



Künstliche Intelligenz im audiovisuellen Sektor

IRIS *Spezial*

Eine Publikation
der Europäischen Audiovisuellen Informationsstelle



IRIS Spezial 2020-2**Künstliche Intelligenz im audiovisuellen Sektor**

Europäische Audiovisuelle Informationsstelle, Straßburg 2020

ISBN 978-92-871-8807-6 (Druckausgabe)

Verlagsleitung – Susanne Nikoltchev, Geschäftsführende Direktorin

Redaktionelle Betreuung – Maja Cappello, Leiterin der Abteilung für juristische Informationen

Redaktionelles Team – Francisco Javier Cabrera Blázquez, Sophie Valais, Juristische Analysten
Europäische Audiovisuelle Informationsstelle

Verfasser (in alphabetischer Reihenfolge)

Mira Burri, Sarah Eskens, Kelsey Farish, Giancarlo Frosio, Riccardo Guidotti, Atte Jääskeläinen, Andrea Pin, Justina Raižytė

Übersetzung

Stefan Pooth, Erwin Rohwer, Sonja Schmidt, Ulrike Welsch, France Courrèges, Julie Mamou, Marco Polo Sarl, Nathalie Sturlèse

Korrektur

Gianna Iacino, Catherine Koleda, Anthony Mills

Verlagsassistenz - Sabine Bouajaja

Presse und PR – Alison Hindhaugh, alison.hindhaugh@coe.int

Europäische Audiovisuelle Informationsstelle

Herausgeber

Europäische Audiovisuelle Informationsstelle

76, allée de la Robertsau F-67000 Strasbourg

France

Tel. : +33 (0)3 90 21 60 00

Fax : +33 (0)3 90 21 60 19

iris.obs@coe.int

www.obs.coe.int

Titellayout – ALTRAN, France

Bitte zitieren Sie diese Publikation wie folgt:

Cappello M. (ed.), *Künstliche Intelligenz im audiovisuellen Sektor*, IRIS Spezial, Europäische Audiovisuelle Informationsstelle, Straßburg, 2020

© Europäische Audiovisuelle Informationsstelle (Europarat), Straßburg, Dezember 2020

Jegliche in dieser Publikation geäußerten Meinungen sind persönlicher Natur und sollten in keiner Weise dahingehend verstanden werden, dass sie die Auffassung der Informationsstelle, ihrer Mitglieder oder des Europarats wiedergeben.

Künstliche Intelligenz im audiovisuellen Sektor

Mira Burri, Sarah Eskens, Kelsey Farish, Giancarlo Frosio, Riccardo Guidotti, Atte Jääskeläinen, Andrea Pin, Justina Raižytė



Vorwort

Laut Elon Musk, Gründer von SpaceX und CEO von Tesla, sollten wir „mit künstlicher Intelligenz sehr vorsichtig sein“, denn es könnte sich um „unsere größte existenzielle Bedrohung“ handeln. Das klingt ziemlich beängstigend. Trotzdem redet alle Welt von KI, und immer mehr Unternehmen setzen sie auch ein. Das ist die Zukunft, so heißt es zumindest.

Doch was ist KI eigentlich? Gewiss nicht die dystopische Vision, die Hollywood in so vielen Filmen ausbreitet, von *2001: Odyssee im Weltraum* über *Blade Runner* bis zu *Terminator*. Zumindest noch nicht.

Die Antwort auf die Frage, was KI ist, dürfte wesentlich prosaischer sein: Auch bei KI tun Computer letztlich, was Computer eben tun: Sie führen Softwarecode aus, der massenweise Daten verschlingt und all diese Rohinformationen nach vorab festgelegten Anweisungen nutzt.

Natürlich ist das Potenzial ehrfurchtgebietend. Ob Medizin, Wirtschaft, Verkehr oder Energie – die praktischen Einsatzmöglichkeiten von KI sind scheinbar grenzenlos. Doch wie jede andere technische Entwicklung ist sie nicht ohne Risiken. Ja, sie hat auch ihre Schattenseite, aber so dramatisch, wie Elon Musk uns glauben machen will, ist es wohl nicht. Zumindest noch nicht.

In der audiovisuellen Industrie, wie auch in anderen Branchen, dürfte der zunehmende Einsatz von künstlicher Intelligenz einen Paradigmenwechsel einläuten, denn er kann die gesamte Wertschöpfungskette verändern: von der Inhaltsproduktion über Programmgestaltung und Werbung bis – angesichts der Fülle von Angeboten und Geräten und der Personalisierung von Inhalten – zu den Erwartungen und Verhaltensweisen der Konsumenten. Die Schattenseite ist jedoch, dass KI zur Verbreitung von Fake News beitragen kann. Und sie wirft Fragen auf, etwa nach dem Informationsrecht der Nutzerinnen und Nutzer, nach Medienvielfalt und Medienpluralismus und nach dem Datenschutz, um nur einige wenige zu nennen.

Im Anschluss an einen Workshop, den die Europäische Audiovisuelle Informationsstelle im Dezember 2019 organisiert hat, um die von KI ausgehenden Chancen und Herausforderungen im audiovisuellen Sektor, insbesondere im journalistischen Bereich und in der Filmwirtschaft, zu diskutieren, haben wir beschlossen, diesen Auswirkungen mit der Publikation des vorliegenden Berichts genauer nachzugehen. Weitere Informationen über die interessante Veranstaltung mit einer Zusammenfassung der Diskussionen und Links zu den Vorträgen der Teilnehmenden finden Sie hier: <https://www.obs.coe.int/de/web/observatoire/-/workshop-artificial-intelligence-in-the-audiovisual-industry->

Dieser Bericht, den die juristische Abteilung der Europäischen Audiovisuellen Informationsstelle während der schwierigen Lockdown-Monate konzipiert, gestaltet und koordiniert hat, untersucht verschiedene Fragen, die eine Analyse aus regulatorischer Sicht erfordern. Er gliedert sich in drei Teile.

Der erste ist übergreifenden Themen gewidmet und beginnt mit **Kapitel 1**, das von IT-Spezialist Riccardo Guidotti (Universität Pisa) verfasst wurde. Er erklärt, was KI ist, und behandelt insbesondere deren Blackbox-Problem, also die Intransparenz ihrer Funktionsweisen und Entscheidungsprozesse, und die Frage, wie eine erklärbare KI ermöglicht werden könnte. Auf diese technische Einführung folgen zwei Überblickskapitel: In **Kapitel 2** erklärt Andrea Pin (Universität Padua), welche regulatorischen Probleme sich aus der

Sammlung und Nutzung des Stoffs ergeben, aus dem KI-Träume sind: Big Data. **Kapitel 3**, verfasst von Sarah Eskens (Universität Amsterdam), gibt einen Überblick über Fragen zu den Auswirkungen von KI auf die Meinungs- und Informationsfreiheit, etwa zum rechtlichen Rahmen für den KI-Einsatz bei den Medien und zu den Folgen für die Freiheit der Meinungsäußerung anderer.

Der zweite Teil der Publikation stellt spezifische Bereiche des Medienrechts und der Medienpolitik vor, in denen KI künftig tiefgreifende Auswirkungen haben könnte. An erster Stelle steht die kulturelle Vielfalt im Zeitalter der Algorithmen. Mira Burri (Universität Luzern) befasst sich in **Kapitel 4** damit, dass KI – von der Personalisierung von Nachrichten bis hin zu Empfehlungsalgorithmen bei Video-on-Demand-Diensten – den Schlüssel zu unseren Informationsbedürfnissen und Unterhaltungswünschen zu haben scheint, welche Auswirkungen dies hat und ob Regulierungsbedarf besteht. Andere heikle Rechtsfragen behandelt Giancarlo Frosio (Center for International Intellectual Property Studies an der Universität Straßburg) in **Kapitel 5**: Steht Maschinen, die Werke „schaffen“, ein Urheberrecht zu? Oder kann eine natürliche oder juristische Person Inhaberin des Urheberrechts an einem Werk sein, das eine Maschine geschaffen hat? In **Kapitel 6** erklärt Justina Raižytė (European Advertising Standards Alliance) die Fülle neuer Möglichkeiten, die KI Werbetreibenden eröffnet, den theoretischen Komfortgewinn für die Kundschaft, aber auch die enormen Datenschutzprobleme, die KI aufwirft. In **Kapitel 7** stellt Kelsey Farish (Anwaltskanzlei DAC Beachcroft LLP, London) verschiedene persönlichkeitsrechtliche Themen vor – Ghost Acting, Persönlichkeitsrechte Lebender und Toter sowie Rechtsverletzungen (insbesondere Deepfakes).

Im dritten Teil der Publikation stellt Atte Jääskeläinen (Technische Universität Lappeenranta und London School of Economics and Political Sciences) in **Kapitel 8** die aus seiner Sicht wichtigsten regulatorischen Herausforderungen vor, die von KI für den audiovisuellen Sektor ausgehen. Im Mittelpunkt stehen dabei mögliche Regulierungsbereiche und potenzielle Risiken.

Die einleitenden Texte und die Schlussbemerkungen, verfasst von Francisco Javier Cabrera Blázquez, Juristischer Analyst bei der Europäischen Audiovisuellen Informationsstelle, versuchen all diese verschiedenen rechtlichen und politischen Fragen einzuordnen.

Diesen großartigen Autorinnen und Autoren gilt mein herzlicher Dank dafür, dass dieser Bericht so informativ und vielseitig geworden ist. Unseren Leserinnen und Lesern kann ich nur sagen: Viel Spaß beim Lesen!

Straßburg, Dezember 2020

Maja Cappello

IRIS-Koordinatorin

Leiterin der Abteilung für juristische Information

Europäische Audiovisuelle Informationsstelle

Der Europarat befasst sich mit KI in den Menschenrechten und anderen spezifischen Kontexten.

Informieren Sie sich unter <https://www.coe.int/en/web/artificial-intelligence/home> über die Arbeit des Ad-hoc-Ausschusses für Künstliche Intelligenz (CAHAI)

Inhaltsverzeichnis

1. Künstliche Intelligenz und Erklärbarkeit	3
1.1. Was ist künstliche Intelligenz?	3
1.1.1. Eine kurze Geschichte der künstlichen Intelligenz	4
1.1.2. Unterschiedliche Ansätze zu künstlicher Intelligenz	7
1.1.3. Anwendungen künstlicher Intelligenz	9
1.2. Was ist erklärbare künstliche Intelligenz?	10
1.2.1. Beweggründe für XAI	12
1.2.2. Die Dimensionen der Interpretierbarkeit	13
1.2.3. Verschiedene Erklärungen und wie man sie liest	17
1.3. KI und XAI im Medienbereich	22
1.3.1. KI-Anwendungen und Erklärbarkeit	23
1.3.2. Video-on-Demand-Dienste in der Praxis	27
1.4. Fazit	28
<hr/>	
2. Der Stoff, aus dem KI-Träume sind – Big Data	31
2.1. Einführung	31
2.2. Datenschutz als Gatekeeper für Big Data	33
2.2.1. Vereinigte Staaten von Amerika	33
2.2.2. Europäische Union	34
2.2.3. China	36
2.2.4. Drei verschiedene Ansätze?	37
2.3. Verzerrung und Diskriminierung durch Big Data	38
2.4. Information der Bevölkerung: Medien, Fehlinformationen und illegale Inhalte	39
2.5. Big-Data-Politik und die politische Blase	43
2.6. Medien als Überwachungsinstanz?	45
2.7. Der Medienmarkt: Big-Data-getriebene Marktstrategien	47
2.8. Regulatorische Ansätze für KI-basierte Systeme	49
2.9. Fazit	50
<hr/>	
3. Auswirkungen des Einsatzes von künstlicher Intelligenz durch Nachrichtenmedien auf die Freiheit der Meinungsäußerung	55
3.1. Einführung	55
3.2. KI-Anwendungen für Nachrichtenmedien	56
3.3. KI-Einsatz durch Nachrichtenmedien als Element der Medienfreiheit	58

3.3.1. Die demokratische Rolle der Nachrichtenmedien	58
3.3.2. Für wen gilt die Medienfreiheit?.....	60
3.3.3. Pflichten und Verantwortung und journalistische Ethikkodizes	62
3.4. Auswirkungen der KI auf das Recht auf freie Meinungsäußerung von Nachrichtennutzerinnen und -nutzern und anderen an öffentlichen Debatten Teilnehmenden	65
3.5. Verpflichtungen von Staaten in Bezug auf die Medienfreiheit.....	68
3.6. Fazit	70

4. Politik für kulturelle Vielfalt im KI-Zeitalter 73

4.1. Einführung	73
4.2. Das veränderte Umfeld der Erstellung, Verbreitung, Nutzung und Wiederverwendung von Inhalten	74
4.2.1. Die neuen Vermittler	74
4.2.2. Implikationen KI-getriebener redaktioneller Akteure.....	76
4.3. Handlungsmöglichkeiten: neue Werkzeuge für den Umgang und Dialog mit digitalen Vermittlern	80
4.3.1. Regulierung von Algorithmen	81
4.3.2. Regulierung durch Algorithmen.....	84
4.4. Schlussbemerkungen.....	88

5. Urheberrecht – Ist die Maschine eine Urheberin? 93

5.1. Einführung	93
5.2. Technologie	95
5.3. Schutz: Ist KI-generierte Kreativität schutzfähig?	97
5.3.1. Persönlichkeit: Kann eine Maschine eine Rechtspersönlichkeit haben?	97
5.3.2. Urheberschaft: Kann eine Maschine Urheberin sein?.....	100
5.3.3. Originalität: Kann eine Maschine Originalität aufweisen?	107
5.4. Politische Lösungsmöglichkeiten: Sind Anreize notwendig?	111
5.4.1. Kein Schutz: Gemeinfreiheit KI-generierter Werke	112
5.4.2. Urheberschaft und Rechtsfiktionen: Soll ein Mensch der Urheber sein?.....	113
5.4.3. Soll ein Roboter der Urheber sein?	119
5.4.4. Schutz sui generis für KI-generierte Kreativität.....	120
5.4.5. Verleihung von Rechten an Verleger und Verbreiter.....	121
5.5. Fazit	122

6. KI in der Werbung: Tor nach *Deadwood* oder sinnvolle Datennutzung?..127

6.1. Einleitung.....	127
6.2. KI in der Werbung: Von der Nachverfolgung digitaler Fußabdrücke zur Erstellung von Werbeskripts	128

6.2.1.	Programmatische Werbung: die Werbe- und Datenbörse.....	129
6.2.2.	Algorithmische Kreativität: Wenn KI die Fantasie spielen lässt.....	132
6.2.3.	Von kreativen Spielen zu Gewinnen.....	134
6.2.4.	Fazit: KI-gestützte intelligente Werbung.....	138
6.3.	Bedenken im Hinblick auf Big Data und KI.....	139
6.3.1.	Bestehender Rechtsrahmen in Europa.....	140
6.3.2.	Fazit: (Meistens) das Gute, das Schlechte und das Hässliche.....	143
6.4.	Nutzung von KI für intelligente Werberegulierung.....	144
6.4.1.	Sinnvolle Datenerfassung durch Avatare.....	145
6.4.2.	KI-Entwicklungen für die Werbe-Compliance in Frankreich.....	146
6.4.3.	Nutzung von Technologie zur Steigerung des Vertrauens in den niederländischen Werbemarkt.....	148
6.4.4.	Technologische Lösungen von einer der Lokomotiven der Werbebranche.....	149
6.4.5.	Künftige Grenzen für die Werbeselbstkontrolle.....	150
6.5.	Fazit: Der große Datenrausch.....	151
6.6.	Danksagung.....	152
6.7.	Interviewverzeichnis.....	154

7. Persönlichkeitsrechte: Von Hollywood bis Deepfakes157

7.1.	Einleitung.....	157
7.2.	Künstliche Intelligenz macht es möglich: Deepfakes und Ghost Actors.....	158
7.2.1.	Deepfakes.....	159
7.2.2.	Ghost Actors.....	159
7.3.	Persönlichkeitsrechte und ihre Auswirkungen.....	160
7.3.1.	Erster Aspekt: Publicity als (geistiges) Eigentum.....	161
7.3.2.	Zweiter Aspekt: Publicity und Schutz eines persönlichen Markenzeichens.....	162
7.3.3.	Dritter Aspekt: Schutz der Privatsphäre.....	163
7.3.4.	Vierter Aspekt: Persönliche Würde und verwandte Schutzrechte.....	165
7.4.	Gesetze in ausgewählten Ländern.....	167
7.4.1.	Deutschland.....	168
7.4.2.	Frankreich.....	169
7.4.3.	Schweden.....	171
7.4.4.	Guernsey.....	172
7.4.5.	Vereinigtes Königreich.....	173
7.4.6.	Kalifornien.....	175
7.5.	Wie geht es weiter mit dem audiovisuellen Sektor in Europa?.....	177

