

Autor: Stefan Papastefanou
Dokumenttyp: Aufsatz
Literaturnachweis: CR 2019,
209-215 **Juris**

Quelle:



Verlag Dr. Otto Schmidt, Köln

Fundstelle: CR 2019, 209-215

Zitiervorschlag: Papastefanou, CR 2019, 209-215

Genetic Breeding Algorithms als Form des „Machine Learning“ im Urheber- und Patentrecht

Rechtliche Herausforderungen beim Schutz von Algorithmen des Genetic Breeding Models

Stefan Papastefanou

Der Beitrag legt den Fokus auf die etwas ins Vergessene geratene Methode des „Genetic Breeding Algorithm“ und zeigt auf, wie sich die technischen Unterschiede und Gemeinsamkeiten zu den populäreren Machine-Learning Ansätzen rechtlich auswirken können.

In Anerkennung der großen Bedeutung von Machine-Learning und KI-Systemen hat das European Patent Office (EPO) Leitlinien für das entsprechende Schutzniveau im Patentrecht veröffentlicht, so dass sich die Frage stellt, ob und durch welche rechtlichen Systeme ein adäquater Schutz möglich ist. Hierbei ist auch das Verhältnis zum Urheberrecht von besonderer Relevanz, da sich die Frage stellt, ob nicht durch eine Kombination ein ausreichendes Schutzniveau erreicht wird.

I. Formen des Machine-Learning

- 1 Künstliche Intelligenz ist ein interdisziplinäres Gebiet der Informatik mit dem Ziel, intelligentes Maschinenverhalten zu kreieren. Frühe Ansätze zur KI konzentrierten sich auf regelbasierte Systeme. Solche Systeme wurden für den Betrieb in Umgebungen mit sehr engen Bedingungen konfiguriert, in denen das Verhalten des KI-Systems zuvor anhand formeller Regeln festgelegt wurde. Wissen wurde als Regelwerk dargestellt, anhand dessen das KI-System die Ergebnisse für bestimmte Probleme bestimmen konnte; etwa als eine Struktur von if-else-Regeln, die durchlaufen werden konnten, um eine Lösung für ein bestimmtes Problem oder eine bestimmte Frage zu finden. Solche regelbasierten Systeme waren jedoch typischerweise nicht in der Lage, über das zur Verfügung gestellte Wissen hinaus zu verallgemeinern.

1. Macht und Transparenz von Algorithmen

- 2 Algorithmen entscheiden über die Angebote der einzelnen Nutzer in modernen Medien-Plattformen und die Darstellung von „News-Feed“ bzw. vergleichbaren Inhalts-Übersichten im Bereich der sozialen Medien. Bei nahezu jeder Tätigkeit im Online-Bereich trifft ein Algorithmus

Entscheidungen im Vorhinein und speichert Daten über die Entscheidungen des Nutzers, um die Angebote anzupassen.

- 3 Diese erhebliche Bedeutung von Algorithmen legt es nahe, dass eindeutig verstanden ist, wie sie funktionieren, insbesondere dann, wenn sie es nicht tun. Allerdings sind die von den Algorithmen bearbeiteten Probleme zu komplex, dass sie nicht mehr als verständliches Netzwerk von if-else-Regeln beschrieben werden können. Durch die Nutzung von Machine Learning ist es nicht mehr möglich, gänzlich nachzuvollziehen, wie diese Algorithmen funktionieren. Dennoch sind sie als geistiges Eigentum von ganz erheblichem Wert für Unternehmen und ihre Funktionsweise mindestens Gegenstand von Geschäfts-

- 209 -

Papastefanou, CR 2019, 209-215

- 210 -

geheimnissen. Das Interesse an einer adäquaten Schutzstruktur für diese bedeutsamen Algorithmen ist daher auch von erheblichem wirtschaftlichem Interesse.

- 4 Es ist fast unmöglich, einen Algorithmus von Hand zu erstellen, der die Nuancen dessen erfasst, was eine Gewebeprobe krebsartig macht, oder eine betrügerische Finanztransaktion beschreibt. Selbst einfache Bilderkennungssoftware (etwa wie sich ein bestimmtes Straßenschild aussieht und sich von sonstiger Umgebung unterscheidet) lässt sich kaum aus dem Stegreif mit einem Algorithmus programmieren. Mit ausreichend gekennzeichnetem Datensatz können Computer jedoch eine statistische Funktion erzeugen, die Eingangsdaten einer Kategorie (z.B. Straßenschild oder Umgebung) zuordnet. Die Klassifizierung und Clusterung von Daten mit diesem Ansatz wird als maschinelles Lernen bezeichnet und kann in ihrer konkreten Ausgestaltung verschiedenen Methoden folgen.
- 5 DeepMind¹, ein 2010 gegründetes und mittlerweile sehr erfolgreiches Forschungsunternehmen im Bereich der künstlichen Intelligenz, hat bereits eine Reihe internationaler Patentanmeldungen eingereicht. Die Anwendungen beziehen sich auf grundlegende Aspekte des modernen Machine-Learning. Seit 2016 hat das Unternehmen eine große Anzahl von internationalen Patentanmeldungen vorgenommen, bei denen die Einbindung der KI einen signifikanten Anteil an der Entwicklung der zu patentierenden Erfindung hatte.² Diese Anmeldungen entsprechen in der Zeit von September 2016 bis Dezember 2016 einer Anmeldequote von fast einer Anmeldung pro Woche.³
- 6 Diese rasante Entwicklung im Bereich der KI-Systeme zeigt, welches erhebliche Potential in der Nutzung von Machine-Learning für die Gewinnung von modernen Software-Lösungen liegt.
- 7 Einer der aktuell mächtigsten Ansätze des Machine-Learning stützt sich zumindest teilweise auf biologischen Prinzipien: das künstliche neuronale Netzwerk oder Neural Network Method. Ein neuronales Netzwerk ist ein verbindungsorientierter Ansatz, der von der Art und Weise, wie die Neuronen des Gehirns verdrahtet sind und interagieren, inspiriert wird. Ein künstliches neuronales Netzwerk besteht aus einer Anzahl von Knoten (*Neuronen*), die durch Kanten verbunden

sind. Jede Kante hat eine entsprechende Gewichtung und jeder Knoten hat eine Aktivierungsfunktion, die auf die ankommenden Kanten des Knotens wirkt. Knoten werden typischerweise in Schichten zusammengefasst, wobei die Ausgaben einer vorhergehenden Schicht in die Eingänge der nächsten Schicht eingehen. Während des Trainings werden die Gewichtungsdatensätze iterativ aktualisiert; so dass die Ausgabe des Netzwerks für eine bestimmte Eingabe der tatsächlichen Ausgabe der Eingabe genau entspricht.⁴

