integrität und gegen den unbefugten Zugang zu diesen geschützt. Die bloße Nutzung ist jedoch nicht strafbar.

[139] vgl. Kraus, Michael: „Datenlizenzverträge“ in: Jürgen Taeger (Hrsg.): Internet der Dinge, 2015, Seite 544.

4. Deliktsrecht

Ob das Deliktsrecht eine Zuordnung von Rechten an Daten ermöglicht, wird immer wieder in der juristischen Literatur diskutiert.[140] Meist konzentriert sich die Diskussion darauf, ob sich ein sonstiges Recht im Sinne des § 823 Abs. 1 BGB auf Daten ableiten lässt.[141] Es bleibt dennoch offen, ob Daten als sonstiges Recht überhaupt in Frage kommen können.
Als sonstiges Recht kommen „zum einen weitere einzelne Persönlichkeitsrechte, zum anderen eigentumsähnliche Rechte, die sowohl eine positive Nutzungsfunktion haben als auch absolute Abwehrbefugnisse gewähren“, in Frage.[142]

Dabei zu beachten ist, dass der BGH entschied, dass ein Datenbestand als ein „selbstständiges vermögenswertes Gut“ im Sinne eines sonstigen Rechts eingeordnet werden kann.[143] Voraussetzung hierfür ist, dass Daten auf einem Datenträger verkörpert sind und die Daten verändert oder gelöscht wurden, ohne dabei die Sachsubstanz des Datenträgers zu beschädigen.[144]

Eine Zuordnung der Daten und etwaige Rechte erfolgt somit nicht über die Daten selbst, sondern über den betreffenden physischen Datenträger. Bei modernen Anwendungen fallen jedoch Daten und physischer Datenträger auseinander, sodass keine Zuordnung über einen Datenträger möglich ist.[145]

Zudem muss die grundlegende Unterscheidung im BGB zwischen sogenannten Stammrechten und sogenannten Rechtsdurchsetzungsrechten beachtet werden. Stammrechte, wie das normierte Eigentumsrecht in § 903 BGB, ermöglicht eine Zuordnung von Gütern.[146] Die im Deliktsrecht enthaltene Rechtsdurchsetzungsrechte, wie der Schadensersatzanspruch in § 823 Abs. 1 BGB, ermöglichen keine Zuordnung von Rechten.[147]

Das Deliktsrecht ermöglicht somit keine Zuordnung von Rechten an Daten ohne entsprechenden Datenträger.[148]

---

[140] vgl. Arkenau, Judith; Wübbelmann, Judith: “Eigentum und Rechte an Daten – Wem gehören die Daten?“ in: Taeger, Jürgen (Hrsg.): Internet der Dinge, 2015, Seite 99; Zech, Herbert: Information als Schutzgut, 2012, Seite 386f.
[141] vgl. Arkenau, Judith; Wübbelmann, Judith: “Eigentum und Rechte an Daten – Wem gehören die Daten?“ in: Taeger, Jürgen (Hrsg.): Internet der Dinge, 2015, Seite 99; Zech, Herbert: Information als Schutzgut, 2012, Seite 386f.
[142] Teichmann, in: Jauering, BGB, §823, Rn.18.
[143] BGH, Urt. v. 02.07.1996 – X ZR 64/94, NJW 1996, 2924; vgl hierzu auch: OLG Karlsruhe, Urt. v. 07.11.1995 – 3 U 15/95, NJW 1996, 200; OLG Oldenburg, Beschluss vom 24.11.2011 – 2 U 98/11, ZD 2012, 177.
[144] vgl. OLG Karlsruhe, Urt. v. 07.11.1995 – 3 U 15/95, NJW 1996, 200; OLG Oldenburg, Beschluss v. 24.11.2011 – 2 U 98/11, ZD 2012, 177.
[145] vgl. Arkenau, Judith; Wübbelmann, Judith: “Eigentum und Rechte an Daten – Wem gehören die Daten?“ in: Taeger, Jürgen (Hrsg.): Internet der Dinge, 2015, Seite 100; vgl. Zech, Herbert: Information als Schutzgut, 2012, Seite 386f.
[146] vgl. Hofmann, Franz: "Absolute Rechte an Daten - immaterialgüterrechtliche Perspektive" in: Pertot, Tereza (Hrsg.): Rechte an Daten, 2020, Seite 12.
[147] vgl. Hofmann, Franz: "Absolute Rechte an Daten - immaterialgüterrechtliche Perspektive" in: Pertot, Tereza (Hrsg.): Rechte an Daten, 2020, Seite 12.
[148] vgl. Zech, Herbert: Daten als Wirtschaftsgut – Überlegungen zu einem „Recht des Datenerzeugers“ in: Computer und Recht, Heft 3, 2015, Seite 143; Arkenau, Judith; Wübbelmann, Judith: “Eigentum und Rechte an Daten – Wem gehören die Daten?“ in: Taeger, Jürgen (Hrsg.): Internet der Dinge, 2015, Seite 100.

Zusammenfassung – Deliktsrecht:
Das Deliktsrecht, welches grundsätzlich mit Rechtsdurchsetzungsrechten ausgestaltet ist, kann, anders als rechtszuordnende Stammrechte (wie § 903 BGB), keine Rechten an Daten zuordnen.
Ein deliktischer Schutz erfolgt somit nicht über den Bezugspunkt der Daten als solches, sondern vielmehr über den des Datenträgers.

B. Ergebnis

Die vorangehende Untersuchung möglicher Rechte an Maschinendaten zeigt, dass Rechte an Daten begründet und einer Person zugeordnet werden können. Dabei ist jedoch zu beachten, dass diese Rechte in ihrer Zielrichtung und Wirkung zu unterscheiden sind. Die Rechte an Daten lassen sich in Ausschließlichkeitsrechte, Abwehrrechte und rein obligatorische Rechte unterscheiden.

Die sachenrechtliche Untersuchung zeigt, dass Daten keine körperlichen Sachen, sondern immaterielle Güter beziehungsweise nicht-körperliche Gegenstände sind. Diese Schlussfolgerung führt dazu, dass weder Eigentum noch Besitz an Daten als solche begründet werden können. Aufgrund der Einstufung als immaterielles Gut bzw. Gegenstand können jedoch obligatorische Rechte in Form von vertraglichen Regelungen bestehen.

Die Analyse aus der immaterialgüterrechtlichen Perspektive zeigt, dass die Schutzpositionen im Urhebergesetz unterschiedlich ausfallen. Maschinendaten stellen kein Computerprogramm im Sinne des § 69a UrhG dar. Ebenso scheidet ein Schutz als Datenbankwerk im Sinne des § 4 Abs. 2 UrhG regelmäßig aus. Einzig der Schutz als Datenbank sui generis durch das Leistungsschutzrecht in § 87a UrhG kann erfüllt werden. Hierbei sind jedoch zwei wichtige Erkenntnisse zu beachten. Zum einen wird nicht ein einzelnes Datum geschützt, sondern nur die Datenbank als solche. Zum anderen stellt der Datenbankschutz einige Hürden auf, die zwingend zu berücksichtigen sind. Problematisch ist regelmäßig die geforderte Unabhängigkeit der Elemente einer Datenbank und die teilweise schwierige Unterscheidung zwischen nicht berücksichtigungsfähigen Investitionen in das Generieren und berücksichtigungsfähigen Investitionen in das Sammeln von Daten.

In Summe kann der Datenbankschutz sui generis in § 87a UrhG Maschinendaten schützen, birgt jedoch ein gewisses Rechtsrisiko, aufgrund der teils schwierigen Handhabung.
Der Datenbankhersteller hat nach § 87b UrhG das ausschließliche Recht, die Datenbank insgesamt oder einen nach Art und Umfang wesentlichen Teil der Datenbank zu vervielfältigen, zu verbreiten und öffentlich wiederzugeben. Unwesentliche Teile einer Datenbank können durch Dritte genutzt werden, ohne das Recht des Datenbankherstellers zu verletzen, sofern deren Nutzung – insbesondere die kumulative Vervielfältigung – nicht darauf gerichtet ist, die Datenbank in ihrer Gesamtheit oder zu einem wesentlichen Teil wieder zu erstellen[149] und solange die genutzten Bestandteile nicht selbst urheberrechtlich geschützt sind.

Der Schutz des GeschGehG ist für Maschinendaten grundsätzlich eröffnet. Es ist zu beachten, dass angemessene Geheimhaltungsmaßnahmen seitens des Geschäftsgeheimnisinhabers

[149] BGH, Urteil v. 22. 6. 2011 - I ZR 159/10.

getroffen werden, da ohne Einhaltung angemessener Maßnahmen zur Geheimhaltung von Maschinendaten der Schutz des GeschGehG erlischt. Der Geschäftsgeheimnisschutz räumt dem Inhaber keine Ausschließlichkeitsrechte, sondern lediglich sogenannte Abwehrrechte ein.

Das Straf- und Deliktsrecht räumt keiner Person Rechte im zivilrechtlichen Sinne ein. Es besteht dennoch ein strafrechtlicher Schutz, sofern die einschlägigen Straftatbestände erfüllt werden. Ebenso können deliktsrechtliche Ansprüche bestehen, sofern eine Rechtezuordnung über einen Datenträger erfolgen kann.

Im Ergebnis sind die entsprechenden Rechte und Schutzpositionen an Maschinendaten sehr heterogen. Daher ist die Durchführung einer Einzelfallprüfung hinsichtlich der möglichen Rechte und Schutzpositionen an Maschinendaten nötig. Aufgrund dieser teils schwierigen Einstufung bieten sich Lösungen auf der Grundlage vertraglicher Regelungen besonders an.

Die sachenrechtliche Untersuchung kommt zum Ergebnis, dass Daten aufgrund fehlender Verkörperung nicht als Sachen gelten, sodass gesetzlich nicht ein Eigentümer oder Besitzer an Daten definiert werden kann.

Umso zwingender ist es, vertraglich festzuhalten, inwiefern Rechte an Daten bestehen und welcher (Rechts-)Person diese zugeordnet werden sollen.

V. Die rechtliche Schutzfähigkeit von dezentral trainierten neuronalen Netzen

Im Folgenden wird die Schutzfähigkeit der verschiedenen Bestandteile neuronaler Netze entlang der Prozessschritte des Federated Learning anhand der dargelegten Rechtslage bewertet. Im Stadium eines untrainierten neuronalen Netzes wird die Schutzfähigkeit der Netzarchitektur, der Lern- und Optimierungsfunktion und der Hyperparameter geprüft. Danach richtet sich der Fokus auf die trainierten und fusionierte Gewichtungsdaten von trainierten bzw. fusionierten neuronalen Netzen, um schließlich die Schutzfähigkeit von Sammlungen trainierter Gewichtungsdaten und fusionierter Gewichtungsdaten darzustellen.

Bei allen Prüfungsschritten dient das Anwendungsbeispiel aus Kapitel III. B. als untersuchungsrelevanter Sachverhalt.

A. Schutz von untrainierten neuronalen Netzen

Untrainierte neuronale Netze bestehen im Wesentlichen aus der Netzarchitektur, der Lern- und Optimierungsfunktion und den Hyperparametern. In genannter Reihenfolge sollen diese Bestandteile auf ihre immaterialgüterrechtliche Schutzfähigkeit untersucht werden.

1. Netzarchitektur

Die Netzarchitektur ist das Fundament eines neuronalen Netzes. Sie beschreibt die Anzahl von Schichten und die Anzahl der Neuronen pro Schicht. Bestehend aus einer Eingabe- und Ausgabeschicht und mehreren verborgenen Schichten, bildet die Netzarchitektur den Prozess der Verarbeitung ab. Die Eingabeschicht wird maßgeblich durch die zu verarbeitenden Daten vorgegeben. Ebenso wird die Ausgabeschicht durch die Ergebnisse oder Kategorien, die das neuronale Netz am Ende der Verarbeitung erzielen soll, bestimmt. Die Ausgestaltung und

Komplexität der verborgenen Schichten sind wiederum sehr stark von der konkreten Verarbeitung und dem jeweiligen Ersteller abhängig. Je komplexer eine Verarbeitung ist, desto mehr verborgene Schichten müssen bei der Erstellung der Netzarchitektur erstellt werden. Darüber hinaus müssen Transferfunktionen zwischen den Neuronenschichten und die Aktivierungsfunktionen pro Schicht festgelegt werden.

Die zu verarbeitenden Daten und die Funktionalität eines neuronalen Netzes sind folglich maßgebliche Einflussfaktoren bei dessen Erstellung und bieten dabei zugleich einen großen Gestaltungsspielraum für den jeweiligen Ersteller.

Diesem Verständnis folgend wird nachstehend der rechtliche Schutz der Netzarchitektur dargestellt.

a) Urhebergesetz

Für einen urheberrechtlichen Schutz für die Netzarchitektur kommt ein Schutz als Computerprogramm (§ 69a UrhG), als Datenbankwerk (§ 4 Abs. 2 UrhG) und als Datenbank (§ 87a UrhG) in Frage.

(1) Computerprogramm – § 69a UrhG

Für die Beurteilung, ob eine in Programmcode gefasste Netzarchitektur (beispielsweise in der Programmiersprache Python, Java oder sonstigen üblichen Sprachen für ML-Modelle) bereits als geschütztes Computerprogramm im Sinne des § 69a UrhG gilt, ist zunächst zu klären, ob eine Netzarchitektur der von der Rechtsprechung angewandten Definition[150] eines Computerprogramms entspricht. Dafür müsste diese eine Folge von Befehlen verkörpern, die eine bestimmte Funktion oder Aufgabe oder ein bestimmtes Ergebnis anzeigt, ausführt oder erzielt.

Das Design eines neuronalen Netzes entsteht im Wesentlichen in zwei Stadien. Im ersten Stadium erfolgt das Lernen, also die Anpassung der Gewichte zwischen den einzelnen Neuronen innerhalb der Netzarchitektur.[151] Das zweite Stadium ist die Anwendung[152] des neuronalen Netzes mit trainierten Gewichtungsdaten, um teilautomatisierte oder auch vollautomatisierte Prozesse auszuführen.[153]

Die untrainierte Netzarchitektur besteht bereits aus einer Folge von Befehlen, die eine Lernaufgabe ausführen kann und weist damit laut einer vertretenen Auffassung den erforderlichen Steuerungscharakter auf.[154] Dieser Ansicht folgend wäre eine zentrale Voraussetzung der Definition eines Computerprogramms erfüllt.

---

[150] s. hierzu bereits in Kapitel VI.A Ziff.2.
[151] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 765.
[152] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 765.
[153] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 765.
[154] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 765.

Entgegen dieser Auffassung wird teilweise die Meinung vertreten, dass neuronale Netze nicht schon im Stadium des Lernens eine Steuerungsfunktion aufweisen.[155] Danach heißt es, dass neuronale Netze „erst durch das Training und Lernen des Netzwerkes und nicht bereits durch den initialen Schaffungsprozess zu einer Funktionalität reifen".[156] Von einem geschützten Computerprogramm in diesem Stadium auszugehen, wird daher abgelehnt.[157]

Ein untrainiertes neuronales Netz und somit die bloße Netzarchitektur ist zwar nicht in der Lage, verlässliche Ergebnisse zu erzielen, da ein zielführendes Training des Netzes noch bevorsteht. Jedoch sind qualitative Kriterien gemäß § 69a Abs. 3 UrhG bei der Beurteilung der Schutzfähigkeit nicht von Belang und folglich außer Acht zu lassen.[158] Entscheidend ist hingegen die Erfüllung einer bestimmten Funktion, einer Aufgabe oder das Erzielen eines bestimmtes Ergebnisses durch die steuernde Funktion des Computerprogramms.[159]

Dies ist bei einem untrainierten neuronalen Netz zu bejahen, da bereits hier die programmierte Netzarchitektur grundsätzlich in der Lage ist, die Gewichte zu trainieren. Ebenso kann das neuronale Netz ein Ergebnis aufgrund der festgelegten Ausgabeschicht anzeigen, sofern die Gewichte randomisiert wurden. Diese Verarbeitung wird mit hoher Wahrscheinlichkeit kein richtiges Ergebnis anzeigen, aufgrund des fehlenden Trainingsprozesses, jedoch ist die Güte des Ergebnisses ein qualitatives Kriterium und damit irrelevant für die Beurteilung der Schutzfähigkeit. Die hier vertretene Meinung schließt sich deshalb der Ansicht an, dass ein neuronales Netz bereits im initialen Schaffungsprozess eine Funktionalität aufweist.[160]

Auch wenn man zum entgegengesetzten Ergebnis hinsichtlich der Bestimmtheit des Computerprogramms gelangen würde, bleibt ein neuronales Netz im untrainierten Status jedenfalls eine Vorstufe eines Computerprogramms in der Entwicklung und folglich ein vom Schutzbereich erfasstes Entwurfsmaterial gemäß § 69a Abs. 1 UrhG.[161]

Damit eine programmierte Netzarchitektur als geschütztes Computerprogramm im Sinne des Gesetzes gilt, bedarf es – wie dargestellt - darüber hinaus einer persönlichen geistigen Schöpfung gemäß § 69a Abs. 3 UrhG. Danach werden Computerprogramme geschützt sofern „sie individuelle Werke in dem Sinne darstellen, dass sie das Ergebnis der eigenen geistigen Schöpfung ihres Urhebers sind". Bei der Erstellung einer Netzarchitektur ist regelmäßig eine persönliche geistige Schöpfung anzunehmen, da ein Entwickler ein neuronales Netz für seine Zwecke programmiert. Die konkrete Programmierung ist somit durch einen Menschen als Urheber geprägt.

---

[155] vgl. Schuster/Hunzinger, in: Schuster/Grützmacher, UrhG, § 69a Rn. 5.
[156] Schuster/Hunzinger, in: Schuster/Grützmacher, UrhG, § 69a Rn. 5.
[157] vgl. Schuster/Hunzinger, in: Schuster/Grützmacher, UrhG, § 69a Rn. 5.
[158] vgl. dazu auch Nebel, Jens; Stiemerling, Oliver: "Aktuelle Programmiertechniken und ihr Schutz durch § 69a UrhG" in: Computer und Recht, Heft 1, 2016, Seite 63-64, die nur darauf abstellen, ob die Programmzeilen überhaupt eine funktionale Bedeutung haben.
[159] vgl. BGH, Urt. v. 9.5.1985 – I ZR 52/83, NJW 1986, 192, 196.
[160] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 765; vgl. dazu auch Papstefanou, Steffan: "Genetic Breeding Algorithms als Form des "Machine Learning" im Urheber- und Patentrecht" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 4, 2019, Seite 212.
[161] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 765; Dreier, in: Dreier/Schulze, UrhG, § 69a, Rn. 14.

Zu diskutieren ist, ob sogenannte Frameworks oder Libraries, die regelmäßig bei der Erstellung von Computerprogrammen und der Erstellung von neuronalen Netzen eingesetzt werden, schädlich sind. Dafür muss zunächst geklärt werden, was ein Framework und eine Library im informationstechnischen Sinne ist und inwiefern sich diese Elemente voneinander unterscheiden.

Allgemein verfolgen Frameworks und Libraries das Ziel, den Funktionsumfang eines Programms zu erweitern. Eine Library ist eine Sammlung von sogenannten Klassen und Funktionen. Über eine Schnittstelle können diese Funktionen abgerufen werden, wodurch der Entwickler leichter programmieren kann, weil eine Funktion nicht durch den Entwickler jedes Mal neu programmiert werden muss, sondern einfach die Hilfsfunktion aus der Library genutzt werden kann.

Ein Framework ist eine besondere Form einer Library. Das Framework ruft, anders als bei normalen Libraries, die Funktionen selbstständig auf.
Der Entwickler behält bei der der Einbindung einer Library die Kontrolle, wann Funktionen aufgerufen werden sollen. Bei einem Framework fehlt diese Freiheit und der Entwickler ordnet sich dem Stil des Frameworks unter. Sowohl Libraries als auch Frameworks sind keine eigenständigen Programme, sondern liefern lediglich Bausteine für die zielführende Programmierung.

Daher ist anzunehmen, dass eine persönliche geistige Schöpfung ebenso vorliegt, wenn ein Entwickler Libraries oder Frameworks bei der Erstellung einer Netzarchitektur verwendet. Der Entwickler behält besonders bei der Verwendung von Libraries die Kontrolle über die genutzten Funktionen und selbst bei Frameworks, bei denen sich der Entwickler dem jeweiligen Programmstil und seiner Struktur unterordnet, hat der Entwickler ausreichenden Gestaltungsspielraum bei der Umsetzung seiner Ideen und Grundsätze.[162]

Im Endeffekt ist danach zu fragen, ob der Entwickler genug Gestaltungsoptionen hinsichtlich seiner konkreten Umsetzung besitzt. Solange dies gegeben ist, kann von einer persönlichen geistigen Schöpfung ausgegangen werden.[163]

So ist der folgenden Meinung hinsichtlich der Nutzung von Frameworks und Libraries im Zusammenhang mit der Erstellung eines neuronalen Netzes zuzustimmen:

> „*Dass bestimmte Befehle in spezialisierten Programmierframeworks […] bereits dem Grunde nach vordefiniert sind, schließt Individualität dabei nicht per se aus.*
> *Sofern diese Befehle im Rahmen der Umsetzung des neuronalen Netzes auf Basis des Programmierframeworks individualisiert und vom Entwickler kombiniert werden, kann man davon ausgehen, dass der Schaffensprozess demnach noch immer von ihm gesteuert wird.*“[164]

---

[162] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 766. Dort wird die Nutzung des Frameworks „Tensorflow“ nicht schädlich eingestuft, da man zwischen der Basissoftware (Tensorflow) und dem vom Programmierer entwickelten Programm unterscheiden müsse.