8. Ansätze für einen nachhaltigen Regulierungsrahmen für die audiovisuelle Industrie in Europa..... 183

8.1. Einführung.....	183
8.1.1. Grundbegriffe der KI.....	185
8.2. Wie wird KI in der audiovisuellen Industrie eingesetzt?.....	187
8.3. Ist KI anders als frühere Technologien?.....	189
8.3.1. Wer ist verantwortlich, wenn die KI Schaden anrichtet?.....	190
8.3.2. Es geht nicht nur um die Wirtschaft.....	191
8.4. Wir sind moralisch verpflichtet, mit der KI Gutes zu tun.....	192
8.5. Regulierung sollte den Menschen in den Mittelpunkt stellen und sich an Zielen orientieren.....	193
8.5.1. Größere Risiken sollten angegangen werden.....	194
8.5.2. Verantwortlich ist der Mensch.....	195
8.5.3. Transparenz als Zwischenlösung?.....	195
8.6. Im Mittelpunkt der Mensch, nicht die Technologie.....	196

Abbildungen

Abbildung 1.	Beispiel für von TREPAN zurückgegebene globale baumbasierte Erklärungen.....	18
Abbildung 2.	Beispiel für von CORELS zurückgegebenen Erklärungen zur Regelliste.....	18
Abbildung 3.	Beispiel einer von LORE zurückgegebenen faktischen und kontrafaktischen regelbasierten Erklärung.....	19
Abbildung 4.	Beispiel für eine Erklärung auf der Grundlage der Merkmalbedeutung durch LIME.....	20
Abbildung 5.	Beispiel für eine Erklärung auf der Grundlage der Merkmalbedeutung durch SHAP.....	20
Abbildung 6.	Beispiel für Salienzarten, die von verschiedenen Erklärungsmethoden zurückgegeben wurden. Die erste Spalte enthält das analysierte Bild und das Label, das durch das Blackbox-Modell b des KI-Systems vergeben wurde.....	21
Abbildung 7.	Beispiel für eine von ABELE zurückgegebene Erklärung von Exemplaren (links) und Kontraexemplaren (rechts). Auf jedem (Kontra-)Exemplar ist das vom Blackbox-Modell b des KI-Systems vergebene Label angegeben.....	22

Tabellen

Tabelle 1.	Glossar zur programmatischen Werbung.....	129
Tabelle 2.	Auf kreative KI-Technologien gestützte Werbe- und Marketingkampagnen.....	134

Verzeichnis der Abkürzungen

AI4SG	AI for Social Good
ANN	Künstliche neuronale Netze (<i>artificial neuronal networks</i>)
AVMD-RL	Richtlinie über audiovisuelle Mediendienste
CDPA	Copyright Designs and Patents Act
CL	Computerlinguistik
CPU	Herkömmliche Prozessoren
DNN	Tiefe neuronale Netze (<i>deep neural networks</i>)
DSGVO	Datenschutz-Grundverordnung
DSP	Demand-Side-Plattform
EASA	European Advertising Standards Alliance (Europäische Dachorganisation für Selbstkontrollorganisationen der Werbebranche)
EDPB	European Data Protection Board
EGE	Europäische Gruppe für Ethik der Naturwissenschaften und der Neuen Technologien (<i>European Group on Ethics in Science and New Technologies</i>)
EGMR	Europäischer Gerichtshof für Menschenrechte
EMRK	Europäische Konvention der Menschenrechte
EPG	Elektronischer Programmführer
EPRS	Wissenschaftlicher Dienst des Europäischen Parlaments
EuGH	Gerichtshof der Europäischen Union
GAN	Generative kontradiktorische Netze (<i>generative adversarial networks - GAN</i>)
GPS	General Problem Solver
GPU	Grafikprozessoren (<i>graphic processing units</i>)
KI	Künstliche Intelligenz
LT	Logic Theorist (erstes denkendes Programm)
ÖRR	Öffentlich-rechtliche Rundfunkveranstalter
ROI	Return-on-Investment
RTB	Real-Time Bidding
SSP	Supply-Side-Plattform
TG	Textgenerierung

VFX	Visual Special-Effects
XAI	Erklärbare künstliche Intelligenz (<i>explainable artificial intelligence</i>)

Die Blackbox

*Wie bereits im Vorwort dieser Publikation erwähnt, ist KI eine faszinierende und zugleich beängstigende Entwicklung. Was sie heute schon kann und was sie künftig noch leisten könnte, ist in der Tat ehrfurchtgebietend, und die verschiedenen Beiträge in dieser Publikation zeugen von den vielen Möglichkeiten, wie die KI den audiovisuellen Sektor revolutionieren kann (oder bereits revolutioniert). KI-Maschinen können Musik und Texte schreiben, Film- und Lesetipps geben und sogar Tote (virtuell) zum Leben erwecken! Vielleicht löst KI darum, wie jede andere bahnbrechende technologische Entdeckung der Vergangenheit, auch Angstgefühle aus. Das ist nur natürlich. Es liegt in der Natur des Menschen, Angst vor Phänomenen zu haben, die er weder begreifen noch kontrollieren kann. Daher ist das wohl drängendste Problem, das eine KI-Regulierung zu lösen hat, das sogenannte „Blackbox-Problem“. Wie **Riccardo Guidotti** in seinem Beitrag zu dieser Publikation erläutert, „sind Blackbox-Modelle Werkzeuge, mit denen die KI Aufgaben erledigt, bei denen die Logik des Entscheidungsprozesses entweder nicht zugänglich oder zwar zugänglich, aber für Menschen nicht verständlich ist“. Es handelt sich also um eine Maschine, die Entscheidungen über das Leben von Menschen trifft, ohne dass ein Mensch die Gründe für diese Entscheidungen kontrolliert oder kennt. Das Problem besteht laut Guidotti „nicht nur im Mangel an Transparenz, sondern auch in möglichen Verzerrungen, die die Blackboxes von versteckten Vorurteilen und Artefakten in den Trainingsdaten übernehmen, mit denen die undurchschaubaren Modelle des maschinellen Lernens der KI-Systeme arbeiten“. Tatsächlich ist Transparenz heute eines der Hauptprobleme beim Einsatz von Algorithmen. Wenn ein Algorithmus, wie es heißt, einem Kochrezept gleicht, müssen die Algorithmen, die bestimmte Firmen einsetzen, wie die Coca-Cola-Formel sein, das bestgehütete Geheimrezept der Welt. Wahr ist aber auch, dass viele Menschen mit Algorithmen umgehen wie mit manchen Gerichten: Solange es schmeckt, ist ihnen das Rezept eigentlich egal, und meist wollen sie auch lieber nicht wissen, was darin ist. Zumindest in Extremfällen gibt es jedenfalls vieles, was Angst machen kann. Daher fordern Fachleute KI-Systeme, deren Mechanismen und Ergebnisse erklärbar sind.*

1. Künstliche Intelligenz und Erklärbarkeit

Riccardo Guidotti, Universität Pisa

Künstliche Intelligenz ist heute einer der wichtigsten wissenschaftlichen und technologischen Bereiche, mit enormen sozioökonomischen Auswirkungen und allgegenwärtiger Anwendung in allen Bereichen der modernen Informationsgesellschaft. Zu den prominenten Anwendungen, die auf künstlicher Intelligenz basieren, gehören Sprachassistenten (zum Beispiel Siri und Alexa), autonome Vehikel (zum Beispiel selbstfahrende Autos, Drohnen, Reinigungsroboter), medizinische Diagnose, Spamfilter und Bilderkennung. Künstliche Intelligenzsysteme erreichen ihre beeindruckende Leistung bei der Nachahmung menschlichen Verhaltens hauptsächlich durch undurchsichtige Modelle maschinellen Lernens. Diese Modelle basieren im Allgemeinen auf tiefen neuronalen Netzen, die die Logik ihrer internen Prozesse verbergen.

Der Mangel an Transparenz in der Art und Weise, wie diese Modelle Entscheidungen treffen, ist ein zentrales ethisches Moment und ein begrenzender Faktor für ihre Anwendung in sozial sensiblen und sicherheitskritischen Kontexten. Tatsächlich besteht das Problem nicht nur im Mangel an Transparenz, sondern auch in den möglichen Verzerrungen (*bias*), die Blackbox-Modelle von Artefakten und Vorurteilen, die in den Trainingsdaten verborgen sind, übernehmen. Des Weiteren kann künstliche Intelligenz zur Erstellung synthetischer realistischer Inhalte genutzt werden. Getragen von Monetisierung verändert künstliche Intelligenz die Medien- und Unterhaltungsindustrie tiefgreifend, von personalisierten Empfehlungen bis hin zur Erstellung von Inhalten.

1.1. Was ist künstliche Intelligenz?

Künstliche Intelligenz (KI) ist die „Intelligenz“, die Maschinen, Technologien oder Softwareprogramme bei der Ausführung einer Tätigkeit zeigen.¹ Der Begriff „künstlich“ wird verwendet, um sie von der „natürlichen“ oder „biologischen“ Intelligenz des Menschen zu unterscheiden. KI ist ein Forschungsgebiet der Informatik, das versucht, den Kern der Intelligenz zu verstehen und intelligente Maschinen herzustellen, die denken und reagieren und damit menschliche Intelligenz simulieren. Das Studium von KI gilt historisch gesehen als das Studium „intelligenter Agenten“, die eine *Umgebung* wahrnehmen und *Handlungen*

¹ S. Russell und P. Norvig, *Artificial intelligence: a modern approach*. Pearson.

ausführen, die ihre Chancen maximieren, ein vordefiniertes Ziel erfolgreich zu erreichen.² Die mit KI verbundenen Theorien und Technologien sind seit ihrer Entstehung immer ausgereifter geworden und die Anwendungsbereiche haben sich erweitert.

1.1.1. Eine kurze Geschichte der künstlichen Intelligenz

Der Begriff „künstliche Intelligenz“ wurde 1956 von John McCarthy während eines Workshops an der Universität Dartmouth³ vorgeschlagen, um KI von *Kybernetik* abzugrenzen.⁴ Der Workshop gilt als die Geburtsstunde der KI.

1.1.1.1. Die frühen Jahre der KI

Die frühen Jahre der KI (1952 - 1969) waren voller Erfolge, die sich auf die primitiven Computer jener Zeit und auf den Glauben beschränkten, dass Computer nicht mehr seien als leistungsfähige Rechner, die lediglich Mathematik beherrschen. Allen Newell und Herbert Simon entwarfen nach *Logic Theorist* (LT), dem ersten denkenden Programm, den *General Problem Solver* (GPS), der, anders als LT, menschliches Problemlösungsverhalten imitieren sollte. Dank GPS formulierten Newell und Simon die berühmte „physikalische Symbolsystemhypothese“, die besagt, dass jedes System, das Intelligenz zeigt, durch die Manipulation von Symbolen funktionieren muss. 1958 entwickelte John McCarthy am Massachusetts Institute of Technology die Programmiersprache LISP, die bis in die 1990er Jahre als die vorherrschende KI-Sprache verwendet wurde.⁵ In den 1960er Jahren gab es viele erfolgreiche neue Forschungsrichtungen⁶ in der KI.

1.1.1.2. Der erste KI-Winter

Von 1970 bis 1980 sah sich die KI mit dem so genannten „KI-Winter“ konfrontiert, in dem sich die Forschung stetig verlangsamte.⁷ Die Fortschrittsversprechen der Forscher in der KI erfüllten sich nicht wegen der technischen Grenzen der eingesetzten Computer zur Realisierung von KI-Programmen, die nur „Spielzeug“-Probleme lösen konnten.⁸ Verarbeitungsgeschwindigkeit oder Speicher reichten nicht aus, um etwas wirklich

² D. Poole, A. Mackworth und R. Goebel, *Computational Intelligence*. Pearson.

³ D. Crevier (1993). *AI: the tumultuous history of the search for artificial intelligence*, Basic Books, Inc.

⁴ KI und Kybernetik sind zwei verschiedene, wenngleich miteinander verbundene Forschungsgebiete, die auf dem gleichen Prinzip binärer Logik basieren. Während es bei KI jedoch darum geht, Maschinen zu erschaffen, die menschliche Intelligenz nachahmen und sich wie Menschen verhalten können, basiert Kybernetik auf einer konstruktivistischen Weltsicht und konzentriert sich auf Mensch-Maschine-Interaktionen, das heißt wie ein System Informationen verarbeitet, auf sie reagiert und sich entsprechend verändert. Die Unterschiede zwischen KI und Kybernetik sind also nicht nur semantischer, sondern auch konzeptueller Natur.

⁵ E. D. Reilly, *Milestones in computer science and information technology*, Greenwood Publishing Group.

⁶ P. McCorduck und C. Cfe, *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. CRC Press.

⁷ S. Russell und P. Norvig, a. a. O.

⁸ D. Crevier, a. a. O.

Nützlich hervorzubringen. Die von McCarthy eingeführten logikbasierten KI-Systeme mit Deduktionsprogrammen waren nicht in der Lage, echte Probleme zu lösen, da sie eine große Anzahl von Schritten erforderten, um sehr einfache Theoreme zu beweisen.⁹ Außerdem benötigen viele KI-Programme in der Praxis enorme Datenmengen. Leider verfügte zu jener Zeit niemand über eine entsprechend große Datenbasis oder war in der Lage, eine solche aufzubauen. In den 1970er Jahren wurde die erfolgreiche logische Programmiersprache Prolog¹⁰ entwickelt, ein ergiebigerer Logikansatz für KI, der steuerbare Berechnungen ermöglichte. Kritiker des logischen Ansatzes begannen eine Debatte, ob es notwendig sei, Maschinen zu haben, die wie Menschen denken, oder Maschinen, die Probleme unabhängig davon lösen können, wie Menschen es tun. Die Stellen, die KI-Forschung finanzierten, waren in der Folge vom mangelnden Fortschritt enttäuscht und stellten nahezu die gesamte Forschungsfinanzierung ein. Gleichzeitig wurde die Forschung zu neuronalen Netzen für fast 10 Jahre unterbrochen, nachdem 1969 das Buch *Perceptrons* veröffentlicht worden war. Ein Perzeptron¹¹ ist eine primitive Form eines neuronalen Netzes; heutzutage sind neuronale Netze ein unverzichtbarer Bestandteil moderner KI-Systeme. Ein (künstliches) neuronales Netz ist ein Modell maschinellen Lernens, das von biologischen neuronalen Netzen inspiriert ist und aus künstlichen Neuronen besteht. Es erhält eine Eingabe, verknüpft die Eingabe mit dem internen Zustand der Neuronen und erzeugt mit Hilfe einer Aktivierungsfunktion eine Ausgabe. Bei den Eingaben handelt es sich um Daten wie Tabellen, Bilder oder Dokumente, und die Ausgabe ist eine Klassifikation. Ein neuronales Netz lernt, wie es auf der Grundlage einer bestimmten Eingabe aus einem annotierten Trainingsdatenbestand eine Ausgabe zurückgibt.

1.1.1.3. Der Boom der KI

In den 1980er Jahren rückte „Wissen“ in den Mittelpunkt der KI-Forschung, und viele Unternehmen begannen, Formen der KI einzusetzen, die als „Expertensysteme“ bezeichnet wurden. Ein Expertensystem ist ein Algorithmus, der durch die Nutzung eines gegebenen Wissens, das durch „Wenn-Dann“-Regeln repräsentiert wird, die Entscheidungsfähigkeit eines menschlichen Experten nachahmt.¹² Es besteht aus einer „Wissensbasis“, die Fakten und Regeln repräsentiert, und aus einer „Inferenzmaschine“, die die Regeln auf die bekannten Fakten anwendet, um neue Fakten abzuleiten. Expertensysteme gehörten zu den ersten erfolgreichen KI-Softwareprogrammen, die in Wirtschaftsunternehmen eingesetzt wurden. Die Forscher erkannten, dass die Leistung von Expertensystemen aus dem Wissen, das sie enthielten, resultierte und dass „... Intelligenz auf der Fähigkeit beruhen könnte, große Mengen unterschiedlichen Wissens auf unterschiedliche Weise zu nutzen.“¹³ Dieser Vertrauensschub für KI brachte die Kreditgeber dazu, wieder in die KI-Forschung zu investieren. Parallel dazu gab es eine „Wiederbelebung“ neuronaler Netze. Hinton und

⁹ P. McCorduck und C. Cfe, a. a. O.