2. Evolutionärer Ansatz im Machine-Learning

- 8 Genetic Breeding Modelle⁵ sind zwar momentan weniger populär als die rekursive Neural Network Method und Deep Learning, allerdings lässt sich diese Methode durch ihre Anlehnung an die Evolution natürlicher Lebewesen einfacher beschreiben und mit zunehmender Rechenkapazität ist zu erwarten, dass Genetic Breeding Modelle als Untergruppe der Evolutionary Algorithms Modelle⁶ an Popularität wieder zunehmen.⁷ Zusätzlich handelt es sich um einen der ältesten Ansätze zur Verwirklichung maschinellen Lernens.⁸ Die Methoden sind außerdem grundsätzlich ähnlich, als dass auch neuronale Netzwerke eine Trainingsmethode aufweisen und einen weiteren Algorithmus, der Änderungen vornimmt, um einen Ziel-Algorithmus zu beschreiben.⁹ Diese funktionalen Unterschiede in der strukturellen Verwirklichung des Programm-Aufbaus stellen auf rechtlicher Ebene verschiedene Probleme dar. Die zentralen Voraussetzungen im Rahmen der Evolutionstheorie, die den Prozess der natürlichen Selektion ermöglichen¹⁰ und im Bereich des Genetic Breeding Algorithmus nachgebildet werden, sind folgende: Variation¹¹, Vererbung¹², begrenzte Überlebensquote¹³ und Wettbewerb¹⁴. Natürliche Selektion funktioniert daher in folgendem Zirkel:

1. Zunächst besteht eine Population an Organismen.
2. Der Eignungstest besteht in Form des Überlebens in der Natur.
3. Die natürliche Selektion tritt durch Versterben der weniger angepassten Organismen ein, während die anderen überleben und sich fortpflanzen.
4. Durch die Fortpflanzung vermischen sich genetischen Eigenschaften.
5. Zufällige Mutationen beeinflussen die genetische Entwicklung.

- 210 -

Papastefanou, CR 2019, 209-215

- 211 -

6. Es besteht eine neue Population von Organismen in Form der zweiten Generation.
- 9 Um diese Umstände nachzubilden, werden bei einer Genetic Breeding Methode im Rahmen der Evolutionary Algorithms folgende zentrale Elemente programmiert:
- Grundlegende Programmierungsumgebung
 - „Builder“-Algorithmus
 - „Test“-Algorithmus

- Ziel-Algorithmus, das ideale Ergebnis der Methode: ein Algorithmus, der das gegebene Problem mit für ausreichend befundener Präzision lösen kann

- 10 Selbst einfache Unterscheidungsalgorithmen¹⁵ lassen sich in der Regel nicht ohne weiteres programmieren, da zwar verschiedene Beschreibungen durch einen Nutzer möglich sind, aber diese nicht adäquat in einfache if-else-Regeln umgesetzt werden können. Auch bei einfachen Abgrenzungsfragen, wie bei einer Bilderkennung zu Straßenverkehrsschildern, lassen sich zwar sprachlich bestimmte Beschreibungen und Anweisungen zur Unterscheidung ausdrücken, die von einem Computer aber ohne eine Umsetzung in Code nicht verstanden werden kann. Aus diesem Grund kann auch nicht der Ziel-Algorithmus selbst programmiert werden, sondern zunächst werden die funktionalen Bestandteile des „Builder“-Algorithmus und des „Test“-Algorithmus definiert.¹⁶ Diese jeweiligen Algorithmen sind einfacher zu beschreiben und in ihrer Funktion auch wiederum einem (mitunter komplexen) Netz von if-else-Regeln zugänglich und damit auch im klassischen Sinne zu programmieren.

a) Der unfähige Baumeister und der unwissende Tester

- 11 Im Rahmen der Genetic Breeding Method wird das maschinelle Lernen nach folgendem Konzept realisiert: Die ersten Algorithmen, die durch den „Builder“-Algorithmus zu Beginn programmiert werden, sind in ihrer Programmierung nahezu zufällig und überhaupt nur grundsätzlich in der Lage irgendwelche Lösungen für das bestehende Problem zu liefern.¹⁷ Eine genauere Programmierung durch den „Builder“-Algorithmus ist nicht möglich, da dieser – wie die menschlichen Programmierer auch – keine Vorstellung davon hat, wie der Ziel-Algorithmus funktionieren wird.¹⁸ Wäre dieses Wissen in irgendeiner Form bekannt, könnte der Ziel-Algorithmus direkt programmiert werden.¹⁹ Anschließend wird diese erste Generation von Algorithmen an den „Test“-Algorithmus verwiesen. Viel mehr kann der „Test“-Algorithmus auch nicht leisten, er selbst nicht im Stande ist, die Zwecke des Ziel-Algorithmus zu erfüllen.
- 12 Er führt lediglich eine Abfrage der Algorithmen durch, welche sich wiederum leichter programmieren und durch menschliche Entwickler überwachen lässt. Diese Abfrage erfolgt anhand von Datenmengen, die der menschliche Programmierer dem „Test“-Algorithmus zur Verfügung stellt. Die Datenmengen bestehen aus bekannten Lösungssätzen zu dem grundsätzlichen Problem, so dass der „Test“-Algorithmus die bestehenden Algorithmen dahingehend testen kann, wie präzise ihre Programmierung durch den „Builder“-Algorithmus die Problembearbeitung ermöglicht.²⁰ Die Algorithmen, die in dieser Test-Phase die besten Ergebnisse erzielen, werden durch den „Builder“-Algorithmus anschließend weiterentwickelt, während alle anderen Algorithmen gelöscht werden. Zur Weiterentwicklung setzt der „Builder“-Algorithmus nun aus der genetischen Evolution bekannte Methoden ein, indem die besten Algorithmen kopiert werden und ihre Programm-Elemente kombiniert oder erweitert werden bzw. auch „mutations-ähnlich“ zufällige Veränderungen vorgenommen werden.²¹ Diese „zweite Generation“ von Algorithmen mit den Veränderungen in ihrer Programmierung wird nun in einer erneuten Runde an den „Test“-Algorithmus verwiesen. Dieser Zirkel lässt sich folgendermaßen darstellen:

1. Zunächst besteht ein Grundbestand an Algorithmen.
2. Der Eignungstest wird durch den „Test“-Algorithmus durchgeführt.

3. Der Test-Algorithmus führt eine Selektion durch; die besten Algorithmen werden behalten, während der Rest gelöscht wird.
 4. Der „Builder“-Algorithmus kombiniert die besten Algorithmen.
 5. Zufällige Mutationen in den Algorithmen beeinflussen die Weiter-Entwicklung.
 6. Es besteht ein neuer Bestand an Algorithmen in Form der zweiten Generation.
- 13 Mithin ist der Ablauf sehr ähnlich zu dem Prozess der natürlichen Selektion. Mit jeder Iteration steigt (zunächst sehr langsam) die Erfolgsquote der entwickelten Algorithmen. Aufgrund der Menge der verschiedenen Algorithmen und der großen Anzahl der Iterationen zwischen „Builder“- und „Test“-Algorithmus, der nur durch die Rechenleistung Grenzen gesetzt sind, entstehen Algorithmen, die dem Ziel-Algorithmus immer näher kommen und keine zufälligen Ergebnisse mehr produzieren.²²

b) Der Ziel-Algorithmus als „Blackbox“

- 14 Der Ziel-Algorithmus ist nun in der Lage auch Datenmengen zu bearbeiten und darin enthaltene Probleme zu lösen, die noch unbekannt sind. Diese beabsichtigte Verallgemeinerungsfähigkeit eines Algorithmus, auch unbekannte Probleme in einem bekannten System (zumindest mit einer bestimmten Präzision, die durch die entsprechenden Software-Entwickler defi-

- 211 -

Papastefanou, CR 2019, 209-215

- 212 -

niert wird) zu lösen, kann so erreicht werden, was durch eine klassische Nutzung von if-else-Regeln nicht möglich gewesen ist.

- 15 Wie der Ziel-Algorithmus genau funktioniert und die Problemlösung erfolgt, kann weder der „Builder“- noch der „Test“-Algorithmus darstellen. Auch der menschliche Entwickler oder der Ziel-Algorithmus selbst kann keine umfassenden Einblicke in seine Funktionsweise geben, selbst wenn einzelne Zeilen Code oder auch Code-Cluster noch dem menschlichen Verständnis zugänglich sind.
- 16 Zur Sicherstellung der erfolgreichen Funktionsweise eines Genetic Breeding Modells ist es erforderlich, dass riesige Datenmengen zur Verfügung stehen, die der „Test“-Algorithmus nutzen kann. Allerdings müssen diese Datenmengen auch Lösungen zu den Problemen enthalten, die jedoch durch eine menschliche Prüfung eingepflegt werden müssen, was in der Regel nicht von Entwicklerteams selbst vorgenommen wird, sondern durch den allgemeinen Internet-Traffic ermöglicht wird. Aus diesem Grund erfreuen sich die „Not a robot“-Test auf verschiedensten Websites großer Beliebtheit, da sie – neben einer tatsächlichen Verifizierung – zur Datensammlung eingesetzt werden. Aus diesen Daten können beispielsweise Tests konstruiert werden, um etwa Straßenverkehrsschilder zu erkennen, die dazu gedacht sind, einen „Test“-Algorithmus für autonome Fahrsysteme mit entsprechenden Testdaten zu versorgen.

- 17 Im Bereich von Entscheidungs-Algorithmen auf Medienplattformen oder innerhalb von sozialen Netzwerken werden die Tests auf andere Art und Weise ermittelt. In dieser Umgebung ist die Lösung eines Problems bestimmtes menschliches Verhalten, so dass allgemeines Nutzer- bzw. Besucherverhalten selbst die Datenmengen bildet.²³ Solange ein Verhalten einem Test zugänglich ist, ist es im Verfahren eines Genetic Breeding Modells möglich, einen entsprechenden Ziel-Algorithmus generieren zu lassen.

II. Rechtliche Herausforderungen zwischen Urheberrecht und Patentrecht

1. Verhältnis zum urheberrechtlichen Schutz nach § 69a UrhG

- 18 Die Frage nach dem Schutz von Software im Urheberrecht ist bereits Gegenstand zahlreicher Diskussion gewesen und die wesentlichen Grundsätze bereits ausführlich an anderer Stelle dargestellt.²⁴ Auch für den Bereich Machine Learning gibt es bereits Analysen, welches Schutzniveau für die Methode der neuronalen Netzwerke im Urheberrecht verfügbar ist.²⁵ Zur Erweiterung der rechtlichen Analyse ist es jedoch notwendig, zu erörtern, dass bereits die funktionalen und strukturellen Unterschiede zwischen den verschiedenen Methoden des Machine Learning zu wesentlichen Abweichungen in der rechtlichen Bewertung führen können. Aus diesem Grund wird im Folgenden die Schutzfähigkeit der Elemente des Genetic Breeding Modells analysiert.