[163] vgl. Grützmacher, in: Wandtke/Bullinger, UrhG, § 69a UrhG, Rn. 35.

[164] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 766.

Unterstützt wird die hier vertretende Meinung von dem Umstand, dass für den Schutz von Computerprogrammen die Grundätze der sogenannten „kleinen Münze“ und damit der niedrigen Schöpfungshöhe Anwendung finden.[165] Solange also programmierte neuronale Netze nicht völlig trivial oder vollkommen banal ausgestaltet wurden, scheitern diese nicht an der geforderten Gestaltungshöhe.[166]

Bei der Beurteilung der Schutzfähigkeit einer Netzarchitektur ist außerdem zu beachten, dass gemäß § 69a Abs. 2 S.1 und S. 2 UrhG ausschließlich die Ausdrucksformen wie beispielsweise der Quell- oder Maschinencode eines Computerprogramms geschützt werden. Gemäß § 69a Abs. 2 UrhG sind zugrundliegende Ideen oder Grundsätze explizit vom Schutz ausgeschlossen.

Zusammenfassen können Netzarchitekturen in Form von Programmcode nach der hier vertretenden Meinung als geschützte Computerprogramme im Sinne des § 69a UrhG angesehen werden.[167]

(2) Datenbankwerk – § 4 Abs. 2 UrhG

Den aufgezeigten Schutzpositionen folgend, stellt sich die Frage, ob eine Netzarchitektur auch als Datenbankwerk geschützt sein kann. Dabei ist zu beachten, dass der Unionsgesetzgeber eine Kollision oder Überschneidung von Datenbankrechten und Rechten an geschützten Computerprogrammen verhindern möchte. Dies ergibt sich bereits aus dem Gesetzestext, welcher in § 4 Abs. 2 UrhG besagt, dass „Ein zur Schaffung des Datenbankwerkes oder zur Ermöglichung des Zugangs zu dessen Elementen verwendetes Computerprogramm (§ 69a UrhG) […] nicht Bestandteil des Datenbankwerkes [ist]“. Diese Unterscheidung ergibt sich ebenso aus dem Erwägungsgrund 23 der EG-Richtlinie zum rechtlichen Schutz von Datenbanken (DB-RL)[168]. Folglich findet eine strenge Unterscheidung zwischen Datenbank und Computerprogramm statt, nach der ein Computerprogramm keine Datenbank darstellen kann. Wiederum enthalten bloße Datenbanken keine Steuerungsbefehle und sind folglich nicht als Computerprogramm geschützt. Unbeschadet bleibt die etwaige Schutzfähigkeit eines Computerprogramms zur Datenbankerstellung oder -abfrage.

Bejaht man den Schutz als Computerprogramm gemäß § 69a UrhG, scheidet ein Schutz als Datenbank dementsprechend aus. Zu diesem Ergebnis gelangt man jedoch nur aufgrund der ergebnisorientierten Herangehensweise, nach der ein doppelter Schutz an immateriellen

---

[165] vgl. Grützmacher, in: Wandtke/Bullinger, UrhG, § 69a UrhG, Rn. 35.

[166] vgl. Grützmacher, in: Wandtke/Bullinger, UrhG, § 69a UrhG, Rn. 37.

[167] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 766; Redeker, Helmut: IT-Recht, 2020, A. Der Schutz von Software, Rn. 15; Hartmann, Frank; Prinz, Matthias: "Immaterialgüterrechtlicher Schutz von Systemen Künstlicher Intelligenz" in: Wettbewerb in Recht und Praxis, Heft 12, 2018, Seite 1436; Hartmann, Matthias; Ohst, Claudia: „Künstliche Intelligenz im Immaterialgüterrecht“ in: KI & Recht kompakt, 2020, Seite 331-338; Söbbing, Thomas: "Künstliche neuronale Netze - Rechtliche Betrachtung von Software- und KI-Lernstrukturen" in: MMR, Heft 2, 2021, Seite 114.

[168] Richtlinie 96/9/EG des Europäischen Parlaments und des Rates vom 11.März 1996 über den rechtlichen Schutz von Datenbanken, Abl.EG Nr. L 77 v. 27.3.1996, Seite 20 ff.; zuletzt geändert durch die Richtlinie 2019/79/EU des Europäischen Parlaments und des Rates vom 17.04.2019 über das Urheberrecht und die verwandten Schutzrechte im digitalen Binnenmarkt und zur Änderung der Richtlinien 96/9/EG und 2001/29/EG, Abl.EG Nr. L 130 v. 17.05.2019, Seite 92 ff.

Gütern in Form einer Datenbank und eines Computerprogramms nicht gewollt ist, und nicht aufgrund der Prüfung der jeweiligen Tatbestandsvoraussetzungen. Man könnte jedoch argumentieren, dass für die Prüfung eines Datenbankwerkes im Sinne des § 4 Abs. 2 UrhG diejenigen Bestandteile einer Datenbank, die bereits ein Computerprogramm darstellen, bei der Beurteilung der Schutzfähigkeit außer Acht zu lassen sind.

Dennoch stellt sich die berechtige Frage, ob eine Netzarchitektur ein geschütztes Datenbankwerk im Sinne des § 4 Abs. 2 UrhG darstellen könnte. Schließlich könnte der Schutz als Computerprogramm aufgrund der oben genannten Gründe ausscheiden.

Als Datenbankwerk kommen zunächst nur Sammelwerke im Sinne der Legaldefinition in § 4 Abs. 1 UrhG in Betracht. Ein Sammelwerk ist eine Sammlung von unabhängigen Elementen, die aufgrund ihrer Auswahl oder Anordnung eine persönliche geistige Schöpfung darstellt.

Zunächst ist zu prüfen, ob eine Netzarchitektur als Sammlung unabhängiger Elemente gilt. Unabhängig sind Elemente, wenn diese von anderen Elementen getrennt werden können, ohne dass diese Trennung den selbstständigen Informationswert der Elemente beeinträchtigt.[169] Der EuGH stellt bei der Beurteilung der Unabhängigkeit darauf ab, ob sich ein Dritter für ein herausgelöstes Element interessieren würde, was zu bejahen sei, sofern das Element einem Dritten eine sachdienliche Information liefern kann. [170]
Dass ein Dritter eine sachdienliche Information durch ein einziges Neuron (inklusive Aktivierungsfunktion) erhält, ist höchstwahrscheinlich zu verneinen, da für die sachdienliche Information weitere Informationen oder Daten (wie weitere Neuronen) erforderlich sein dürften.

Da es einzelnen Neuronen einer Netzarchitektur meist an der Unabhängigkeit fehlen dürfte, scheidet ein Schutz als Datenbankwerk gemäß § 4 Abs. 2 UrhG regelmäßig aus.

(3) Datenbank – § 87a UrhG

Auch der Datenbankschutz nach § 87a UrhG setzt eine Sammlung von unabhängigen Elementen voraus, weshalb ein Schutz von Neuronen als Datenbank im Sinne des §87a UrhG regelmäßig ebenfalls aufgrund der fehlenden Unabhängigkeit der einzelnen Neuronen innerhalb der Netzarchitektur ausscheiden dürfte.

[169] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 12; vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 6; vgl. Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 21.
[170] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187, Rn. 28, 29.

b) Geschäftsgeheimnisgesetz

Damit eine Netzarchitektur als Geschäftsgeheimnis im Sinne des GeschGehG anzusehen ist, bedarf es der Erfüllung der in § 2 Nr. 1 GeschGehG enthaltenen Voraussetzungen. Dafür müsste eine Netzarchitektur eine Information mit einem Geheimnischarakter und wirtschaftlichem Wert darstellen. Des Weiteren müsste der Geheimnisinhaber ein berechtigtes Interesse an der Geheimhaltung und angemessene Geheimhaltungsmaßnahmen getroffen haben.

Die Schutzvoraussetzungen des berechtigten Interesses und der Umsetzung angemessener Geheimhaltungsmaßnahmen werden aufgrund ihrer Subjektivität und damit ihrer Abhängigkeit vom jeweiligen Inhaber nur kurz betrachtet.

Damit eine Netzarchitektur als geheime Information anzusehen ist, müsste diese zum einen eine Information darstellen und zum anderen geheim sein. Eine Information im Sinne des GeschGehG wird als „Wissen um eine Tatsache“ definiert.[171] Das Kriterium dürfte hinsichtlich einer Netzarchitektur eher erfüllt sein, da es nicht zuletzt niedrig anzusetzen ist, sodass es keiner bestimmten Form oder Individualität, Neuheit, Kreativität oder Originalität bedarf.[172]

Für die Beurteilung, ob eine Information als geheim anzusehen ist, ist der Bezugspunkt der Geheimhaltung, der Grad der Geheimhaltung und die Bestimmung des maßgeblichen Verkehrskreises entscheidend.[173]

Es ist also zu untersuchen, inwiefern die Netzarchitektur oder dessen Bestandteile geheim sind. Neuronale Netze sind beliebte Modelle im ML-Bereich, weshalb zwar die einzelnen Bestandteile oft nicht als geheim anzusehen sind. Die konkrete Zusammensetzung der Neuronen innerhalb der Netzarchitektur dürfte hingegen meist nicht allgemein bekannt sein, da sich die Lösungswege für die Realisierung insbesondere komplexer Verarbeitungen von Entwickler zu Entwickler unterscheiden dürften.

Ob diese genaue Struktur dennoch allgemein bekannt ist, ist von einem bestimmten Personenkreis abhängig. Gemäß § 2 Nr. 1 GeschGehG ist der Geheimnischarakter nach dem Personenkreis, der üblicherweise mit solch einer Information umgeht, zu beurteilen. Im Falle einer Netzarchitektur ist also auf ein Fachpublikum, das regelmäßig mit neuronalen Netzen zu tun hat, abzustellen.[174]

Aufgrund der informationsspezifischen Bestimmung des vertraulichkeitsbegründenden Personenkreises[175] wäre im konkreten Fall von Netzarchitekturen auf internationale Datenwissenschaftler und Informatiker abzustellen, die sich mit neuronalen Netzen auseinandersetzen.
Einfache, vorgefertigte oder branchenbekannte Netzarchitekturen (bspw. Standardarchitekturen zweiklassiger neuronaler Netzwerke[176]) dürften als allgemein bekannt

[171] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 2,3.
[172] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 2,3.
[173] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 7.
[174] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 9.
[175] Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG § 2 Rn. 33.
[176] s. sog. „fully connected case” im Azure Machine Learning-Designer, Microsoft: “Komponente „Two-Class Neural Network“ (Neuronales Netz mit zwei Klassen)”, online URL:

gelten, sodass diese keine geheime Information darstellen. Ebenso scheidet ein Schutz für Netzarchitekturen aus, sofern diese mittels einfacher Rekonstruktionen in wenigen Stunden selbst erstellt werden können.[177]
Je komplexer also eine Netzarchitektur ausgestaltet ist, desto wahrscheinlicher ist es, dass diese nicht allgemein bekannt und somit als geheim anzusehen ist. „Tiefe“ neuronale Netze abseits vorgefertigter Netzarchitekturen können somit regelmäßig eine geheime Information darstellen.

Darüber hinaus müsste die Netzarchitektur einen wirtschaftlichen Wert verkörpern. Ob eine Information einen wirtschaftlichen Wert besitzt, kann durch einen möglichen Handelswert oder durch den möglichen Schaden einer unbefugten Nutzung oder Offenlegung festgestellt werden.[178] Eine unbefugte Offenlegung der Netzarchitektur im Quellcode im Internet könnte somit die Interessen des Entwicklers eines neuronalen Netzes beeinträchtigen. Eine wirtschaftliche Verwertung wäre durch diese unbefugte Offenlegung immens eingeschränkt, da sich zum einen die Offenlegung nicht rückgängig machen lässt und zum anderen die Offenlegung des Quellcodes jedermann die Nutzung ermöglicht.

Im Übrigen ist das Vorliegen eines wirtschaftlichen Wertes als niedrige Hürde anzusehen, zumal sich dieser laut Gesetzestext bereits aus dem Geheimnischarakter ergibt.[179]

Ein wirtschaftlicher Wert ist somit für eine Netzarchitektur regelmäßig anzunehmen.

Des Weiteren bedarf es für die Erfüllung des Geschäftsgeheimnisschutzes eines berechtigten Interesses an der Geheimhaltung und der Umsetzung angemessener Geheimhaltungsmaßnahmen.

Ein berechtigtes Interesse liegt grundsätzlich vor, sofern nicht gegen legitime Interessen verstoßen wird. Ein solcher Verstoß läge beispielsweise bei der Behinderung von investigativem Journalismus[180] oder der Unterstützung illegaler Aktivitäten vor.[181] Ebenso liegt kein berechtigtes Interesse an der Geheimhaltung jeder Information vor, sodass die Einstufung jeder Information als Geschäftsgeheimnis einen Verstoß gegen das berechtigte Interesse darstellt.[182] Diese Ausnahmen dürften auf neuronale Netze im Bereich der Fertigungsindustrie regelmäßig nicht zutreffen.

Grundsätzlich hat der Ersteller einer Netzarchitektur ein Interesse an dessen Geheimhaltung, zumal eine wirtschaftliche Verwertung bei immateriellen Gütern deren Geheimhaltung voraussetzt. Das ergibt sich bereits aus der Natur der Sache und der Tatsache, dass regelmäßig Computerprogramme in Maschinencode übermittelt werden. Eine Netzarchitektur erfüllt somit regelmäßig die Tatbestandsvoraussetzungen eines Geschäftsgeheimnisses. Der jeweilige Inhaber hat jedoch Sorge zu tragen, dass den Umständen nach angemessene Geheimhaltungsmaßnahmen getroffen werden.

---

https://docs.microsoft.com/de-de/azure/machine-learning/component-reference/two-class-neural-network, Abruf 12.09.2022.

[177] vgl. Hiéramente, in: BeckOK GeschGehG, § 2, Rn. 9.

[178] vgl. Erwägungsgrund 14 der TS-RL.

[179] vgl. § 2 Nr. 1 lit. a GeschGehG.

[180] vgl. Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG, § 2, Rn. 73.

[181] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 444.

[182] vgl. Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG, § 2, Rn. 73.

Zusammenfassung – Netzarchitektur:

Die Netzarchitektur eines neuronalen Netzes kann nach der hier vertretenen Meinung die Anforderungen eines Computerprogrammes im Sinne des § 69a UrhG erfüllen.
Es wird dennoch darauf hingewiesen, dass es unterschiedliche Rechtsauffassungen hinsichtlich der Auslegung des Begriffs Computerprogramm im Zusammenhang mit der Netzarchitektur eines neuronalen Netzes gibt. Unabhängig davon lässt sich die Netzarchitektur eines neuronalen Netzes als Vorstufe eins Computerprogramms in der der Entwicklung einordnen und stellt somit ein geschütztes Entwurfsmaterial gemäß § 69a Abs. 1 UrhG dar.[183]
Der Schutz als Datenbankwerk im Sinne des § 4 Abs. 2 UrhG und als Datenbank im Sinne des § 87a UrhG scheidet in der Regel aufgrund fehlender Unabhängigkeit der einzelnen Neuronen aus.

Eine Netzarchitektur kann als Geschäftsgeheimnis geschützt sein, sofern diese in ihrer Anordnung geheim und durch angemessene Geheimhaltungsmaßnahmen geschützt ist.
Ein wirtschaftlicher Wert und ein berechtigtes Interesse an der Geheimhaltung eines neuronalen Netzes dürften nur ausnahmsweise zu verneinen sein.

2. Lern- und Optimierungsfunktion

Die Lern- und Optimierungsfunktion ist für das Anpassen der Gewichte innerhalb der Netzarchitektur verantwortlich und deshalb entscheidend für die Erstellung funktionsfähiger neuronaler Netze.

Durch den Abgleich der Labels der Trainingsdaten werden mithilfe der Lern- und Optimierungsfunktion die Gewichte angepasst, sodass diese bei weiteren Verarbeitungen erfolgversprechende Ergebnisse liefern. Die Lern- und Optimierungsfunktion ermöglicht also die Erstellung verlässlicher ML-Modelle. Dabei ist zu beachten, dass die Lern- und Optimierungsfunktion unterschiedlich ausgestaltet werden kann. Das jeweilige neuronale Netz, die zu Verfügung stehende Menge und Qualität an Trainingsdaten, der Arbeitsspeicher des ausführenden Geräts und weitere Faktoren beeinflussen die Auswahl der jeweiligen Lern- und Optimierungsfunktion.[184]

Ein beliebtes Verfahren einer Lern- und Optimierungsfunktion ist die sogenannte Fehlerrückführung (auch „Backpropagation“ genannt). Dabei werden die Gewichte des neuronalen Netzes, vom Ergebnis ausgehend, rückwärts angepasst. Es werden also vom Output-Layer Richtung Input-Layer alle Gewichte der Hidden-Layer konfiguriert. Ziel ist es, die sogenannte Fehlerquote der falschen Verarbeitung zu minimieren.

[183] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 765; Dreier, in: Dreier/Schulze, UrhG, § 69a, Rn. 14.

[184] vgl. Nguyen, Giang; et al.: "Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey" in: Artificial Intelligence Review, Ausgabe 52, Heft 1, 2019, Seite 81.

a) Urhebergesetz

Bei der Prüfung eines urheberrechtlichen Schutzes kommt für die Lern- und Optimierungsfunktion grundsätzlich nur ein Schutz als Computerprogramm im Sinne des § 69a UrhG in Frage.

Um zu verstehen, ob eine Lern- und Optimierungsfunktion als Computerprogramm geschützt ist, muss zuallererst die Abgrenzung zwischen einem Computerprogramm und einem Algorithmus veranschaulicht werden. Grundsätzlich ist eine Lern- und Optimierungsfunktion ein Algorithmus. Allgemein werden Algorithmen als „präzise Verarbeitungsvorschriften, die von einem mechanisch oder elektronisch arbeitenden Gerät durchgeführt werden können", definiert.[185] Alle Computerprogramme in Form eines Programmcodes stellen somit Algorithmen dar.[186] Dennoch sind nicht alle Computerprogramme oder Algorithmen durch das Urheberrecht geschützt.
Algorithmen, die allgemein bekannte mathematische Rechenregeln darstellen, sind beispielsweise nicht schutzfähig.[187] Sobald Algorithmen sogenannte Routinen darstellen, die zum Standardrepertoire oder der Programmiertechnik gehören, scheiden sie als schutzfähige Computerprogramme aus.[188] Zum Standardrepertoire der Programmiertechniken zählen alle Routinen, die sich bei der Lösung bestimmter Aufgaben bewährt haben und allgemein üblich verwendet werden. Diese Auslegung widerspricht auch dem Erwägungsgrund 13 der Computerprogramm-RL nicht, die lediglich die Ideen und Grundsätze, welche einem Algorithmus zugrunde liegen, explizit vom Urheberrechtsschutz ausklammert.[189]
Die Einräumung von urheberrechtlichen Schutzpositionen an solchen Routinen würde zur Monopolisierung logischer Abfolgen und Methoden führen, die der Entwicklung neuer Ideen erheblich im Weg stehen würde.[190]
Die Abgrenzung zwischen nicht schutzfähiger Routine und schutzfähigem Computerprogramm erfolgt nach dem möglichen Gestaltungsspielraum, den ein gewährter Schutz eines Computerprogramms übrig lassen würde. Solange also genügend Gestaltungsspielraum verbleibt, sodass es „anderen Urhebern die Schaffung ähnlicher oder sogar identischer Programme [ermöglicht]" [191], soll den Urhebern dieser Werken Schutz gewährt werden „sofern sie die Werke anderer nicht kopieren."[192] Die Gewährung eines Schutzes für einzelne Algorithmen, die grundlegenden Routinen darstellen, würde für andere keinen Gestaltungsraum überlassen.

Algorithmen, die bei einem gewährten Schutz Gestaltungsraum für andere belassen, können unter den Schutz eines Computerprogramms fallen.[193] Dies kann dazu führen, dass ein geschütztes Computerprogramm in Form von Programmcode Bestandteile ausweist, die nicht schutzfähige Routinen verkörpern.[194]

---

[185] Kaboth/Spies, in: BeckOK, UrhG, § 69a, Rn. 12.
[186] vgl. Kaboth/Spies, in: BeckOK, UrhG, § 69a, Rn. 12.
[187] vgl. Dreier, in: Dreier/Schulze, UrhG, § 69a, Rn. 22.
[188] vgl. Spindler, in: Schricker/Loewenheim, UrhG, § 69a, Rn. 12a.
[189] vgl. Erwägungsgrund 11 der Computerprogramm-RL, vom 23. April 2009, 2009/24/EG.
[190] vgl. Spindler, in: Schricker/Loewenheim, UrhG, § 69a, Rn. 12a.
[191] EuGH, Urt. v. 2.5.2012 – C-406/10, MMR 2012, 468, 469.
[192] EuGH, Urt. v. 2.5.2012 – C-406/10, MMR 2012, 468, 469.
[193] vgl. Dreier, in: Dreier/Schulze, UrhG, § 69a, Rn. 22.
[194] vgl. Dreier, in: Dreier/Schulze, UrhG, § 69a, Rn. 22; vgl. Grützmacher, in: Wandtke/Bullinger, UrhG, § 69a UrhG, Rn. 29.