¹⁰ D. Crevier, a. a. O.

¹¹ P.-N. Tan et al., *Introduction to data mining*. Pearson Education India.

¹² P. Jackson, *Introduction to expert systems*. Addison-Wesley Longman Publishing Co.

¹³ P. McCorduck und C. Cfe, a. a. O.

Rumelhart machten die „Rückpropagierung“,¹⁴ eine effiziente Methode zum Training neuronaler Netze, populär. Diese Trainingsmethode setzte effizient künstliche neuronale Netze (*artificial neuronal networks* - ANN) ein, Systeme maschinellen Lernens, die von den biologischen neuronalen Netzen des menschlichen Gehirns inspiriert sind.¹⁵ ANN „lernen“ aus Beispielen, die in einem Wissensdatenbestand enthalten sind, wie eine Aufgabe zu beurteilen ist, ohne dass sie jedoch bestehende aufgabenspezifische Regeln benötigen. Sie können zum Beispiel erkennen, ob ein Bild einen Fußgänger oder ein Auto enthält, indem sie aus Bildern mit gelabeltem Inhalt und ohne jegliche Vorkenntnisse über die untersuchten Objekte lernen.

1.1.1.4. Der zweite Winter der KI

Die Forschungsinvestitionen in KI stiegen und fielen während des „zweiten Winters der KI“ (1987 - 1993). Desktop-Computer von Apple und IBM, die keine Form der KI benötigten, steigerten langsam Leistung und Geschwindigkeit und wurden 1987 effizienter als die teuren LISP- und Prolog-Maschinen. Doch trotz der Kritik einiger Investoren und Regierungen machte KI weiter Fortschritte. In diesen Jahren wurde das Konzept der „intelligenten Agenten“ dank der Definition eines „rationalen Agenten“ von Ökonomen vollendet. Ein intelligenter Agent ist ein System, das Maßnahmen ergreift, die die Erfolgchancen in Bezug auf ein vordefiniertes Ziel maximieren. Darüber hinaus wurde KI zu einer „rigorosen“ wissenschaftlichen Disziplin, da KI-Forscher bei der Entwicklung von KI-Programmen vermehrt anspruchsvolle mathematische Tools einsetzten. Zum Beispiel wurde die Wahrscheinlichkeits- und Entscheidungstheorie durch das Buch von Judea Pearl in die KI eingebracht.¹⁶ Trotz dieser offensichtlichen Fortschritte erhielt KI als theoretisches wissenschaftliches Forschungsfeld jedoch wenig Aufmerksamkeit, da eine Nutzung von Algorithmen, die ursprünglich für KI entwickelt wurden, als Teil größerer Systeme in der Technologiebranche einsetzte, zum Beispiel *Data Mining*, medizinische Diagnose, Spracherkennung, Suchmaschinen, Bankensoftware, Industrierobotik usw.

1.1.1.5. Big Data, Deep Learning und KI

Ungeachtet der oben erwähnten Fortschritte war der eigentliche Wendepunkt vor allem auf die enorme Leistungssteigerung der Computer in den 1990er Jahren zurückzuführen. Sehr berühmte Erfolgsbeispiele aufgrund dieser technologischen Fortschritte in der KI sind Deep Blue¹⁷ und Watson.¹⁸ IBMs Deep Blue war das erste schachspielende KI-System, das 1997 gegen einen Schachweltmeister, Garry Kasparov,¹⁹ gewann. Im Jahr 2011 schlug IBMs Fragenbeantwortungssystem Watson die Champions der TV-Quizshow „Jeopardy“ um Längen. Über die Fortschritte bei der Computerleistung hinaus trat KI ab 2010 dank des

¹⁴ D. Rumelhart, G. Hinton und R. Williams, „Learning representations by back-propagating errors“, *Nature* 323, 533–536 (1986), <https://doi.org/10.1038/323533a0>.

¹⁵ P.-N. Tan et al., a. a. O.

¹⁶ J. Pearl (1988), *Probabilistic reasoning in intelligent systems*.

¹⁷ Verfügbar unter <https://www.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/>.

¹⁸ Verfügbar unter <https://www.ibm.com/ibm/history/ibm100/us/en/icons/watson/>.

¹⁹ S. Russell und P. Norvig, a. a. O.

technologischen Fortschritts in Bezug auf Speicherkapazität, den einfachen Zugriff auf große Datenmengen („Big Data“) und fortgeschrittene Techniken des maschinellen Lernens wie tiefe neuronale Netze in ein neues Zeitalter ein.

- *Big Data* bezeichnet eine riesige Sammlung von Daten, die mit herkömmlicher Software nicht gespeichert, verwaltet und verarbeitet werden kann. Die Ära der großen Datenmengen hat ihren Ursprung in zwei großen Feldern:
 - (i) den Industriesektoren, die Informationen vom Tätigkeitsprotokoll bis zu den Einkäufen von Kunden speichern;
 - (ii) der weit verbreiteten Sammlung personenbezogener Nutzerinformationen durch Smartphones und mobile Geräte aus verschiedenen Quellen wie Online-Posts in sozialen Netzwerken, E-Mails, Mobilitätsspuren, Gesundheitsdaten usw.
- Tiefe neuronale Netze („deep neural networks“ - DNN) sind Modelle, die als eine Weiterentwicklung der traditionellen ANN durch die Verbindung vieler (tiefer) Verarbeitungsschichten realisiert sind. *Deep Learning* ist der Zweig des maschinellen Lernens, der DNN untersucht. DNN können für die Beurteilung von Aufgaben eingesetzt werden, die viel komplexer sind als diejenigen, die mit ANN gelöst werden können, zum Beispiel Bilderkennung, Spracherkennung, Verarbeitung natürlicher Sprache usw. Die jüngste Popularität von DNN ist jedoch hauptsächlich auf neuartige Grafikprozessoren („graphic processing units“- GPU) zurückzuführen. GPU ermöglichen im Vergleich zu herkömmlichen Prozessoren (CPU) eine deutliche Beschleunigung des Lernprozesses von DNN und dessen effiziente Ausführung. Leider leiden DNN, wie im nächsten Kapitel erörtert wird, unter einem entscheidenden Nachteil, nämlich fehlender Interpretierbarkeit.²⁰

1.1.2. Unterschiedliche Ansätze zu künstlicher Intelligenz

Historisch gesehen sind im Umgang mit KI vier verschiedene Begriffe²¹ in Bezug auf zwei Betrachtungsebenen anerkannt:

1. Beobachten der künstlichen Art des „Denkens“ gegenüber dem Beobachten künstlichen „Verhaltens“;
2. Modellierung von „Menschen“ oder Modellierung eines „idealen“ Standards (genannt Rationalität).

Die vier verschiedenen Begriffe sind daher „menschlich denken“, „rational denken“, „menschlich handeln“ und „rational handeln“. Verschiedene Forscher mit unterschiedlichen Ansätzen haben sich diesen vier Begriffen gewidmet. Infolgedessen wurde die KI-Forschung in Teilbereiche aufgeteilt, die häufig nicht miteinander kommunizieren. Diese Teilbereiche lassen sich im Hinblick auf philosophische Variationen und Begriffe, die Aufgabenstellung

²⁰ R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti und D. Pedreschi (2018), "A survey of methods for explaining black box models", *ACM computing surveys (CSUR)*, 51(5):1-42.

²¹ S. Russell und P. Norvig, a. a. O.

zur Erreichung bestimmter Ziele und die Anwendung bestimmter technischer Methoden differenzieren.

In Bezug auf philosophische Unterschiede findet man den menschenzentrierten Ansatz und den rationalistischen Ansatz. Der menschenzentrierte Ansatz legt nahe, dass KI natürliche Intelligenz simulieren sollte. Auf der anderen Seite beinhaltet ein rationalistischer Ansatz eine Kombination aus Mathematik und Ingenieurwesen und suggeriert, dass die menschliche Biologie irrelevant sei. Unter diesem Blickwinkel lässt sich KI entweder durch einfache, elegante Prinzipien wie Logik oder Optimierung gestalten, oder es erfordert die Lösung vieler eindeutiger und komplexer Probleme. In Anbetracht der verschiedenen Herausforderungen in der KI wurde das allgemeine Problem der Schaffung einer Intelligenz in Teilprobleme unterteilt, die spezifische Fähigkeiten umfassen, über die ein intelligentes System verfügen sollte. Die wichtigsten Teilprobleme sind maschinelles Lernen, Planung, Schlussfolgern, Problemlösung, Wissensrepräsentation, Wahrnehmung, Robotik, Verarbeitung natürlicher Sprache und soziale Intelligenz.²² Jedes Teilproblem entspricht einem Teilgebiet der Informatik.

Schließlich lässt sich in der Geschichte der KI eine breite Palette an Methoden erkennen, die drei verschiedenen Kategorien angehören:

1. „Kybernetik“ erforscht die Verbindungen zwischen Neurobiologie und Informationstheorie und versucht, Maschinen zu entwerfen, die elektronische Netze zur Darstellung rudimentärer Intelligenz nutzen.²³
2. „Symbolische KI“ basiert auf der Annahme, dass es durch die Manipulation von Symbolen möglich ist, viele Aspekte der menschlichen Intelligenz nachzubilden.²⁴
3. „Statistische lernbasierte KI“ beruht auf starken mathematischen Ansätzen.

Bekannte Methoden, die in KI-Systemen verwendet werden, sind „Logik“, die zur Wissensrepräsentation und zur Problemlösung verwendet wird, „probabilistische Methoden“, die beim Schlussfolgern, Planen, Lernen, Wahrnehmen und in der Robotik verwendet werden, „Such- und Optimierungsmethoden“, die für die Planung und die Robotik verwendet werden, „Methoden maschinellen Lernens“ wie Entscheidungsbaumklassifikatoren und Unterstützungsvektormaschinen sowie „tiefe neuronale Netze“, die für fast jede Aufgabenstellung eingesetzt werden. Ein Nachteil einiger dieser leistungsfähigen statistischen Lernmethoden ist, dass sie nicht interpretierbar sind, das heißt ein Mensch kann die Logik dieser Systeme bei der Entscheidungsfindung nicht nachvollziehen.

²² D. Poole, A. Mackworth und R. Goebel, a. a. O.

²³ N. Wiener, *Cybernetics (or control and communication in the animal and the machine)*, Cambridge (Massachusetts).

²⁴ J. Haugeland, *Artificial intelligence: the very idea*.

1.1.3. Anwendungen künstlicher Intelligenz

Was kann KI heute leisten und in welchen Bereichen wird sie eingesetzt? Eine erschöpfende Antwort auf diese Frage ist nicht einfach, da KI heutzutage auf einer Vielzahl von Gebieten und Aufgaben angewandt wird. Im folgenden Abschnitt finden sich kurze Darstellungen zu einigen KI-Anwendungen, die für die Leser dieser Publikation bemerkenswert oder interessant sein könnten.

- **Roboterfahrzeuge.** Selbstfahrende autonome Fahrzeuge sind dank der Fortschritte in der KI möglich geworden. Individuelle KI-Komponenten, die in Assistenzsysteme wie Kollisionsschutz, Spurwechsel, Bremsen usw. integriert sind, tragen zur Gesamtfunktion von autonomen Fahrzeugen bei. KI-Unternehmen, die sich mit Roboterfahrzeugen befassen, sind Tesla, Google und Apple.²⁵
- **Gesundheitswesen.** KI im Gesundheitswesen wird zur Unterstützung von Ärzten genutzt. Beispielsweise können KI-Systeme für die Diagnose von Krankheiten, die Analyse des Verhältnisses von Therapie und deren Ergebnis, das Erkennen von Dosierungsproblemen, die Unterstützung von Chirurgen bei Operationen, die Unterstützung von Radiologen bei der Interpretation von Bildern und die Entwicklung neuer Medikamente eingesetzt werden.²⁶
- **Marketing, Wirtschaft und Finanzen.** Unternehmen und Finanzinstitutionen waren die Ersten, die KI-Systeme für Marktanalysen, Abwanderungsvorhersagen, Preisprognosen, Aktienbeobachtung, Portfoliomanagement, algorithmischen Handel usw. übernommen haben. KI wird darüber hinaus effizient zur Bekämpfung von Betrug und Finanzkriminalität eingesetzt.
- **Medien.** Spezielle KI-Anwendungen sind in der Lage, Medieninhalte wie Fernsehsendungen, Werbung, Filme und Videos zu analysieren. Typische Anwendungsbereiche sind Gesichts- oder Objekterkennung, automatische Untertitelung, Erkennung relevanter Szenen und Zusammenfassung von Inhalten. Medienanalyse auf der Grundlage von KI ermöglicht die Erstellung beschreibender Schlüsselwörter für einen Medienbeitrag, um die Mediensuche zu vereinfachen. Eine weitere Anwendung besteht darin, Medieninhalte auf ihre Eignung hin zu überwachen oder geeignete/unangemessene Logos und Produkte im Zusammenhang mit Werbung automatisch zu erkennen.
- **Nachrichten und Veröffentlichungen.** Heutzutage verwenden viele Unternehmen KI-Techniken, um Nachrichten und Berichte automatisch zu generieren. Durch KI sind Unternehmen zudem in der Lage, Texte zu verfassen. Ein Beispiel für eine Anwendung ist die Erstellung personalisierter Zusammenfassungen von Sportveranstaltungen.²⁷ Eine andere Anwendung wandelt strukturierte Daten in Kommentare in natürlicher Sprache um.
- **Musik.** KI ermöglicht es, bis zu einem gewissen Grad menschenähnliche Kompositionen nachzuahmen, und sie hilft Menschen beim Musizieren oder

²⁵ CBInsights, *33 Corporations working on autonomous vehicles*. Abgerufen am 16. März 2017.

²⁶ E. Coiera, *Guide to Medical Informatics, the Internet and Telemedicine*, Chapman & Hall, Ltd., GBR, 1st edition.

²⁷ Abrufbar unter <https://www.barrons.com/articles/big-data-and-yahoos-quest-for-mass-personalization-1377938511>.

Singen.²⁸ Computergestützte Begleitungssysteme sind in der Lage, einem menschlichen Interpreten zuzuhören und ihm zu folgen, so dass sie synchron dazu spielen können. Interaktive Kompositionstechnologien ermöglichen es KI, mit einer Musikkomposition auf die Darbietung eines Live-Musikers zu reagieren. Schließlich sind Projekte wie Google Magenta, Sony Flow Machines oder IBM Watson Beat in der Lage, Musik in jedem beliebigen Stil zu komponieren, nachdem sie große Datenbanken mit Liedern analysiert haben. Andere KI-Anwendungen für Musik umfassen darüber hinaus Musikmarketing und Musikhören.

- **Deepfakes.** *Deepfakes* sind synthetische Medieninhalte, die durch KI-Techniken erstellt werden und für den Menschen real erscheinen.²⁹ Im Allgemeinen handelt es sich dabei um Bilder oder Videos, in denen eine Person mithilfe von *Deep-Learning*-Methoden durch eine andere Person ersetzt wird. Die Hauptmethoden zur Erstellung von *Deepfakes* beinhalten das Training generativer Ansätze wie zum Beispiel *Generative Adversarial Networks* (GAN)³⁰ oder Autoencoder.³¹ *Deepfakes* können zwar für scherzhafte Zwecke verwendet werden, besser kennt man sie jedoch als Hoaxes, gefälschte Nachrichten, pornografische Videos von Prominenten und Finanzbetrug. Folglich arbeiten sowohl Regierungen als auch die Industrie an der Entwicklung von KI-Tools, um sie aufzuspüren und einzudämmen. Der Grund dafür ist, dass KI in absehbarer Zukunft wahrscheinlich nicht nur realistische Bilder und Videos, sondern auch vollständige Medieninhalte wie Filme, Fernsehserien und Fernsehsendungen wie Reality- und Quizshows erstellen kann.