a) Grundstruktur und Funktionsalgorithmen

- 19 Die Grundstruktur und die jeweiligen Funktionen eines neuronalen Netzwerks entsprechen bei der Genetic Breeding Variante dem „Builder“- und „Test“-Algorithmus. Bei diesen Funktionen handelt es sich ebenso um Programme im klassischen Sinne der if-else Regelwerke, so dass die Grundstruktur mit den jeweiligen Funktionen ein Computerprogramm darstellt, das von § 69a UrhG erfasst ist.²⁶ Richtigerweise wird auch argumentiert, dass ein untrainierter Algorithmus selbst zwar noch keine entsprechende, praktisch-anwendbare Funktion erfüllen kann, was aber für die Einordnung der Grundstruktur als Computerprogramm oder wenigstens als Zwischenergebnis irrelevant ist.²⁷
- 20 Nicht schutzfähig ist dagegen die grundsätzlich Idee eines Computerprogramms oder eines Algorithmus, solange dieser nicht in einer konkretisierten Folge von Steuerungsbefehlen festgehalten ist, was gewisse Schutzlücken in der komplexen Struktur eines neuronalen Netzwerkes aufweist.²⁸ Im Bereich des Genetic Breeding Modells stellt sich dieselbe Gefahr, da auch hier die Grundstruktur auf einer einfachen Programmierung basiert und sich die eigentlich beabsichtigte Wirkung erst durch Anwendung und das damit verbundene „Wachsen“ des Ziel-Algorithmus zeigt.
- 21 Die Anforderungen an die Schöpfungshöhe des jeweiligen Codes sind dadurch gewahrt, dass die jeweiligen Grundstrukturen des Genetic Breeding Modells wie bei der Neural Network Methode keine allgemein verwendeten Leistung darstellen, sondern vielmehr eine eigene Programmierleistung, die nicht auf einer computergenerierten Umgebung basieren.²⁹

b) Der Ziel-Algorithmus als Werk

- 22 Der Schutz des Trainingsergebnisses in neuronalen Netzwerken birgt grundsätzlich die Gefahr, dass aufgrund des Ergebnis-Typs unberechtigtweise in Form von Einstellungsdaten eine funktionale Kopie angefertigt werden kann, ohne dass der Schutzbereich des § 69 Abs. 1 UrhG dadurch betroffen ist.³⁰ In Abgrenzung dazu besteht bei der Variante des Genetic Breeding Models dieses Problem nicht, da es zahlreiche Vorläufer des Ziel-Algorithmus gibt, die für sich genommen selbst konkretisierte Steuerungselemente aufweisen und einen klaren Anweisungsmodus von Befehlen und Folgen haben, so dass bereits

- 212 -

Papastefanou, CR 2019, 209-215

- 213 -

die der enge³¹ Definition des Computerprogramms nach § 69 Abs. 1 UrhG erfüllt ist.

- 23 Eine Teilung des Programms in „Gewichtungsinformationen“ und Steuerungsbefehle wie dies beim Neural Networking Ansatz möglich ist, ist bei der Genetic Breeding Variante nicht möglich, da der konkrete Aufbau des Codes auf einer teilweise zufälligen „Mutation“ und teilweise gezielten Kombination von Code-Teilen beruht – eine Trennung würde ein tatsächliches und umfassendes Verständnis des Ziel-Algorithmus voraussetzen, das so jedoch in diesem Vorgang nicht existiert.

c) Der Ziel-Algorithmus als geistige Schöpfung

- 24 Viel stärker ausgeprägt als in der Frage des neuronalen Netzwerkes ist im Rahmen der Genetic Breeding Methode das Problem der geistigen Schöpfung nach § 69a Abs. 3 UrhG . Während es bei neuronalen Netzwerken lediglich um die automatische Generierung von Gewichtungsinformationen geht³², besteht bei der Genetic Breeding Methode das gesamte Ergebnis in Form des Zielalgorithmus aus dem „Evolutionsprozess“, auf den ein menschlicher Nutzer keinen Einfluss mehr hat, da der Vorgang nach Definition und Überwachung der „Builder“- und „Test“-Algorithmen vollständige automatisiert abläuft. Während eine menschliche Schöpfung unter Zuhilfenahme von Zufall-Prozessen noch zulässig wäre, ist dies im vorliegenden Fall jedoch nicht gegeben.³³
- 25 Ein Schutz als Datenbank nach § 4 Abs. 2 UrhG oder § 87a UrhG scheidet aufgrund der bereits dargestellten Unmöglichkeit der Teilung in durch den Trainingsablauf gewonnene Gewichtungsinformationen und das konkretisierte Programm für die Genetic Breeding Methode aus, wohingegen für das Neural Networking System eine Schutzmöglichkeit an dieser Stelle noch angedacht werden kann.³⁴

2. Patentierbarkeit von KI- und Machine Learning-Erfindungen

- 26 Die grundsätzliche Schwierigkeit bei der Patentierung von Software liegt im Begriff der „Erfindung“. Die Definition als „Lehre zum technischen Handeln“³⁵ setzt die Lösung eines technischen

Problems mit technischen Mitteln voraus und nicht mit rein mathematischen Mitteln, wozu Algorithmen in der Regel gehören.

- 27 Aufgrund der Mannigfaltigkeit von Errungenschaften im Bereich des Machine-Learning ist es daher sinnvoll, diese zunächst grob in drei Kategorien zu gruppieren: grundlegende Algorithmen, Plattformen und Anwendungen.

a) Grundsätzliche Unterscheidungen im Bereich der Software-Patente

- 28 Errungenschaften innerhalb der Kategorie der grundlegenden Algorithmen beziehen sich auf die KI- und maschinellen Lernalgorithmen selbst, ohne die Anwendung auf ein bestimmtes Problem zu berücksichtigen. Patent-Anmeldungen, die in diese Kategorie fallen, werden in Europa und insbesondere Deutschland am geringsten mit Patentschutz versehen. Daher können reine Algorithmen für Machine-Learning *per se* nicht innerhalb des europäischen Systems patentiert werden, da sie als mathematische Methode angesehen werden und mathematische Methoden als solche als Nicht-Erfindungen gem. Art. 52 Abs. 2, 3 EPÜ bzw. § 1 Abs. 3 Nr. 1 PatG gelten.
- 29 Als Plattformen gelten solche Software-Realisierungen, die über die bloßen Algorithmen hinausgehen und versuchen, eine Plattform bereitzustellen, von der ein Problem gelöst werden kann, ohne den Umfang der Lösung explizit auf einen bestimmten Anwendungsbereich zu beschränken.