Daher bleiben Anwendungen und Verknüpfungen von Algorithmen in einem Programm sowie die Art und Weise der Implementierung und Zuordnung zueinander schutzfähig.[195]

Mithin dürften sehr einfache Lern- und Optimierungsfunktionen, welche sich für die Lösung bestimmter Aufgaben bewährt haben und auch ebenso üblich verwendet werden, nicht schutzfähige Routinen darstellen. Oftmals werden Lern- und Optimierungsfunktionen durch das Framework vordefiniert und ein Entwickler kombiniert lediglich verschiedene Lernalgorithmen.[196] Für viele Lern- und Optimierungsfunktionen wird ein separater Schutz als Computerprogramm auch deshalb regelmäßig ausscheiden.[197]
Hingegen kann eine selbst erstellte hochkomplexe Lern- und Optimierungsfunktion durch Verknüpfung mit weiteren Algorithmen durchaus als Computerprogramm geschützt sein[198], wobei die genaue Abgrenzung immer vom konkreten Einzelfall abhängig ist.

b) Geschäftsgeheimnisgesetz

Das GeschGehG schützt geheime Informationen vor dem Zugang unbefugter Dritter. Standardisierte Lern- und Optimierungsfunktionen sind jedoch nicht geheim, da dem betreffenden Fachkreis diese Lern- und Optimierungsfunktionen überwiegend bekannt sein dürften. Folglich scheiden diese aus dem Schutzbereich des GeschGehG aus.

Für komplexe Lern- und Optimierungsfunktionen könnte dagegen der Schutzbereich eröffnet sein. Dafür dürften diese dem Fachkreis, bestehend aus internationalen Datenwissenschaftlern und Informatikern, die sich regelmäßig mit neuronalen Netzen auseinandersetzen, nicht allgemein bekannt oder ohne Weiteres zugänglich sein. Zu beachten ist hierbei die schnelle Innovationsgeschwindigkeit im Bereich des ML. [199] Eine geheime Lern- und Optimierungsfunktion kann somit in der Zukunft zum Stand der Technik gehören, aufgrund paralleler Entwicklung und Offenlegung.

Sollte eine Lern- und Optimierungsfunktion geheim sein, besitzt diese grundsätzlich einen wirtschaftlichen Wert, aufgrund der niedrigen Hürde eines potenziellen Wertes.[200] Ein berechtigtes Interesse an der Geheimhaltung ist grundsätzlich anzunehmen und die Umsetzung angemessener Geheimhaltungsmaßnahmen ist vom konkreten Inhaber und Einzelfall abhängig.

Im Ergebnis können Lern- und Optimierungsfunktionen abseits standardisierter Funktionen ein Geschäftsgeheimnis darstellen.

---

[195] vgl. BGH, Urt. v. 04.10.1990 – I ZR 139/89, Betriebssysteme, GRUR 1991, 449, 453; vgl. Wiebe, in: Spindler/Schuster, UrhG, § 69a, Rn. 25.
[196] vgl. Hartmann, Matthias; Ohst, Claudia: „Künstliche Intelligenz im Immaterialgüterrecht“ in: KI & Recht kompakt, 2020, Seite 327.
[197] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 766.
[198] vgl. BGH, Urt. v. 04.10.1990 – I ZR 139/89, Betriebssysteme, GRUR 1991, 449, 453; vgl. Wiebe, in: Spindler/Schuster, UrhG, § 69a, Rn. 25.
[199] vgl. Döbel, Inga; et al.: "Maschinelles Lernen - Kompetenzen, Anwendungen und Forschungsbedarf" in: Fraunhofer-Allianz Big Data und Künstliche Intelligenz online, 2018, URL: https://www.bigdata- ai.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/BMBF_Fraunhofer_ML-Ergebnisbericht_Gesamt.pdf, Abruf 14.02.2022, Seite 14, 16.
[200] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 443.

Zusammenfassung – Lern- und Optimierungsfunktion:

Vordefinierte Lern- und Optimierungsfunktionen scheiden aus dem urheberrechtlichen Schutz aus.
Abseits davon können eigenentwickelte komplexe und in Programmcode formulierte Lern- und Optimierungsfunktionen ein schutzfähiges Computerprogramm im Sinne des § 69a UrhG darstellen. Zu beachten ist dabei, dass es sich um keine Routinen, also Algorithmen, die allgemein bekannte mathematische Rechenregeln darstellen, handeln darf.
Ebenso können Lern- und Optimierungsfunktionen, die nicht allgemein bekannt oder ohne Weiteres zugänglich sind, unter den Geschäftsgeheimnisschutz fallen.

3. Hyperparameter

Eine zentrale Rolle innerhalb der Lern- und Optimierungsfunktion stellen die sogenannten Hyperparameter dar. Diese werden vor der Verarbeitung festgelegt und bestimmen, wie schnell ein Gewicht angepasst werden soll oder wie viele Durchläufe die Lern- und Optimierungsfunktion tätigen soll. Ohne die richtigen Hyperparameter ist eine erfolgsversprechende Erstellung eines neuronalen Netzes nur schwer zu erreichen.

Eine urheberrechtliche Einordnung scheidet aus: Hyperparameter, die grundsätzlich viel Wissen und Erfahrung benötigen, sind losgelöst nur einzelne Werte ohne weiterreichende Programmieraufwand.
Sie könnten lediglich über den urheberrechtlichen Schutzbereich der Netzarchitektur mitgeschützt werden, sofern eine Entnahme aus dem Programmcode der Netzarchitektur mit einer zustimmungspflichtigen Umarbeitung gemäß § 69c Nr. 2 UrhG verbunden wäre.[201]

Hyperparameter können jedoch unter den Schutzbereich des GeschGehG fallen. Dafür müssten diese ein Geschäftsgeheimnis im Sinne des § 2 Nr. 1 GeschGehG sein. Demnach müssten Hyperparameter eine geheime Information mit wirtschaftlichem Wert darstellen. Ebenso müsste ein berechtigtes Interesse an der Geheimhaltung bestehen und der Inhaber muss angemessene Geheimhaltungsmaßnahmen treffen.

Diskutierbar ist, ob Hyperparameter überhaupt eine Information darstellen. Unter einer Information im Sinne des GeschGehG versteht man das „Wissen um eine Tatsache“.[202] Somit handelt es sich um eine niedrige Hürde.
Hyperparameter stellen Wissen über das Training neuronaler Netze dar. Folglich erfüllen Hyperparameter die Voraussetzung einer Information.

Ob Hyperparameter ebenfalls eine geheime Information darstellen, ist vom konkreten Einzelfall abhängig. Hyperparameter werden in aller Regel nicht in ihren Bestandteilen insgesamt geheim sein, jedoch wird die genaue Anordnung und Zusammensetzung für einen bestimmtes neuronales Netz und dessen Trainingsprozess regelmäßig nicht einem entsprechenden Fachpublikum zugänglich oder bekannt sein.

[201] vgl. im Zusammenhang mit der Entnahme von Gewichtungsdaten: Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 767.
[202] vgl. Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 2,3.

Sobald sich jedoch bestimmte Hyperparameter für gewisse Netzarchitekturen und Trainingsprozesses in den hiermit befassten Fachkreisen etabliert haben, scheiden sie unzweifelhaft aus dem Schutzbereich eines Geschäftsgeheimnisses aus, da diese allgemein bekannt sind. Inwiefern sich deren Geheimnischarakter darlegen lässt, ist durchaus zu berücksichtigen. Hyperparameter müssten sich eindeutig von anderen Hyperparametern abgrenzen lassen. Was für deren Geheimsein spricht, ist die Tatsache, dass viele Hyperparameter das Ergebnis mehrerer Trainingsmodelle darstellen. Vorstellbar sind auch Hyperparameter, die durch experimentelle Versuche komplexer tiefer neuronaler Netze entwickelt wurden.

Damit können Hyperparameter auch einen wirtschaftlichen Wert besitzen, da diese Wissen und Erfahrung über erfolgreiche Trainingsprozesse bündeln können und ein wirtschaftlicher Wert bereits bei einem potenziellen Wert besteht, wie es beispielsweise bei neuen technischen Erkenntnissen der Fall ist.[203]

Unter der Annahme, dass ein legitimes berechtigtes Interesse an der Geheimhaltung besteht und der Inhaber den Umständen nach angemessene Geheimhaltungsmaßnahmen getroffen hat, können Hyperparameter ein Geschäftsgeheimnis im Sinne des GeschGehG darstellen.

Zusammenfassung – Hyperparameter:

Für Hyperparameter scheidet ein urheberrechtlicher Schutz nach hier vertretener Auffassung aus. Jedoch können Hyperparameter abseits bekannter Heuristiken ein Geschäftsgeheimnis darstellen. Die Abgrenzung zwischen geheimen und allgemein bekannten Hyperparametern könnte sich jedoch mit zunehmender Kenntnis wichtiger Parameter in der Praxis als schwierig herausstellen. Hinzukommt, dass derjenige, welcher sich auf den Schutz des Geschäftsgeheimnisgesetzes berufen möchte, darlegen und ggf. beweisen muss, dass der als Geheimnis geschützten Information am Markt nicht bekanntes Wissen zugrunde liegt.[204]

[203] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 443.
[204] vgl. Fuhlrott, in: BeckOK, GeschGehG, § 2, Rn. 68.

B. Schutz von trainierten neuronalen Netzen

Nachdem die rechtliche Schutzfähigkeit von untrainierten neuronalen Netzen untersucht wurde, soll nachfolgend der Schutz von trainierten neuronalen Netzen geprüft werden.
Den maßgeblichen Unterschied zwischen untrainierten und trainierten neuronalen Netzen bilden trainierte Gewichtungsdaten. Um trainierte Gewichtungsdaten jedoch zu erhalten, bedarf es einer Vielzahl an Test- und Trainingsdaten. Daher sollen ebenso die möglichen Schutzrechtspositionen an Test- und Trainingsdaten aufgezeigt werden.

1. Test- und Trainingsdaten

Test- und Trainingsdaten sind ein integraler Bestandteil eines ML-Learning Prozesses. Ohne Trainingsdaten mit hoher Qualität und Güte oder mit zu wenigen Trainingsdaten können keine sicheren und zuverlässigen ML-Modelle entwickelt werden.[205] Bei der Aufbereitung von Test- und Trainingsdaten können verschiedene Handlungen vollzogen werden, wie die Bereinigung von falschen Daten, die Aggregation und Modifikation verschiedener Daten oder die Beschriftung von Daten mit einem Zielwert bzw. Label.

Obwohl eine Aussage über die Schutzfähigkeit von Test- und Trainingsdaten immer eine Frage des Einzelfalls ist, soll unter bestimmten Annahmen untersucht werden, inwiefern diesen rechtlicher Schutz zukommen kann.

a) Urhebergesetz

Aus urheberrechtlicher Perspektive kommen drei Schutzrechtspositionen hinsichtlich Test- und Trainingsdaten in Frage. Eine geschützte Darstellung wissenschaftlicher oder technischer Art gemäß § 2 Abs. 1 Nr. 7 UrhG, ein Datenbankwerk gemäß § 4 Abs. 2 UrhG und eine Datenbank gemäß § 87a UrhG.

Nachfolgend sollen Test- und Trainingsdaten auf die genannten Schutzrechtspositionen untersucht werden.

(1) Darstellung wiss. oder techn. Art – § 2 Abs. 1 Nr. 7 UrhG

Die Hürde einer Darstellung wissenschaftlicher oder technischer Art ist eher niedrig anzusetzen. Es genügen bereits einfachste wissenschaftliche Erkenntnisse in einer Darstellung, die eine persönliche geistige Schöpfung verkörpern.[206] Darstellungen wissenschaftlicher oder technischer Art sind beispielsweise Konstruktionszeichnungen, Stadtpläne, Karten, Skizzen, Hinweisschilder, Tabellen, statistische Übersichten oder

[205] vgl. Döbel, Inga; et al.: "Maschinelles Lernen - Kompetenzen, Anwendungen und Forschungsbedarf" in: Fraunhofer-Allianz Big Data und Künstliche Intelligenz online, 2018, URL: https://www.bigdata- ai.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/BMBF_Fraunhofer_ML-Ergebnisbericht_Gesamt.pdf, Abruf 14.02.2022, Seite 14, 47.; Europäische Kommission: "Weißbuch zur künstlichen Intelligenz", COM (2020) 65 final vom 19.2.2020, Seite 22; Bundestag: Bericht der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale, in: Bundestag online, Drucksache 19/23700, 2020, URL: https://dserver.bundestag.de/btd/19/237/1923700.pdf, Abruf 14.02.2022, Seite 53.

[206] vgl. BGH, Urt. v. 1. 6. 2011 – I ZR 140/09, GRUR 2011, 803, Rn. 43.

Schaubilder.[207] Davon zu unterscheiden sind Aufzählungen bloßer Fakten oder Darstellungen, die auf einer DIN-Norm basieren.[208] Diese Darstellungen überschreiten nicht die Hürde einer persönlichen geistigen Schöpfung im Sinne des § 2 Abs. 1 Nr. 7 UrhG.

Damit Test- und Trainingsdaten die Voraussetzung erfüllen, müssten diese wissenschaftliche Erkenntnisse verkörpern, die sich durch ihre Art und Form der Auswahl, Einteilung, übersichtliche Anordnung oder durch andere Zusammenhänge auszeichnen.[209]

Bei der Beurteilung, ob eine persönliche geistige Schöpfung vorliegt, müsste bei Test- und Trainingsdaten darauf abgestellt werden, ob sich diese von herkömmlichem Trainingsmaterial aufgrund der Behandlung und Zusammenstellung abweichen.[210] Unter der Annahme, dass Test- und Trainingsdaten beispielsweise unter Anwendung spezifischen Fachwissens einzeln aufbereitet und zusammengestellt wurden, könnte dieses Schutzvoraussetzung erfüllt sein. So wird angenommen, dass Test- und Trainingsdaten, die eine Darstellung von radiologischen oder dermatologischen Bildern mit medizinischen Befunden verkörpern, eine persönliche geistige Schöpfung darstellen können.[211] Derjenige, der diese Bilder aufbereitet, müsste dafür eine Darstellung verwenden, die von standardisieren Vorgaben abweichen. Bei komplexem Material, wie radiologischen oder dermatologischen Befunden inklusive Bilder, könnte der Ersteller eine Darstellung verwenden, welche die Voraussetzung einer persönlichen geistigen Schöpfung erfüllt.

Bei der Beurteilung einer persönlichen geistigen Schöpfung genügt bereits ein geringes Maß an individueller Prägung[212] (kleine Münze des Programmschaffens[213]). Danach versteht man ein Minimum an Gestaltungshöhe, die gerade noch urheberechtsschutzfähig ist.[214]
Sollte sich jedoch der Ersteller bei der Darstellung von komplexem Material dennoch an standardisierten Vorgaben orientieren, scheidet ein Schutz aufgrund fehlender persönlicher geistiger Schöpfung aus. [215]

Bei Test- und Trainingsdaten, die durch Maschinen erstellt und dargestellt werden, kann eine persönliche geistige Schöpfung regelmäßig ausgeschlossen werden. Die Maschine wird in der Regel Test- und Trainingsdaten nach standardisierten Vorgängen aufbereiten. Es handelt sich hierbei eher um eine Fleißarbeit mit kaum Spielraum für individuelle Leistung und somit auch ohne Akt der persönlichen geistigen Schöpfung.[216]

---

[207] vgl. Schulze, in: Dreier/Schulze, UrhG, § 2, Rn. 222.
[208] vgl. Schulze, in: Dreier/Schulze, UrhG, § 2, Rn. 222, 228.
[209] vgl. Schulze, in: Dreier/Schulze, UrhG, § 2, Rn. 235.
[210] vgl. Hacker, Philipp: "Immaterialgüterrechtlicher Schutz von KI-Trainingsdaten" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 10, 2020, Seite 1027.
[211] vgl. Hacker, Philipp: "Immaterialgüterrechtlicher Schutz von KI-Trainingsdaten" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 10, 2020, Seite 1027; vgl. Maron, Roman C.; et al.: "Robustness of convolutional neural networks in recognition of pigmented skin lesions" in: European Journal of Cancer, Ausgabe 145, 2021, Seiten 81-91; vgl. dazu auch Klinikum Universität Heidelberg: „Künstliche Intelligenz schlägt Hautärzte“ in: dieselbe online, URL: https://www.klinikum.uni-heidelberg.de/newsroom/kunstliche-intelligenz-schlagt-hautarzte/, Abruf 14.02.2022.
[212] Explizit geregelt in Art. 1 Abs. 3 Software-RL.
[213] vgl. BGH v. 03.03.2005 – I ZR 111/02.
[214] vgl. Loewenheim/Leistner, in: Schricker/Loewenheim, UrhG, § 2, Rn. 61.
[215] vgl. dazu Landkarten und Stadtpläne: Ahlberg, in: BeckOK, UrhG, § 2, Rn. 151.
[216] vgl. dazu Landkarten und Stadtpläne: Ahlberg, in: BeckOK, UrhG, § 2, Rn. 151.

Im Ergebnis können Test- und Trainingsdaten eine Darstellung wissenschaftlicher und technischer Art darstellen. Zu beachten ist, dass nicht entscheidend ist, was dargestellt wird, sondern wie etwas dargestellt wird.[217] Somit ist die Beurteilung einer möglichen Schutzfähigkeit auch immer eine Entscheidung des Einzelfalles.

(2) Datenbankwerk – § 4 Abs. 2 UrhG

Für Test- und Trainingsdaten könnte ebenso das Schutzrecht eines Datenbankwerks gemäß § 4 Abs. 2 UrhG in Frage kommen. Danach müssten die Daten ein Sammelwerk darstellen. Zusätzlich müssten die enthaltenden Elemente unabhängig, systematisch oder methodisch angeordnet und einzeln mit Hilfe elektronsicher Mittel oder auf andere Weise zugänglich sein.

Gemäß der Legaldefinition in § 4 Abs. 1 UrhG ist ein Sammelwerk eine Sammlung von Werken, Daten oder anderen unabhängigen Elementen, die aufgrund der Auswahl oder Anordnung der Elemente eine persönliche geistige Schöpfung darstellen. Ein Sammelwerk kann danach in eine Sammlung und in ein Werk im urheberrechtlichen Sinne unterteilt werden. Eine Sammlung wird angenommen, sobald eine Zusammenstellung mehrerer unabhängiger Elemente vorliegen.[218] Wie viele Elemente benötigt werden, lässt sich dabei nicht auf eine bestimmte Anzahl beziffern. Vielmehr muss ein sogenannter Sammlungscharakter vorliegen, bei dem die Anzahl der Elemente erkennbar ist.[219] Dieses Merkmal sollte aber regelmäßig bei Sammlungen von Test- und Trainingsdaten nicht zum Ausscheiden führen, da in der Regel eine Vielzahl an Test- und Trainingsdaten verwendet werden, welche die Voraussetzung einer Sammlung grundsätzlich erfüllen dürften.

Des Weiteren müsste diese Sammlung ein Werk im Sinne des § 2 Abs. 2 UrhG darstellen. Dabei bedarf es gemäß § 4 Abs. 1 UrhG einer persönlichen geistigen Schöpfung hinsichtlich der Auswahl und Anordnung der enthaltenen Elemente.

Das OLG Frankfurt sah die Voraussetzung einer persönlichen geistige Schöpfung bereits als erfüllt an, sofern ein anderer Urheber möglicherweise eine andere Auswahl oder Anordnung getroffen hätte.[220] Entgegen dieser Auffassung entschied der EuGH später die Auslegung einer persönlichen geistigen Schöpfung anders und interpretierte die Anforderung strenger.[221] Im Falle „Football Dataco“ entschied der EuGH, dass eine persönliche geistige Schöpfung vorliegt, wenn der Schöpfer bei der Auswahl oder Anordnung eine freie und kreative Entscheidung trifft.[222] Des Weiteren liegt keine persönliche geistige Schöpfung, wenn „die Erstellung der Datenbank durch technische Erwägungen, Regeln oder Zwänge bestimmt wird, die für künstlerische Freiheit keinen Raum lassen“.[223]

Test- und Trainingsdaten, die durch einen Menschen ausgewählt und angeordnet werden, um ein ML-Modell zu trainieren, könnten eine persönliche geistige Schöpfung verkörpern. Für diese Ansicht spricht, dass die Daten durch eine Person teilweise einzeln aufbereitet,

[217] vgl. Schulze, in: Dreier/Schulze, UrhG, § 2, Rn. 223.
[218] vgl. Marquardt, in: Wandtke/Bullinger, UrhG, § 4 UrhG, Rn. 4.
[219] vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 4; Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 3, 4; Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 11.
[220] vgl. OLG Frankfurt/M., Urt. v. 19.6.2001 – 11 U 66/00, MMR 2002, 687.
[221] vgl. EuGH, Urt. v. 1. 3. 2012 – C-604/10, GRUR 2012, 386, Rn. 38.
[222] vgl. EuGH, Urt. v. 1. 3. 2012 – C-604/10, GRUR 2012, 386, Rn. 38.
[223] vgl. EuGH, Urt. v. 1. 3. 2012 – C-604/10, GRUR 2012, 386, Rn. 39.

ausgewählt und angeordnet werden. Dagegen könnte man einwenden, dass die Auswahl und Anordnung dieser Test- und Trainingsdaten durch technische Erwägungen, Regeln oder Zwänge, die ein ML-Modell vorgeben, geprägt sind. Somit wäre kein Freiraum für eine persönliche geistige Schöpfung, die auch der EuGH zwingend voraussetzt, übrig.[224] Für den Schutz als Datenbankwerk kommt es mithin entscheidend darauf an, ob die konkret vorliegenden Test- und Trainingsdaten in ihrer Auswahl und Anordnung auf einer freien und kreativen Entscheidung basieren, oder ob technische Erwägungen, Regeln oder Zwänge den Schaffungsprozess prägten.