1.2. Was ist erklärbare künstliche Intelligenz?

Heutzutage sind KI-Systeme nicht nur in der Lage, den Informationsprozess des menschlichen Denkens und Lernens zu simulieren, sie können bei der Lösung einiger Aufgaben die menschliche Intelligenz sogar übertreffen. Dies ist möglich, weil künstliche Intelligenz keine menschliche Intelligenz ist und sich künstliche Intelligenz aufgrund der umfassenden Anwendung komplexer Methoden wie zum Beispiel *Deep Learning* nicht wie menschliche Intelligenz verhält oder zumindest nach einem vom Menschen nicht immer nachvollziehbaren Entscheidungsprozess handelt. In der Tat hat sich im letzten Jahrzehnt etwas entwickelt, das Frank Pasquale die „Blackbox-Gesellschaft“³² nennt, in der KI-Systeme undurchsichtige Entscheidungsmodelle einsetzen, um ihre Entscheidungsprozesse auszuführen. Ausschlaggebend dafür ist die hohe Leistung in Bezug auf die Genauigkeit,³³ die diese Blackbox-Modelle erreichen. Beispiele hierfür sind neuronale Netze und tiefe

²⁸ C. Roads, „Research in music and artificial intelligence“, *ACM Computing Surveys (CSUR)*, 17(2):163–190.

²⁹ J. Kietzmann, L. W. Lee, I. P. McCarthy und T. C. Kietzmann, „Deepfakes: Trick or treat?“, *Business Horizons*, 63(2):135-146.

³⁰ I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville und Y. Bengio, „Generative adversarial nets“, in *Advances in neural information processing systems*, Seiten 2672–2680.

³¹ A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow und B. Frey, „Adversarial autoencoders“, *arXiv preprint arXiv:1511.05644*.

³² F. Pasquale, *The black box society*, Harvard University Press.

³³ P.-N. Tan et al., a. a. O.

neuronalen Netze, Unterstützungsvektormaschinen („support vector machine“ - SVM) und Ensembleklassifikatoren, aber auch Kombinationen aus Expertensystemen, *Data Mining* und hartcodierter Software, die die Logik ihrer internen Entscheidungsprozesse vor dem Menschen „verstecken“.³⁴ Somit sind Blackbox-Modelle Tools, die von KI zur Bewältigung einer Aufgabe eingesetzt werden, für die entweder die Logik des Entscheidungsprozesses nicht zugänglich oder zwar zugänglich, für den Menschen aber nicht nachvollziehbar ist.

Der Mangel an Erklärungen, wie diese Blackbox-Modelle Entscheidungen treffen, stellt ein Problem für ihre Anwendung in sicherheitskritischen Kontexten und sozial sensiblen Bereichen wie Gesundheitswesen und Recht dar. Das Problem ist nicht nur der Mangel an Transparenz, es sind auch mögliche Verzerrungen, welche die Blackboxen von Vorurteilen und Artefakten übernehmen, die in den Trainingsdaten, die von den undurchsichtigen Modellen maschinellen Lernens der KI-Systeme verwendet werden, versteckt sind. Algorithmen maschinellen Lernens erstellen nach einer Lernphase Modelle. Dies wird durch große Datenmengen ermöglicht, welche aus Aufzeichnungen von Geschäftsvorgängen und den digitalen Spuren stammen, die Menschen bei ihren täglichen Aktivitäten (zum Beispiel Einkäufe, Bewegungen, Posts in sozialen Netzwerken usw.) hinterlassen. Diese riesigen Datenmengen können menschliche Verzerrungen und Vorurteile enthalten. Entscheidungsmodelle, die aus ihnen gelernt haben, können somit solche Verzerrungen übernehmen, was zu ungerechten und falschen Entscheidungen führen kann. Folglich hat die Forschung im Bereich erklärbarer KI („explainable artificial intelligence“ - XAI) in letzter Zeit viel Aufmerksamkeit erhalten.³⁵

Darüber hinaus sieht die Datenschutz-Grundverordnung (DSGVO)³⁶ ein allgemeines Recht auf Erklärung vor, um „aussagekräftige Informationen über die involvierte Logik“ zu erhalten, wenn automatisierte Entscheidungsfindung stattfindet. Trotz widersprüchlicher Meinungen unter Rechtsgelehrten bezüglich der tatsächlichen Tragweite dieser Klauseln,³⁷ besteht Einvernehmen darüber, dass die Umsetzung eines solchen Prinzips unerlässlich ist und dass sie heute eine große offene wissenschaftliche Herausforderung darstellt.

XAI ist über zahlreiche Industriesektoren und wissenschaftliche Disziplinen hinweg das Herzstück einer verantwortungsvollen Wissenschaft. Wie können Unternehmen ihren KI-Produkten vertrauen, ohne das logische Grundprinzip ihrer Komponenten maschinellen

³⁴ Der interessierte Leser findet Details zu neuronalen Netzen, SVM und Ensembleklassifikatoren in P.-N. Tan et al., a. a. O.

³⁵ R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti und D. Pedreschi (2018), a. a. O., T. Miller, „Explanation in artificial intelligence: Insights from the social sciences“, *Artificial Intelligence*, 267:1–38. A. Adadi und M. Berrada, *Peeking inside the black-box: A survey on explainable artificial intelligence (xai)* IEEE Access, 6:52138–52160.

³⁶ Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung), <https://eur-lex.europa.eu/legal-content/DE/TXT/?qid=1532348683434&uri=CELEX%3A02016R0679-20160504>.

³⁷ G. Malgieri und G. Comandé, „Why a right to legibility of automated decision-making exists in the General Data Protection Regulation“, *International Data Privacy Law*, 7(4):243–265. B. Goodman und S. Flaxman, „EU regulations on algorithmic decision-making and a 'right to explanation'“, in *CML workshop on human interpretability in machine learning (WHI 2016)*, New York, NY, <http://arxiv.org/abs/1606.08813v1>. S. Wachter, B. Mittelstadt und L. Floridi, „Why a right to explanation of automated decision-making does not exist in the general data protection regulation“, *International Data Privacy Law*, 7(2):76–99.

Lernens zu verstehen? Und wie können wiederum die Nutzer den KI-Diensten vertrauen? Es wird unmöglich sein, das Vertrauen der Menschen in KI zu stärken, ohne die von undurchsichtigen Modellen verfolgte Logik zu erklären.

1.2.1. Beweggründe für XAI

Neben theoretischen, ethischen und rechtlichen Beweggründen, die hinter der Notwendigkeit einer erklärbaren KI stehen, gibt es reale Fälle, in denen Diskriminierung oder Fehler hätten vermieden werden können, wenn die KI nicht undurchsichtig gewesen wäre. Zugang zu den Gründen für KI-Entscheidungen ist insbesondere bei sicherheitskritischen KI-Systemen wie selbstfahrenden Autos und Medikamenten entscheidend, wo eine mögliche Fehlentscheidung sogar zum Tod von Menschen führen kann. Im Fall eines selbstfahrenden Uber-Autos, das im Jahr 2018 in Tempe, Arizona, einen Fußgänger überfahren und getötet hat, hätte beispielsweise die Verwendung interpretierbarer Modelle Uber geholfen, die Gründe für die Entscheidung zu verstehen und seine Haftung zu regeln.

Ein weiteres inhärentes Risiko von Blackbox-Komponenten, die von KI-Systemen verwendet werden, ist, dass möglicherweise falsche Entscheidungen aufgrund erlernter sinnwidriger Korrelationen oder Artefakte, die in den Trainingsdaten enthalten sind, getroffen werden. Zum Beispiel zeigen Ribeiro et al.,³⁸ dass ein Klassifikator, der darauf trainiert war, Wölfe und Schlittenhunde zu unterscheiden, seine Vorhersagen zur Erkennung eines Wolfes allein auf das Vorhandensein von Schnee im Hintergrund stützte. Die KI traf diese Wahl, weil alle Trainingsbilder mit Wölfen Schnee im Hintergrund hatten. In einem anderen Beispiel schloss 2016 die KI-Software, die von Amazon verwendet wurde, um die Gebiete der Vereinigten Staaten festzulegen, in die Amazon am selben Tag kostenlos liefert, unbeabsichtigt Wohnbezirke von Minderheiten von der Teilnahme an dem Programm aus (oft, während gleichzeitig alle angrenzenden Wohnbezirke möglich waren).³⁹ Kürzlich zeigten die Journalisten von ProPublica, dass die COMPAS-Bewertung, ein Vorhersagemodell für das „Risiko erneuter Straffälligkeit“ (Geschäftsgeheimnis von Northpointe), eine starke ethnische Verzerrung aufweist. Tatsächlich wurden nach dieser Bewertung schwarze Personen, die nicht erneut straffällig geworden waren, doppelt so oft als „hohes Risiko“ eingestuft wie Weiße, die nicht erneut straffällig geworden waren. Auf der anderen Seite wurden weiße Wiederholungstäter doppelt so oft als „geringes Risiko“ eingestuft wie schwarze Wiederholungstäter.⁴⁰

³⁸ M. T. Ribeiro, S. Singh und C. Guestrin (2016), "Why should I trust you?: Explaining the predictions of any classifier", in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Seiten 1135-1144. ACM.

³⁹ Abrufbar unter <http://www.techinsider.io/how-algorithms-can-be-racist-2016-4>.

⁴⁰ Abrufbar unter <http://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

1.2.2. Die Dimensionen der Interpretierbarkeit

„Interpretieren“ bedeutet, bestimmten Konzepten einen Sinn zu geben oder sie zu erklären und sie in verständlichen Begriffen darzustellen.⁴¹ Daher wird die „Interpretierbarkeit“ von KI als die Fähigkeit definiert, Entscheidungen in für einen Menschen verständlichen Begriffen zu „erklären“ oder ihnen Sinn zu geben.⁴² Diese Definition setzt voraus, dass die Konzepte, aus denen eine Erklärung besteht, schlüssig sind und keiner weiteren Erklärung bedürfen. Im Grunde genommen ist eine Erklärung eine „Schnittstelle“ zwischen Mensch und KI, und sie ist gleichzeitig sowohl für den Menschen verständlich als auch ein genauer Proxy der KI. Es lässt sich eine Reihe von „Dimensionen“ ausmachen, um die Interpretierbarkeit von KI-Systemen zu analysieren, die ihrerseits die bestehenden unterschiedlichen Arten von Erklärungen widerspiegeln.⁴³

1.2.2.1. Blackbox-Erklärung versus konzeptionsbedingte Erklärung

Man unterscheidet zwischen Blackbox-Erklärung und konzeptionsbedingter Erklärung. Im ersten Fall geht es darum, eine KI mit Blackbox-Modell mit einer Erklärungsmethode zu verknüpfen, die in der Lage ist, die Entscheidungen der Blackbox zu interpretieren. Im zweiten Fall besteht die Strategie darin, das undurchsichtige Modell durch ein transparentes Modell zu ersetzen, in dem der Entscheidungsprozess konzeptionsbedingt zugänglich ist. Genauer gesagt besteht die Idee der Blackbox-Erklärung darin, die hohe Leistung des von der KI verwendeten undurchsichtigen Modells beizubehalten und eine XAI-Technik zu verwenden, um die Erklärungen zu erhalten.⁴⁴ Dieser Ansatz ist heutzutage der im Bereich der XAI-Forschung am häufigsten anzutreffende. Auf der anderen Seite besteht die „konzeptionsbedingte Erklärung“ darin, direkt ein transparentes und interpretierbares Modell zu konzipieren und die Blackbox-Komponente im KI-System durch das neue interpretierbare Modell zu ersetzen.⁴⁵ In der Literatur gibt es verschiedene Modelle, die als interpretierbar anerkannt sind. Beispiele sind „Entscheidungsbaum“, „Entscheidungsregeln“ und „lineare Modelle“.⁴⁶ Diese Modelle gelten als für den Menschen leicht nachvollziehbar und interpretierbar. Fast alle von ihnen opfern zugunsten der Interpretierbarkeit jedoch Leistung. Darüber hinaus lassen sie sich nicht effizient auf

⁴¹ Abrufbar unter <https://www.merriam-webster.com/>.

⁴² F. Doshi-Velez und B. Kim, „Towards a rigorous science of interpretable machine learning“, *arXiv preprint arXiv:1702.08608*. A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins et al., „Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges towards responsible ai“, *Information Fusion*, 58:82-115.

⁴³ R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti und D. Pedreschi (2018), a. a. O.

⁴⁴ M. Craven und J. W. Schawlik, „Extracting tree-structured representations of trained networks“, in *Advances in neural information processing systems*, Seiten 24-30. M. T. Ribeiro, S. Singh und C. Guestrin (2016), a. a. O. S. M. Lundberg und S.-I. Lee, „A unified approach to interpreting model predictions“, in *Advances in neural information processing systems*, Seiten 4765-4774.

⁴⁵ C. Rudin, „Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead“, *Nature Machine Intelligence*, 1(5):206-215. C. Rudin und J. Radin, „Why are we using black box models in ai when we don't need to? a lesson from an explainable ai competition“, *Harvard Data Science Review*, 1(2).

⁴⁶ A. A. Freitas, „Comprehensible classification models: a position paper“, *ACM SIGKDD explorations newsletter*, 15(1):1-10.

Datentypen wie Bilder oder Text anwenden, sondern nur auf tabellarische, relationale Daten, also Tabellen.

1.2.2.2. Globale *versus* lokale Erklärungen

Man unterscheidet zwischen einer globalen und einer lokalen Erklärung, je nachdem, ob die Erklärung es erlaubt, die gesamte Logik eines von einem KI-System verwendeten Modells zu verstehen, oder ob sie sich auf einen spezifischen Fall bezieht, das heißt nur eine einzige Entscheidung interpretierbar ist. Eine „globale“ Erklärung bedeutet, eine Möglichkeit zu bieten, jede denkbare Entscheidung eines Blackbox-Modells zu interpretieren. Im Allgemeinen wird das Verhalten der Blackbox mit einem transparenten Modell, das darauf trainiert ist, das Verhalten der Blackbox nachzuahmen und auch für den Menschen nachvollziehbar zu sein, näherungsweise dargestellt. Mit anderen Worten, das interpretierbare Näherungsmodell der Blackbox bietet eine globale Interpretation. Globale Erklärungen sind recht schwierig zu erreichen und bislang nur für KI möglich, die mit tabellarischen Daten arbeiten. Eine lokale Erklärung bedeutet, die Gründe für das „Ergebnis“ abzurufen, das von einem Blackbox-Modell in Bezug auf die Entscheidung für eine bestimmte Instanz zurückgegeben wird. In diesem Fall ist es nicht erforderlich, die gesamte der KI zugrunde liegende Logik zu erklären, sondern lediglich den Grund für die Vorhersage in Bezug auf eine bestimmte Eingabeinstanz. Ein interpretierbares Näherungsmodell für das KI-Blackbox-Verhalten wird daher nur in der „Nachbarschaft“ der analysierten Instanz, das heißt nur in Bezug auf ähnliche Instanzen verwendet. Die Idee ist, dass es in einer solchen Nachbarschaft leichter ist, mit einem einfachen und verständlichen, interpretierbaren Modell eine Näherung der KI zu erreichen. In den folgenden Abschnitten werden mehrere lokale Erklärungsansätze analysiert.

1.2.2.3. Interpretierbare Modelle zur Erklärung von KI

Im folgenden Abschnitt werden kurz die interpretierbaren Modelle beschrieben, die am häufigsten zur Erklärung undurchsichtiger KI-Systeme oder zum Ersatz von Blackbox-Komponenten verwendet werden.

- Ein „Entscheidungsbaum“ nutzt einen Graphen, der wie ein Baum strukturiert ist und sich aus internen Knoten, die Überprüfungen auf Merkmale oder Attribute repräsentieren (zum Beispiel ob eine Variable einen Wert kleiner, gleich oder größer als ein Schwellenwert hat), und Blattknoten, die eine Entscheidung repräsentieren, zusammensetzt. Jeder Zweig stellt ein mögliches Ergebnis dar.⁴⁷ Die Pfade von der Wurzel bis zu den Blättern stellen die Klassifikationsregeln dar. Die gebräuchlichsten Regeln sind „Wenn-Dann-Regeln“, wobei die „Wenn“-Klausel eine Kombination von Bedingungen bezüglich der Eingabevariablen darstellt. Ist die Klausel verifiziert, gibt der „Dann“-Teil die KI-Aktion an.

⁴⁷ J. R. Quinlan, *C4.5: Programs for Machine Learning*. Elsevier.

