

Lisa Käde

Kreative Maschinen und Urheberrecht

Die Machine Learning-Werkschöpfungskette vom Training
über Modellschutz bis zu Computational Creativity



Nomos

Datenrecht und neue Technologien

Herausgegeben von

Prof. Dr. Matthias Cornils

Prof. Dr. Martin Ebers

Prof. Dr. Mario Martini

Prof. Dr. Dr. Frauke Rostalski

Prof. Dr. Giesela Rühl

Prof. Dr. Björn Steinrötter

Band 2

Lisa Käde

Kreative Maschinen und Urheberrecht

Die Machine Learning-Werkschöpfungskette vom Training
über Modellschutz bis zu Computational Creativity



Nomos

The book processing charge was funded by the Baden-Württemberg Ministry of Science, Research and Arts in the funding programme Open Access Publishing and the University of Freiburg.

Dissertation an der rechtswissenschaftlichen Fakultät
der Albert-Ludwigs-Universität Freiburg
Dekan: Prof. Dr. Jan von Hein

Mündliche Prüfung 10./11.05.2021 in Freiburg
Erstgutachter: Prof. Dr. Thomas Dreier, M.C. J.
Zweitgutachter: Prof. Dr. Maximilian Haedicke

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in
der Deutschen Nationalbibliografie; detaillierte bibliografische
Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Zugl.: Freiburg, Univ., Diss., 2021

1. Auflage 2021

© Lisa Käde

Publiziert von
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden
www.nomos.de

Gesamtherstellung:
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden

ISBN (Print): 978-3-8487-7231-5

ISBN (ePDF): 978-3-7489-1245-3

DOI: <https://doi.org/10.5771/9783748912453>



Onlineversion
Nomos eLibrary



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung –
Weitergabe unter gleichen Bedingungen 4.0 International Lizenz.

Meinen Eltern

Vorwort

Die vorliegende Arbeit wurde im Wintersemester 2020/2021 von der rechtswissenschaftlichen Fakultät der Albert-Ludwigs-Universität Freiburg als Dissertation angenommen. Sie entstand während meiner Zeit als akademische Mitarbeiterin am Zentrum für Angewandte Rechtswissenschaft (ZAR) am Karlsruher Institut für Technologie. Rechtsprechung und Literatur sind bis Juni 2021 berücksichtigt.

Mein besonderer Dank gilt Herrn Prof. Dr. Thomas Dreier für die hervorragende Betreuung und Unterstützung. Er hat mich stets mit Anregungen und Informationen zu neuen Ideen inspiriert, und mich dazu motiviert, diese Arbeit deutlich interdisziplinärer zu gestalten als vielleicht für juristische Dissertationen üblich. Herrn Prof. Dr. Maximilian Haedicke danke ich für die zügige Erstellung des Zweitgutachtens.

Ganz herzlich danken möchte ich auch meinem Korrekturleseteam Franziska Delleske und Julia Tabori für ihre Bereitschaft, sich auf die Arbeit einzulassen und ihre hilfreichen Anmerkungen, sowie Mark Keinhörster für den „Technik-Check“. Stephanie von Maltzan danke ich für das immer offene Ohr für jegliche Ideen und Gedankenexperimente und die produktiven Diskussionen über verschiedenste Fragestellungen.

Meinen Kolleginnen und Kollegen am ZAR danke ich für die vielen spannenden Unterhaltungen und die tolle gemeinsame Zeit.

Den Herausgebern und dem Verlag Nomos danke ich für die Aufnahme dieser Arbeit in die Schriftenreihe „Datenrecht und Neue Technologien“.

Von Herzen danken möchte ich meiner Familie, allen voran meinen Eltern Karin und Andreas Käde: Danke für Eure unendliche Geduld, Motivation und Inspiration nicht nur bei der Durchsicht und während der Entstehung dieser Dissertation, sondern während meines gesamten Ausbildungsweges und darüber hinaus. Euch ist diese Arbeit gewidmet.