KI-generierte Test- und Trainingsdaten würden aus diesem Grund beispielsweise regelmäßig ausscheiden. Genauso eine Sammlung mehrerer Datenpunkte in alphabetischer oder chronologischer Reihenfolge, oder eine Zusammenstellung nach einer Logik oder Zweckmäßigkeit. Insbesondere Letzteres lässt keinen Freiraum für eine persönliche geistige Schöpfung.[225]

Damit ein Sammelwerk vorliegt, müssen die Elemente innerhalb der Sammlung unabhängig sein. Elemente besitzen eine Unabhängigkeit, wenn sich diese nach Aufnahme in ein Sammelwerk erkennen und herauslösen lassen.[226] Die Voraussetzung eines unabhängigen Elements unterscheidet das Datenbankwerk von einer Darstellung wissenschaftlicher oder technischer Art.[227]
Unter Heranziehung der Literaturmeinung eines unabhängigen Elements im Sinne des § 87a UrhG liegt ein unabhängiges Element vor, sofern ein Element von anderen Elementen getrennt werden kann, ohne dass der selbstständige Informationswert hierdurch beeinträchtigt wird.[228]

Auf Test- und Trainingsdaten angewendet wird man feststellen, dass diese sehr heterogen sind und eine eindeutige Einstufung nicht erfolgen kann. Test- und Trainingsdaten können sowohl Bilddateien als auch einzelne Werte von Sensoren darstellen. Bei Bilddateien wird man viel eher eine Unabhängigkeit bejahen können, als bei einzelnen Werten.
Der EuGH hat in einem Urteil den Einwand, dass der Wert einzelner Bestandteile von Elementen sich auf annähernd null reduzieren, sofern man diese auf den einzelnen Bestandteil herunterbricht, unberücksichtigt gelassen und nicht bei der Beurteilung der Elemente einbezogen.[229] Bei der Beurteilung von Kombinationen von Zahlen und Werten sind diese somit nicht auf deren kleinste Bestandteile zu reduzieren. Diesem Einwand ist zuzustimmen, da das Reduzieren von Elementen auf deren einzelnen Bestandteile zur Folge hätte, dass jedes Element keine Unabhängigkeit besitzen würde. Vielmehr hätte man nur einzelne Zahlen und Werte.

[224] vgl. EuGH, Urt. v. 1. 3. 2012 – C-604/10, GRUR 2012, 386, Rn. 39.
[225] vgl. OLG Hamburg, Urt. v. 06.05.1999 – 3 U 246/98, GRUR 2000, 319, 320; OLG Hamburg, Beschluss vom 3. Mai 1996 – 3 W 53/96, ZUM 1997, 145, 146.
[226] vgl. Ahlberg, in: BeckOK, UrhG, § 4, Rn. 15.
[227] vgl. Hacker, Philipp: "Immaterialgüterrechtlicher Schutz von KI-Trainingsdaten" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 10, 2020, Seite 1028.
[228] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 12; vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 6; vgl. Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 21.
[229] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187, Rn. 18, 28.

Diese Ansicht wird auch durch weitere höchstrichterliche Rechtsprechung untermauert.[230] Ebenso würde es dem „Willen des Unionsgesetzgebers zuwider[laufen]“,[231] der bewusst dem Begriff der Datenbank eine weite Bedeutung verleihen wollte.[232]

Bei der Beurteilung der Unabhängigkeit von Daten ist darauf abzustellen, ob sich ein Dritter für ein herausgelöstes Element interessiert, da dieses dem Dritten eine sachdienliche Information liefert.[233] Sollte dies der Fall sein, wäre eine Unabhängigkeit bei entsprechenden Daten erfüllt.
Test- und Trainingsdaten können zwar unterschiedlich ausgestaltet werden, jedoch kann bei Daten, die sich durch einen Funktionsbereich und einen Zielwert auszeichnen, Unabhängigkeit angenommen werden. Das gilt sowohl für Test- und Trainingsdaten, die durch einen Menschen einzeln bestimmt wurden, als auch für maschinengenerierte Daten. Bei maschinengenerierten Datensammlungen ist jedoch zu beachten, dass diese nicht das Kriterium einer persönlichen geistigen Schöpfung erfüllen.

Test- und Trainingsdaten können somit regelmäßig das Kriterium von unabhängigen Elementen erfüllen.

Der Gesetzgeber verlangt noch weitere Voraussetzungen, um einen Schutz als Datenbankwerk zu erhalten. Zum einen wird verlangt, dass die enthaltenden Elemente systematisch oder methodisch angeordnet sind und zum anderen müssen die Elemente einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sein.

Systematisch ist eine Anordnung, „wenn sie sich an einem System, einer Klassifizierung oder einem Ordnungsschema orientiert“.[234] Methodisch ist eine Anordnung, „wenn sie einer bestimmten ordnenden Handlungsanweisung oder einem bestimmten Plan folgt.“[235]

Ob bestimmte Test- und Trainingsdaten systematisch oder methodisch angeordnet sind, ist vom jeweiligen Einzelfall abhängig. Aufgrund der bereits bestehenden inneren Ordnung von Test- und Trainingsdaten sollte das Kriterium der systematischen und methodischen Anordnung regelmäßig erfüllt sein.

Des Weiteren müssen Datenbankwerke einzeln oder mit Hilfe elektronsicher Mittel oder auf andere Weise zugänglich sein. Maßgeblich ist dabei die Möglichkeit des Zugangs und nicht die konkrete Form des Zugangs. Folglich kommen sowohl elektronische als auch nicht-elektronische Datenbanken in Frage.[236]
Dabei liegt eine Einzelzugänglichkeit vor, sofern einzelne Elemente abgerufen werden können, ohne die nachträgliche Hinzufügung eines Abfragesystems.[237]

---

[230] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187; vgl. OLG München, Urt. v. 9.5.2019 – 29 U 1048/18, GRUR-RR 2020; Österreichischer Oberster Gerichtshof, Beschluss vom 10. 7. 2001 – 4 Ob 155/01, GRUR Int. 2002, 452; BGH, Urt. v. 6.5.1999 – I ZR 199/96, MMR 1999, 470.
[231] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187, Rn. 26.
[232] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187, Rn. 26.
[233] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187, Rn. 28, 29.
[234] vgl. Dreier, in: Dreier/Schulze, UrhG, § 4, Rn. 17.
[235] vgl. Dreier, in: Dreier/Schulze, UrhG, § 4, Rn. 17.
[236] vgl. Marquardt, in: Wandtke/Bullinger, UrhG, § 4 UrhG, Rn. 10.
[237] vgl. Wiebe, in: Spindler/Schuster, UrhG, § 4, Rn. 12.

Für Test- und Trainingsdaten, die grundsätzlich elektronisch vorliegen, ist diese Voraussetzung in der Regel unproblematisch; aufgrund ihrer inneren Struktur sind sie für Abfragesysteme einzeln abrufbar.

Im Ergebnis können Datensammlungen, die für den Test- und Trainingsprozess verwendet werden, den Schutz als Datenbankwerk erwerben. Entscheidend ist dabei insbesondere das Vorliegen einer persönlichen geistigen Schöpfung hinsichtlich der Auswahl und Anordnung der Daten. Inwiefern dies im konkreten Fall vorliegt, ist zu prüfen. Test- und Trainingsdaten, die durch technische Erwägungen, Regeln oder Zwänge geprägt sind, scheiden aus dem Schutzbereich eines Datenbankwerks aus.

In einem Federated Learning Prozess sind standardisierte Test- und Trainingsdaten aufgrund der Vergleichbarkeit der Modelle nicht unüblich. Standardisierte Test- und Trainingsdaten sprechen jedoch gegen eine persönliche geistige Schöpfung und gegen einen Schutz als Datenbankwerk.

(3) Datenbank – § 87a UrhG

Der Datenbankschutz in § 87a UrhG schützt nicht eine persönliche geistige Schöpfung, sondern eine Investition in die Beschaffung, Überprüfung und Darstellung einer Datenbank. Die Schutzposition eines Datenbankwerks und einer Datenbank teilen einige Voraussetzungen. Beide Schutzrechte fordern eine Sammlung von unabhängigen Elementen, die systematisch oder methodisch angeordnet sind und einzeln mit Hilfe von elektronischen oder anderen Mitteln zugänglich und abrufbar sind.
Die gemeinsamen Tatbestandsmerkmale wurden bereits im Rahmen der Prüfung als Datenbankwerk geprüft. Daher soll nachfolgend nur auf den zentralen Unterschied eingegangen werden.

Der zentrale Unterschied ist das Vorliegen einer wesentlichen Investition in die Beschaffung, Überprüfung und Darstellung der Datenbank. Eine Investition ist wesentlich, wenn eine Leistung nach objektiver Betrachtung nicht von jedermann leicht zu erbringen ist.[238] Als Beispiel für eine unwesentliche Investition gelten Beschaffungsleistungen in öffentlich leicht zugängliche Daten.[239]
Welche Form eine Investitionsleistung hat, ist dabei grundsätzlich zweitrangig. So kommt sowohl der Einsatz menschlicher, finanzieller als auch technischer Ressourcen – quantifizierbar wie auch qualifizierbar – in Frage.[240]
Entscheidend dagegen ist, dass die Investition in Leistungen für die Beschaffung, Überprüfung oder Darstellung der Elemente fließt.

Wie in Kapitel IV. festgestellt, können wesentliche Investitionen in die Beschaffung, Überprüfung oder Darstellung von Daten als berücksichtigungsfähige Investitionsleistung angesehen werden, Investitionen in die Generierung neuer Daten bleiben hingegen außer Betracht.[241] Wann von einer nicht berücksichtigungsfähigen Investition in die

[238] vgl. BGH, Urt. v. 1. 12. 2010 – I ZR 196/08, GRUR 2011, 724, Rn. 23; vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 54; Schack, Haimo: „Urheberrechtliche Gestaltung von Webseiten unter Einsatz von Links und Frames“ in MMR, 2001, Heft 1, Seite 12.
[239] vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 15.
[240] vgl. EuGH, Urt. v. 9.11.2004 – C-338/02, GRUR 2005, 252, Rn. 28.
[241] vgl. Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 341-

Datengenerierung auszugehen ist, lässt sich der bisherigen Rechtsprechung allerdings nicht klar entnehmen. In diesem Zusammenhang ist zu beachten, dass sogenannte Sole-Source Datenbanken vom Schutz als Datenbank ausgeschlossen sind[242], wann diese vorliegen, ist weder in der Literatur noch durch einschlägige Rechtsprechung eindeutig geklärt.

Aufgrund der Vielfältigkeit von Test- und Trainingsdaten und der nicht eindeutigen Rechtslage ist eine Einzelfallprüfung erforderlich. Dennoch lassen sich gewisse Ergebnisse für bestimmte Test- und Trainingsdaten festhalten.

So sind beispielsweise Kosten, Zeit- und Arbeitsaufwände in das Sammeln, Überprüfen und Darstellen radiologischer Bilder, die für das Training eines ML-Modells bestimmt sind, regelmäßig eine zu berücksichtigende Beschaffungsleistung, weil für Dritte unter Einsatz vergleichsweiser Aufwendungen eine Beschaffung möglich wäre.

Im Falle maschinengenerierter Test- und Trainingsdaten ist außerdem zu fragen, ob Dritte mit ähnlichen Aufwendungen in der Lage wären, ähnliche Daten zu erzeugen.[243] Sollte dies möglich sein, liegt nach der hier vertretenen Meinung eine berücksichtigungsfähige Investitionsleistung vor. Sofern diese Daten jedoch durch niemanden oder nur mit unverhältnismäßigen Aufwendungen beschafft werden können, scheidet der Schutz als Datenbank, aufgrund der Problematik einer sogenannten Sole-Source Datenbank, aus.[244]

Bei der Beurteilung von Überprüfungsleistungen sind solche Mittel als Investitionsleistung zu berücksichtigen, die der Zuverlässigkeit, der Richtigkeitsprüfung, der Feststellung von vorhandenen Elementen oder dem Korrigieren von unzutreffenden oder veralteten Daten dienen.[245] Zu beachten ist, dass Überprüfungsleistungen, die zeitgleich oder dem Stadium einer Datenerzeugung zugehörig sind, sofern sie als nicht berücksichtigungsfähige Beschaffungsleistung gelten.[246]

Prüft oder labelt beispielsweise ein Data Scientist Test- und Trainingsdaten, so ist die Trennung zwischen Überprüfung und Beschaffung möglich und folglich auch berücksichtigungsfähig. Verarbeitet und prüft hingegen ein KI-System Test- und Trainingsdaten, ist zu beachten, dass diese Leistung nicht unmittelbar im Stadium der Datenerzeugung stattfinden darf, um als berücksichtigungsfähig zu gelten. Sollte jedoch die

342; Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 276.

[242] vgl. EuGH, Urt. v. 9.11.2004 – C-338/02, GRUR 2005, 252, Rn. 31.

[243] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 271-272; Leistner, Matthias: "Datenbankschutz - Abgrenzung zwischen Datensammlung und Datengenerierung" in: Computer und Recht, Heft 1, 2018, Seite 202.

[244] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 271; vgl. Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 341.

[245] vgl. Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 47, 48; Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 37.

[246] vgl. EuGH, Urt. v. 9. 11. 2004 – C-203/02, GRUR 2005, 244, 247; Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 37; vgl. Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 44; Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 277.

Leistung getrennt ausgewiesen sein[247], so können beispielsweise finanzielle Mittel für den Kauf oder die Erstellung eines KI-Systems anerkannt sein.

Ebenso kommt eine Investitionsleistung in Form einer Darstellungsleistung in Frage. Dies sind Aufwände, welche einer Datenbank ihre Funktionen verleihen, also ihre systematische oder methodische Anordnung oder auch die Möglichkeit der Zugänglichkeit der enthaltenen Elemente.[248] Als berücksichtigungsfähige Investitionsleistungen kommen beispielsweise Abfragesysteme, Indizes,[249] Nutzungsrechte an Software[250] oder auch Kosten für die technische Infrastruktur inklusive Pflege und Wartung in Betracht.[251]

Zusammenfassend können Test- und Trainingsdaten den Schutz einer Datenbank im Sinne des § 87a UrhG grundsätzlich nach Maßgabe der vorangestellten Differenzierungen erwerben.

b) Geschäftsgeheimnisgesetz

Damit für Test- und Trainingsdaten der Schutz als Geschäftsgeheimnis gilt, müssen Test- und Trainingsdaten die Legaldefinition eines Geschäftsgeheimnisses nach § 2 Nr. 1 GeschGehG erfüllen.

Ein Geschäftsgeheimnis ist eine Information, die weder insgesamt noch in der genauen Anordnung und Zusammensetzung ihrer Bestandteile geheim ist, einen wirtschaftlichen Wert hat und an der ein berechtigtes Geheimhaltungsinteresse besteht. Ebenso fordert der Gesetzgeber, dass den Umständen nach angemessene Geheimhaltungsmaßnahmen durch den Inhaber getroffen werden.

Test- und Trainingsdaten stellen unzweifelhaft eine Information dar. Allerdings ist zu diskutieren, ob Test- und Trainingsdaten als geheim anzusehen sind. Dafür müssten diese entweder insgesamt oder in ihrer genauen Anordnung und Zusammensetzung einer entsprechenden Fachgruppe nicht allgemein bekannt oder ohne Weiteres zugänglich sein. Bei der Bestimmung der Fachgruppe müssen aufgrund der informationsspezifischen Bestimmung[252] Datenwissenschaftler und Informatiker mit entsprechendem Hintergrund herangezogen werden, die sich regelmäßig mit Test- und Trainingsdaten in einem ML-Prozess auseinandersetzen.

Das alle Test- und Trainingsdaten in Summe geheim sind, wird nur in Ausnahmefällen gegeben sein. Da Test- und Trainingsdaten in der Regel für ein bestimmtes ML-Modell aufbereitet werden, können diese, sofern sie nicht bereits anderweitig veröffentlich wurden, geheim sein. Wahrscheinlicher wird jedoch die genaue Anordnung und Zusammensetzung von Test- und Trainingsdaten in solchen Fällen geheim sein.
Öffentlich zugängliche Test- und Trainingsdatensammlungen können kein Geschäftsgeheimnis darstellen. Auch Test- und Trainingsdaten, welche ohne Weiteres zugänglich sind i.S.v. §2 Nr.1 lit. a GeschGehG, können nicht Gegenstand eines

[247] vgl. Sattler in: Sassenberg/Faber, Rechtshandbuch I 4.0, S.47, 2. Aufl. 2020.
[248] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 38.
[249] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 38.
[250] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 39.
[251] vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 13.
[252] Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG § 2 Rn. 33.

Geschäftsgeheimnisses sein. Wie bereits in Kapitel IV.A Ziffer 2 ausgeführt, dürfte es bei der Beurteilung, ob eine Information als zugänglich gilt oder nicht, insbesondere darauf ankommen, wie viel Zeit und Kosten ein Dritter aufwenden müsste, um den gleichen Kenntnisstand über die betreffende Information zu erlangen.[253]

Ebenso könnte der Schutz als Geschäftsgeheimnis zu einem späteren Zeitpunkt erlöschen, sofern Dritte Zugang zu diesen erhalten. Vorstellbar ist der Zugang mittels der Möglichkeit des sogenannten Reveres Engineering. Es konnten bereits Test- und Trainingsdaten aus dezentral trainierten ML-Modellen rekonstruiert werden.[254] Es ist also nicht unwahrscheinlich, dass allgemein Test- und Trainingsdaten durch Dritte selbst rekonstruiert werden, was zur Folge hätte, dass der Schutz als Geschäftsgeheimnis verloren geht.

Unter Ausschluss dieser Fälle können Test- und Trainingsdaten die Voraussetzungen einer geheimen Information im Sinne des GeschGehG erfüllen.
Test- und Trainingsdaten bilden des Weiteren einen zentralen Bestandteil eines ML-Prozesses, sodass ein wirtschaftlicher Wert anzunehmen ist.

Der Ersteller von Test- und Trainingsdaten dürfte in der Regel auch ein berechtigtes Interesse an der Geheimhaltung haben. Inwiefern angemessene Geheimhaltungsmaßnahmen durch den Geschäftsgeheimnisinhaber getroffen wurden, ist vom jeweiligen Einzelfall abhängig.
Im Ergebnis können Test- und Trainingsdaten den Schutz als Geschäftsgeheimnis im Sinne des GeschGehG erwerben.

Der Geschäftsgeheimnisinhaber erhält keine Ausschließlichkeitsrechte an den Test- und Trainingsdaten, sondern lediglich die im GeschGehG normierten Abwehrrechte. Mit diesen Abwehrrechten kann der Inhaber die unerlaubte Erlangung, Nutzung und Offenlegung seiner geschützten Test- und Trainingsdaten unterbinden.

Zusammenfassung – Test- und Trainingsdaten:

Die Beurteilung der möglichen Schutzrechte an Test- und Trainingsdaten ist aufgrund ihrer Vielfalt nicht allgemeingültig zu beantworten.
Unter gewissen Voraussetzungen können Sammlungen von Test- und Trainingsdaten als Darstellung wissenschaftlicher und technischer Art (§ 2 Abs. 1 Nr. 7 UrhG), als Datenbankwerk (§ 4 Abs. 2 UrhG) oder als Datenbank (§ 87a UrhG) geschützt sein.
Test- und Trainingsdaten können eine geheime Information von wirtschaftlichem Wert darstellen. Sollte der Inhaber zusätzlich ein berechtigtes Interesse haben und angemessene Geheimhaltungsmaßnahmen treffen, können Test- und Trainingsdaten im Ergebnis auch durch das GeschGehG geschützt sein.

[253] Hauck, in: Heermann/Schlingloff, MüKoUWG, GeschGehG § 2 Rn. 9.
[254] vgl. Zhu, L., et. al.: "Deep leakage from gradients", in: arXiv online, 2019, URL: https://arxiv.org/pdf/1906.08935.pdf, Abruf 14.02.2022, Seite 1-5.

2. Trainierte Gewichtungsdaten

Gewichtungsdaten sind ein zentraler Bestandteil neuronaler Netze. Sie machen es möglich, dass neuronale Netze „lernen“ können, und sie stellen das Produkt eines aufwendigen Trainingsprozesses dar. Ein neuronales Netz mit derselben Netzarchitektur kann durch die Übertragung trainierter Gewichtungsdaten unmittelbar eingesetzt werden.

Sofern ein neuronales Netz verlässliche Ergebnisse erzielt, verkörpern trainierte Gewichtungsdaten zum einen die richtige Erstellung eines neuronalen Netzes und zum anderen bilden sie – wirtschaftlich betrachtet – das Produkt einer Investition. Schließlich werden trainierte Gewichtungsdaten unter dem Einsatz einer Vielzahl an Test- und Trainingsdaten erzeugt. Trainierte Gewichtungsdaten stellen somit ein Immaterialgut innerhalb von neuronalen Netzen dar, an dessen Schutz der Inhaber regelmäßig ein hohes Interesse haben dürfte.

Sofern trainierte Gewichtungsdaten durch gesetzlich Rechte geschützt werden, stellt sich in einem Federated Learning Prozess ebenso die Frage, wem Rechte an den trainierten Gewichtungsdaten zugeordnet werden.