- Bei einer „Regelliste“ gibt die KI bei einem geordneten Regelsatz als Entscheidung das Ergebnis der ersten verifizierten Regel zurück.⁴⁸
- Schließlich ermöglichen „lineare Modelle“ die Visualisierung der Merkmalbedeutung, das heißt sowohl Vorzeichen als auch Größenordnung des Beitrags der Attribute für eine bestimmte Vorhersage.⁴⁹ Wenn das Vorzeichen eines Attributwertes positiv ist, ist der Beitrag eine Erhöhung der Ausgabe, anderenfalls eine Verringerung. Höhere Größenordnungen von Attributwerten weisen auf einen höheren Einfluss auf die Vorhersage des Modells hin.

1.2.2.4. Desiderate der Interpretierbarkeit

Da interpretierbare Modelle erforderlich sind, um zu Erklärungen zu gelangen, sollten einige Desiderate bei ihrer Übernahme berücksichtigt werden,⁵⁰ um das Vertrauen in ein gegebenes Modell zu erhöhen.

- „Interpretierbarkeit“ besteht darin zu bewerten, inwieweit eine gegebene Erklärung für den Menschen nachvollziehbar ist. Ein häufig verwendeter Ansatz zur Messung der Interpretierbarkeit ist die „Komplexität“ des interpretierbaren Surrogatmodells. Die Komplexität wird im Allgemeinen anhand der „Größe“ des interpretierbaren Modells geschätzt. Zum Beispiel kann die Komplexität einer Regel an der Anzahl der Klauseln in der Bedingung gemessen werden; bei linearen Modellen ist es möglich, die Anzahl der Nicht-Null-Gewichte zu zählen, während es bei Entscheidungsbäumen um die Tiefe des Baumes geht.
- „Wiedergabetreue“ besteht darin zu bewerten, inwieweit das interpretierbare Surrogatmodell in der Lage ist, die Entscheidung der KI entweder global oder lokal genau zu „imitieren“. Die Wiedergabetreue kann in Bezug auf die Entscheidungen, die durch das Blackbox-Modell getroffen werden, praktisch in Form von Genauigkeitswerten, F1-Werten usw. gemessen werden.⁵¹ Darüber hinaus sollte ein interpretierbares Modell andere wichtige allgemeine Desiderate befriedigen. So sollte es zum Beispiel eine hohe Genauigkeit bei der Bewertung der Fähigkeit des interpretierbaren Surrogatmodells haben, Entscheidungen in Bezug auf neuartige Instanzen zu treffen.
- „Fairness“ und „Privatsphäre“ sind grundlegende Desiderate, um den Schutz von Gruppen vor Diskriminierung zu gewährleisten⁵² und um sicherzustellen, dass das interpretierbare Modell keine sensiblen Informationen preisgibt.⁵³

⁴⁸ X. Yin und J. Han, „Cpar: Classification based on predictive association rules“, in *Proceedings of the 2003 SIAM International Conference on Data Mining*, Seiten 331-335. SIAM.

⁴⁹ M. T. Ribeiro, S. Singh und C. Guestrin (2016), a. a. O.

⁵⁰ A. A. Freitas, a. a. O.

⁵¹ P.-N. Tan et al., a. a. O.

⁵² A. Romei und S. Ruggieri, „A multidisciplinary survey on discrimination analysis“, *The Knowledge Engineering Review*, 29(5):582-638.

⁵³ Y. A. A. S. Aldeen, M. Salleh und M. A. Razzaque, *A comprehensive review on privacy preserving data mining*. SpringerPlus, 4(1):694.

- „Nützlichkeit“ ist eine weitere Eigenschaft, die das Vertrauen in ein Modell beeinflussen kann. Zum Beispiel kann eine interaktive Erklärung nützlicher sein als eine textliche und starre Erklärung.

1.2.2.5. Modellspezifische *versus* modellagnostische Erklärungsmethoden

Man unterscheidet zwischen modellspezifischen und modellagnostischen Erklärungsmethoden, je nachdem, ob die zum Abrufen der Erklärung angewandte Technik auf ein bestimmtes, von einem KI-System übernommenes Modell oder auf jede Art von KI angewendet werden kann. Der am häufigsten verwendete Ansatz zur Erklärung von KI-Blackboxen ist das so genannte „*Reverse Engineering*“. Der Begriff rührt daher, dass die Erklärung durch Beobachtung dessen erlangt wird, was mit der Ausgabe, das heißt der KI-Entscheidung geschieht, wenn die Eingabe kontrolliert geändert wird.

- Eine Erklärungsmethode ist „modellspezifisch“ oder nicht verallgemeinerbar,⁵⁴ wenn sie nur zur Interpretation bestimmter Typen von Blackbox-Modellen verwendet werden kann. Wenn ein Erklärungsansatz beispielsweise darauf ausgelegt ist, einen zufälligen Wald⁵⁵ zu interpretieren und intern mit dem Abstand zwischen den Bäumen arbeitet, kann ein solcher Ansatz nicht zur Erklärung der Vorhersagen eines neuronalen Netzes verwendet werden.
- Andererseits ist eine Erklärungsmethode „modellagnostisch“ oder verallgemeinerbar, wenn sie unabhängig von dem zu erklärenden Blackbox-Modell verwendet werden kann. Die internen Charakteristika der KI werden nicht verwendet, um ein zur Näherung des Blackbox Verhaltens interpretierbares Modell zu erstellen.

1.2.2.6. Nutzerhintergrund

Unterschiedliche Niveaus an Hintergrundwissen und ungleiche Erfahrungen in verschiedenen Aufgabenbereichen führen zu unterschiedlichen Vorstellungen und Anforderungen an den Gebrauch von Erklärungen. Fachexperten verstehen gegebenenfalls komplexe Erklärungen, während gewöhnliche Nutzer einfache und effiziente Erläuterungen benötigen. In der Tat hängen die Sinnhaftigkeit und der Nutzen einer Erklärung vom jeweiligen Beteiligten ab.⁵⁶ So kann zum Beispiel, wenn man den oben erwähnten COMPAS-Fall nimmt, eine spezifische Erklärung für eine Bewertung für einen Richter sinnvoll sein, der den Vorschlag des KI-Unterstützungssystems verstehen und überprüfen möchte und möglicherweise feststellt, dass es gegen Schwarze voreingenommen ist. Auf der anderen Seite ist dieselbe Erklärung für einen Strafgefangenen, der die Realität des Schwarzseins nicht ändern kann, nicht von Nutzen. Der Strafgefangene kann jedoch die Annahme nützlich

⁵⁴ D. Martens, B. Baesens, T. Van Gestel und J. Vanthienen, "Comprehensible credit scoring models using rule extraction from support vector machines", *European journal of operational research*, 183(3):1466-1476.

⁵⁵ P.-N. Tan et al., a. a. O.

⁵⁶ U. Bhatt, A. Xiang, S. Sharma, A. Weller, A. Taly, Y. Jia, J. Ghosh, R. Puri, J. M. Moura und P. Eckersley, "Explainable machine learning in deployment", in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, S. 648-657.

und daher für sich sinnvoll finden, dass er, wenn er älter ist, ein geringeres Rückfallrisiko haben und ihm eher Hausarrest gewährt werden wird.

1.2.2.7. Zeitliche Einschränkungen

Die Zeit, die sich der Nutzer für das Verstehen einer Erklärung lassen kann oder die ihm dafür zur Verfügung steht, ist ein entscheidender Aspekt. Die Zeitverfügbarkeit eines Nutzers ist offenkundig eng mit dem Szenario verbunden, in dem das Vorhersagemodell anzuwenden ist. In einigen Kontexten, in denen der Nutzer schnell eine Entscheidung treffen muss, zum Beispiel bei einer Operation oder im Falle einer drohenden Katastrophe, ist es vorzuziehen, eine einfache und effiziente Erklärung zu haben. In Kontexten hingegen, in denen die Zeit bei der Entscheidung keine entscheidende Rolle spielt, zum Beispiel während eines Verfahrens zur Freigabe eines Darlehens, wäre vielleicht eine komplexere und erschöpfendere Erklärung vorzuziehen.

1.2.3. Verschiedene Erklärungen und wie man sie liest

Der aufstrebende Bereich der XAI erschafft eine breite Palette von Alternativen zur Erklärung der Blackbox-Komponenten von KI-Systemen. Es ist in der Tat nicht möglich, eine einzige Art von Erklärung zu definieren, die für jede Anwendung geeignet wäre. Die folgenden Abschnitte veranschaulichen die am häufigsten verwendeten Arten von Erklärungen.

1.2.3.1. Globale Erklärungen

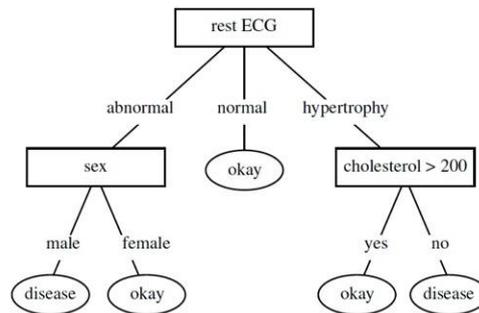
1.2.3.1.1. Baumbasierte Erklärungen

Die Näherung einer undurchsichtigen KI-Komponente mit einem Baum war einer der ersten verwendeten Ansätze.⁵⁷ Die TREPAN-Methode ist in der Lage, alle möglichen Entscheidungen, die von einem neuronalen Netz, das auf tabellarischen Daten basiert, getroffen werden, in einem einzigen Entscheidungsbaum darzustellen. TREPAN erstellt einen Entscheidungsbaum zur Näherung der durch die Netze repräsentierten Konzepte, indem es ein Gewinnverhältnis („gain ratio“)⁵⁸ maximiert, das auf der Wiedergabetreue des Baumes in Bezug auf die Entscheidung des neuronalen Netzes berechnet wird. TREPAN-Ergebnisse ermöglichen die globale Untersuchung eines neuronalen Netzes durch eine Baumstruktur, die ausgehend von einer Wurzel für jeden Pfad die Bedingungen aufzeigt, die den Entscheidungsprozess des KI-Systems bestimmen.

⁵⁷ M. Craven und J. W. Schawlik, a. a. O.

⁵⁸ P.-N. Tan et al., a. a. O.

Abbildung 1. Beispiel für von TREPAN zurückgegebene globale baumbasierte Erklärungen



1.2.3.1.2. Regelliste

Wie bereits erwähnt, besteht eine Alternative zur Erklärung von Blackbox-Klassifikatoren darin, direkt transparente Modelle für die KI-Systeme zu entwerfen. Die CORELS-Methode⁵⁹ erstellt eine Regelliste und bietet eine optimale Lösung für tabellarische Daten. Ein Beispiel für eine Regelliste ist in Abbildung 2 dargestellt. Die Regeln werden nacheinander gelesen, und die KI nimmt die Entscheidung der ersten Regel, für die die Bedingungen verifiziert sind.

Abbildung 2. Beispiel für von CORELS zurückgegebenen Erklärungen zur Regelliste

```

if (age=23-25) and (priors=2-3) then predict yes
else if (age = 18 - 20) then predict yes
else if (sex = male) and (age = 21 - 22) then predict yes
else if (priors > 3) then predict yes
else predict no
  
```

1.2.3.1.3. Lokale Erklärungen

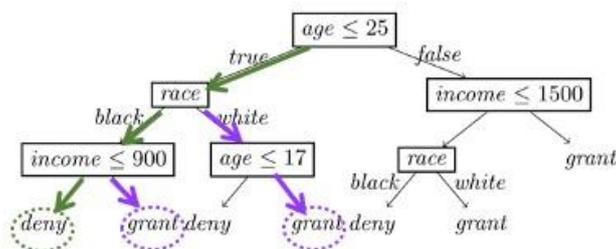
Die obigen Erklärungen sind globale Erklärungen. Wenn die zu erklärenden undurchsichtigen KI-Modelle jedoch zu kompliziert sind, ist es besser, eine lokale XAI-Methode anzuwenden und die Gründe für die Entscheidungen für die verschiedenen Instanzen separat abzurufen. Daher konzentriert sich die XAI-Forschung heutzutage mehr auf lokale Erklärungen. Die repräsentativsten lokalen Erklärungen werden in den folgenden Abschnitten beschrieben.

⁵⁹ E. Angelino, N. Larus-Stone, D. Alabi, M. Seltzer und C. Rudin, "Learning certifiably optimal rule lists", in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Seiten 35-44. ACM.

1.2.3.1.4. Regelbasierte Erklärungen

In den Erklärungen mit Wenn-Dann-Regeln nach dem Muster „Wenn - Bedingungen, Dann - Folge“ entspricht die „Folge“ der Entscheidung der KI, während die „Bedingungen“ die „faktischen Gründe“ für die „Folge“ erklären. Beispielsweise könnte die Erklärung für die Ablehnung des Kreditantrags eines Kunden mit „Alter=22, Rasse=schwarz und Einkommen=800“ von einer Bank, die eine KI verwendet, die faktische Regel „Wenn $\text{Alter} \leq 25$ und $\text{Rasse} = \text{schwarz}$ und $\text{Einkommen} \leq 900$ Dann ablehnen“ sein. Die LORE-Methode baut einen lokalen Entscheidungsbaum in der Nachbarschaft der analysierten Instanz auf⁶⁰ und extrahiert dann aus dem Baum eine einzige Regel, die die Gründe für die Entscheidung in Bezug auf die spezifische Instanz aufzeigt (siehe den grünen Pfad in Abbildung 3). ANCHOR⁶¹ ist ein weiterer XAI-Ansatz zur lokalen Erklärung der KI durch Entscheidungsregeln, die als Anker bezeichnet werden. Ein Anker enthält eine Reihe von Merkmalen mit den Werten, die für das Zustandekommen einer bestimmten Entscheidung grundlegend sind.

Abbildung 3. Beispiel einer von LORE zurückgegebenen faktischen und kontrafaktischen regelbasierten Erklärung



1.2.3.1.5. Merkmalbedeutung

Lokale Erklärungen können auch in Form von Merkmalbedeutung zurückgegeben werden. Abbildung 4 zeigt die von LIME⁶² zurückgegebene Merkmalbedeutung mit positiven und negativen Beiträgen zum Ergebnis der Blackbox und Zuordnung ihrer Bedeutung. LIME nimmt ein lineares Modell als interpretierbares lokales Surrogat an und gibt unter Nutzung der Regressionskoeffizienten die Bedeutung der Merkmale als Erklärung zurück. Abbildung 5 zeigt die von SHAP⁶³ zurückgegebene Merkmalbedeutung. SHAP liefert die lokale einzigartige additive Merkmalbedeutung für einen bestimmten Datensatz. Je höher ein SHAP-Wert ist, desto höher ist der Beitrag des Merkmals. In geeigneten Umfeldern können

⁶⁰ R. Guidotti, A. Monreale, F. Giannotti, D. Pedreschi, S. Ruggieri und F. Turini (2019a), "Factual and counterfactual explanations for black box decision making", *IEEE Intelligent Systems*.

⁶¹ M. T. Ribeiro, S. Singh und C. Guestrin (2018), "Anchors: High-precision model-agnostic explanations", in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI)*.

⁶² M. T. Ribeiro, S. Singh und C. Guestrin (2016), a. a. O.

⁶³ S. M. Lundberg und S.-I. Lee, a. a. O.

LIME und SHAP auch zur Erklärung von KI-Systemen verwendet werden, die mit Text arbeiten.