b) EPO-Richtlinien

- 30 Am 1.11.2018 hat das EPO Richtlinien³⁶ zur Patentierbarkeit für Machine Learning und allgemeine KI-Systeme veröffentlicht.
- 31 Zunächst geben die Richtlinien den klassischen Ansatz wieder, dass Machine Learning und KI-Systeme auf Berechnungsmodellen und Algorithmen basieren, die zur Klassifizierung, Clusterbildung, Regression und Verringerung der Dimensionen genutzt werden und nennt neben den hier bereits erwähnten neuronalen Netzwerken und genetische Algorithmen zusätzlich Support-Vektormaschinen, k-Merans, Kernregression und Diskriminanzanalyse. Diese Modelle und Algorithmen weisen einen „per se abstrakten mathematischen Charakter“³⁷ auf und sind daher grundsätzlich von der Patentierbarkeit ausgeschlossen, unabhängig von dem konkret genutzten Modell. Der Technizität kann jedoch durch einen Bezug zu einem technischen Verfahren, in dessen Rahmen der Algorithmus verwendet wird, genüge getan werden.³⁸
- 32 Das EPA warnt: „Bei der Prüfung, ob der beanspruchte Gegenstand insgesamt technischen Charakter hat (Art. 52 [1], [2] und [3]), werden Ausdrücke wie ‚Support Vector Machine‘, ‚Reasoning Engine‘ oder ‚neuronales Netzwerk‘ sorgfältig betrachtet, da sie sich normalerweise auf abstrakte Modelle ohne technischen Charakter beziehen.“
- 33 Die Leitlinien nennen als Beispiel für Klassifizierungsalgorithmen insbesondere „die Verwendung eines neuronalen Netzwerks in einem Herzüberwachungsgerät zur Identifizierung unregelmäßiger Herzschläge einen technischen Beitrag. Die Klassifizierung von digitalen Bildern, Videos,

Audio- oder Sprachsignalen auf der Grundlage von Merkmalen auf niedriger Ebene (z.B. Kanten oder Pixelattributen für Bilder) sind weitere ty-

- 213 -

Papastefanou, CR 2019, 209-215

- 214 -

pische technische Anwendungen von Klassifizierungsalgorithmen.“³⁹

- 34 Dies ist eine bedeutende Entwicklung, da hierdurch möglicherweise die Möglichkeit eines europäischen Patentschutzes für Methoden für das Training von AI- oder Machine-Learning-Algorithmen sowie Mechanismen zur Generierung von Trainingsdatensätzen, die in diesem Training verwendet werden, eröffnet werden.

c) Machine Learning in der EPO-Praxis

- 35 Erfolgreiches Beispiel für eine Einbindung eines Machine-Learning Algorithmus in einem patentfähigen Plattform-Format ist etwa ein Verfahren zum Abrufen von Bildern basierend auf Bildsegmentierung und Merkmalsdeskriptoren⁴⁰. Hier wurde jedoch nicht offenbart, dass die Erfindung sich auf einen spezifischen Anwendungsbereich bezieht, wie etwa das Abrufen medizinischer Bilder. Auch ein Verfahren zum verteilten Training einer Support Vector Machine (SVM) beschreibt ein entsprechendes Plattformformat⁴¹. Der technische Charakter der offenbarten Erfindung ergibt sich aus der Tatsache, dass die SVM über mehrere verteilte lokale Sätze trainiert wird. Daher ist diese Variante kein einfacher Algorithmus mehr.
- 36 Zur Kategorie der „Anwendungen“ gehöre solche, die zur Lösung eines bestimmten Problems ein Machine-Learning Konzept anwenden, ohne die Lösung notwendigerweise auf einen bestimmten Algorithmus zu beschränken. Aktuelle Beispiele von gelungenen Patent-Anmeldungen solcher Konzepte vor dem EPO sind die Verwendung eines Machine-Learning Algorithmus, um anomale Muster in Videodaten über lange Zeiträume zu erkennen⁴² sowie ein Verfahren zum Klassifizieren der Ursachen von Maschinenausfällen unter Verwendung von Machine-Learning und entsprechenden Sensordaten.⁴³ von Sensoren erhalten werden. Methoden innerhalb dieser Kategorie sind typischerweise dadurch gekennzeichnet, dass sie sich verstärkt auf einen spezifischen Anwendungsbereich konzentrieren, als dies bei den allgemeinen Algorithmen Machine-Learning Algorithmen der Fall ist.
- 37 Wie bei jeder computerbezogenen Erfindung muss insofern auch beim Machine-Learning ein relevanter technischer Effekt erzeugt oder eine technische Lösung für ein technisches Problem bereitgestellt werden.
- 38 Wird dagegen der Machine-Learning Ziel-Algorithmus nur dazu genutzt, um eine automatische Klassifizierung von Daten vorzunehmen, ist eine Patentanmeldung aufgrund von mangelnder Technizität oder weitergehender technischer Verwendung nicht möglich.⁴⁴ Gleiches gilt bei der automatischen Steuerung einer Druckmaschine, die sich durch die Verwendung einer linearen

Regressionsanalyse auszeichnet, um subjektive und objektive harmonische Analysedaten in Bezug auf die Druckqualität zu korrelieren und Regressionsparameter zu erhalten. Sobald diese Parameter einmal gelernt sind, können sie verwendet werden, um die subjektiven Daten basierend auf den erhaltenen objektiven Daten vorherzusagen. Nur bei der Anwendung dieser Daten auf ein konkretes Problem ist diese Methode erfinderisch.⁴⁵ Aber auch eine solche Anwendung kann von der Patentierung ausgeschlossen sein, wenn die beschriebenen Merkmale zur Klassifizierung von Datensätzen als eine offensichtliche Lösung eingeordnet werden.⁴⁶