Da trainierte Gewichtungsdaten entweder im Programmcode enthalten sind oder in Form einer separaten Datei gespeichert werden und bei einer entsprechenden Verarbeitung durch das jeweilig ausführende Programm geladen werden, soll nachfolgend der Schutz als Computerprogram (§ 69a UrhG), Datenbankwerk (§ 4 Abs. 2 UrhG) und als Datenbank (§ 87a UrhG) untersucht werden. Ebenso soll der mögliche Schutz als Geschäftsgeheimnis geprüft werden.

a) Computerprogramm – § 69a UrhG

Der Urheberrechtsschutz nach § 69a UrhG setzt das Vorliegen eines Computerprogramms und einer persönlichen geistigen Schöpfung voraus. Um jedoch die gesetzliche Schutzfähigkeit als Computerprogramm für trainierte Gewichtungsdaten eindeutig aufzeigen zu können, bedarf es der Darstellung zweier Verständnisse eines geschützten Computerprogramms.

Der europäische und deutsche Gesetzgeber haben sich bei der Normierung des Computerprogrammschutzes bewusst gegen eine gesetzliche Definition entschieden.[255]

In Anlehnung an die Definition der WIPO definiert der BGH ein Computerprogramm in seiner Rechtsprechung als „eine Folge von Befehlen […], die nach Aufnahme in einen maschinenlesbaren Träger fähig sind zu bewirken, dass eine Maschine mit informationsverarbeitenden Fähigkeiten eine bestimmte Funktion oder Aufgabe oder ein

[255] vgl. EG-Kommission: „Vorschlag für eine Richtlinie des Rates über den Rechtsschutz von Computerprogrammen“, KOM (88) 89/C 91/95 vom 12.04.1989, Seite 6; Bundestag: Entwurf eines Zweiten Gesetzes zur Änderung des Urheberrechtsgesetzes, in: Bundestag online, Drucksache 12/4022, 1992, URL: https://dserver.bundestag.de/btd/12/040/1204022.pdf, Abruf 14.02.2022, Seite 9.

bestimmtes Ergebnis anzeigt, ausführt oder erzielt".[256] Die Schutzfähigkeit eines Computerprogramms erfordert mithin eine eindeutige Steuerungsfunktion.[257]

Dagegen stellt das weite Verständnis eines Computerprogramms nicht auf eine Steuerungsfunktion ab, sondern vielmehr auf einen steuernden Einfluss einer bestimmten Information auf ein Computerprogramm.[258] Ein steuernder Einfluss liegt danach vor, wenn einer Information eine funktionale Bedeutung für das Verhalten eines Computers zukommt.[259]

Diese beiden Verständnisse wirken – je nach Anwendung – stark begrenzend oder erweiternd auf den Anwendungsbereich des Computerprogrammschutzes. Für den Schutz der Gewichtungsdaten als Computerprogramm im Sinne des § 69a UrhG ist dies entscheidend.

Unter der Anwendung des engen Verständnisses scheiden trainierte Gewichtungsdaten aus dem Schutzbereich aus. Gewichtungsdaten sind zwar für die Datenverarbeitung durch ein neuronales Netz entscheidend, sie stellen losgelöst jedoch nur eine Sammlung von einzelnen Werten bzw. Daten dar. Diese sind nicht ohne Weiteres fähig, eine Maschine beispielsweise zur Ausführung einer bestimmten Funktion oder Aufgabe zu bringen.[260] Eine eindeutige Steuerungsfunktion liegt somit bei trainierten Gewichtungsdaten nicht vor.
Sollten Gewichtungsdaten jedoch fest in den Programmcode integriert sein, wie beispielsweise in Form von Maschinencode, stellen Gewichtungsdaten einen Bestandteil einer Folge von Befehlen dar. Danach wäre das neuronale Netz samt Gewichtungsdaten vom Schutz als Computerprogramm erfasst.[261] Jedoch bleiben die einzelnen Gewichtungsdaten innerhalb des Programmcodes gemeinfrei, da sie keine Folge von Steuerungsbefehlen im engeren Sinne darstellen.[262] Zu beachten ist dennoch, dass die Vervielfältigung oder Änderung der trainierten Gewichtungsdaten nur gestattet sein könnte, sofern „der Vorgang der Extrahierung [der

[256] vgl. BGH, Urt. v. 9.5.1985 – I ZR 52/83, NJW 1986, 192, 196; dazu auch OLG Hamburg, Urt. v. 12.3.1998 – 3 U 226/97, MMR 1999, 230, 230.
[257] vgl. Söbbing, Thomas: "Algorithmen und urheberrechtlicher Schutz" in: Computer und Recht, Heft 4, 2020, Seite 227.
[258] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 767; Nebel, Jens; Stiemerling, Oliver: "Aktuelle Programmiertechniken und ihr Schutz durch § 69a UrhG" in: Computer und Recht, Heft 1, 2016, Seite 65; Hartmann, Frank; Prinz, Matthias: "Immaterialgüterrechtlicher Schutz von Systemen Künstlicher Intelligenz" in: Wettbewerb in Recht und Praxis, Heft 12, 2018, Seite 1438.
[259] vgl. Nebel, Jens; Stiemerling, Oliver: "Aktuelle Programmiertechniken und ihr Schutz durch § 69a UrhG" in: Computer und Recht, Heft 1, 2016, Seite 63-64.
[260] vgl. Ehinger, Patrick: "Urheberrechtlicher Schutz von neuronalen Netzen und Erzeugnissen von K.I.-Software" in: Kommunikation & Recht, Heft 7, 2019, Seite 13; vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, H12, 2018, Seite 766-767.
[261] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 767; Hartmann, Matthias; Ohst, Claudia: „Künstliche Intelligenz im Immaterialgüterrecht" in: KI & Recht kompakt, 2020, Seite 338; hinsichtlich Parametereinstellungen im Programmcode differenziert: OLG Hamburg, Urt. v. 13.04.2012 – 5 U 11/11.
[262] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 767; vgl. auch: vgl. OLG Hamburg, Urt. v. 12. 3. 1998 – 3 U 226–97, NJW-RR 1999, 483, 484; vgl. Grützmacher, in: Wandtke/Bullinger, UrhG, § 69a UrhG, Rn. 17.

Gewichtungsdaten aus dem Programmcode] nicht mit einer zustimmungsbedürftigen Umarbeitung i.S.d. § 69c Nr. 2 UrhG verbunden“ ist.[263]

Ebenso scheitern trainierte Gewichtungsdaten an dem Tatbestandsmerkmal einer persönlichen geistigen Schöpfung. Zwar kann das Auswählen von Trainingsdaten[264] oder die Bestimmung der Anzahl an Neuronen pro Schicht menschliche Handlungen mit einer persönlichen geistigen Schöpfung darstellen.

Bei der Beurteilung, ob eine persönliche geistige Schöpfung vorliegt, ist allerdings auf den eigentlichen Trainingsprozess und nicht auf Handlungen vor dem eigentlichen Trainingsprozess abzustellen. Der eigentliche Trainingsprozess und somit die Anpassung der jeweiligen Gewichtungsdaten werden bei ML-Modelle voll automatisiert ohne die Einwirkung eines Menschen vollzogen.[265] Gegen das Vorliegen einer persönlichen geistigen Schöpfung bei Gewichtungsdaten spricht vor allem die Tatsache, dass weder der Entwickler noch derjenige, der die Test- und Trainingsdaten auswählt, ein zielführendes Anpassen der einzelnen Gewichte bewirken kann. Daraus folgt auch der Begriff der „Blackbox“. Der Entwickler kann nicht oder nur teilweise nachvollziehen, wie ein neuronales Netz die einzelnen Gewichte im Wege des Trainings anpasst.[266]

Trainierte Gewichtungsdaten als solche stellen somit kein geschütztes Computerprogramm im Sinne des § 69a UrhG dar.

b) Datenbankwerk – § 4 Abs. 2 UrhG

Ausgehend von der fehlenden Schutzfähigkeit als Computerprogramm, sollen trainierte Gewichtungsdaten auf den möglichen Schutz als Datenbankwerk gemäß § 4 Abs. 2 UrhG untersucht werden. Diese Untersuchung liegt nahe, zumal trainierte Gewichtungsdaten regelmäßig als separate Datei gespeichert werden.

Damit ein Datenbankwerk vorliegt, bedarf es eines Sammelwerks, dessen Elemente systematisch oder methodisch angeordnet sind und einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sind. Die Voraussetzung einer Sammlung ist niedrig anzusetzen, sie ist bereits bei einer Zusammenstellung mehrerer Werke, Daten oder anderer unabhängiger Elemente erfüllt.[267]

---

[263] Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 767; s. zur Abgrenzung zwischen Programmablauf und Programmsubstanz: LG Hamburg, Urt. v. 14.01.2022 – 308 O 130/19.

[264] vgl. Hartmann, Frank; Prinz, Matthias: "Immaterialgüterrechtlicher Schutz von Systemen Künstlicher Intelligenz" in: Wettbewerb in Recht und Praxis, Heft 12, 2018, Seite 1437.

[265] vgl. hierzu auch: Stechern, David: "Künstliche Intelligenz - Auf der Suche nach dem Zuordnungsobjekt" in: IP-Rechtsberater, Heft 1, 2020, Seite 24-25.

[266] vgl. Körner, Sven: „Nachvollziehbarkeit von KI-basierten Entscheidungen“ in: Braegelmann, Tom; Kaulartz, Markus (Hrsg.): Rechtshandbuch Artificial Intelligence und Machine Learning, 2020, Seite 126-127; Apel, Simon; Kaulartz, Markus: "Rechtlicher Schutz von Machine Learning-Modellen" in: Recht Digital, Heft 1, 2020, Seite 26; Hartmann, Frank; Prinz, Matthias: "Immaterialgüterrechtlicher Schutz von Systemen Künstlicher Intelligenz" in: Wettbewerb in Recht und Praxis, Heft 12, 2018, Seite 1435.

[267] vgl. Marquardt, in: Wandtke/Bullinger, UrhG, § 4 UrhG, Rn. 4.

Ein Werk liegt nur vor, sofern eine persönliche geistige Schöpfung hinsichtlich der Auswahl und Anordnung der Elemente innerhalb der Sammlung vorhanden ist.[268] Eine persönliche geistige Schöpfung bedingt, dass ein Werkstück das Ergebnis eines menschlichen Urhebers und dessen Gedankeninhalt verkörpert.[269]

Die Auswahl und Anordnung der Gewichtungsdaten stellen jedoch nicht das Ergebnis eines menschlichen Urhebers und dessen Gedankeninhalt dar. Vielmehr werden Gewichtungsdaten durch die Lern- und Optimierungsfunktion des neuronalen Netzes angepasst. Folglich findet keine Auswahl oder Anordnung durch einen Urheber statt.

Trainierte Gewichtungsdaten scheiden aus dem Datenbankwerkschutz gemäß § 4 Abs. 2 UrhG aus, da sie keine persönliche geistige Schöpfung darstellen.

c) Datenbank – § 87a UrhG

Für den Schutz als Datenbank gemäß § 87a UrhG bedarf es keiner persönlichen geistigen Schöpfung.

Inwiefern Gewichtungsdaten die weiteren Voraussetzungen als Datenbank im Sinne des § 87a UrhG erfüllen, soll nachfolgend untersucht werden.

Wie bereits bei der Analyse als Datenbankwerk festgestellt, können trainierte Gewichtungsdaten eine Sammlung von Elementen darstellen. Der Gesetzgeber verlangt jedoch, dass die in der Datenbank enthaltenden Elemente auch unabhängig sind.
Nach der Rechtsprechung des EuGH gelten Elemente als unabhängig, sofern sich ein Dritter für ein herausgelöstes Elemente interessiert, da das Element dem jeweiligen Dritten eine sachdienliche Information liefert.[270]
Unabhängig, ob man ein einzelnes Gewichtungsdatum pro Neuron oder mehrere Gewichtungsdaten pro Neuronenschicht als Element klassifiziert, können diese ein einzelnes Element darstellen. Datenkombinationen können einzelne Elemente bilden.[271] Ein Dritter wird ohne weitere Informationen über ein einzelnes ihm vorliegendes Gewichtungsdatum, oder auch mehrere vorliegende Gewichtungsdaten pro Schicht, aus diesen Gewichtungsdaten an sich keine für ihn sachdienliche Information ableiten können.

Gewichtungsdaten liefern erst durch ihre Einbindung in eine entsprechende Netzarchitektur und Verknüpfung mit weiteren Neuronen in der Netzarchitektur eine sachdienliche Information.[272]

---

[268] vgl. EuGH, Urt. v. 1. 3. 2012 – C-604/10, GRUR 2012, 386, Rn. 33, 42.
[269] vgl. Schulze, in: Dreier/Schulze, UrhG, § 2, Rn. 8, 12.
[270] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187, Rn. 28, 29.
[271] vgl. OLG München, Urt. v. 9.5.2019 – 29 U 1048/18, GRUR-RR 2020, 1; Österreichischer Oberster Gerichtshof, Beschluss vom 10. 7. 2001 – 4 Ob 155/01, GRUR Int. 2002, 452; BGH, Urt. v. 6.5.1999 – I ZR 199/96, MMR 1999, 470.
[272] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 768-769; Ehinger, Patrick: "Urheberrechtlicher Schutz von neuronalen Netzen und Erzeugnissen von K.I.-Software" in: Kommunikation & Recht, Heft 7, 2019, Seite 13-14; Apel, Simon; Kaulartz, Markus: "Rechtlicher Schutz von Machine Learning-Modellen" in: Recht Digital, Heft 1, 2020, Seite 29; Hartmann, Frank; Prinz, Matthias: "Immaterialgüterrechtlicher Schutz von Systemen Künstlicher Intelligenz" in: Wettbewerb in Recht und Praxis, Heft 12, 2018, Seite 1437; Hetmank,

Einzelnen Gewichtungsdaten fehlt es somit an der erforderlichen Unabhängigkeit; ohne diese scheiden Sammlungen von trainierten Gewichtungsdaten aus dem Datenbankschutz aus.

d) Geschäftsgeheimnisgesetz

Nachdem trainierte Gewichtungsdaten regelmäßig aus dem urheberrechtlichen Schutzbereich ausscheiden, stellt sich weiterhin die Frage, ob das GeschGehG einen alternativen Schutz bieten könnte.

Damit eine Information als Geschäftsgeheimnis anzusehen ist, muss diese geheim sein und einen wirtschaftlichen Wert besitzen. Ebenso muss der Inhaber ein Geheimhaltungsinteresse haben und angemessene Geheimhaltungsmaßnahmen treffen. Ob trainierte Gewichtungsdaten diese Voraussetzungen erfüllen, wird nachfolgend geprüft.

Beginnend mit dem Geheimsein von trainierten Gewichtungsdaten, ist auf den jeweiligen Personenkreis, der üblicherweise mit Gewichtungsdaten und neuronalen Netzen umgeht, abzustellen. Sofern Gewichtungsdaten mit einer eigenentwickelten Netzarchitektur und eigens aufbereiteten Test- und Trainingsdaten erstellt wurden, werden diese grundsätzlich nicht in ihrer genauen Anordnung einem entsprechenden Personenkreis bekannt sein oder ohne Weiteres zugänglich sein. Ein Personenkreis bräuchte für die Erstellung ähnlicher Gewichtungsdaten die zugrundliegende Netzarchitektur und die Trainingsdaten.

Bei der Verwendung öffentlich zugänglicher Netzarchitekturen und Trainingsdaten wäre somit eine Erstellung ähnlicher Trainingsdaten möglich. Folglich wären Trainingsdaten nicht als geheim anzusehen und ein Schutz als Geschäftsgeheimnis ausgeschlossen.

Unter der Prämisse, dass weder die Netzarchitektur noch die verwendeten Trainingsdaten jedermann öffentlich zugänglich sind, können trainierte Gewichtungsdaten als geheime Information angesehen werden. Diese sind weder dem entsprechenden Personenkreis allgemein bekannt noch ohne Weiteres zugänglich.

Aufgrund dessen können trainierte Gewichtungsdaten auch einen wirtschaftlichen Wert besitzen. Trainierte Gewichtungsdaten wurden zwar für eine besonderes neuronales Netz erstellt und können nicht ohne Weiteres auf in anderen Netzarchitekturen verwendet werden, jedoch verkörpern sie das Ergebnis eines aufwendigen Trainingsprozesses. Die Erkenntnisse aus trainierten Gewichtungsdaten lassen sich somit sehr wohl auch auf fremde neuronale Netze anwenden.

Die Hürde eines wirtschaftlichen Werts wird zudem niedrig angesetzt, sodass nur belanglose Informationen aus dem Schutzbereich des GeschGehG ausscheiden.[273] Bereits ein kommerzielles Potential genügt zur Bejahung eines wirtschaftlichen Wertes.[274] Trainierte

Sven; Lauber-Rönsberg, Anne: "Künstliche Intelligenz – Herausforderungen für das Immaterialgüterrecht" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 6, 2018, Seite 579.

[273] vgl. Erwägungsgrund 14 der TS-RL; vgl. Schur, Nico: Die Lizenzierung von Daten, 2020, Seite 68.

[274] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 443; vgl. auch Hiéramente, in: BeckOK, GeschGehG, § 2, Rn. 16.

Gewichtungsdaten können unstreitig ein kommerzielles Potential aufweisen (s. etwa Kapitel III. B zur Anwendung im Rahmen des Predictive Maintenance Service) und somit einen wirtschaftlichen Wert besitzen.

Ein berechtigtes Interesse des Inhabers an der Geheimhaltung kann angenommen werden. Die Angemessenheit der Geheimhaltungsmaßnahmen ist vom konkreten Einzelfall abhängig und soll nachfolgend angenommen werden. Trainierte Gewichtungsdaten können folglich ein Geschäftsgeheimnis verkörpern. Der Inhaber kann somit den Schutz des GeschGehG in Anspruch nehmen. [275]

Das GeschGehG klärt indes nicht die Zuordnung von Geschäftsgeheimnissen bei Beteiligung mehrerer Parteien (bspw. im Rahmen des Federated Learning). § 2 Nr. 2 GeschGehG definiert den Geschäftsgeheimnisinhaber lediglich als „jede natürliche oder juristische Person, die die rechtmäßige Kontrolle über ein Geschäftsgeheimnis hat“. Wer rechtmäßige Kontrolle hat, ist dabei nicht eindeutig geklärt und bedarf weiterer Rechtsfortbildung durch Rechtsprechung.[276] Besonders im Falle von Mehrpersonenkonstellationen kann nach derzeitigem Stand lediglich festgehalten werden, dass Personen, welche durch eine verbotene Handlung gem. § 4 GeschGehG die Kontrolle über ein Geschäftsgeheimnis erlangt haben, von einer solchen Zuordnung ausgeschlossen sind.
Den Parteien in einem Federated Learning Prozess ist daher zu empfehlen, die Stellung als Geschäftsgeheimnisinhaber vertraglich zu regeln.[277]

Bei fehlender vertraglicher Klarstellung kommt eine Gemeinschaft nach Bruchteilen gemäß §§ 741 ff. BGB in Frage.[278] Danach wären alle beteiligten Parteien als Geschäftsgeheimnisinhaber anzusehen. Diese gesetzliche Zuordnung birgt jedoch das Risiko einer fehlerhaften Verwertung der trainierten Gewichtungsdaten und dem Verlust des Geschäftsgeheimnisschutzes. Sollte ein rechtmäßiger Inhaber beispielsweise nicht angemessene Geheimhaltungsmaßnahmen treffen, würde der Geschäftsgeheimnisschutz für alle beteiligten Inhaber erlöschen.

Zusammenfassung – Trainierte Gewichtungsdaten:

Trainierte Gewichtungsdaten sind nicht als Computerprogramm (§ 69a UrhG) oder Datenbankwerk (§ 4 Abs. 2 UrhG) geschützt, da sie keine persönliche geistige Schöpfung verkörpern. Darüber hinaus erfüllen trainierte Gewichtungsdaten nicht das von der Rechtsprechung angewendete enge Verständnis eines Computerprogramms, da ihnen die erforderliche Steuerungsfunktion fehlt.
Ein *sui generis* Schutz als Datenbank (§ 87a UrhG) scheidet aufgrund fehlender Unabhängigkeit der einzelnen Neuronen oder der Neuronenschichten aus.
Trainierte Gewichtungsdaten können jedoch die Voraussetzungen eines Geschäftsgeheimnisses im Sinne des GeschGehG erfüllen. Bei Mehrpersonenkonstellationen

[275] Zum selben Ergebnis gelangen auch: Sassenberg, Thomas; Kuß, Christian: „Künstliche Intelligenz und Machine Learning“ in: Sassenberg, Thomas; Faber, Tobias (Hrsg.): Rechtshandbuch Industrie 4.0 und Internet of Things, 2020, Seite 449.
[276] vgl. Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 445; vgl. Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG, § 2, Rn. 104-108.
[277] vgl. Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG, § 2, Rn. 107.
[278] vgl. Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG, § 2, Rn. 107.

wie dem Federated Learning Prozess besteht jedoch vertragsrechtlicher Handlungsbedarf hinsichtlich der Bestimmung des rechtmäßigen Inhabers der trainierten Gewichtungsdaten.

C. Schutz von fusionierten neuronalen Netzen

Durch den Federated Learning Prozess werden mehrere neuronale Netze dezentral trainiert. Diese dezentral trainierten Netze enthalten unterschiedlich trainierte Gewichtungsdaten.
Um den Trainingserfolg zwischen den neuronalen Netzen zu verbinden, werden die trainierten Gewichtungsdaten der teilnehmenden Partner durch die zentrale Stelle fusioniert, wobei dieser eine Vielzahl an trainierten Gewichtungsdaten zufließt.