Abbildung 4. Beispiel für eine Erklärung auf der Grundlage der Merkmalbedeutung durch LIME

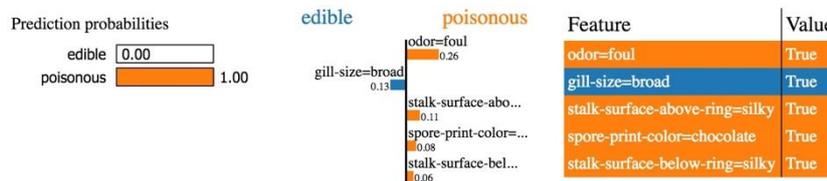
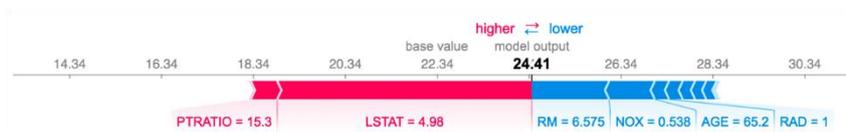


Abbildung 5. Beispiel für eine Erklärung auf der Grundlage der Merkmalbedeutung durch SHAP



1.2.3.1.6. Salienzkarten

In der Bildverarbeitung bestehen typische Erklärungen aus „Salienzkarten“, das heißt Bildern, die den positiven (oder negativen) Beitrag jedes Pixels zum Ergebnis der Blackbox zeigen. Salienzkarten werden erstellt, um DNN-Modelle mit gradienten- und perturbationsbasierten Attributionsmethoden lokal zu erklären. Diese XAI-Ansätze suchen derart nach den wichtigsten Pixeln des Bildes, dass die Wahrscheinlichkeit maximiert wird, dass die KI dieselbe Antwort zurückgibt, ohne irrelevante Pixel zu berücksichtigen. Bei geeigneten Bildtransformationen, die das Konzept der „Super-Pixel“ nutzen, können auch Methoden wie LORE und LIME verwendet werden, um an Bildern arbeitende KI zu erklären. Die ABELE-Methode⁶⁴ verwendet generative Modelle, um eine Salienzkarte auszugeben, die die angrenzenden Gebiete hervorhebt, die unter Beibehaltung derselben Entscheidung der von der KI verwendeten Blackbox unterschiedlich sein können. Abbildung 6 ist ein Vergleich von Salienzkarten zur Klassifikation der handschriftlichen Ziffern 9 und 0 nach den Erklärungsmethoden ABELE,⁶⁵ LIME,⁶⁶ SALiency,⁶⁷ GRADInput,⁶⁸ INTGrad⁶⁹ und ELRP.⁷⁰

⁶⁴ R. Guidotti, A. Monreale, S. Matwin und D. Pedreschi (2019b), "Black box explanation by learning image exemplars in the latent feature space", in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Seiten 189-205. Springer.

⁶⁵ R. Guidotti, A. Monreale, S. Matwin und D. Pedreschi (2019b), a. a. O.

⁶⁶ M. T. Ribeiro, S. Singh und C. Guestrin (2016), a. a. O.

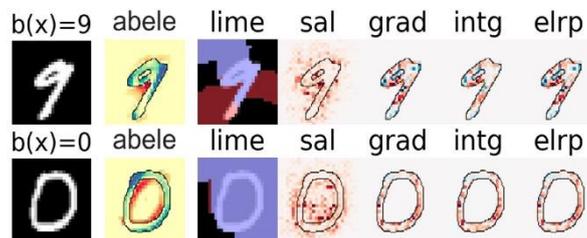
⁶⁷ K. Simonyan, A. Vedaldi und A. Zisserman, "Deep inside convolutional networks: Visualising image classification models and saliency maps", *arXiv preprint arXiv:1312.6034*.

⁶⁸ A. Shrikumar et al., "Not just a black box: Learning important features through propagating activation differences", *arXiv:1605.01713*.

⁶⁹ M. Sundararajan et al., "Axiomatic attribution for dnn", in *IKML*. JMLR. P.-N. Tan et al.

⁷⁰ S. Bach, A. Binder et al., "On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation", *PloS one*, 10(7):e0130140.

Abbildung 6. Beispiel für Salienzkarten, die von verschiedenen Erklärungsmethoden zurückgegeben wurden. Die erste Spalte enthält das analysierte Bild und das Label, das durch das Blackbox-Modell b des KI-Systems vergeben wurde.



1.2.3.1.7. Prototypbasierte Erklärungen

Eine Erklärung auf der Grundlage von „Prototypen“ gibt ähnliche Muster wie die analysierte Instanz zurück, was die Gründe für die Entscheidung des KI-Systems verdeutlicht. Prototypen werden als Grundlage für die Repräsentation einer Kategorie oder eines Konzepts verwendet.⁷¹ Prototypbasierte Erklärungen können sich auf tabellarische Daten, Bilder und Text beziehen. Bei Li et al.⁷² und Chen et al.⁷³ werden Bildprototypen als Grundlage für das Konzept der Interpretierbarkeit verwendet.⁷⁴ Kim et al.⁷⁵ diskutieren das Konzept der „Kontraprototypen“ für tabellarische Daten, das heißt Prototypen, die zeigen, was anders sein sollte, um eine andere Entscheidung zu erhalten. Exemplare und Kontraexemplare werden von ABELE⁷⁶ verwendet, um die Nutzbarkeit der Erklärung auf der Grundlage einer Salienzkarte zu erhöhen. Exemplare (links) und Kontraexemplare (rechts) für 9 und 0 sind in Abbildung 7 dargestellt.

⁷¹ M. Frixione und A. Lieto, „Prototypes vs exemplars in concept representation“, in *KEOD*, Seiten 226-232.

⁷² O. Li, H. Liu, C. Chen und C. Rudin, „Deep learning for case based reasoning through prototypes: A neural network that explain its predictions“, in *Thirty-second AAAI conference on artificial intelligence*.

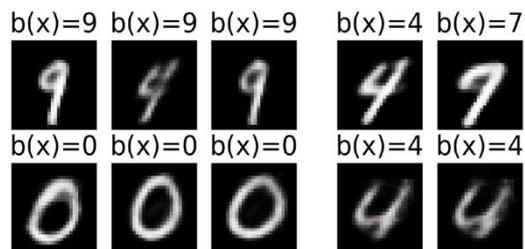
⁷³ C. Chen, O. Li, A. Barnett, J. Su und C. Rudin, „This looks like that: deep learning for interpretable image recognition“, *arXiv:1806.10574*.

⁷⁴ J. Bien und R. Tibshirani, „Prototype selection for interpretable classification“, *The Annals of Applied Statistics*, 5(4):2403-2424.

⁷⁵ B. Kim, O. O. Koyejo und R. Khanna, „Examples are not enough, learn to criticize! criticism for interpretability“, in *Advances In Neural Information Processing Systems*, Seiten 2280-2288.

⁷⁶ R. Guidotti, A. Monreale, S. Matwin und D. Pedreschi (2019b), a. a. O.

Abbildung 7. Beispiel für eine von ABELE zurückgegebene Erklärung von Exemplaren (links) und Kontraexemplaren (rechts). Auf jedem (Kontra-)Exemplar ist das vom Blackbox-Modell b des KI-Systems vergebene Label angegeben.



1.2.3.1.8. Kontrafaktische Erklärungen

Eine „kontrafaktische“ Erklärung zeigt, was anders sein müsste, um die Entscheidung des Blackbox-Modells zu ändern. Die Bedeutung von Kontrafakten liegt darin, dass sie den Menschen bei der Argumentation zu Ursache-Wirkungs-Beziehungen zwischen beobachteten Merkmalen und Klassifikationsergebnissen unterstützen.⁷⁷ Während faktische, direkte Erklärungen wie Entscheidungsregeln und Merkmalbedeutung für das Verständnis der Gründe für ein bestimmtes Ergebnis entscheidend sind, zeigt eine kontrafaktische Erklärung, was sich in einer bestimmten Instanz ändern sollte, um ein anderes Klassifikationsergebnis zu erhalten.⁷⁸ Die oben erwähnte LORE-Methode⁷⁹ bietet zusätzlich zu einer faktischen Erklärungsregel eine Reihe *kontrafaktischer Regeln*. Bei Abbildung 3 ist der Satz kontrafaktischer Regeln violett hervorgehoben und zeigt „Wenn Einkommen ≥ 900 Dann gewähren, oder Wenn Rasse = weiß Dann gewähren“, und stellt damit klar, welche Änderungen die Entscheidung umkehren würden. Die ABELE-Erklärungsmethode⁸⁰ sieht Kontraexemplarbilder vor, die die Ähnlichkeiten und Unterschiede zwischen Instanzen derselben Klasse und anderer Klassen hervorheben.

1.3. KI und XAI im Medienbereich

KI-Technologien verändern und erfinden die Medienindustrie und ihr Marketing neu, insbesondere um die Monetarisierung von Inhalten zu erleichtern und den Endnutzern superpersonalisierte Dienste und Werbung anzubieten. Vor allem im Kino, Fernsehen und Hörfunk, in den Printmedien und der Werbung werden KI-Anwendungen in großem Umfang

⁷⁷ R. M. Byrne, "Counterfactuals in explainable artificial intelligence (xai): evidence from human reasoning", in *Proceedings of the Twenty-Aighth International Joint Conference on Artificial Intelligence, IJCAI-19*, Seiten 6276-6282. A. Apicella, F. Isgrò, R. Prevete und G. Tamburrini, "Contrastive explanations to classification systems using sparse dictionaries" in *International Conference on Image Analysis and Processing*, Seiten 207-218. Springer.

⁷⁸ S. Wachter, B. Mittelstadt und L. Floridi, a. a. O.

⁷⁹ R. Guidotti, A. Monreale, F. Giannotti, D. Pedreschi, S. Ruggieri und F. Turini (2019a), a. a. O.

⁸⁰ R. Guidotti, A. Monreale, S. Matwin und D. Pedreschi (2019b), a. a. O.

genutzt. Laut Chan-Olmsted⁸¹ lassen sich zwei Charakteristika ausmachen, um KI-Anwendungen im Medienbereich zu unterscheiden:

- Einige Anwendungen sind relevanter für das Medienpublikum, die „Nachfrageseite“, andere konzentrieren sich mehr auf interne Strategien der Medienanbieter, die „Angebotsseite“. Manche davon sind wiederum auf beide Gruppen anwendbar, zum Beispiel Publikumsbeteiligung, erweiterte Erlebniserfahrung und Botschaftsoptimierung.
- Einige KI-Anwendungen beziehen sich auf „Inhalte-Ersteller“, während andere zur Analyse und Entdeckung von Inhalten eher für „Inhalte-Distributoren“ relevant sind. Die bei Weitem aktivsten Unternehmen bei der Einführung von KI-Technologien sind Online-Nachrichtendienste von Unternehmen wie *New York Times* und Video-on-Demand-Dienste wie Netflix oder Prime Video. Beispiele für solche KI-Anwendungen sind Empfehlungen, Personalisierung, Monitoring und Zuhören in sozialen Netzwerken, emotionales Tracking und Zugänglichkeit, Videoerstellung und -nachbearbeitung, Informationsverifizierung, prädiktive Erfolgsanalyse, Kundenbeziehungen, automatisierte Entwurfserstellung und Sprachassistenten.

1.3.1. KI-Anwendungen und Erklärbarkeit

Technologieunternehmen wie Amazon, Netflix, Facebook und Google sind führend bei der KI-Expansion im Mediensektor. Zum Beispiel basieren die „Empfehlungssysteme“ von Amazon Prime, Netflix und Spotify auf KI-Methoden.⁸² Eine kürzlich durchgeführte Erhebung zeigt, dass neue Medien KI am häufigsten dazu nutzen, ihre Empfehlungsdienste zu verbessern. Eine weitere Anwendung von KI könnte darin bestehen, die Verbindung zwischen Medien und Publikum neu zu erfinden, das heißt KI könnte eingesetzt werden, um die Gefühle, Vorlieben und Unterhaltungen des Publikums zu verstehen. Dies würde es ermöglichen, das Interesse des Publikums in Echtzeit abzugleichen, um durch personalisierte Medieninhalte ein besseres Nutzungserlebnis zu ermöglichen. Schließlich könnte KI Medienunternehmen dabei helfen, neue Geschäftsmöglichkeiten zu identifizieren: Handlungsstränge oder Charaktere könnten auf der Grundlage der Vorlieben und des Geschmacks der Nutzer, ihrer Meinungen in sozialen Netzwerken, Unterhaltungen usw. geschaffen werden.⁸³ Im folgenden Abschnitt diskutieren wir einige spezifische Anwendungen der KI in der Medienindustrie.

1.3.1.1. Empfehlung

Die auffälligste Anwendung von KI im Medienbereich ist die Empfehlung von Inhalten. Ziel eines „Empfehlungssystems“ ist es, die „Bewertung“ oder „Präferenz“ vorherzusagen, die ein

⁸¹ S. M. Chan-Olmsted, "A review of artificial intelligence adoptions in the media industry", *International Journal on Media Management*, 21(3-4):193- 215.

⁸² S. M. Chan-Olmsted, a. a. O.

⁸³ J. Kietzmann, J. Paschen und E. Treen, "Artificial intelligence in advertising: How marketers can leverage artificial intelligence along the consumer journey", *Journal of Advertising Research*, 58(3):263-267.

Nutzer bestimmten Inhalten im Vergleich zu anderen geben würde. Empfehlungssysteme nutzen in der Regel „kollaboratives Filtern“ und „inhaltsbasiertes Filtern“ sowie andere, zum Beispiel wissensbasierte Systeme.⁸⁴ Empfehlungssysteme sind in vielen Bereichen weit verbreitet, der Medienbereich ist für ihre Nutzung jedoch besonders geeignet. Die Idee dieser Ansätze besteht darin, das bisherige Verhalten eines Nutzers aufgrund zuvor „ausgewählter“ Medieninhalte oder numerischer Bewertungen dieser Inhalte zu modellieren. Außerdem können Empfehlungssysteme auch ähnliche Verhaltensweisen anderer Nutzer berücksichtigen.

Die weit verbreitete Verwendung von Empfehlungssystemen und die Notwendigkeit, das Vertrauen der Nutzer in KI-Systeme zu gewinnen, impliziert, dass jeder Nutzer Zugang zu erklärbaren Empfehlungen haben muss.⁸⁵ Mit anderen Worten, die Empfehlungen müssen nicht nur genau und nützlich, sondern auch verständlich sein. Die relevantesten Arten erklärbarer Empfehlungen für Medien sind nutzerbasierte Erklärungen, merkmalsbasierte Erklärungen und objektbasierte Erklärungen:

- Nutzerbasierte Erklärungen könnten in etwa so aussehen: „Dieser Inhalt wird Ihnen empfohlen, weil ähnliche Nutzer ihn zuvor ausgewählt haben“; sie stützen sich auf eine Reihe (anonymisierter) ähnlicher Nutzer und die von ihnen ausgewählten Inhalte.
- Eine merkmalsbasierte Erklärung wäre: „Dieser Inhalt wird Ihnen empfohlen, weil er durch diese Merkmale (zum Beispiel Merkmale zu Themen, Schauspielern, Musik usw.) beschrieben wird, die Ihnen gefallen“, und bietet die Merkmale entsprechend den Bewertungen, die Sie ihnen zugeordnet haben.
- Schließlich würde eine objektbasierte Erklärung lauten: „Dieser Inhalt wird Ihnen empfohlen, weil er diesen anderen Inhalten, die Ihnen bisher gefallen haben, ähnlich ist“.⁸⁶

Frühe Empfehlungsmodelle wie objekt-/nutzerbasierte Modelle sind transparent und erklärbar. Die Verstärkung von Transparenz wurde als ein entscheidender Aspekt erkannt, wenn es um die Erhöhung der Vertrauenswürdigkeit, Effektivität, Überzeugungskraft, Effizienz und Zufriedenheit des Endnutzers geht.⁸⁷ Jüngste Fortschritte in der KI und die Verwendung von ANN haben dazu beigetragen, die Präzision der Empfehlungen zu verbessern, haben jedoch durch die Verwendung komplexer, undurchsichtiger Modelle wie ANN Transparenz vollständig eliminiert. Die mangelnde Erklärbarkeit von Empfehlungssystemen in der Medienindustrie kann zu vielen Problemen führen. Ohne die Nutzer wissen zu lassen, warum bestimmte Ergebnisse bereitgestellt werden, sinkt gegebenenfalls die Effizienz des Systems, die Nutzer zu einem bestimmten Inhalt zu lenken, was die Vertrauenswürdigkeit des Systems weiter verringern kann.

⁸⁴ C. D. Manning, P. Raghavan und H. Schütze, *Introduction to information retrieval*, Cambridge university press.

⁸⁵ Y. Zhang und X. Chen, „Explainable recommendation: A survey and new perspectives“, *arXiv preprint arXiv:1804.11192*.

⁸⁶ Der interessierte Leser kann in Zhang Y. und Chen X., a. a. O., weitere relevante Arten erklärbarer Empfehlungen finden.

⁸⁷ N. Tintarev und J. Masthoff, „A survey of explanations in recommender systems“, in *2007 IEEE 23rd international conference on data engineering workshop*, Seiten 801-810. IEEE.