Karlsruhe, im Juni 2021

Lisa Käde

Inhaltsübersicht

Erster Teil: Einführung	25
§ 1 Einleitung	27
§ 2 Grundlagen	33
Zweiter Teil: Die Maschine lernt – auch anhand urheberrechtlich geschützter Daten?	61
§ 3 Urheberrechtlich relevante Vorgänge im maschinellen Lernprozess	63
§ 4 Mögliche Ausnahmetatbestände	81
§ 5 Ergebnis des zweiten Teils	93
Dritter Teil: Schutzmöglichkeiten für Modelle maschinellen Lernens im Urheberrecht	95
§ 6 Technische Bestandsaufnahme	99
§ 7 Schutzmöglichkeiten im Urheberrecht und den verwandten Schutzrechten	117
§ 8 Ergebnis des dritten Teils	169
Vierter Teil: Schutz von KI-Erzeugnissen	171
§ 9 Schutzvoraussetzungen im Urheberrecht im Kontext von KI-Erzeugnissen	173
§ 10 Wem gebührt ein möglicher Schutz?	193

INHALTSÜBERSICHT

§ 11 Ergebnis des vierten Teils	213
Fünfter Teil: Computational Creativity – Kreative Künstliche Intelligenz?	215
§ 12 Intelligenz, Kreativität und ihre Schnittstellen zueinander und zum Recht	217
§ 13 Computational Creativity	245
§ 14 Ergebnis des fünften Teils	261
Sechster Teil: Gesamtergebnis und Ausblick	263
§ 15 Gesamtergebnis	265
§ 16 Ausblick	269
Mail-Austausch mit <i>Mike Cook</i> , Projekt <i>ANGELINA</i>	283

Inhaltsverzeichnis

Erster Teil: Einführung	25
§ 1 Einleitung	27
A. Problemstellungen	28
B. Gang der Arbeit	29
C. Grenzen der Arbeit	30
§ 2 Grundlagen	33
A. Künstliche Intelligenz	33
I. Begriff der künstlichen Intelligenz	33
1. Ursprung	34
2. KI heute	37
a) Definitionsansätze	38
aa) „Klassische“ KI-Definitionen	38
bb) Alternative Definition	40
cc) Hier: Strukturelle Definition	40
b) Funktionsmäßige Kategorisierung	41
II. Zusammenfassung	42
B. Maschinelles Lernen	42
I. Definitionen	43
1. Algorithmus	43
2. Lernen	43
3. Machine Learning	44
4. Modell	44
II. Untersuchte Technologien	45
1. Entscheidungsbäume bzw. Random Forest-Modelle	46
2. Künstliche neuronale Netze	47
3. Sonderfall: Generative Adversarial Networks	50
III. Einige grundlegende Begriffe	52
1. Klassifikation, Klassifizierung	52
2. Regression	53
3. Überwachtes Lernen (Supervised Learning)	53
4. Unüberwachtes Lernen (Unsupervised Learning)	54

5. Lernen durch Verstärkung (Reinforcement Learning)	54
6. Parameter, Gewichte, Biases, Hyperparameter, Topologie, Schichten, Netzstruktur, Architektur	54
7. Trainingsprozess, Trainingsdaten, Trainingsergebnisse	55
IV. Technische Einordnung	57
1. Verhältnis des Machine Learning zu Text und Data Mining	58
2. Verhältnis des Machine Learning zu Expertensystemen	59
V. Zusammenfassung	59
Zweiter Teil: Die Maschine lernt – auch anhand urheberrechtlich geschützter Daten?	61
§ 3 Urheberrechtlich relevante Vorgänge im maschinellen Lernprozess	63
A. Technische Beleuchtung des maschinellen Lernprozesses	63
I. Einlesevorgang und Analyse bzw. Aufbereitung der Trainingsdaten	64
1. Einlesen einzelner oder aller Datensätze	65
2. Umwandlung der Daten in ein maschinenlesbares Format	65
II. Trainingsvorgang	67
B. Urheberrechtliche Relevanz des Einlese- und des Trainingsvorgangs	68
I. Vervielfältigung, § 16 UrhG	68
II. Bearbeitung oder Umgestaltung, § 23 S. 1 UrhG	70
1. Dem Original dienende Funktion, Anpassung an andere Nutzungsformen	71
2. Freie Benutzung?	71
3. Einwilligungserfordernis bei Veröffentlichung oder Verwertung	72
a) Bearbeitung der Trainingsdatenbank?	72
b) Veröffentlichung oder Verwertung der Bearbeitung	73
aa) Veröffentlichung	73
bb) Verwertung	73
c) Zwischenergebnis	74
III. Zusammenhang zwischen Trainingsdaten und Output generativer Modelle	74
IV. Exkurs: Manipulierte ML-Modelle	75
1. Beispiel: Least Significant Bit Attack	76