Fraglich ist, ob diese Vielzahl an trainierten Gewichtungsdaten eine schutzfähige Sammlung im Sinne des Datenbankschutzes darstellt. Zum anderen soll untersucht werden, ob fusionierte Gewichtungsdaten gesetzlich geschützt sind.

1. Sammlungen von trainierten Gewichtungsdaten

Im Folgenden wird zunächst die Schutzfähigkeit von trainierten Gewichtungsdaten analysiert.

a) Datenbankwerk – § 4 Abs. 2 UrhG

Eine Datenbank ist gemäß § 4 Abs. 2 UrhG ein Sammelwerk. Es bedarf demnach einer Sammlung und einer persönlichen geistigen Schöpfung im Sinne eines Werkes (§ 2 Abs. 2 UrhG).
Solange der sogenannte Sammlungscharakter durch eine Menge der enthaltenen Elemente erkennbar ist, liegt eine Sammlung im Sinne des § 4 Abs. 2 UrhG vor.[279] Es ist damit regelmäßig anzunehmen, dass eine Sammlung von trainierten Gewichtungsdaten die Voraussetzungen einer Sammlung im Sinne des UrhG erfüllen wird.

Ob eine Sammlung von trainierten Gewichtungsdaten eine persönliche geistige Schöpfung darstellt ist allerdings fraglich.

Bei der Beurteilung, ob ein Datenbankwerk vorliegt, muss sich die persönliche geistige Schöpfung auf die Auswahl und Anordnung der Elemente innerhalb der Sammlung richten.[280] Es ist grundsätzlich anzunehmen, dass ein Mensch die trainierten Gewichtungsdaten in Form einer Sammlung anlegt.

Dennoch ist danach zu fragen, ob es sich bei einer Sammlung von trainierten Gewichtungsdaten um das Ergebnis eines menschlichen Urhebers und dessen Gedankeninhalt handelt,[281] oder ob nicht technische Erwägungen, Zwänge oder Regeln die Sammlung prägen und infolgedessen auch keine persönliche geistige Schöpfung vorliegt.[282]

---

[279] vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 4; Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 3, 4; Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 11; Marquardt, in: Wandtke/Bullinger, UrhG, § 4 UrhG, Rn. 4.
[280] vgl. EuGH, Urt. v. 1. 3. 2012 – C-604/10, GRUR 2012, 386, Rn. 33, 42.
[281] vgl. Schulze, in: Dreier/Schulze, UrhG, § 2, Rn. 8, 12.
[282] vgl. EuGH, Urt. v. 1. 3. 2012 - C-604/10, GRUR 2012, 386, Rn. 33, 42.

Hierbei wird man regelmäßig zu dem Ergebnis kommen, dass eine Sammlung von trainierten Gewichtungsdaten keinen Gestaltungsspielraum für persönliche geistige Schöpfungen offen lässt, sondern vielmehr technische Erwägungen und Zwänge die Handlungen eines Erstellers sehr einschränken. Zumal bei derartigen Sammlungen nicht die kreative Auswahl und Anordnung im Vordergrund stehen wird, sondern vielmehr die einheitliche Darstellung zur möglichst effektiven Auswertung.

Ein Schutz als Datenbankwerk im Sinne des § 4 Abs. 2 UrhG scheidet daher aus.

b) Datenbank – § 87a UrhG

Ohne Vorliegen einer persönlichen geistigen Schöpfung kommt jedoch der Datenbankschutz gemäß § 87a UrhG in Frage. Danach wird anstelle einer persönlichen geistigen Schöpfung eine Investitionsleistung in die Beschaffung, Überprüfung und Darstellung einer Datenbank geschützt.

Um eine Datenbank darzustellen, bedarf es einer Sammlung von unabhängigen Elementen. Eine Sammlung konnte bereits bei der Prüfung eines Datenbankwerkschutzes festgestellt werden.

Zu prüfen ist, ob die Elemente auch die erforderliche Unabhängigkeit aufweisen. Ein Element ist unabhängig, wenn es von anderen Elementen in der Datenbank getrennt werden kann, ohne dass dieses Element seinen selbstständigen Informationswert verliert bzw. dieser beeinträchtigt wird.[283] Ob sich das Herauslösen eines Elements auf dessen Informationswert auswirkt, hängt davon ab, ob ein Dritter sich für ein herausgelöstes Element interessiert, weil dieses ihm sachdienliche Informationen liefert.[284]
Um die Unabhängigkeit eines Elements zu bestimmen, ist zu fragen, was ein Element bildet und von anderen Elementen abgrenzt.
Anders als bei der Beurteilung der Unabhängigkeit bei trainierten Gewichtungsdaten wird bei einer Sammlung mehrerer trainierter neuronaler Netze nicht auf ein einzelnes Neuron abgestellt. Vielmehr bildet ein Versionsstand eines trainierten neuronalen Netzes inklusive Metadaten eines Clients ein Element innerhalb der Sammlung. Diese Zusammenfassung oder auch Kombination mehrerer Daten ist nach der hier vertretenen Meinung zulässig.[285]

Dieses Element könnte ein Dritter nunmehr auswerten und für eigene Zwecke nutzen. Beispielsweise könnte der Forschungsbereich, welcher sich mit dem Thema der „*Explainable AI*“ auseinandersetzt, die jeweiligen Elemente einzeln untersuchen und für ihre Zwecke auswerten. Folglich liefern diese Elemente einem Dritten sachdienliche Informationen.

Die einzelnen Elemente innerhalb dieser Sammlung erfüllen somit das Kriterium der Unabhängigkeit.

Nachdem eine Sammlung aus unabhängigen Elementen bejaht werden konnte, stellt sich die Frage nach der berücksichtigungsfähigen Investitionsleistung des Datenbankherstellers.

---

[283] vgl. Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 12; vgl. Dreier, in: Dreier/Schulze, UrhG, § 87a, Rn. 6; vgl. Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 21.
[284] vgl. EuGH, Urt. v. 9. 11. 2004 – C-444/02, GRUR 2005, 254 Rn. 34.
[285] vgl. EuGH, Urt. v. 29.10.2015 – C-490/14, GRUR 2015, 1187, Rn. 20.

Als Investitionsleistung kommen Aufwendungen in Frage, welche für die Beschaffung, Überprüfung oder Darstellung der Elemente einer Datenbank erforderlich sind.

Bei der Erstellung einer Sammlung von trainierten Gewichtungsdaten wird voraussichtlich die Leistung in die Beschaffung der Elemente innerhalb der Datenbank die größte Relevanz haben. Ebenso eröffnen sich bei der Beurteilung einer Beschaffungsleistung einige Fragestellungen.

Die Rechtsprechung unterscheidet die Beschaffungsleistung in das Sammeln bereits vorhandener Elemente und in das Erfinden neuer Elemente.[286] Nach der Rechtsprechung des EuGH ist das Erfinden neuer Elemente keine berücksichtigungsfähige Beschaffungsleistung.[287]

Daher ist zu prüfen, ob das Übermitteln trainierter Gewichtungsdaten von den Clients an die zentrale Stelle als berücksichtigungsfähige Beschaffungsleistung anzusehen ist. Für eine berücksichtigungsfähige Sammelleistung spricht, dass die zentrale Stelle nicht die trainierten Gewichtungsdaten selbst erzeugt, sondern von den Clients erhält. Somit findet die eigentliche Generierung bei den Clients statt.
Gegen diese Ansicht spricht die Tatsache, dass die zentrale Stelle das Computerprogramm vorgibt, mit welchem die Clients die trainierten Gewichtungsdaten erzeugt. Somit könnte lediglich eine Verschiebung der eigentlichen Generierung vorliegen, die faktisch nur einer Person, der zentralen Stelle, zugutekommt.

Bei der Beurteilung ist des Weiteren zu berücksichtigen, dass sogenannte Sole-Source Datenbanken für einen Datenbankschutz gemäß § 87a UrhG nicht in Frage kommen.[288]
Bei Sole-Source Datenbanken kennt lediglich eine Person die enthaltenen Daten, da diese nicht frei verfügbar oder durch eigene Bemühungen selbst zusammengestellt werden können.[289] Es besteht somit eine Monopolgefahr durch die mögliche Gewährung eines Datenbankschutzes.[290] Der EuGH hat deshalb solche Datenbanken als nicht schutzfähig eingestuft.[291]

Aufgrund der vorangegangenen Problematik wird man bei Sammlungen von trainierten Gewichtungsdaten wahrscheinlich zum Ergebnis kommen, dass der Inhalt dieser Sammlungen weder frei verfügbar ist, noch durch eigene Bemühungen selbst zusammengestellt werden kann. Hierfür wäre vielmehr gerade die genaue Kenntnis über die

---

[286] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 271; vgl. Leistner, Matthias: "Datenbankschutz - Abgrenzung zwischen Datensammlung und Datengenerierung" in: Computer und Recht, Heft 1, 2018, Seite 20.
[287] vgl. EuGH, Urt. v. 9.11.2004 – C-338/02, GRUR 2005, 252, Rn. 31.
[288] vgl. EuGH, Urt. v. 9.11.2004 – C-338/02, GRUR 2005, 252, Rn. 31; Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 271; vgl. Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 341.
[289] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 271.
[290] vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 271.
[291] vgl. EuGH, Urt. v. 9.11.2004 – C-338/02, GRUR 2005, 252, Rn. 31.

genaue Netzarchitektur, die verwendeten Hyperparameter und die verwendeten Trainingsdaten erforderlich. Da dies in diesem Fall jedoch höchstwahrscheinlich nicht möglich ist, wird man den Datenbankschutz verneinen müssen.[292] Auch andere Leistungen in die Überprüfung oder Darstellung des Datenbankinhalts werden, aufgrund der Monopolproblematik, nichts an der fehlenden Schutzfähigkeit ändern können.[293]

Eine Sammlung von trainierten Gewichtungsdaten dürfte mithin regelmäßig nicht als Datenbank gemäß § 87a UrhG geschützt sein.

c) Geschäftsgeheimnisgesetz

Da trainierte Gewichtungsdaten als solche regelmäßig die Voraussetzungen eines Geschäftsgeheimnisses in § 2 Nr. 1 GeschGehG erfüllen, kann ebenso eine Sammlung dieser trainierten Gewichtungsdaten ein Geschäftsgeheimnis verkörpern.

Eine Sammlung von trainierten Gewichtungsdaten ist grundsätzlich nicht den jeweiligen Personenkreis allgemein bekannt oder ohne Weiteres zugänglich. Ebenso besitzt eine Sammlung von trainierten Gewichtungsdaten einen wirtschaftlichen Wert. Zudem kann grundsätzlich davon ausgegangen werden, dass der Inhaber ein berechtigtes Interesse an der Geheimhaltung hat. Der Inhaber der Sammlung muss außerdem angemessene Geheimhaltungsmaßnahmen treffen. Diese sind wiederrum vom jeweiligen Geschäftsgeheimnis und vom jeweiligen Inhaber abhängig.

Eine Sammlung von trainierten Gewichtungsdaten kann somit den Schutz als Geschäftsgeheimnis erwerben.

Zusammenfassung – Sammlungen von Gewichtungsdaten:

Sammlungen von Gewichtungsdaten können eine Wissensbasis darstellen.
Ungeachtet dessen wird man zu dem Ergebnis kommen, dass ein Schutz durch den Datenbankwerkschutz (§ 4 Abs. 2 UrhG) und den Datenbankschutz sui generis (§ 87a UrhG) ausscheidet.
Sammlungen von Gewichtungsdaten fehlt es an einer persönlichen geistigen Schöpfung hinsichtlich der Auswahl und Anordnung. Ebenso dürften sie in der Regel sogenannte Sole-Source Datenbanken darstellen, welche nach der Rechtsprechung des EuGH keinen Schutz als Datenbank im Sinne des §87a UrhG genießen.
Sammlungen von Gewichtungsdaten können jedoch als Geschäftsgeheimnis im Sinne des GeschGehG geschützt sein.

[292] vgl. EuGH, Urt. v. 9.11.2004 – C-338/02, GRUR 2005, 252, Rn. 31; vgl. Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 271; vgl. Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 341.

[293] vgl. EuGH, Urt. v. 9. 11. 2004 – C-203/02, GRUR 2005, 244, 247; Hermes, in: Wandtke/Bullinger, UrhG, § 87a UrhG, Rn. 37; vgl. Vohwinkel, in: BeckOK, UrhG, § 87a UrhG, Rn. 44; Sagstetter, Thomas: "Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, 2019, Seite 277.

2. Fusionierte Gewichtungsdaten

Nachdem die zentrale Stelle eine Vielzahl an trainierten Gewichtungsdaten gesammelt hat, werden diese zentral fusioniert. Dabei entscheidet eine Person, welche trainierten Gewichtungsdaten auf welche Art und Weise fusioniert werden sollen. In diesem Schritt finden regelmäßig Fusionsalgorithmen Anwendung.
Ob und inwieweit sich der gesetzliche Schutz fusionierter Gewichtungsdaten von trainierten Gewichtungsdaten unterscheidet wird im Folgenden geprüft.

a) Computerprogramm – § 69a UrhG

Bei der Beurteilung, ob ein geschütztes Computerprogramm im Sinne des § 69a UrhG hinsichtlich fusionierter Gewichtungsdaten vorliegt, gelten die Ausführungen und Ergebnisse zu trainierten Gewichtungsdaten in Kapitel V. B. 2. entsprechend. Im Folgenden soll dennoch kurz auf die Besonderheiten fusionierter Gewichtungsdaten eingegangen werden.

Fusionierte Gewichtungsdaten weisen zwar eine funktionale Bedeutung für das Verhalten eines neuronalen Netzes und daher auch für ein Computerprogramm auf, allerdings besitzen sie keine Steuerungsfunktion.[294]

Auch das Vorliegen einer persönlichen geistigen Schöpfung wird man bei fusionierten Gewichtungsdaten verneinen müssen, da Fusionsalgorithmen im Fusionsschritt zum Einsatz kommen. Diese Fusionsalgorithmen nutzen bekannte gängige Matrixoperationen, wie beispielweise die Durchschnittsbildung, um trainierte Gewichtungsdaten zu kombinieren. Somit basiert der Fusionsschritt nicht auf der Gedankenwelt eines möglichen Urhebers, sondern vielmehr auf übliche Gestaltungsmittel und informatische und mathematische Erwägungen. Diese vorgegebenen Algorithmen prägen den Fusionsschritt wesentlich, sodass nicht von einer persönlichen geistigen Schöpfung auszugehen ist.[295]

Davon getrennt zu beachten ist der Fusionsalgorithmus als solches. Die konkrete Ausdrucksform eines solchen Fusionsalgorithmus könnte den Schutz als Computerprogramm im Sinne des § 69a UrhG erlangen (vgl. Ausführungen in Kapitel V. A.2).

Im Ergebnis scheitern auch fusionierte Gewichtungsdaten daran, den Schutz als Computerprogramm zu erwerben.

b) Datenbankwerk – § 4 Abs. 2 UrhG

Ein Datenbankwerk ist ein Sammelwerk. Ein Sammelwerk ist laut der Legaldefinition in § 4 Abs. 1 UrhG eine Sammlung von Werken, Daten oder anderen unabhängigen Elementen, die aufgrund ihrer Auswahl oder Anordnung der Elemente eine persönliche geistige Schöpfung darstellen. Um als Datenbankwerk urheberrechtlichen Schutz zu genießen, müssten

[294] vgl. Ehinger, Patrick: "Urheberrechtlicher Schutz von neuronalen Netzen und Erzeugnissen von K.I.-Software" in: Kommunikation & Recht, Heft 7, 2019, Seite 13; vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 766-767.
[295] vgl. Schulze in Dreier/ Schulze, UrhG, § 2, Rn. 33.

fusionierte Gewichtungsdaten eine persönliche geistige Schöpfung hinsichtlich der Auswahl oder Anordnung unabhängiger Elemente darstellen.

Im Gegensatz zu trainierten Gewichtungsdaten, sind fusionierte Gewichtungsdaten nicht das Produkt eines vollautomatisierten Computerprogramms, sondern vielmehr das Ergebnis menschlichen Handels. Eine Person wählt trainierte Gewichtungsdaten, die fusioniert werden sollen, aus. Mithin ist zu prüfen, ob die spätere Sammlung von fusionierten Gewichtungsdaten das Ergebnis einer persönlichen geistigen Schöpfung hinsichtlich der Auswahl oder Anordnung verkörpert.
Gegen eine persönliche geistige Schöpfung spricht eindeutig, dass im Fusionsschritt die handelnde Person regelmäßig einen Fusionsalgorithmus nutzt, um die jeweiligen Gewichtungsdaten zu fusionieren. Diese basieren auf üblichen informatischen und mathematischen Erwägungen. Ebenso hat die handelnde Person keinen großen Gestaltungsspielraum hinsichtlich der Auswahl der Fusionierung. Am Ende sollen schließlich fusionierte Gewichtungsdaten erstellt werden, die vielversprechende Ergebnisse erzeugen sollen.
Des Weiteren stoßen auch fusionierte Gewichtungsdaten auf das Problem der fehlenden Unabhängigkeit. Genau wie bei trainierte Gewichtungsdaten, bilden bei fusionierten Gewichtungsdaten einzelne Neuronen ein Element innerhalb der Sammlung. Eine sachdienliche Information lässt sich hier ohne weiteren Kontext jedoch nicht herausziehen. Erst durch die Einbindung in eine Netzarchitektur und durch Verknüpfung mit weiteren „benachbarten“ Neuronen lassen sich sachdienliche Information ableiten.[296]

Im Ergebnis stellen fusionierte Gewichtungsdaten kein Sammelwerk im Sinne des § 4 Abs. 1 UrhG dar. Zudem fehlt fusionierten Gewichtungsdaten das Merkmal der gesetzlich vorausgesetzten Unabhängigkeit.

Folglich scheidet ein Schutz als Datenbankwerk im Sinne des § 4 Abs. 2 UrhG aus.

c) Datenbank – § 87a UrhG

Der Datenbankschutz sui generis in § 87a UrhG setzt zwar keine persönliche geistige Schöpfung voraus, jedoch müssen die enthaltenen Elemente unabhängig sein. Wie durch Vorangegangenes bereits geprüft, erfüllen fusionierte Gewichtungsdaten die Voraussetzung eines unabhängigen Elements allerdings nicht. Des Weiteren besteht die Problematik einer Sole-Source Datenbank bei Sammlungen von fusionierten Gewichtungsdaten, schließlich ist der genaue Inhalt solcher Sammlungen nur für die fusionierende Person bekannt.

Fusionierte Gewichtungsdaten erfüllen somit nicht den Datenbankschutz in § 87a UrhG.

d) Geschäftsgeheimnisschutz

[296] vgl. Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 768-769; Ehinger, Patrick: "Urheberrechtlicher Schutz von neuronalen Netzen und Erzeugnissen von K.I.-Software" in: Kommunikation & Recht, Heft 7, 2019, Seite 13-14; Apel, Simon; Kaulartz, Markus: "Rechtlicher Schutz von Machine Learning-Modellen" in: Recht Digital, Heft 1, 2020, Seite 29; Hartmann, Frank; Prinz, Matthias: "Immaterialgüterrechtlicher Schutz von Systemen Künstlicher Intelligenz" in: Wettbewerb in Recht und Praxis, Heft 12, 2018, Seite 1437.

Abwehrrechte aus dem GeschGehG erfordern, dass die zu schützende Information ein Geschäftsgeheimnis im Sinne der Legaldefinition in § 2 Nr. 2 GeschGehG darstellt. Ein geschütztes Geschäftsgeheimnis liegt nur vor, wenn eine Information geheim ist und einen wirtschaftlichen Wert aufweist. Darüber hinaus muss der Inhaber ein berechtigtes Interesse an der Geheimhaltung besitzen und angemessene Geheimhaltungsmaßnahmen getroffen haben. Im Folgenden wird geprüft, ob fusionierte Gewichtungsdaten diese Voraussetzungen erfüllen.

Fusionierte Gewichtungsdaten müssen zunächst eine geheime Information darstellen. Was als geheim anzusehen ist, hängt vom entsprechenden Personenkreis ab. Der Personenkreis bestimmt sich dabei nach der konkreten Information. In diesem Fall wäre somit zu erörtern, ob die Information einem Personenkreis, der üblicherweise mit dieser Art von Information umgeht, allgemein bekannt oder ohne Weiteres zugänglich ist.[297] Hier wären die Informationen, auf die abgestellt wird, die fusionierten Gewichtungsdaten selbst und der entsprechende Personenkreis wären nunmehr Fachleute, die sich mit neuronalen Netzen, insbesondere Federated Learning, auseinandersetzen.

Dabei müssen die fusionierten Gewichtungsdaten insgesamt oder in ihrer genauen Anordnung und Zusammensetzung geheim sein. Unter der Annahme, dass die zugrundliegenden trainierten Gewichtungsdaten bereits geheim sind und niemand außer der zentralen Stelle Zugang zu diesen trainierten Gewichtungsdaten hat, wird man zum Ergebnis gelangen, dass fusionierte Gewichtungsdaten geheim sind.
Andernfalls werden zumindest die genaue Anordnung und Zusammensetzung der fusionierten Gewichtungsdaten regelmäßig geheim und nicht ohne Weiteres zugänglich sein.
Eine einfache Rekonstruktion wird für den entsprechenden relevanten Personenkreis, ohne Zugang zu zahlreichen weiteren Informationen, nicht ohne weiteres möglich sein.
Fusionierte Gewichtungsdaten können somit als geheim angesehen werden.