1.3.1.2. Personalisierung und Kundenbezogenheit

Personalisierung beim Vorschlagen und Kuratieren von Medieninhalten ist ein grundlegender Aspekt, der in der Medienindustrie mit KI angegangen wird. KI-Systeme in der Medienindustrie zeichnen sich in der Tat dadurch aus, dass sie dank Empfehlungssystemen die Strategie zur Inhalteverbreitung präzise auf den Kunden zuschneiden. Beispielsweise können KI-Systeme Trends in sozialen Netzwerken analysieren, um die besten zu verbreitenden Inhalte zu ermitteln. Eine weitere Anwendung ist Publikumsanalyse für eine automatische Generierung von Titeln/Zusammenfassungen/Illustrationen mit Schlüsselwörtern, die eine höhere Sichtbarkeit des Inhalts garantieren. Darüber hinaus kann KI die Generierung und Kuratierung von Medieninhalten automatisieren, themenbasierte Wiedergabelisten regelmäßig aktualisieren und Nutzerprofile erstellen, um kundenbezogene Empfehlungen zu geben. Auf diese Weise können die Medieninhalte, die jedem Nutzer angeboten werden, unterschiedlich sein und auf das Profil jedes Nutzers, seine Bewegungen oder den Zeitpunkt und Ort der Mediennutzung zugeschnitten werden. Andere Anwendungen sind darauf ausgerichtet, den Nutzer dazu zu bringen, sich mit dem richtigen Inhalt zu befassen, der im richtigen Format zum richtigen Zeitpunkt und auf vollständig personalisierte Weise angeboten wird. Es ist, als hätte man für jeden Einzelnen einen persönlichen Lektor, der das perfekte Leseerlebnis kuratiert.

1.3.1.3. Erstellung von Inhalten

Wie bereits erwähnt, ist eine der jüngsten Anwendungen von KI die Erstellung von Nachrichten, Musik und Videos. Insbesondere geht es hier um den Begriff „Roboterjournalismus“ oder „automatisierter Journalismus“. Dabei verwenden KI-Systeme Algorithmen zur Erzeugung natürlicher Sprache, um Daten und Wissen in Nachrichten, Bilder und Videos zu verwandeln. Beispielsweise können KI-Systeme auf der Grundlage früherer Artikel und verfügbarer Daten leicht Artikel schreiben, die für Menschen relativ langweilig zu verfassen sind, wie zum Beispiel Artikel über das Wetter oder Finanzberichte. Im Hinblick auf Videos kann KI durch die Nutzung der Bilderkennung kohärente Videomontagen produzieren. Die meisten großen Herausgeber von Schnittsoftware haben bereits automatische Videoverarbeitungsfunktionen hinzugefügt, um Redakteuren Zeit zu sparen. Andere Softwareprogramme wie Gingalab⁸⁸ verwenden KI, um aufgrund vordefinierter redaktioneller Richtungen (zum Beispiel Humor, Spannung, Fokus auf einen Protagonisten usw.) automatisierte „Best-of“-Videos zu erstellen. Im September 2018 strahlte die BBC eine Sendung aus, die vollständig von einem Roboter erstellt wurde.⁸⁹

⁸⁸ Abrufbar unter <https://gingalab.com/>.

⁸⁹ Abrufbar unter <https://www.bbc.co.uk/programmes/b0bhwk3p#:~:text=Made%20by%20Machine%3A%20When%20AI%20Met%20the%20Archive.Dokumentarfilm.>

1.3.1.4. Erkennung gefälschter Inhalte

Der Schwachpunkt dieser unglaublichen Errungenschaft der KI, das heißt ihrer „Schöpfungskraft“, ist das *Deepfake*-Phänomen (perfekte Täuschung).⁹⁰ KI kann zwar gefälschte Medieninhalte erzeugen, sie kann aber glücklicherweise auch zur Aufdeckung gefälschter Inhalte beitragen. Tatsächlich kann KI ein entscheidender Trumpf bei der Bekämpfung von Fehlinformationen sein, da dieselbe Technologie, die zur Herstellung einer Fälschung verwendet wird, auch zu deren Aufdeckung genutzt werden kann. Durch umfangreiche analytische Fähigkeiten und Algorithmen maschinellen Lernens kann KI die Verifizierung von Medieninhalten wie Nachrichten, Bildern und Videos teilweise automatisieren. Das Hauptproblem besteht darin, dass die Qualität der Erkennung aus der Erfahrung der KI resultiert, die mit der Verfügbarkeit von Datenquellen zur Unterscheidung zwischen echten und gefälschten Inhalten zusammenhängt. Außerdem muss eine solche Informationsquelle von Menschen generiert werden, die Medieninhalte manuell als gefälscht oder nicht annotieren. Dieser manuelle Schritt kann zu einer Verzerrung der Daten führen, da Menschen möglicherweise nicht in der Lage sind, alle Medieninhalte zu verifizieren, die für das Training eines voll funktionsfähigen KI-Systems erforderlich sind, und sich auf ihre Gefühle verlassen müssen, bei der Beurteilung was echt und was gefälscht ist. XAI kann in dieser Phase des Trainings aus zwei Gründen entscheidend sein: Erstens wollen die Nutzer von KI zur Erkennung von gefälschten Inhalten sicher sein, dass die Logik, die bei der Erkennung des gefälschten Inhalts verfolgt wird, für den Menschen nachvollziehbar ist. Die Erwartung lautet ungefähr: „Diese Nachricht ist gefälscht, weil sich Sätze zu oft wiederholen und die angezeigten Bilder von bestehenden Websites stammen.“ Zweitens dürfen sich KI-Systeme bei der Bereitstellung von Vorschlägen nicht auf einen verzerrten Datenbestand stützen. Wenn alle realen Nachrichten aus ein und derselben Quelle stammen, könnte die Erklärung ungefähr lauten: „Diese Nachricht ist gefälscht, weil sie von der *New York Times* nicht verbreitet wird.“

1.3.1.5. Weitere Anwendungen

KI und XAI können für viele andere Anwendungen in der Medienindustrie eingesetzt werden. Im folgenden Abschnitt werden einige von ihnen genannt, ohne auf Einzelheiten einzugehen. KI kann als Tool zur Verbesserung von Unterhaltungen im Internet eingesetzt werden, mit anderen Worten, um Hassreden, Diskriminierung, Trolle usw. zu erkennen. Auch in diesem Fall ist es von entscheidender Bedeutung, die Gründe zu erfahren, aus denen unangemessene Posts als solche erkannt werden. KI zur Spracherkennung ist eine Grundlage für viele moderne Dienste und Sprachassistenten wie Amazons Alexa, Google Home oder Apples Siri, die in jedem intelligenten Gerät vorhanden sind. Sie nutzen KI und natürliche Sprachverarbeitung, um unsere Fragen zu beantworten und unsere Anweisungen auszuführen. Schließlich ist es erwähnenswert, dass KI strategische Auswirkungen auf die Monetarisierung und die Vorhersage des Erfolgs von Medieninhalten hat. Gleichzeitig wird XAI in Bezug auf Medienethik entscheidend, um dem Publikum auf transparente Weise die Logik der KI-Systeme zu kommunizieren, die mit den Nutzern interagieren oder

⁹⁰ Für eine Definition von *Deepfakes* siehe oben 1.1.3.

Entscheidungen für sie treffen. Bestimmte Fragen könnten auftauchen, und über XAI können Nutzer diese Fragen möglicherweise beantwortet bekommen, zum Beispiel: Was ist das richtige Verhältnis zwischen Personalisierung und Entdeckung von Inhalten? Welchen Grad an Empfehlung wollen wir? Warum werden diese Medieninhalte als real angesehen? Im Rahmen der DSGVO besteht der erste Schritt der Medienindustrie darin, deutlich zu machen, welche Inhalte von einer KI empfohlen/erstellt werden.

1.3.2. Video-on-Demand-Dienste in der Praxis

Video-on-Demand-Dienste haben die Art und Weise verändert, wie wir Medieninhalte schauen, von Fernsehserien und Comedy-Shows bis hin zu Spiel- und Zeichentrickfilmen. Diese Dienste passen das Nutzererlebnis algorithmisch durch eine starke Personalisierung an, die auf einem großen Bestand an Metadaten (einschließlich Genre, Kategorien, Besetzung und Erscheinungsdatum), auf Daten zum Nutzerverhalten (wie Suche, Browsen, Bewertung und Gerätetyp), aber auch auf den für die Homepage ausgewählten Vorschlagsleisten, den für diese Leisten ausgewählten Titeln, dem Bildmaterial für jeden Film, den Filmen in der Wiedergabeliste usw. basiert. Das verwendete KI-System entwickelt sich weiter, sammelt ständig die personenbezogenen Daten jedes Nutzers und bietet immer eine maßgeschneiderte Visualisierung der Optionen, die der Benutzer je nach Nutzung und Kontext am ehesten anklicken wird. Das Endziel besteht darin, die beste Kombination von Inhalten zu finden, die die Nutzer zufriedenstellen kann, anstatt Inhalte zu finden, die einfach den meisten Nutzern entsprechen. Algorithmen untermauern somit eher Kreativität und Vielfalt und nicht Standardisierung. Diese bemerkenswerten Funktionen werden auf der Grundlage der KI-Empfehlungssysteme angeboten, welche die Daten von Millionen von Nutzern sammeln, die die Inhalte auf diesen Plattformen schauen und bewerten.

Niemand weiß genau, wie dieses Empfehlungssystem funktioniert. Video-on-Demand-Dienste stellen in der Regel eine Beschreibung ihres Empfehlungssystems in Klartext zur Verfügung, Details der Entscheidungsfindung werden jedoch nicht preisgegeben. In diesem Sinne sind die KI-Empfehlungssysteme Blackbox-Modelle *par excellence*. Theoretisch ist ein Nutzer vielleicht gar nicht daran interessiert, wie die Empfehlungen ablaufen, weil er sich einfach bei einigen angenehmen Medieninhalten entspannen möchte. Die Empfehlungen sind jedoch möglicherweise nicht gänzlich personenbezogen, sondern werden durch Marketingstrategien oder - noch schlimmer - durch Verzerrungen in den für die Modelle des maschinellen Lernens verwendeten Daten kanalisiert. Wenn sie durch die Anwendung von KI, die undurchsichtige, auf maschinellen Lernmodellen basierende Empfehlungsmethoden nutzt, unterstützt werden, beruhen Entscheidungen oder Vorhersagen auf dem durch automatisierte Prozesse erzielten Lernen, und der verfügbare oder ausgewählte Trainingsdatenbestand repräsentiert gegebenenfalls nicht die Bevölkerung, für deren Bewertung er konzipiert wurde. Beispielsweise könnten Statistiken, die auf Menschen basieren, die Filme mit asiatischen Helden meiden, zu einer Diskriminierung dieser Filmkategorie führen und sie fälschlicherweise mit einer niedrigen Punktzahl für eine Bevölkerung bewerten, die dennoch an dieser Art von Medieninhalten interessiert sein könnte. Daher ist die Verwendung von XAI zum Verständnis des Trainingsdatenbestands und zur Analyse, wie die Daten die Ergebnisse für verschiedene

Bevölkerungen beeinflussen, von entscheidender Bedeutung für die Identifizierung von Verzerrungen. Für die Video-on-Demand-Dienste könnte eine globale Erklärung beschreiben, wie sich der Algorithmus im Allgemeinen verhält. Es lässt sich zum Beispiel erkennen, dass das KI-Empfehlungssystem an einem Wochentag nicht kurz vor Mitternacht einen dreistündigen Film vorschlägt. Auf der anderen Seite könnte eine lokale Erklärung beschreiben, wie sich die KI bei einer bestimmten Person verhält. Wenn die analysierte Kundin beispielsweise in ihrer Mittagspause in der Regel von 12:00 bis 14:00 Uhr Video-on-Demand-Inhalte schaut, wird ihr in diesem Zeitfenster kein dreistündiger Film vorgeschlagen. Der Grund dafür ist, dass der Video-on-Demand-Dienst möglicherweise zu dem Schluss gekommen ist, dass dies die beste, auf Routine basierende Vorgehensweise ist, auch wenn der dreistündige Film perfekt den Interessen der Nutzerin entspricht. Andererseits möchte die Nutzerin vielleicht nicht, dass der Dienst diese Art von persönlichen Informationen nutzt, um Empfehlungen auszusprechen. Ausführliche Erklärungen würden Video-on-Demand-Dienste also helfen, größeres Vertrauen ihrer Nutzer zu gewinnen. Theoretisch sollten bei jeder Anwendung im Medienbereich, die KI nutzt, Vorschläge auf unvoreingenommenen Empfehlungen und einer vertrauensvollen Beziehung zwischen dem Dienst und seinen Nutzerinnen und Nutzern beruhen.

1.4. Fazit

Künstliche Intelligenz kann nicht die endgültige Lösung für jede Anwendung sein, und insbesondere im Medienbereich muss sie an einen Menschen gebunden sein, sowohl bei der Erstellung und Überprüfung von Medieninhalten als auch bei der Betrachtung empfohlener Medieninhalte. KI ist in der Tat von grundlegender Bedeutung für die Nachfrageseite, für den Zugang zu Inhalten und für die Monetarisierung. KI hat ein großes Potenzial für das Gemeinwohl, wenn sie dabei hilft, durch Massen von Inhalten zu navigieren, indem sie Suchvorgänge und personalisierte Empfehlungen optimiert und Manipulation verhindert. Mit den geeigneten XAI-Tools und dem nötigen Vertrauen seitens des Publikums und der Anbieter könnte KI die Medienindustrie und alle damit zusammenhängenden Sektoren und Anwendungen wirksam ankurbeln.

Big Data

Auch die Sammlung und Verwendung von – bewusst oder versehentlich zur Verfügung gestellten – personenbezogenen Daten durch Dritte kann drastische Folgen für das Privatleben der Betroffenen haben. Darüber hinaus gibt es Situationen, in denen der Staat oder Private einen Einblick in das Leben von Nutzerinnen oder Nutzern verlangen, der über das hinausgeht, was diese zu akzeptieren bereit sind. Andrea Pin stellt in seinem Beitrag zu dieser Publikation fest, dass „sich der mediale Raum aufgrund der heutigen weiten Verbreitung der KI seiner einzigartigen Rolle bewusst werden muss“ und dass sich der Mediensektor „um einen rechtmäßigen, ethischen und robusten KI-Einsatz bemühen“ sollte. Ein besonderes Anliegen ist die angemessene Rolle der Medienplattformen bei der Verwaltung ihrer Inhalte. Es wird diskutiert, inwieweit sie sich „nicht mit einer rein passiven Rolle begnügen“ sollten, um das „lohnende ethische Ziel“ zu verfolgen, „ihre Inhalte zu überwachen“. In diesen Fällen ist gerade ihr „Mangel an Menschlichkeit“ einer der größten Nachteile der KI. Filteralgorithmen können potenziell schädliche Inhalte äußerst effizient finden und entfernen, aber differenzierte Entscheidungen zu komplexen Rechtsgebieten können nur Menschen treffen.⁹¹

⁹¹ Barker A., Murphy H., „YouTube reverts to human moderators in fight against misinformation“, *Financial Times*, 20. September 2020, <https://www.ft.com/content/e54737c5-8488-4e66-b087-d1ad426ac9fa>.

2. Der Stoff, aus dem KI-Träume sind – Big Data

Andrea Pin, außerordentlicher Professor für vergleichendes öffentliches Recht, Universität Padua

2.1. Einführung

Immer wieder hört man, Big Data sei das Öl der KI-Revolution.⁹² Seitdem Datenwissenschaft und Technologieentwicklung ihre Kräfte gebündelt haben, wird die Welt von einem massiven Informationsfluss geflutet, der Auswirkungen auf unser Leben und auf unser Verständnis von Politik, Wirtschaft und Kultur hat. Dank der Fähigkeiten der KI hat das Phänomen Big Data enormen und wohl auch dauerhaften Einfluss darauf, wie Einzelne und Gruppen Pläne machen, Informationen über sich und die Welt beschaffen, Unterhaltung genießen und Geselligkeit pflegen.

Die heutigen Computer sind technisch sehr leistungsstark. Ihre Algorithmen sind extrem ausgeklügelt. Ihre neuronalen Netze bilden die geistigen Prozesse des Menschen nach und ermöglichen ihnen komplexe Analysen. Durch die Verarbeitung von Big Data können Unternehmen die Entscheidungen und Präferenzen der Kundinnen und Kunden so früh antizipieren, dass sie deren Wünsche vorhersagen können, bevor sie ihnen selbst überhaupt bewusst werden. Dank Big Data erfolgen Geschäftsprozesse nun immer öfter „proaktiv“ statt „reaktiv“.⁹³

Das Internet spielt in diesem Szenario eine grundlegende Rolle. Da Menschen das Internet praktisch pausenlos nutzen, um Informationen zu teilen, sogar über sich selbst und ihr Leben, sammelt das Web das Rohmaterial, mit dessen Hilfe die KI dann Schlüsse zieht, Vermutungen anstellt und Antworten auf Anfragen findet. Der Oxford-Philosoph Luciano Floridi prägte den Begriff „Onlife“, um zu beschreiben, wie häufig und unbewusst Menschen zwischen der realen Welt und der Online-Welt hin- und herwechseln.⁹⁴

⁹² Pan S. B., „Get to know me: Protecting privacy and autonomy under big data’s penetrating gaze“, *Harvard Journal of Law and Technology* 30, 2016, S. 239,

<https://jolt.law.harvard.edu/assets/articlePDFs/v30/30HarvJLTech239.pdf>; Surden H., „Artificial intelligence and law: An overview“, *Georgia State University Law Review* 35, 2019, S. 1311 und 1315.