2. Urheberrechtliche Bewertung	78
V. Zusammenfassung	78
§ 4 Mögliche Ausnahmetatbestände	81
A. Situation de lege lata	81
I. Zulässigkeit sowohl der Vervielfältigung als auch der Bearbeitung gem. § 60d UrhG?	81
II. Zulässigkeit der Vervielfältigung gem. § 44a UrhG?	83
1. Flüchtige Vervielfältigungen	84
2. Rechtmäßige Nutzung	84
3. Vorübergehende Vervielfältigung rechtswidriger Quellen	84
4. Zwischenergebnis	85
B. Situation nach zu erwartenden Änderungen durch die Umsetzung der DSM-Richtlinie	86
I. Was ist umzusetzen?	86
II. Wie erfolgt die Umsetzung?	87
1. Welche Auswirkungen ergeben sich für den ML-Prozess?	90
C. Zusammenfassung	91
§ 5 Ergebnis des zweiten Teils	93
Dritter Teil: Schutzmöglichkeiten für Modelle maschinellen Lernens im Urheberrecht	95
§ 6 Technische Bestandsaufnahme	99
A. Grundlegende Begriffe	99
I. Frameworks und Bibliotheken	100
II. API	100
III. Objekte und Funktionen	100
B. Grundbausteine für ML-Modelle: Frameworks, Bibliotheken, APIs	103
C. Quellcode	104
D. Trainiertes Modell in <i>Python</i>	104
I. Vorab: Einsatz eines trainierten Modells	105
II. Begriff des trainierten Modells	107
1. „Trainiertes Modell“ in <i>TensorFlow</i>	107
2. „Trainiertes Modell“ in <i>Keras</i>	109
3. „Trainiertes Modell“ in <i>PyTorch</i>	110
4. „Trainiertes Modell“ in <i>Scikit-Learn</i>	111

III. Zusammenfassung und Definition	112
E. Trainierter Random Forest in <i>R</i>	113
F. Trainierte Parameter	114
G. Hyperparameter	115
H. Zusammenfassung	116
§ 7 Schutzmöglichkeiten im Urheberrecht und den verwandten Schutzrechten	117
A. Einleitung	117
I. Forschungsstand	117
II. Hier gewählter Lösungsansatz	117
B. Trainiertes ML-Modell in <i>Python</i> (insb. trainiertes KNN)	118
I. Schutz als Datenbankwerk gem. § 4 Abs. 2 UrhG	118
1. Datenbankwerk	119
a) Sammlung	119
aa) <i>TensorFlow</i>	119
bb) <i>Keras</i>	120
cc) <i>PyTorch</i>	120
dd) <i>Scikit-Learn</i>	121
ee) Zusammenfassung und Subsumtion	121
b) Unabhängige Elemente	122
c) Systematische oder methodische Anordnung	124
d) Zugänglichkeit der Elemente	125
e) Zwischenergebnis	125
2. Persönliche geistige Schöpfung	127
a) Persönliche oder eigene geistige Schöpfung?	127
b) Schöpfung in Auswahl oder Anordnung	128
aa) Auswahl	128
bb) Anordnung	128
cc) Anordnung für Datenbankwerke i. d. R. programmseitig vorgegeben	128
dd) Schöpfungsspielraum in der Auswahl	129
3. Ergebnis	131
4. Wer ist der Urheber?, oder: Schutzzumfang und Folgen	131
II. Investitionsschutz gem. §§ 87a ff. UrhG	131
1. Datenbank	132
2. Investitionsgegenstand	132
3. Wesentlichkeit der Investition	133
4. Ergebnis	135

5. Schutzzumfang und Folgen	135
III. Schutz als Computerprogramm gem. § 69 a UrhG	135
1. Trainiertes Modell als Computerprogramm	136
a) Begriffsklärung Computerprogramm	136
b) (Keine) Einordnung von ML-Modellen als Computerprogramm in der Literatur	137
c) Berücksichtigung der Bestandteile eines trainierten Modells für den Computerprogrammbegriff	138
aa) Quellcode	140
bb) Hyperparameter und Parameter	141
cc) Kombination zur Laufzeit	142
dd) Sonstige Schutzgegenstände	143
ee) Zusammenfassung: infrage kommende Schutzgegenstände	144
2. Eigene geistige Schöpfung	144
a) Menschlich-gestalterische Tätigkeit	144
b) Geistiger Gehalt	145
c) Wahrnehmbarkeit	145
d) Individualität	145
3. Ergebnis	147
IV. Zusammenfassung	147
C. Trainierter Random Forest in <i>R</i>	148
D. Trainierte Parameter	150
I. Schutz als Datenbankwerk gem. § 4 Abs. 2 UrhG	150
1. Datenbankwerk	150
a) Künstliches neuronales Netz	151
b) Random Forest	151
c) Zwischenergebnis	152
2. Persönliche geistige Schöpfung	152
a) ML-Modell in <i>Python</i>	152
b) Random Forest in <i>R</i>	153
3. Ergebnis	154
II. Investitionsschutz gem. §§ 87a ff. UrhG	155
1. Vorliegen einer Datenbank	155
2. Investition	155
a) Gegenstand der Investition	155
aa) Berechnung bzw. Optimierung der Parameter	156
bb) Andere Investitionsgegenstände	159
cc) Zwischenergebnis	159