Damit fusionierte Gewichtungsdaten zudem einen wirtschaftlichen Wert besitzen, dürfen diese keine belanglosen Informationen darstellen.[298] Davon ist jedoch nicht auszugehen, zumal fusionierte Gewichtungsdaten das Ergebnis zahlreicher Investitionen darstellen. Fusionierte Gewichtungsdaten weisen somit einen wirtschaftlichen Wert im Sinne des GeschGehG auf.

Unter der Voraussetzung, dass der Inhaber sowohl ein berechtigtes Interesse an der Geheimhaltung besitzt, als auch angemessene Geheimhaltungsmaßnahmen trifft, fallen fusionierte Gewichtungsdaten unter den Geschäftsgeheimnisschutz.
Dabei wird die zentrale Stelle, welche die rechtmäßige Kontrolle hat, als einziger rechtmäßiger Inhaber der fusionierten Gewichtungsdaten betrachtet werden.

Zusammenfassung – fusionierte Gewichtungsdaten:

Fusionierte Gewichtungsdaten scheiden ähnlich wie trainierte Gewichtungsdaten von einem urheberrechtlichen Schutz aus.
Sie stellen kein geschütztes Computerprogramm (§ 69a UrhG), kein Datenbankwerk (§ 4 Abs. 2 UrhG) und keine Datenbank (§ 87a UrhG) dar.

[297] vgl. § 2 Nr. 1 a GeschGehG.
[298] vgl. Erwägungsgrund 14 der TS-RL; vgl. Schur, Nico: Die Lizenzierung von Daten, 2020, Seite 68.

Das Merkmal einer persönlichen geistigen Schöpfung wird aufgrund informatischer und mathematischer Erwägungen regelmäßig nicht erfüllt.
Zudem weisen einzelne Neuronen nicht die erforderliche Unabhängigkeit auf, welche zentral für die Erfüllung eines Datenbankschutzes ist.
Fusionierte Gewichtungsdaten können jedoch eine geheime Information mit wirtschaftlichem Wert und somit ein Geschäftsgeheimnis darstellen.

VI. Regelungsansätze zur Lösung entstehender Zuordnungsfragen

Bevor im Folgenden auf potenzielle Regelungsansätze für die Lösung von Zuordnungsfragen in Federated Learning Prozessen eingegangen wird, sei darauf hingewiesen, dass Federated Learning Prozesse unterschiedlich ausgestaltet sein können.
Die rechtliche Prüfung in den Kapiteln IV- und V. dieses Gutachtens bezog sich immer auf einen konkreten Federated Learning Prozess, bei welchem der zentralen Stelle eine übergeordnete Rolle zukommt.

Bei dieser Konstellation ergeben sich je nach Stadium des Federated Learning Prozesses Zuordnungsfragen. Im Stadium eines untrainierten neuronalen Netzes stellen sich keine Zuordnungsfragen hinsichtlich der Netzarchitektur, Lern- und Optimierungsfunktion und Hyperparameter. Diese Rechtsobjekte dürften im betrachteten Fall einer Predictive Maintenance Anwendung (vgl. Kapitel III. B.) der zentralen Stelle zuzuordnen sein, sei es als Urheberin oder als rechtmäßige Inhaberin im Sinne des GeschGehG.

Dagegen treten in der Trainingsphase eines neuronalen Netzes jedoch Zuordnungsfragen hinsichtlich der Nutzung von Test- und Trainingsdaten und der Stellung als rechtmäßiger Inhaber der trainierten Gewichtungsdaten auf. Im Stadium von fusionierten neuronalen Netzen beschränkt sich die Zuordnungsfrage bezüglich der Nutzung von fusionierten Gewichtungsdaten.

Nachfolgend werden mögliche vertragliche Regelungsansätze für Federated Learning Konstellationen skizziert. Zunächst soll die Konstellation aufgegriffen werden, in welcher die zentrale Stelle eine übergeordnete Position einnimmt. Daraufhin soll die Möglichkeit des Cross-Licensing beziehungsweise der Netzwerkverträge thematisiert werden. Abschließend werden mögliche Treuhandmodelle im Zusammenhang mit Federated Learning dargestellt.

A. Zentrale Stelle mit übergeordneter Position

Nimmt die zentrale Stelle im Federated Learning Prozess eine übergeordnete Position gegenüber den teilnehmenden Clients ein, bietet sich eine Regelung auf vertraglicher Basis an.

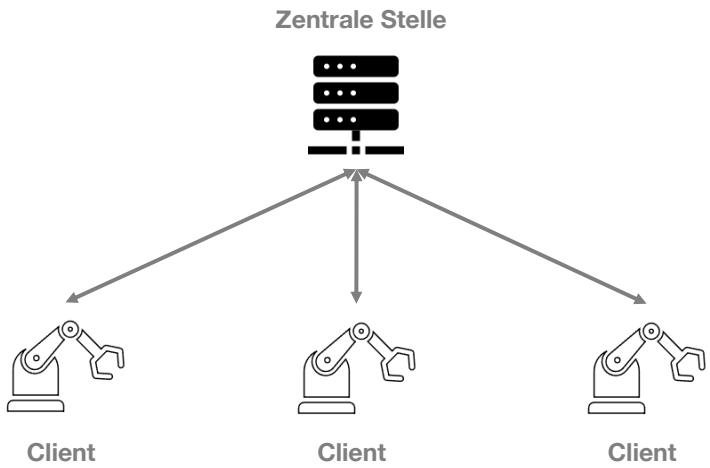

Abbildung 2: Federated Learning Konstellation (1/3)

Diese vertragliche Regelung wird zwischen der zentralen Stelle und den einzelnen Clients festgelegt. Die einzelnen Clients stehen nicht in einem vertraglichen Verhältnis zueinander.

Unter Verwendung des Maschinenhersteller-Beispiels (vgl. Kapitel III. B.), in welchem dieser die zentrale Stelle darstellt, ist insbesondere zu beachten, dass die Daten oder Datensammlungen der einzelnen Clients je nach Einzelfall geschützt sein können.[299]

Fraglich ist, ob es im jeweiligen Einzelfall überhaupt einer Einräumung von Nutzungsrechten seitens der Clients an die zentrale Stelle bedarf, da schließlich das ML-Modell auf den Maschinen der Clients ablaufen und die zentrale Stelle keinerlei direkten Zugriff auf die Daten der Clients erhält. Der gesamte Trainingsprozess läuft bei den jeweiligen Clients ab, wobei möglicherweise Geschäftsgeheimnisse der Clients durch den Trainingsprozess erlangt werden.
Für die zentrale Stelle und die Ausgestaltung der Vertragsbeziehung zu den Clients ist dies relevant, da das GeschGehG nicht nur den unbefugten Zugang zu Geschäftsgeheimnissen verbietet, sondern ebenso den unbefugten Zugang, die Aneignung oder das Kopieren von Dateien, aus welchen sich ein Geschäftsgeheimnis ableiten lässt.[300] Durch die Übertragung der trainierten Gewichtungsdaten lassen sich unter gewissen Bedingungen die zugrundliegenden Test- und Trainingsdaten rekonstruieren,[301] sodass es für die zentrale Stelle durchaus möglich ist, Geschäftsgeheimnisse aus trainierten Gewichtungsdaten abzuleiten.

Ebenso verbietet § 4 Abs. 1 Nr. 2 GeschGehG „jedes sonstige Verhalten, das unter den jeweiligen Umständen nicht dem Grundsatz von Treu und Glauben unter Berücksichtigung der anständigen Marktgepflogenheit entspricht.“ Die Wortwahl des Gesetzgebers lässt gewollt einen weiten Auslegungsspielraum, um Ausnahmefälle durch die Rechtsprechung berücksichtigen zu können.[302] Die Regelung ermöglicht eine „Gesamtbetrachtung aufgrund einer Abwägung der berechtigten Interessen unter Berücksichtigung der jeweiligen Umstände des Einzelfalls“ durchzuführen.[303] Demnach sind „allgemeine Erwartungen an die Redlichkeit und Fairness des Verhaltens von Unternehmern“ [304] einzubeziehen.

Unabhängig davon, ob im Einzelfall ein Urheber-, Leistungsschutzrecht oder ein Zugangsschutz aus dem Geschäftsgeheimnisrecht besteht, sollte für alle Beteiligten Klarheit hinsichtlich der Verwendung der Daten der Clients und hieraus abgeleiteter Ergebnisse bestehen.
Es ist daher ratsam, dass die zentrale Stelle transparent über die Verwendung der Daten des Clients aufklärt (beispielsweise in der Präambel oder sonstigen Vorbemerkungen des Regelwerks) und die Parteien dedizierte Datenzuweisungs- und Zugangsregelungen auf den wesentlichen Ergebnisebenen vorsehen.

---

[299] Siehe Kapitel V. B. 1. Test- und Trainingsdaten.
[300] vgl. § 4 Abs. 1 Nr. 1 GeschGehG.
[301] vgl. Zhu, L., et. al.: "Deep leakage from gradients", in: arXiv online, 2019, URL: https://arxiv.org/pdf/1906.08935.pdf, Abruf 14.02.2022, Seite 1-5.
[302] vgl. Bundestag: Entwurf eines Gesetzes zur Umsetzung der Richtlinie (EU) 2016/943 zum Schutz von Geschäftsgeheimnissen vor rechtswidrigem Erwerb sowie rechtswidriger Nutzung und Offenlegung, in: Bundestag online, Drucksache 19/4724, 2018, URL: https://dserver.bundestag.de/btd/19/047/1904724.pdf, Abruf 14.02.2022, Seite 27.
[303] Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG, § 4, Rn. 32.
[304] Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG, § 4, Rn. 32.

Das GeschGehG sieht – anders als das Urheberrecht – keine umfassenden Regelungen über die Arten oder den Inhalt eines Rechtsgeschäfts über Geschäftsgeheimnisse vor, sodass neben zweiseitigen oder mehrseitigen Verträgen auch einseitige Verträge, wie eine Einverständniserklärung, genügen können.[305] Dadurch wäre die Nutzung der Daten durch ein Rechtsgeschäft gemäß § 3 Abs. 2 GeschGehG rechtmäßig.

Des Weiteren sollte die zentrale Stelle hinsichtlich der Rechte an den trainierten Gewichtungsdaten darauf achten, dass diese ihr vertraglich zugeordnet werden. Schließlich verfolgt in einem derart einseitigen Trainingsprozess nur die zentrale Stelle das Trainieren eines ML-Modells. Ohne eine vertragliche Zuordnung besteht die Gefahr, dass die trainierten Gewichtungsdaten sowohl den jeweiligen Clients als auch der zentralen Stelle zugeordnet werden könnten.

Wie in Kapitel V. B.2. festgestellt, enthält das GeschGehG keine Angaben zur Lösung von Mehrpersonenkonstellationen. Daher werden bei fehlender vertraglicher Klarstellung die jeweiligen Rechte über eine Gemeinschaft nach Bruchteilen gemäß §§ 741 ff. BGB gesetzlich zugeordnet.[306] Danach wäre die zentrale Stelle und die jeweiligen Clients als rechtmäßiger Inhaber der jeweils trainierten Gewichtungsdaten anzusehen. Dadurch wäre beispielsweise eine wirtschaftliche Verwertung und die Aufrechterhaltung des Geschäftsgeheimnisschutzes nur schwer vorstellbar.
Es sollte daher vertraglich festgehalten werden, dass die zentrale Stelle als rechtmäßiger Inhaber im Sinne des § 2 Nr. 2 GeschGehG hinsichtlich der trainierten Gewichtungsdaten anzusehen ist.
Zudem sollte vereinbart werden, dass die jeweiligen Clients die trainierten Gewichtungsdaten im Rahmen des Federated Learning Prozesses nutzen dürfen, unter der Voraussetzung, dass die Clients, die von der zentralen Stelle festgelegten angemessenen Geheimhaltungsmaßnahmen, umsetzen.

Bei fusionierten Gewichtungsdaten wird in solchen Konstellationen die zentrale Stelle als rechtmäßiger Inhaber anerkannt. Dennoch ist auch hier die Überlassung der fusionierten Gewichtungsdaten nur unter Festlegung entsprechender Regelungen empfehlenswert. Ähnlich wie bei trainierten Gewichtungsdaten, sollte die zentrale Stelle darauf achten, dass die Clients entsprechende angemessene Geheimhaltungsmaßnahmen treffen.

B. Cross-Licensing/ Netzwerkverträge

In einem Federated Learning Prozess sind auch Konstellationen vorstellbar, in welchen alle beteiligten Clients gleichermaßen Rechte an den trainierten Modellen erhalten. Faktisch bilden dann alle teilnehmenden Clients gemeinsam die zentrale Stelle.

Eine Rechteeinräumung kommt daher nicht nur zwischen einer zentralen Stelle und den jeweiligen Clients in Frage, vielmehr räumen alle Clients sich gegenseitig Rechte ein, um die trainierten Modelle zu nutzen.

[305] vgl. Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG, § 3, Rn. 65-66.
[306] vgl. Alexander, in: Köhler/Bornkamm/Feddersen, GeschGehG, § 2, Rn. 107.

Diese Art der Rechteeinräumung ähnelt dem „cross-licensing“, welches insbesondere im Patentrecht genutzt wird, um gegenseitig benötigte Patentrechte zu erhalten.[307]

Ebenso ähneln solche Konstellationen sogenannten Netzwerkverträgen. Als Netzwerkverträge bezeichnet man Vertragskonstellationen, die sich durch eine Vielzahl von Vertragsparteien auszeichnet, in der jede Partei auf einer gleichrangigen vertraglichen Ebene steht.[308]

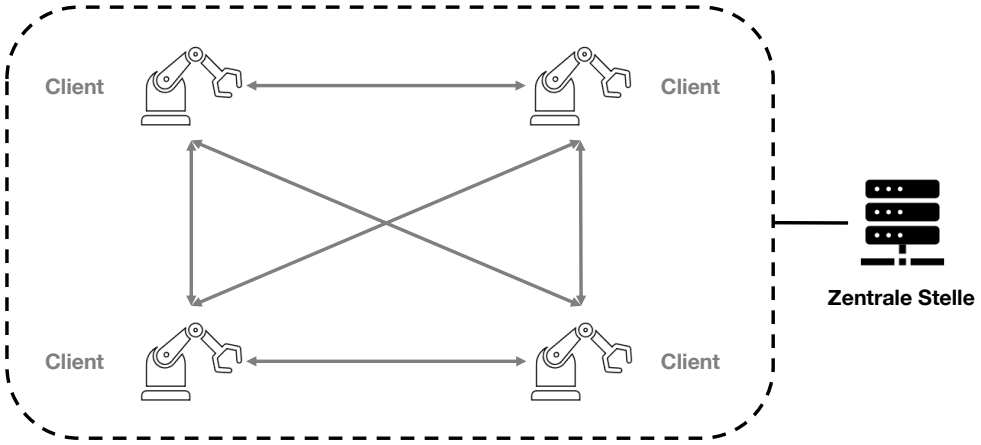

Abbildung 3: Federated Learning Konstellation (2/3)

C. Datentreuhand

Datentreuhandkonstellationen sind derzeitig immer wieder im Gespräch. Sei es im Rahmen der Gaia-X Initiative oder durch den geplanten Data Governance Act der Europäischen Kommission (DGA-E).[309]

Solche Datentreuhandmodelle kommen ebenfalls für Federated Learning Konstellationen in Frage. Treuhandkonstellationen sollen Fälle lösen, „in denen der technisch-faktische Dateninhaber nicht der (einzig) legitime Dateninhaber ist.“[310]
Durch die Einbindung eines neutralen Dritten lassen sich Zuordnungsfragen zu verschiedenen Rechtsgütern regeln und im Sinne der Parteien ausgestalten.

Nach einem Ansatz können Datentreuhandmodelle in einer Management-Funktion, einer freiwilligen Selbstbeschränkung und in der Lösung eines Datenzugangsproblems getrennt werden.[311]

---

[307] vgl. Grindley, Peter: „Cross-Licensing“ in: Augier. Mie; Teece, David (Hrsg.): The Palgrave Encyclopedia of Strategic Management, 2016, Seite 380.
[308] vgl. Börding, Andreas; Jülicher, Tim; Röttgen, Charlotte; v. Schönfeld, Max: „Neue Herausforderungen der Digitalisierung für das deutsche Zivilrecht“ in: Computer und Recht, Heft 2, 2017, Seite 136-138.
[309] Europäische Kommission: "Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates über europäische Daten-Governance (Daten-Governance-Gesetz)", COM (2020) 767 final vom 25.11.2020.
[310] vgl. Specht-Riemenschneider, Louisa; Blankertz, Aline; Sierek, Pascal; et. al.: “Die Datentreuhand“ in: MMR, Heft 6, 2021, Seite 30.
[311] vgl. Wendehorst, Christiane; Schwamberger, Sebastian; Grinzinger, Julia: „Datentreuhand – wie hilfreich sind sachenrechtliche Konzepte?“ in: Pertot, Tereza (Hrsg.): Rechte an Daten, 2020,

Datentreuhandmodelle mit einer Management-Funktion mitteln Daten im Interesse des Datenbereitstellers, wobei diese insbesondere die Zugangskonditionen im Interesse des Datenbereitstellers verhandeln. Daher zieht man auch regelmäßig eine Parallele zu Verwaltungsgesellschaften für Urheberrechte, wie beispielsweise die GEMA oder VG Wort.[312] Die wirtschaftliche Verwertung steht dabei im Fokus.

Das Datentreuhandmodell mit freiwilliger Selbstbeschränkung orientiert sich an dem Konstrukt eines Software-Escrow. Anstelle von Software soll der neutrale Datentreuhändler Daten hinterlegen und bei der Erfüllung gewisser Voraussetzungen oder dem Eintreten besonderer Umstände, wie beispielsweise Insolvenz oder Gesellschafterwechsel („Change of Control“) den Zugang zu diesen Daten ermöglichen.

Bei dem Datentreuhandmodell zur Lösung eines Datenzugangsproblems spricht man auch von einem sogenannten Data Trustee.[313] Dieser soll Dritten kontrollierten Zugang zu Daten gewähren und als eigenständige Partei mit allen Markteilnehmer Voraussetzungen für den Zugang verhandeln.[314] Ebenso soll der Data Trustee die Einhaltung dieser Voraussetzungen für teilnehmende Parteien kontrollieren.[315] Vorteil eines neutralen Data Trustees ist der Ausgleich von Machtasymmetrien zwischen den einzelnen Marktteilnehmer.[316]

Eine konkrete Möglichkeit wäre der Einsatz von Daten-Mittlern wie ihn derzeitig der Gesetzesentwurf der Europäischen Kommission über eine geplanten Data Governance Act vorsieht.[317]
Rechtssystematisch würde der DGA-E nach etwaiger Verabschiedung als EU-Verordnung unmittelbar in allen EU-Mitgliedsstaaten ohne Umsetzung in nationales Recht gelten.

Der DGA-E legt neben der Bereitstellung von Daten, die im Besitz öffentlicher Stellen sind, und der freiwilligen Eintragung von Einrichtungen, die Daten für altruistische Zwecke sammeln und verarbeiten, auch Regelungen für Datenmittler fest.[318] Konkret soll ein Anmelde- und Aufsichtsrahmen für die Erbringung von Diensten für die gemeinsame Datennutzung geregelt werden.[319] Der geplante DGA-E, würde somit auch den Rechtsrahmen für mögliche Datenmittler-Konstellationen bei Federated Learning Prozessen bilden.

Seite 112ff.; vgl. hierzu auch: Specht-Riemenschneider, Louisa; Blankertz, Aline; Sierek, Pascal; et. al.: “Die Datentreuhand“ in: MMR, Heft 6, 2021, Seite 25ff.

[312] vgl. Wendehorst, Christiane; Schwamberger, Sebastian; Grinzinger, Julia: „Datentreuhand – wie hilfreich sind sachenrechtliche Konzepte?“ in: Pertot, Tereza (Hrsg.): Rechte an Daten, 2020, Seite 108ff.

[313] vgl. Specht-Riemenschneider, Louisa; Blankertz, Aline; Sierek, Pascal; et. al.: “Die Datentreuhand“ in: MMR, Heft 6, 2021, Seite 28.

[314] vgl. Specht-Riemenschneider, Louisa; Blankertz, Aline; Sierek, Pascal; et. al.: “Die Datentreuhand“ in: MMR, Heft 6, 2021, Seite 28.

[315] vgl. Specht-Riemenschneider, Louisa; Blankertz, Aline; Sierek, Pascal; et. al.: “Die Datentreuhand“ in: MMR, Heft 6, 2021, Seite 28.

[316] vgl. Wendehorst, Christiane: „Of Elephants in the Room and Paper Tigers – How to Reconcile Data Protection and the Data Economy“ in: Lohsse, Sebastian; Schulze, Reiner; Staudenmayer, Dirk: Trading Data in the Digital Economy: Legal Concepts and Tools, 2017, Seite 352

[317] vgl. Europäische Kommission: "Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates über europäische Daten-Governance (Daten-Governance-Gesetz)", COM (2020) 767 final vom 25.11.2020.

[318] vgl. Art. 1 Abs. 1 des DGA-E.

[319] vgl. Art. 1 Abs. 1 lit. b DGA-E.