3. Machine-Learning Umgebungen als Patent-Subjekt

- 39 Nach dem derzeitigen Patentrecht müssen die genannten Erfinder einer Patentanmeldung eine natürliche Person sein. Um als „Erfinder“ im Sinne des Patentrechts zu gelten, muss diese Person tatsächlich einen Teil des erfinderischen Konzepts verwirklicht haben. Diese Einschränkung ergibt sich begriffstechnisch bereits daraus, dass die Erfindung die Lösung eines technischen Problems voraussetzt, was wiederum die Fähigkeit zur Problem-Identifizierung erfordert, die eine künstliche Intelligenz ohne vorherige Definition von „Problemen“ durch einen menschlichen Entwickler nach aktuellem Stand der Wissenschaft nicht aufweisen kann.⁴⁷ Deswegen wird auch davon gesprochen, dass die Erfindung das Ergebnis „menschlicher schöpferischer Tätigkeit“ ist.⁴⁸ Insofern besteht eine Parallele zur Einschränkung im Urheberrecht durch die „persönliche geistige Schöpfung“.
- 40 Das bloße Anwenden eines Machine-Learning Prozesses zur Lösung eines bestimmten Datensatzes für ein bestimmtes Problem kann unter diesen Voraussetzungen noch als reiner Hilfs-Algorithmus ausgelegt werden, der nicht zum erfinderischen Konzept beiträgt. Wurde ein Ziel-Algorithmus wie bei dem beschriebenen Modell des Genetic Breeding ohne jedes menschliche Zutun außerhalb der Umgebungsschaffung entwickelt, ist eine Patentierung – unabhängig von der Frage nach der Patentierfähigkeit des Algorithmus – bereits wegen der Erfinder-Untauglichkeit ausgeschlossen.⁴⁹ Wenn jedoch das Machine-Learning tatsächlich nur automatisiert einen Teil des erfinderischen Konzepts eigenständig entwickelt hat, besteht derzeit ein Mangel an Klarheit hinsichtlich des Eigentums von künstlich erzeugten Erfindungen. Eine solche Kombination kann sich etwa daraus ergeben, dass der menschliche Entwickler den künstlich-erzeugten Ziel-Algorithmus auf ein technisches Problem in technischer Umgebung anwendet, um so überhaupt erst die Patentierfähigkeit zu erreichen.
- 41 Bei kollaborativen Innovationsansätzen wird das Thema noch komplizierter, da hier beliebig viel Parteien zu einer Lösung Beiträge leisten können und hierbei zusätzlich unterschiedliche Entwicklungsmethoden aus dem Bereich des Machine-Learning angewendet werden können.