b) Wesentlichkeit	160
3. Ergebnis	160
4. Bewertung und praktische Relevanz	160
III. Zusammenfassung	161
E. Hyperparameter	161
I. Schutz als Datenbankwerk gem. § 4 Abs. 2 UrhG	162
1. Datenbank	162
2. Persönliche geistige Schöpfung	162
3. Ergebnis	163
II. Investitionsschutz gem. §§ 87a ff. UrhG	164
1. Datenbank	164
2. Wesentliche Investition	164
3. Ergebnis	165
III. Zusammenfassung	165
F. Untrainiertes Modell	165
§ 8 Ergebnis des dritten Teils	169
Vierter Teil: Schutz von KI-Erzeugnissen	171
§ 9 Schutzvoraussetzungen im Urheberrecht im Kontext von KI- Erzeugnissen	173
A. Urheber ist stets ein Mensch	173
B. Persönliche geistige Schöpfung	174
I. Menschlich-gestalterische Tätigkeit	175
1. Bekanntes Terrain: Zufallskunst bzw. Computerkunst	176
a) Zurechnungszusammenhang	177
b) Auswahltheorie in Abgrenzung zur Präsentationslehre	178
c) Strenge Auswahltheorie	178
d) Kombinierte Auswahltheorie	179
e) Externe Auswahltheorie	179
f) Zusammenfassung	180
2. Vom Bekannten zum Unbekannten: KI-Kunst	181
3. Wann liegt beim Einsatz von KI menschlich-gestalterisches Handeln vor?	182
a) Problemstellung	182
b) Andernorts am Beispiel von <i>ANGELINA</i> vorgesichlagene Herangehensweise	184

c) Kritik an dieser Vorgehensweise	186
d) Lösungsvorschlag	186
II. Geistiger Gehalt	189
III. Wahrnehmbarkeit, Individualität	190
C. Zusammenfassung	190
§ 10 Wem gebührt ein möglicher Schutz?	193
A. Forschungsstand: Wer ist warum Urheber?	193
B. Vorgeschlagene Vorgehensweise zur Ermittlung des Urhebers	195
C. Verwendung des Flussdiagramms anhand von Beispielen	204
I. Google DeepDream	204
II. Google Bach Doodle	206
III. Edmond de Belamy	207
IV. GPT-3	209
D. Zusammenfassung	210
§ 11 Ergebnis des vierten Teils	213
Fünfter Teil: Computational Creativity – Kreative Künstliche Intelligenz?	215
§ 12 Intelligenz, Kreativität und ihre Schnittstellen zueinander und zum Recht	217
A. Intelligenz	217
I. Definition der Intelligenz?	218
II. Relevanz der Intelligenz im Recht	219
1. Zivilrecht	219
2. Strafrecht	220
III. Folgerung	221
B. Kreativität	222
I. Definition?	223
1. Etymologischer Ursprung von „Kreativität“	223
2. Psychologie	224
3. Kognitionswissenschaft	225
a) Boden: P-Kreativität und H-Kreativität	225
b) McCormack/d’Inverno	226
c) Sloman: Intrinsische, expressive und standardtestbasierte Kreativität	227

d) Miller: Little-c- und Big-C-Kreativität	228
4. Andere Ansätze	229
5. Schlussfolgerung: Einführung von Kreativitätsprofilen	229
II. Relevanz der Kreativität im Recht	231
1. Analogien	232
2. Auslegungsmethodik	233
III. Folgerung	234
C. Verhältnis von Kreativität und Intelligenz und deren Bedeutung für das Urheberrecht	234
I. Grundsätzliche Überlegungen	234
1. Beschreibbarkeitsproblem	235
2. Konsequenzen	235
3. Versuch einer Gegenüberstellung	236
II. Bedeutung von Intelligenz und Kreativität im Urheberrecht	237
1. Kreativität und Intelligenz als Grundvoraussetzung im Urheberrecht?	238
2. Von künstlicher Intelligenz über künstliche Kreativität zur hypothetischen Urheberschaft	239
a) Menschlicher vs. maschineller Schöpfungsprozess	240
b) Entstehung einer Idee	240
c) Autonomie im Schöpfungsprozess	241
D. Zusammenfassung	242
§ 13 Computational Creativity	245
A. The Painting Fool	247
B. ANGELINA	249
C. Vergleich von ANGELINA und The Painting Fool anhand ihrer Kreativitätsprofile	250
D. Kritische und kreative urheberrechtliche Überlegungen zu Computational Creativity	251
I. Zu den Kriterien in § 2 Abs. 2 UrhG	251
II. Mögliche Ausgestaltung eines „Urheberrechts“ für nichtmenschliche Urheber	254
1. Durchsetzung durch gesetzlichen Vertreter	254
2. Leistungsschutzrecht de lege ferenda	255
3. Schlussfolgerung und Zusammenfassung	255
III. Hypothetische Urheberschaft	256
1. The Painting Fool	256
2. ANGELINA	257