Datenmittler sind dabei grob neutrale Anbieter von Diensten für die gemeinsame Datennutzung.[320] Datenmittler soll dabei Dateninhaber und Datennutzer vermitteln,[321] ohne dabei die Daten für andere Zwecke zu verwenden.[322] Unter der Vermittlung ist die Herstellung geschäftlicher, rechtlicher oder technischer Beziehungen zwischen dem Dateninhaber und dem Datennutzer zu verstehen.[323]

Unternehmen, die als Datenmittler auftreten möchten, würden dabei einem Anmeldeverfahren bei einer nationalen Aufsichtsbehörde unterliegen. Nach erfolgreicher Anmeldung müssten diese Datenmittler Bedingungen für die Erbringung von Diensten für die gemeinsame Datennutzung einhalten. Die gesetzlichen Vorschriften schreiben dem Datenmittler konkret vor welche Bedingungen im Rahmen solch eines Dienstes eingehalten werden müssen. Darunter fällt beispielsweise, wie der Datenmittler mit den Daten des Dateninhabers oder den anfallenden Metadaten umzugehen hat.[324] Dabei erwähnenswert ist auch, dass die Verfahren der Datenmittler sowohl für den Dateninhaber als auch für den Datennutzer fair, transparent und nichtdiskriminierend gestaltet sein müssen.[325] Explizit wird dabei auch die Preisgestaltung des Datenzugangs genannt.[326]

Sollte ein Unternehmen im Rahmen eines Federated Learning Prozesses die Rolle eines Datenmittlers im Sinne des DGA-E einnehmen, müsste dieser, sofern der DGA-E verabschiedet werden sollte, die betreffenden Vorschriften einhalten. Die teilnehmenden Parteien innerhalb des Federated Learnings müssen hierfür vertragliche Regelungen für eine gemeinsame Datennutzung vereinbaren. Welche genauen Regelungen, die Parteien treffen, bleibt den jeweiligen Parteien offen. Sofern ein Datenmittler bei der gemeinsamen Datennutzung eingeschaltet werden sollte, müssten die vertraglichen Regelungen den Bedingungen des DGA-E entsprechen.

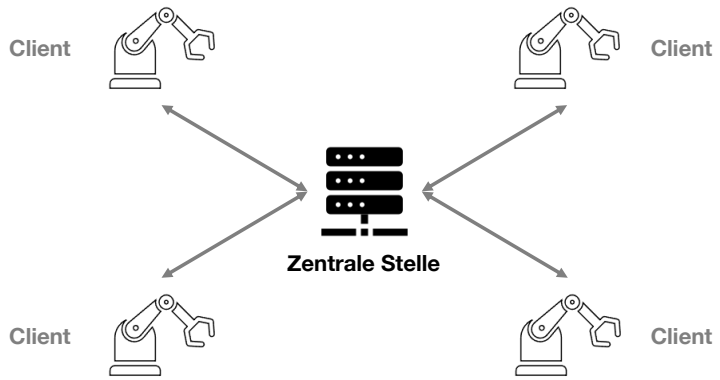

Abbildung 4: Federated Learning Konstellation (3/3)

Im Ergebnis kommen verschiedene Regelungsansätze für Federated Learning Prozesse in Frage. Von bilateralen über multilaterale Vertrags-Konstrukte können die Interessen der beteiligten Parteien, sei es als zentrale Stelle oder als Client, oder als beides, erfüllt werden.

[320] vgl. Art. 2 Nr. 7 DGA-E.
[321] vgl. Erwägungsgrund 22 des DGA-E.
[322] vgl. Erwägungsgrund 26 des DGA-E.
[323] vgl. Erwägungsgrund 22 des DGA-E.
[324] vgl. Art. 11 DGA-E.
[325] vgl. Art. 11 Nr. 3 DGA-E.
[326] vgl. Art. 11 Nr. 3 DGA-E.

Welcher Lösungsansatz Anwendung findet, müssen die jeweiligen Parteien entscheiden. Ausschlaggebend werden hierbei nicht-rechtliche Themen, wie Machtposition, technische Umsetzung und auch Kooperationsbereitschaft der jeweiligen Parteien sein.

Aus der rechtlichen Perspektive ist dabei zu beachten, dass man vor dem Beginn von Federated Learning Aktivitäten die Rechtsobjekte identifiziert, rechtlich auf deren gesetzlichen Schutzfähigkeit prüft und vertraglich im Interesse der Parteien zuordnet. Ansonsten besteht die Gefahr, dass mögliche Rechtsobjekte wie geplant nicht genutzt werden können oder der gesamte Federated Learning Prozess scheitert.

VII. Weiterführende Rechtsfragen

Im Rahmen der Untersuchung ergaben sich weitere Rechtsfragen, welche vom Prüfungsumfang dieses Gutachtens explizit ausgeschlossen sind:

- Wie werden andere dezentral trainierte KI-Systeme außerhalb von künstlichen neuronalen Netzen de lege lata geschützt?
- Werden dezentral trainierte KI-Systeme durch das Patentrecht de lege lata geschützt?
- Inwiefern bedarf es einer gesetzlichen Schutzfähigkeit von KI-Systemen de lege ferenda? Wie könnte eine gesetzliche Schutzfähigkeit von KI-Systemen ausgestaltet werden?
- Inwiefern werden Ergebnisse von KI-Systemen geschützt?
- Wie sind dezentral trainierte KI-Systeme datenschutzrechtlich einzuordnen und inwiefern können diese datenschutzkonform umgesetzt werden?
- Wie können dezentral trainierte KI-Systeme vertraglich geschützt werden?
- Wer haftet in Schadensfällen beim Einsatz von KI-Systemen?
- Inwiefern gestaltet sich die gesetzliche Gewährleistung bei KI-Systemen oder Datenüberlassungen?

Hinweis:
Kurz vor Redaktionsschluss dieses Gutachtens wurde ein inoffizieller (EU-) Kommissionsentwurf über eine Verordnung zur Gewährleistung einer gerechten Verteilung des Wertes in der Datenwirtschaft (Data Act) bekannt. Dieser inoffizielle Gesetzesentwurf sieht insbesondere Rechte und Pflichten von Beteiligten an Data Sharing Sachverhalten (beispielsweise Produkthersteller, Service-Anbieter, Nutzer, Datenhalter und Datenempfänger) vor, welche Auswirkungen auf Federated Learning Prozesse und deren Ergebnisse haben könnten. Die Untersuchung dieser Auswirkungen ist ausdrücklich vom Prüfungsumfang dieses Gutachtens ausgeschlossen.

Literaturverzeichnis

ACR Data Science Institute: "AI-LAB" in: dieselbe online, URL: https://ailab.acr.org, Abruf 14.02.2022

Ahlberg, Hartwig; Götting, Horst-Peter; Lauber-Rönsberg, Anne: Beck'sche Online-Kommentare zum Urheberrecht, 31. Auflage, C.H. Beck Verlag, München 2021 [*Bearbeiter, in: BeckOK, UrhG*]

Apel, Simon; Kaulartz, Markus: "Rechtlicher Schutz von Machine Learning-Modellen" in: Recht Digital, Heft 1, 2020, Seite 24-34

Apple: "Designing for privacy" in Apple WWDC online, URL: https://developer.apple.com/videos/play/wwdc2019/708, Abruf 14.02.2022

Arkenau, Judith; Wübbelmann, Judith: "Eigentum und Rechte an Daten – Wem gehören die Daten?" in: Taeger, Jürgen (Hrsg.): Internet der Dinge, Oldenburger Verlag, Edewecht 2015, Seite 95-110

Bartsch, Michael: "Software als Rechtsgut" in: Computer und Recht, Heft 9, 2010, Seite 553-559

Becker, Maximilian: „Schutzrechte an Maschinendaten und die Schnittstelle zum Personendatenschutz" in: Büscher, Wolfgang; et al. (Hrsg.): Festschrift für Karl-Heinz Fezer zum 70. Geburtstag, C.H. Beck Verlag, München 2016, Seite 815-834

Börding, Andreas; Jülicher, Tim; Röttgen, Charlotte; v. Schönfeld, Max: „Neue Herausforderungen der Digitalisierung für das deutsche Zivilrecht" in: Computer und Recht, Heft 2, 2017, Seite 134-140

Bundesregierung: Entwurf eines Gesetzes zur Umsetzung der Richtlinie über bestimmte vertragsrechtliche Aspekte der Bereitstellung digitaler Inhalte und digitaler Dienstleistungen, in: Bundesministerium der Justiz und für Verbraucherschutz online, vom 13.01.2021, URL: https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/RegE_Bereitstellung digitalerInhalte.pdf?__blob=publicationFile&v=3, Abruf 14.02.2022

Bundestag: Bericht der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale, in: Bundestag online, Drucksache 19/23700, 2020, URL: https://dserver.bundestag.de/btd/19/237/1923700.pdf, Abruf 14.02.2022

Bundestag: Entwurf eines Zweiten Gesetzes zur Änderung des Urheberrechtsgesetzes, in: Bundestag online, Drucksache 12/4022, 1992, URL: https://dserver.bundestag.de/btd/12/040/1204022.pdf, Abruf 14.02.2022

Döbel, Inga; et al.: "Maschinelles Lernen - Kompetenzen, Anwendungen und Forschungsbedarf" in: Fraunhofer-Allianz Big Data und Künstliche Intelligenz online, 2018, URL: https://www.bigdata-ai.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/BMBF_Fraunhofer_ML-Ergebnisbericht_Gesamt.pdf, Abruf 14.02.2022

Schuster, Fabian; Grützmacher, Malte: IT-Recht Kommentar, 1. Auflage, C.H. Beck Verlag, München 2020 [*Bearbeiter, in: Schuster/Grützmacher, Gesetz*]

Dreier, Thomas; Schulze, Gernot: Kommentar zum Urheberrechtsgesetz, 6. Auflage, C.H. Beck Verlag, München 2018 [*Bearbeiter, in: Dreier/Schulze, UrhG*]

EG-Kommission: „Vorschlag für eine Richtlinie des Rates über den Rechtsschutz von Computerprogrammen“, KOM (88) 89/C 91/95 vom 12.04.1989

Ehinger, Patrick; Stiemerling, Oliver: "Die urheberrechtliche Schutzfähigkeit von Künstlicher Intelligenz am Beispiel von Neuronalen Netzen" in: Computer und Recht, Heft 12, 2018, Seite 761-769

Ehinger, Patrick: "Urheberrechtlicher Schutz von neuronalen Netzen und Erzeugnissen von K.I.-Software" in: Kommunikation & Recht, Heft 7, 2019, Seite 12-14

EU-Kommission: „Evaluation of Directive 96/9/EC on the legal protection of databases“, SWD (2018) 146 final vom 25.4.2018

Europäische Kommission: "Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates über europäische Daten-Governance (Daten-Governance-Gesetz)", COM (2020) 767 final vom 25.11.2020

Europäische Kommission: "Weißbuch zur künstlichen Intelligenz", COM (2020) 65 final vom 19.2.2020

Fraunhofer-Institut für Software- und Systemtechnik: „Daten- und Dienste-Ökosystem für KI-orientierte Forschung und Entwicklung“ in: Fraunhofer-Institut für Software- und Systemtechnik online, 2021, URL: https://www.isst.fraunhofer.de/de/geschaeftsfelder/datenwirtschaft/projekte/Gaia-X-4-KI.html, Abruf 14.02.2022

Fuhlrott, Michael; Hiéramente, Mayeul: Beck’sche Online-Kommentare zum Geschäftsgeheimnisgesetz, 8. Auflage, C.H. Beck Verlag, München 2021 [*Bearbeiter*, in: BeckOK, GeschGehG]

Gaster, Jens: „Sui-generis Recht der Datenbankrichtlinie“ in: Hoeren, Thomas; Sieber, Ulrich; Holznagel, Bernd (Hrsg.): Handbuch Multimedia-Recht, 55. Auflage, C.H. Beck Verlag, München 2021, Teil 7.6

Google: “Your chats stay private while Messages improves suggestions” in Google support online URL: https://support.google.com/messages/answer/9327902, Abruf 14.02.2022

Grindley, Peter: „Cross-Licensing“ in: Augier. Mie; Teece, David (Hrsg.): The Palgrave Encyclopedia of Strategic Management, Palgrave Macmillan, London 2016

Hacker, Philipp: "Immaterialgüterrechtlicher Schutz von KI-Trainingsdaten" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 10, 2020, Seite 1025-1033

Hartmann, Frank; Prinz, Matthias: "Immaterialgüterrechtlicher Schutz von Systemen Künstlicher Intelligenz" in: Wettbewerb in Recht und Praxis, Heft 12, 2018, Seite 1431-1438

Hartmann, Matthias; Ohst, Claudia: „Künstliche Intelligenz im Immaterialgüterrecht“ in: Hartmann, Matthias: KI & Recht kompakt (ebook), Springer Verlag, Wiesbaden 2020, Seite 324-361

Heermann, Peter; Schlingloff, Jochen: Münchener Kommentar zum Lauterkeitsrecht. Band 2, 3. Auflage 2022, C.H. Beck Verlag, München 2022 [*Bearbeiter*, in: Heermann/Schlingloff, *Gesetz*]

Hetmank, Sven; Lauber-Rönsberg, Anne: "Künstliche Intelligenz – Herausforderungen für das Immaterialgüterrecht" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 6, 2018, Seite 574-582

Heydn, Truiken J.: "Identitätskrise eines Wirtschaftsguts: Software im Spannungsfeld zwischen Schuldrecht und Urheberrecht" in: Computer & Recht, Heft 12, 2010, Seite 765-776

Hoeren, Thomas: "Dateneigentum - Versuch einer Anwendung von § 303a StGB im Zivilrecht" in: MMR, Heft 8, 2013, Seite 486-491

Hofmann, Franz: "Absolute Rechte an Daten - immaterialgüterrechtliche Perspektive" in: Pertot, Tereza (Hrsg.): Rechte an Daten, Mohr Siebeck, Tübingen 2020, Seite 9-31

Jauernig, Othmar: Bürgerliches Gesetzbuch, Kommentar, 18. Auflage, C.H. Beck Verlag, München 2021 [*Bearbeiter*, in: Jauernig, *Gesetz*]

Joecks, Wolfgang; Miebach, Klaus: Münchener Kommentar zum StGB, 4. Auflage, C.H. Beck Verlag, München 2021 [*Bearbeiter*, in: MüKo, StGB]

Klinikum Universität Heidelberg: „Künstliche Intelligenz schlägt Hautärzte“ in: dieselbe online, URL: https://www.klinikum.uni-heidelberg.de/newsroom/kunstliche-intelligenz-schlagt-hautarzte/, Abruf 14.02.2022

Köhler, Helmut; Bornkamm, Joachim; Feddersen, Jörn: Beck´sche Kurzkommentare, Band 13a, Gesetz gegen unlauteren Wettbewerb, 39. Auflage, C.H. Beck Verlag, München 2021 [*Bearbeiter, in: Köhler/Bornkamm/Feddersen, Gesetz*]

Körner, Sven: „Nachvollziehbarkeit von KI-basierten Entscheidungen“ in: Braegelmann, Tom; Kaulartz, Markus (Hrsg.): Rechtshandbuch Artificial Intelligence und Machine Learning, 1. Auflage, C.H. Beck Verlag, München 2020, Seite 123-134

Kraus, Michael: „Datenlizenzverträge“ in: Jürgen Taeger (Hrsg.): Internet der Dinge, Oldenburger Verlag, Edewecht 2015, Seite 537-550

Krüger, Stefan; Wiencke, Julia; Koch, André: "Der Datenpool als Geschäftsgeheimnis" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 6, 2020, Seite 578.
Leistner, Matthias: "Datenbankschutz - Abgrenzung zwischen Datensammlung und Datengenerierung" in: Computer und Recht, Heft 1, 2018, Seite 17-22

Loewenheim, Ulrich; Leistner, Matthias; Ohly, Ansgar: Urheberrecht Kommentar, 6. Auflage, C.H. Beck Verlag, München 2020 [*Bearbeiter*, in: Schricker/ Loewenheim, UrhG]

Maron, Roman C.; et al.: "Robustness of convolutional neural networks in recognition of pigmented skin lesions" in: European Journal of Cancer, Ausgabe 145, 2021, Seite 81-91

Microsoft: “Komponente „Two-Class Neural Network“ (Neuronales Netz mit zwei Klassen)”, online URL: https://docs.microsoft.com/de-de/azure/machine-learning/component-reference/two-class-neural-network, Abruf 12.09.2022.

Nebel, Jens; Stiemerling, Oliver: "Aktuelle Programmiertechniken und ihr Schutz durch § 69a UrhG" in: Computer und Recht, Heft 1, 2016, Seite 61-69

Nguyen, Giang; et al.: "Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey" in: Artificial Intelligence Review, Ausgabe 52, Heft 1, 2019, Seite 77–124

Ohly, Ansgar: "Das neue Geschäftsgeheimnisgesetz im Überblick" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 5, 2019, Seite 441-451

Papstefanou, Steffan: "Genetic Breeding Algorithms als Form des "Machine Learning" im Urheber- und Patentrecht" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 4, 2019, Seite 209-215

Prenode: "Decentralized Machine Learning für Industriemaschinen" in: dieselbe online, URL: https://prenode.de/de/industriemaschinen/, Abruf 14.02.2022

Redeker, Helmut: IT-Recht, 7. Auflage, C.H. Beck Verlag, München 2020

Säcker, Franz Jürgen; Rixecker, Roland; Oetker, Hartmut: Münchener Kommentar zum Bürgerlichen Gesetzbuch, 8. Auflage, C.H. Beck Verlag, München 2020 [*Bearbeiter*, in: MüKo, BGB]

Sagstetter, Thomas: „Neue Regeln für Big Data & Co.?" in: Husemann, Tim; et al. (Hrsg.): Digitaler Strukturwandel und Privatrecht, Nomos Verlag, Baden-Baden 2019, Seite 249-292

Sassenberg, Thomas; Kuß, Christian: „Künstliche Intelligenz und Machine Learning“ in: Sassenberg, Thomas; Faber, Tobias (Hrsg.): Rechtshandbuch Industrie 4.0 und Internet of Things, 2. Auflage, C.H. Beck Verlag, München 2020, Seite 433-496

Schmidt, Kirsten Johanna; Zech Herbert: „Datenbankherstellerschutz für Rohdaten?", in: Computer und Recht, Heft 7, 2017, Seite 417-426

Schulze, Reiner, Kommentar zum Bürgerliches Gesetzbuch, 10. Auflage, Nomos Verlag, Baden-Baden 2019 [*Bearbeiter, in: BeckOK BGB*]

Schur, Nico: Die Lizenzierung von Daten, Mohr Siebeck, Tübingen 2020

Schuster, Fabian; Grützmacher, Malte: IT-Recht Kommentar, 1. Auflage, C.H. Beck Verlag, München 2020 [*Bearbeiter, in: Schuster/Grützmacher, Gesetz*]

Söbbing, Thomas: "Algorithmen und urheberrechtlicher Schutz" in: Computer und Recht, Heft 4, 2020, Seite 223-228

Söbbing, Thomas: "Künstliche neuronale Netze - Rechtliche Betrachtung von Software- und KI-Lernstrukturen" in: MMR, Heft 2, 2021, Seite 111-116

Specht-Riemenschneider, Louisa; Blankertz, Aline; Sierek, Pascal; et. al.: “Die Datentreuhand“ in: MMR, Heft 6, 2021, Seite 25-48

Spindler, Gerhald; Schuster, Fabian: Recht der elektronischen Medien Kommentar, 4. Auflage, C.H. Beck Verlag, München 2019 [*Bearbeiter*, in: Spindler/Schuster, *Gesetz*]

Stechern, David: "Künstliche Intelligenz - Auf der Suche nach dem Zuordnungsobjekt" in: IP-Rechtsberater, Heft 1, 2020, Seite 23-26

Von Oelffen, Sabine: „Eigentum an Daten“ in: Ballestrem, Johannes; et al. (Hrsg.): Künstliche Intelligenz - Rechtsgrundlagen und Strategien in der Praxis, 1.Auflage, Springer Verlag, Wiesbaden 2020, Seite 77-80

Wandtke, Artur-Axel; Bullinger, Winfried: Praxiskommentar Urheberrecht, 5. Auflage C.H. Beck Verlag, München 2019 [*Bearbeiter, in: Wandtke/Bullinger, UrhG*]

Wendehorst, Christiane; Schwamberger, Sebastian; Grinzinger, Julia: „Datentreuhand – wie hilfreich sind sachenrechtliche Konzepte?“ in: Pertot, Tereza (Hrsg.): Rechte an Daten, 2020, Seite 103-122

Wendehorst, Christiane: „Of Elephants in the Room and Paper Tigers – How to Reconcile Data Protection and the Data Economy“ in: Lohsse, Sebastian; Schulze, Reiner; Staudenmayer, Dirk: Trading Data in the Digital Economy: Legal Concepts and Tools, 2017, Seite 327-355

Wiebe, Andreas: "Landkarten als Datenbanken: Der Informationswert von Daten" in: GRUR-Prax, Heft 3, 2016, Seite 49-50

Wiebe, Andreas: "Schutz von Maschinendaten durch das sui-generis-Schutzrecht für Datenbanken" in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR), Heft 4, 2017, Seite 338-345

Zech, Herbert: „Industrie 4.0 – Rechtsrahmen für eine Datenwirtschaft im digitalen Binnenmarkt“ in GRUR 2015, Heft 12, Seite 1151-1158

Zech, Herbert: Daten als Wirtschaftsgut – Überlegungen zu einem „Recht des Datenerzeugers“ in: Computer und Recht, Heft 3, 2015, Seite 137-146

Zech, Herbert: Information als Schutzgut, Mohr Siebeck, Tübingen 2012

Zhu, L., et. al.: "Deep leakage from gradients", in: arXiv online, 2019, URL: https://arxiv.org/pdf/1906.08935.pdf, Abruf 14.02.2022