- 214 -

Papastefanou, CR 2019, 209-215

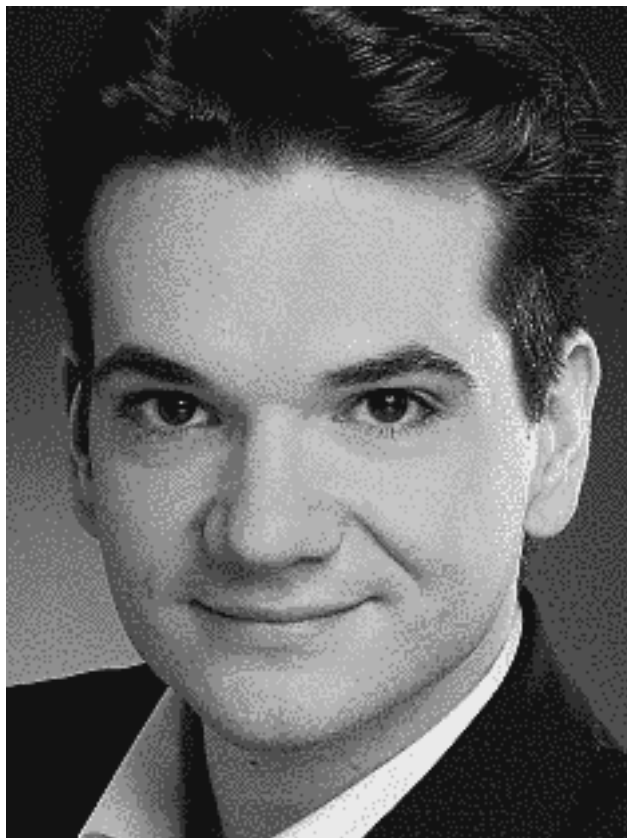
- 215 -

4. Plausibilität im Patentrecht?

- 42 Die Richtlinien des EPO besagen, wenn ein Klassifizierungsverfahren einem technischen Zweck dient, dass die Schritte des Generierens des „Lernprozesses“ und des Bearbeitungsprozesses des Ziel-Algorithmus auch zum technischen Charakter der Erfindung beitragen können, wenn sie das Erreichen dieses technischen Zwecks unterstützen.
- 43 Mithin macht es den Eindruck, dass zumindest im Rahmen einer europäischen Patentanmeldung grundsätzlich für eine Methode zum Entwickeln eines Machine-Learning Algorithmus oder für eine Methode zur Erzeugung von Gewichtungssatzen im Rahmen der Neuronal Network Methode, eine Patentierfähigkeit gegeben ist, wenn eine glaubwürdige Verknüpfung der Methode mit einem zuverlässigen und wiederholbaren technischen Verfahren möglich ist.
- 44 Ein „Test“-Algorithmus etwa, der bewirkt, dass ein neuronales Netzwerk schneller konvergiert oder ein kleinerer Datensatz verwendet wird, um ein technisches Problem zu lösen, kann sich durch diese Verknüpfung für den europäischen Patentschutz qualifizieren.
- 45 Dies scheint den Aspekt der Plausibilität in die Prüfung der Patentierfähigkeit einzuführen, der eigentlich nur in den Bereichen Pharmazie und Biotechnologie anzutreffen ist. Es besteht nun die Möglichkeit aufzuzeigen, dass ein bestimmter „ungelernter“ Algorithmus schneller konvergiert, wenn er z.B. mit einer bestimmten Methode und mit einem bestimmten Gewichtungssatz „trainiert“ wurde. Dieser Vorgang allein dürfte jedoch als Beweis nicht ausreichen, um den allgemeinen Gebrauch des Gewichtungssatzes in jedem vergleichbaren Verfahrenspatent als „plausibel“ zu bezeichnen, wenn es um die Verknüpfung von einem generischen „ungelernten“ Algorithmus und einem bestimmten Gewichtungssatz geht.
- 46 Das Erfordernis glaubwürdig darzustellen, dass die Trainingsmethode einen technischen Effekt im Rahmen eines Verfahrenspatents durch Kombination eines generischen Gewichtungssatzes oder eines generischen „untrainierten“ Algorithmus erreicht wird, scheint auf den ersten Blick schwierig. Es bleibt abzuwarten, ob dieser Beweis in der Praxis geführt werden kann. Für die Patentanmeldung einer Trainingsmethode bzw. die funktionalen Algorithmen eines Machine-Learning Konzepts bedeutet dies, dass es erforderlich sein kann, in die Anmeldung selbst eine Reihe von Arbeitsbeispielen aufzunehmen, um den Anspruch in seiner Gesamtheit zu unterstützen, damit dieser weit genug reicht, um die beabsichtigte Reichweite des Schutzbereichs für den Antragsteller zu gewährleisten.
- 47 Eine ähnliche Überlegung gilt, wenn ein Anspruch auf die Erzeugung eines Gewichtungssatzes gerichtet wird. Die Frage, welche Merkmale dieser Gewichtungssatz grundsätzlich enthalten muss, um glaubwürdig die technische Wirkung zu erzielen, auf die sich die erforderliche erfinderische Tätigkeit stützen lässt, ist nicht ohne weiteres zu beantworten. Die Identifizierung des Kerns der Erfindung kann in solchen Fällen sehr schwierig oder unmöglich sein, und das Plausibilitätsgebot des EPO kann den Geltungsbereich des Antragstellers, den der Anmelder unter solchen Umständen erhalten kann, einschränken.
- 48 Die Einführung eines Plausibilitätsmerkmals lässt daher zumindest in Ansätzen die Befürchtung zu, dass dem EPO eine Wertungsmöglichkeit bei der Beurteilung von Erfindungen im Bereich des Machine Learning eingeräumt wird, die die Rechtssicherheit erheblich einschränken kann.

III. Fazit

- 49 Die Darstellung des urheberrechtlichen und patentrechtlichen Schutzniveaus zeigt auf, dass diese jeweils nur einen lückenhaften Schutz zulassen. Auch ist eine Kombination aufgrund der unterschiedlichen Schutzbereiche und spezifischen Anforderungen keine ausreichende Lösung für die offenen Fragen. Die Richtlinien und die Praxis des EPO zeigen, dass der weiterhin vorausgesetzte Bezug zur Technizität im Patentrecht nur bedingt den Eigenheiten des Machine-Learning gerecht wird. Die Verknüpfung von Schutzwürdigkeit und technischer Lösung mag zwar den gegenwärtigen Voraussetzungen des Patentrechts entsprechen, führt aber zu einer Ungleichbehandlung von Errungenschaften im Bereich des Machine-Learning. Die eigentliche und prägende Leistung, die Datensätze zu sammeln und die Gewichtungsdaten zu bestimmen bzw. den Ziel-Algorithmus entwickeln zu lassen, erfährt durch diese Kriterien in vergleichbaren Fällen nicht ein vergleichbares Schutzniveau.
- 50 Daher liegt es nahe, eine Anforderung an den Gesetzgeber zu stellen, entsprechende Änderungen vorzunehmen, um computergenerierte Erzeugnisse einheitlich zu schützen. Rechteinhaber sollte das Unternehmen sein, das die erforderlichen Aufwendungen getätigt hat, um die Trainingsergebnisse⁵⁰ in Form des Ziel-Algorithmus oder von Gewichtungsdatensätzen zu generieren.



Wissenschaftlicher Mitarbeiter bei Latham & Watkins LLP, und Wissenschaftlicher Mitarbeiter bei Center for Transnational IP, Media and Technology Law and Policy, Bucerius Law School

Fußnoten

- 1) Im Jahr 2017 entwickelte DeepMind medienwirksam eine KI, die als erste fähig war, einen Weltmeister im Spiel „Go“ zu besiegen.
- 2) WO 2018/048934, WO 2018/048945, WO 2018064591, WO 2018071392, WO 2018/081089, WO 2018/083532, WO 2018/083667, WO 2018/083668, WO 2018/083669, WO 2018083670, WO 2018083671 und WO 2018/083672.
- 3) Spätere Anmeldungen wurden zum aktuellen Zeitpunkt noch nicht veröffentlicht.
- 4) Ausführlicher zur Technik und mit entsprechender Übersicht: *Ehinger/Stiemerling*, CR 2018, 761 ff.
- 5) Zur Übersicht: *de Jong/Spears/Gordon*, Using Genetic Algorithms for Concept Learning, Genetic Algorithms for Machine Learning, 5 ff. und *Bäck*, Evolutionary Computation 1: Basic Algorithms and Operators 2000, 59 ff.
- 6) Die Idee zur Nutzung von der natürlichen Auslese inspirierten Codierung von Mutations-, Rekombinations- und Selektionsoperatoren ist bereits seit 1950 bekannt, setzt aber große Mengen an Rechenleistung und Datenmengen zur effektiven Nutzung voraus.
- 7) Yampolskiy, Evolutionary Bioinformatics Online 2018, 16 f.; *Shulman/Bostrom*, Journal of Consciousness Studies 2012, 103 ff.; *Koza*, Genetic Programming and Evolvable Machines 2010, 251, 284; *Orlov/Sipper/Finch*, Genetic Programming Theory and Practice VIII 2011, 1 ff.;
- 8) *Fogel*, Evolutionary Computation: Toward a New Philosophy of Machine Intelligence. 3. Aufl. (2005), S. 59.
- 9) Der „Builder“-Algorithmus besteht bei der Nutzung von Neuronalen Netzwerken aus einer „Anpassungsmethode“, die in der Regeln bei einem konkreten Netzwerk angewandt wird, während bei den Evolutionary Algorithms eine Vielzahl von verschiedenen Algorithmus-Ansätzen erforderlich ist.
- 10) *Darwin*, On the Origin of Species, 1859, S. 126 f.
- 11) Ein Organismus innerhalb der Gesamt-Population muss sich zumindest geringfügig von anderen Organismen unterscheiden.
- 12) Eigenschaften eines Vorgänger-Organismus müssen an die nachfolgende Generation weitergegeben werden können.
- 13) Nur ein bestimmter Teil einer bestimmten Generation von Organismen darf die Möglichkeit haben, zu überleben und seine Eigenschaften an die Nachfolge-Generation weiterzugeben.
- 14) Solche Organismen, die besser an die Umstände angepasst sind, müssen höhere Überlebenschancen haben.
- 15) Etwa die Unterscheidung von Verkehrsstraßen und Bürgersteigen.
- 16)