IV. Schlussfolgerung	258
V. Zusammenfassung	258
E. Fazit	259
§ 14 Ergebnis des fünften Teils	261
Sechster Teil: Gesamtergebnis und Ausblick	263
§ 15 Gesamtergebnis	265
A. Urheberrechtliche Relevanz des ML-Prozesses	265
B. Urheberrechtlicher Schutz von ML-Modellen	265
C. Urheberrechtlicher Schutz von Erzeugnissen generativer ML-Modelle	267
D. Computational Creativity	267
§ 16 Ausblick	269
Mail-Austausch mit <i>Mike Cook</i> , Projekt <i>ANGELINA</i>	283

Abbildungsverzeichnis

2.1	Entscheidungsbaum	47
2.2	Random Forest	48
2.3	Künstliches neuronales Netz mit vier Schichten	48
3.1	Phasen des Machine Learning-Prozesses	64
3.2	Bild mit 4x4 Pixel, stark vergrößerte Darstellung	66
4.1	Änderungen der TDM-Regelungen	91
6.1	„Objektorientierung“ im HGB	101
6.2	Entstehung eines ML-Modells	106
6.3	Überblick über Speichervarianten in den den untersuchten <i>Python</i> -Frameworks und Bibliotheken	113
7.1	Einfaches Beispiel in Keras: Laden eines Modells aus .h5-Datei	139
7.2	Die ersten 20 Einträge eines Baumes in einem Random Forest	148
7.3	Beispiel eines einfachen Quellcodes zur Erstellung eines Modells mit Keras	166
8.1	Ergebnisübersicht Modellschutz	170
10.1	Flussdiagramm Urheberschaft	196
10.2	<i>DeepDream</i> : Input und Output	205
12.1	Kognitionswissenschaft im Kontext	222
12.2	Kategorien der Kreativität	230
12.3	Kreativitätsprofil Patentrecht	231
12.4	Kreativitätsprofile Analogie und Auslegung	233
12.5	Kreativitätsprofil Urheberrecht	238
13.1	Kreativitätsprofile <i>ANGELINA</i> und <i>The Painting Fool</i>	250

Abkürzungsverzeichnis

AI	Artificial Intelligence
Computerprogramm-RL	Richtlinie 2009/24/EG des Europäischen Parlaments und des Rates vom 23. April 2009 über den Rechtsschutz von Computerprogrammen
Datenbank-RL	Richtlinie 96/9/EG des Europäischen Parlaments und des Rates vom 11. März 1996 über den rechtlichen Schutz von Datenbanken
DSM-RL	Richtlinie (EU) 2019/790 des Europäischen Parlaments und des Rates vom 17. April 2019 über das Urheberrecht und die verwandten Schutzrechte im digitalen Binnenmarkt und zur Änderung der Richtlinien 96/9/EG und 2001/29/EG
GeschGeh-RL	Richtlinie (EU) 2016/943 des Europäischen Parlaments und des Rates vom 8. Juni 2016 über den Schutz vertraulichen Know-hows und vertraulicher Geschäftsinformationen (Geschäftsgeheimnisse) vor rechtswidrigem Erwerb sowie rechtswidriger Nutzung und Offenlegung
InfoSoc-RL	Richtlinie 2001/29/EG des Europäischen Parlaments und des Rates vom 22. Mai 2001 zur Harmonisierung bestimmter Aspekte des Urheberrechts und der verwandten Schutzrechte in der Informationsgesellschaft
KI	Künstliche Intelligenz
KNN	Künstliches neuronales Netz
ML	Maschinelles Lernen / Machine Learning
MLaaS	Machine Learning as a Service
NLP	Natural Language Processing
TDM	Text- und Data-Mining
WIPO	World Intellectual Property Organization

Zudem wird hinsichtlich der Verwendung anderer Abkürzungen in dieser Arbeit verwiesen auf *Kirchner*, Abkürzungsverzeichnis der Rechtssprache, 9. Auflage 2018.