- Vgl. *Prins*, Computers & Operations Research 2004, 1985 ff.
- 17) Mit einer grundsätzlichen Übersicht zur Funktionsweise: *Eiben/Smith*, Introduction to Evolutionary Computing, 2. Aufl. 2015, 25 ff.
 - 18) Vgl. *Bentley/Corne*, Creative Evolutionary Systems 2002, 1 ff.
 - 19) Die grundlegende Programmierleistung bezüglich des „Builder“-Algorithmus liegt darin, zumindest rudimentär funktionierende Programme mit ausreichender Varianz programmieren zu können. Eine zufällige Anordnung von Code würde den Prozess über mehrere Magnituden verlangsamen.
 - 20) Vgl. *Bentley/Corne*, Creative Evolutionary Systems 2002, 6 f.
 - 21) Um eine extrem rechen-intensive Entwicklung zu vermeiden, wird auf eine gänzlich zufällige Anpassung der Algorithmen verzichtet. Die weitere Programmierleistung liegt insbesondere darin, eine Form von „intelligent design“ abzubilden und erfolgreiche Elemente beizubehalten und zu verändern.
 - 22) Vgl. *Prins*, Computers & Operations Research 2004, 1999 f.
 - 23) Hierbei testet der „Test“-Algorithmus etwa, welcher der Algorithmen die besten Ergebnisse liefert, wenn es darum geht, die längste Aufmerksamkeit eines Nutzers auf eine bestimmte Website oder ein bestimmtes Produkt zu lenken.
 - 24) *Dreier* in *Dreier/Schulze*, UrhG, § 69a Rz. 6 ff.; *Melullis* in *Benkard*, EPÜ Art. 52, Rz. 247; *Hössle* in *BeckOK PatR*, Rz. 22 ff.
 - 25) *Ehinger/Stiemerling*, CR 2018, 761 ff.
 - 26) Gleiches gilt für die Grundstruktur der neuronalen Netzwerke, *Ehinger/Stiemerling*, CR 2018, 765.
 - 27) *Ehinger/Stiemerling*, CR 2018, 765.
 - 28) *Ehinger/Stiemerling*, CR 2018, 765 ff.
 - 29) *Ehinger/Stiemerling*, CR 2018, 766, wo sich bezüglich der konkret verwendeten und durch Google entwickelte Basissoftware Tensorflow diese Fragen speziell stellen.
 - 30) Mit Einzelheiten weiteren Einzelheiten *Ehinger/Stiemerling*, CR 2018, 766.
 - 31) *Ehinger/Stiemerling*, CR 2018, 767.
 - 32) *Ehinger/Stiemerling*, CR 2018, 768.
 - 33) Vgl. *Schulze* in *Dreier/Schulze*, UrhG, § 2 UrhG Rz. 8, 6.
 - 34) *Ehinger/Stiemerling*, CR 2018, 768 f., wo jedoch die Schutzmöglichkeit aufgrund der fehlenden „Unabhängigkeit“ (für § 87a UrhG) und der fehlenden „persönlichen geistigen Schöpfung“ (für § 4 Abs. 2 UrhG) richtigerweise abgelehnt wird.
 - 35) EPO Guidelines for Examination Part G – Patentability, 3.3.1. Artificial intelligence and machine learning.

- 36) G-II 3.3.1.
- 37) G-II 3.3.
- 38) „Wenn sich ein Anspruch jedoch auf ein Verfahren bezieht, bei dem technische Mittel (z.B. ein Computer) verwendet werden, oder auf ein Gerät, so hat sein Gegenstand einen technischen Charakter als Ganzes und ist daher nicht von der Patentierbarkeit gem. Art. 52 (2) und (3) ausgeschlossen.“
- 39) G-II 3.3.1.
- 40) EP2948877B1.
- 41) EP1770612B1.
- 42) EP2377044B1.
- 43) EP2930578B1.
- 44) T1784/06; ebenso für die Klassifizierung von Text-Dateien: T1358/09 und T1316/09.
- 45) T0297/86.
- 46) T1148/05.
- 47) Mit einem grundlegenden philosophischen Ansatz: *Bahke*, GRUR 1985, 596 ff.
- 48) *Melullis* in Benkard PatG § 6 Rz. 32; *Milde*, JPOS 1969, 378 ff.; *Schickedanz*, GRUR 73, 343, 347.
- 49) Für die Einordnung des Ergebnisses der Computer-Tätigkeit als Entdeckung s. *Zipse*, Mitt. 72, 41 ff.
- 50) *Ehinger/Stiernerling*, CR 2018, 770.

Der Zeitschriftenbeitrag wird von folgenden Dokumenten zitiert

Kommentare

Schuster/Grützmacher, IT-Recht

- Grützmacher, § 4 Sammelwerke und Datenbankwerke; Vorab
- Grützmacher, § 4 Sammelwerke und Datenbankwerke; II. Norminhalt; 1. Schutzvoraussetzungen; a) Datenbanken (§ 4 Abs. 2 i.V.m. Abs. 1)

Sonstiges

Frank Hartmann/Matthias Prinz, Immaterialgüterrechtlicher Schutz von Systemen Künstlicher Intelligenz; Schrifttum

Frank Hartmann/Matthias Prinz, Immaterialgüterrechtlicher Schutz von Systemen Künstlicher Intelligenz; VI. Aktueller Diskussionsstand