⁹³ Microsoft Dynamics 365, *Delivering personalized experiences in times of change*, 2007, S. 3,

<https://www.hso.com/wp-content/uploads/2020/03/Digitally-transforming-customer-experiences-ebook.pdf>.

⁹⁴ Floridi L., „Soft ethics and the governance of the digital“, *Philosophy & Technology* 31, 1, 2018, S. 1.

Dieses Phänomen eskaliert. Im Jahr 2023 wird es schätzungsweise mehr als fünf Milliarden Internetnutzerinnen und -nutzer und 3,6 Geräte pro Kopf geben, und 70% der Weltbevölkerung werden über eine mobile Verbindung verfügen.⁹⁵ Je stärker die Welt vernetzt ist, desto mehr Daten werden anfallen. Es ist kein Zufall, dass eines der derzeit umstrittensten Themen die Einführung von 5G-Netzen ist, denn sie können ihren Eigentümern einen erheblichen Informationsvorteil verschaffen.

Der Medienbereich und die Medienwirtschaft sind in diesem Szenario wichtige Akteure. Ihre Aufgabe besteht seit jeher darin, Informationen zu sammeln, zu verarbeiten und zu verbreiten. Dank Big Data können sie jetzt ein Profil ihres Publikums erstellen und erfahren, was es erwartet, wie man Nachrichten formuliert oder eine Geschichte erzählt oder wie ein gutes Finale für einen bestimmten Film aussähe. Big Data ermöglicht durch die Identifizierung potenzieller Nachrichtenleser oder Kinobesucher eine individuelle Anpassung des Angebots, da „Computer präziser als Menschen sind, wenn es darum geht, anhand von ‚digitalen Fingerabdrücken‘ Charaktereigenschaften [oder] politische Einstellungen [...] vorherzusagen“.⁹⁶

Die Neuerungen, die Big Data mit sich bringt, verändern auch die Medienlandschaft. „... Fernsehen/Filme/Musik in digitaler Form und eine Unzahl von Online-Vertriebsmodellen stellen die etablierten Verbreitungswege (CDs, Kabel) seit Jahren vor eine Herausforderung ... Online-Publisher werten aus, was Konsumentinnen und Konsumenten lesen, wo sie sich aufhalten und welche sozialen Signale sie senden – beispielsweise geteilte Artikel und Trendthemen auf Facebook und Twitter –, um personalisierte, relevante Inhalte anzubieten, die aber nicht zu repetitiv und vorhersehbar sein dürfen. Damit automatisieren und übertreffen sie, was menschliche Redakteure leisten können“.⁹⁷ Traditionelle Medien konkurrieren heute bei der Nachrichtengenerierung mit nicht-professionellen Informationsanbietern, die das Web auf Nachrichten durchsuchen, oder mit Bloggern, die ihre Ansichten über Social-Media-Plattformen teilen, wo Verbreitung und Konsum von Inhalten praktisch nicht zu unterscheiden sind.⁹⁸

Dieses Kapitel behandelt die wichtigsten rechtlichen Konsequenzen eines solchen globalen Wandels in der Medienwelt. Zunächst geht es auf die entscheidende Frage nach dem Schutz der Privatsphäre ein. Anschließend befasst es sich mit den potenziellen Diskriminierungen und Verzerrungen, die bei einer auf Big Data gestützten Strategie drohen, und mit den Risiken von Fehlinformation, politischer Polarisierung und der

⁹⁵ Cisco, *Cisco Annual International Report (2018–2023) White Paper*, 9 March 2020,

https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html?fbclid=IwAR31-e732ws1p1cIW5PYHQjVOJkPSzV0dGt3sq_qkX_P8wb9O4Yn0Ez0a0Y.

⁹⁶ Europäischer Datenschutzbeauftragter, Stellungnahme 7/2015, Bewältigung der Herausforderungen in Verbindung mit Big Data, 19. November 2015, S. 16, https://edps.europa.eu/sites/edp/files/publication/15-11-19_big_data_de.pdf.

⁹⁷ Byers A., „Big data, big economic impact“, 10, 2015,

https://kb.osu.edu/bitstream/handle/1811/75420/ISJLP_V10N3_757.pdf?sequence=1&isAllowed=y. Siehe auch Bruckner M. A., „The promise and perils of algorithmic lenders’ use of big Data“, *Chicago-Kent Law Review* 93, 2018, S. 8, <https://scholarship.kentlaw.iit.edu/cklawreview/vol93/iss1/1/> oder Ambrose M. L., „Lessons from the Avalanche of Numbers: Big Data in Historical Perspective“, *ISJLP*, 11, 2015, S. 213 (“Netflix predicts our movies”).

⁹⁸ Perritt H. H. Jr., „Technologies of storytelling: New models for movies“, *Virginia Sports & Entertainment Law Journal*, 10, 2010, S. 153, http://blogs.kentlaw.iit.edu/perrittseminar/files/2016/07/perritt-technologies-of-storytelling-Westlaw_Document_05_56_44.pdf.

Entwicklung des Medienbereichs zu einem Massenüberwachungssystem. Es folgt ein Überblick über die Veränderung der Medienmärkte und -strategien vor dem Hintergrund der Big-Data-Dynamik. Abschließend wird kurz auf die Debatten über den korrekten regulatorischen Umgang mit Big Data eingegangen.

Insgesamt hat die Notwendigkeit einer KI-Regulierung über die Jahre immer mehr Zuspruch gefunden. Obwohl Technologien global sind und keine Grenzen kennen, gehen Zweck, Ansatz und Rahmen der Regulierung zwischen den großen rechtlichen Akteuren in diesem Szenario – USA, EU und China – weit auseinander. Die USA wollen mit ihrem Ansatz sicherstellen, dass von massivem KI-Einsatz geprägte Märkte offen und effizient bleiben, das vorrangige Anliegen der EU scheint in der Wahrung der Menschenwürde zu bestehen, und China ist vor allem an sozialem Frieden, Stabilität und einer geordneten Entwicklung seiner Wirtschaft gelegen. Zu jedem dieser Ansätze gehört ein spezifischer rechtlicher Umgang mit Big Data.

2.2. Datenschutz als Gatekeeper für Big Data

Es mehren sich die Befürchtungen, dass sich von Big Data getriebene Tools zu einem allgegenwärtigen System der Massenüberwachung und -manipulation verbinden könnten. Einer der wichtigsten Schutzmechanismen gegen diese Bedrohung ist der Datenschutz. Viele Länder und supranationale Rechtssysteme haben Vorschriften erlassen, die begrenzen und überwachen, wie welche Informationen gesammelt und verarbeitet werden – auch mit dem Ziel, Big-Data-Analysen einzuschränken und soziale Verwerfungen zu verhindern. In dieser Hinsicht dienen Datenschutzgesetze als Schutzschild gegen Grenzüberschreitungen durch Big Data.

2.2.1. Vereinigte Staaten von Amerika

Die westliche Welt ist beim Verständnis und beim Schutz der Privatsphäre gespalten. Die Ansätze von USA und EU sind bei Weitem nicht auf einer Linie. Trotz ihrer historischen Sensibilität für den Schutz der Privatsphäre fehlt in den USA eine umfassende Regulierung der Sammlung und Erfassung von Informationen im Internet. Es bestehen mehrere Rechtssysteme nebeneinander, die jeweils einen bestimmten Bereich regeln, aber keine umfassende landesweite Regulierung.⁹⁹ Der US-Ansatz sieht Informationen jedoch in der Regel als riesigen neuen Markt, der sich positiv auf die Volkswirtschaft auswirkt. Einige Bundesstaaten, allen voran Kalifornien, haben begonnen, Gesetze zum Schutz und zur Regulierung der Privatsphäre umzusetzen, doch die Sammlung und Erfassung personenbezogener Daten ist weitgehend erlaubt und wird sogar gefördert. Ein recht allgemeiner rechtlicher Ausgangspunkt ist, dass die betroffenen Personen, die ihre Daten übermitteln, lediglich *wissen* müssen, dass ihre Informationen auf verschiedene Weise

⁹⁹ Houser K. A. & Voss W. G., „The end of Google and Facebook or a new paradigm in data privacy“, *Richmond Journal of Law and Technology*, 25, 2018, S. 18, https://jolt.richmond.edu/files/2018/11/Houser_Voss-FE.pdf.

verarbeitet werden, auch für die Erstellung von Profilen und den Handel mit ihren Präferenzen. Da die meisten Protagonisten der KI-basierten globalen Industrie in den USA ansässig sind, können sie die Vorteile des „neuen Öls“ mit diesem günstigen Regulierungsrahmen voll ausschöpfen.

2.2.2. Europäische Union

Der Schutz der Privatsphäre richtet sich in der Europäischen Union nach der Datenschutz-Grundverordnung (DSGVO),¹⁰⁰ die am 27. April 2016 verabschiedet wurde und am 25. Mai 2018 in Kraft trat. Die DSGVO selbst ist der Höhepunkt eines längeren Prozesses, der den Schutz personenbezogener Daten im Laufe der Jahrzehnte verbessert hat. Sie steht für einen ganz anderen Weg als der US-Ansatz. Obwohl sich die Europäische Union dazu bekennt, „Unternehmen und Behörden den Zugang zu hochwertigen Daten zu erleichtern, um das Wachstum zu steigern und Wert zu schaffen“,¹⁰¹ beruht ihre Haltung insgesamt darauf, dass personenbezogene Daten keine Handelsware sein sollen.¹⁰² Der rechtliche Ausgangspunkt der DSGVO ist, dass eine betroffene Person ihre *Einwilligung* zur Datenverarbeitung geben muss.¹⁰³ Die Einwilligung selbst muss unmissverständlich, freiwillig und in informierter Weise erfolgen:¹⁰⁴ Die betroffene Person muss über den Umfang und den Zweck der Verarbeitung informiert werden.¹⁰⁵ Der Schutz der DSGVO erstreckt sich auf die Daten von EU-Bürgerinnen und -Bürgern wie auch von anderen natürlichen Personen, sofern die Verarbeitung innerhalb der EU erfolgt. Mit anderen Worten, sie schützt jeden Menschen innerhalb ihres Gebiets.¹⁰⁶

Die Kluft zwischen dem US- und dem EU-Ansatz hat zu einem Bruch im transatlantischen Datenaustausch geführt. Die DSGVO ist sehr konservativ, was den Austausch von Informationen betrifft, die innerhalb der Europäischen Union gesammelt werden, und verlangt, dass jeder Datentransfer über die EU-Grenzen nach außen den EU-

¹⁰⁰ Konsolidierter Text: Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung) (Text von Bedeutung für den EWR), <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32016R0679>.

¹⁰¹ European Data Protection Supervisor, Opinion 3/2020 on the European strategy for data, 16 June 2020, S. 4, https://edps.europa.eu/sites/edp/files/publication/20-06-16_opinion_data_strategy_en.pdf. Siehe auch Rat der Europäischen Union, Gestaltung der digitalen Zukunft Europas – Schlussfolgerungen des Rates, 9. Juni 2020, <https://data.consilium.europa.eu/doc/document/ST-8711-2020-INIT/de/pdf>.

¹⁰² Europäischer Datenschutzausschuss, Leitlinien 2/2019 zur Verarbeitung personenbezogener Daten gemäß Artikel 6 Absatz 1 Buchstabe b DSGVO im Zusammenhang mit der Bereitstellung von Online-Diensten für betroffene Personen, Version 2.0, 8. Oktober 2019, Nr. 54, https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_de.pdf

¹⁰³ Artikel 6 DSGVO.

¹⁰⁴ Manheim K. & Kaplan L., „Artificial intelligence: Risks to privacy and democracy“, *Yale Journal of Law & Technology*, 106, 2019, S. 1069, https://yjolt.org/sites/default/files/21_yale_j.l.tech.106_0.pdf.

¹⁰⁵ Artikel 6 Absatz 4 und Artikel 7 DSGVO.

¹⁰⁶ Europäischer Datenschutzbeauftragter, Stellungnahme 3/2018 Stellungnahme des EDSB zu Online-Manipulation und personenbezogenen Daten, 19. März 2018, S. 14, https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_de.pdf.

Standards entspricht.¹⁰⁷ Die EU-Regulierungsphilosophie wird als so datenschutzfreundlich empfunden, dass viele Nicht-EU-Bürger dazu neigen, Unternehmen mit Sitz in der EU gegenüber Unternehmen zu bevorzugen, die nicht der Rechtshoheit der Europäischen Union unterworfen sind. Die Einhaltung der DSGVO ist daher für Unternehmen, die im KI-Bereich tätig sind, auch außerhalb der Europäischen Union zu einem Reputationsplus geworden, das sie dazu drängt, die Vorschriften zum Schutz der Privatsphäre von sich aus umzusetzen.¹⁰⁸

Ein derart hoher Schutz der Privatsphäre durch die DSGVO ist jedoch mit Kosten verbunden. Dass Unternehmen die Einwilligung der Internetnutzerinnen und -nutzer einholen müssen, die ihre Website besuchen, führt zu einer Fülle wiederkehrender und manchmal obskurer Einwilligungsaufforderungen, die traditionell auftauchen, sobald eine Website angezeigt wird.¹⁰⁹ Dieses Phänomen überschwemmt das Internet so sehr, dass die meisten Nutzerinnen und Nutzer einfach auf „Ja“ klicken und mit der Navigation auf der Website fortfahren, ohne darauf zu achten, wie ihre Informationen gesammelt, verarbeitet und verbreitet werden.¹¹⁰ Dieses Vorgehen ist natürlich riskant, aber verständlich. Manche schätzen, dass ein normaler Mensch – weder eine qualifizierte Rechtsanwältin noch ein obsessiv penibler Internetnutzer – 76 Arbeitstage pro Jahr damit verschwenden würde, alle Datenschutzhinweise zu lesen, die ihm eingeblendet werden.¹¹¹ Zu viel Datenschutz kann kontraproduktiv sein, denn ein Mensch kann den gesamten Schutz verlieren, wenn er seine Einwilligung vorschnell gibt und damit eine massive Abschöpfung seiner Daten erlaubt.

Darüber hinaus können die Möglichkeiten der Big-Data-Analyse den Schutz der Privatsphäre, den die DSGVO gewährt, an vielen Fronten schwächen. Erstens sieht die DSGVO für anonymisierte Daten geringere Einschränkungen vor, da die Anonymisierung die Privatsphäre angeblich schützt. Dank zunehmender KI-Fähigkeiten wird der „Rückschluss auf die Identität einer Person durch das Kombinieren von angeblich ‚anonymen‘ Daten mit anderen Datensätzen, einschließlich öffentlich, beispielsweise auf sozialen Medien, zugänglicher Informationen, [...] immer einfacher.“¹¹² „Je größer und umfassender“ eine Datensammlung wird, desto wahrscheinlicher ist es, dass eine Person, deren Daten anonymisiert wurden, wieder identifizierbar wird.¹¹³

Darüber hinaus verlangen die EU-Datenschutzbestimmungen, dass ein Mensch detaillierte Informationen über Zweck und Umfang der Verarbeitung der von ihm übertragenen Daten erhält. Durch neuronale Netze und Deep Learning ziehen KI-basierte

¹⁰⁷ Artikel 45 DSGVO.

¹⁰⁸ Moerel L. & Lyon C., „Commoditization of data is the problem, not the solution – Why placing a price tag on personal information may harm rather than protect consumer privacy, *Future of Privacy Forum*, 24. Juni 2020, <https://fpf.org/2020/06/24/commoditization-of-data-is-the-problem-not-the-solution-why-placing-a-price-tag-on-personal-information-may-harm-rather-than-protect-consumer-privacy>.

¹⁰⁹ Europäischer Datenschutzbeauftragter, Stellungnahme 7/2015, Bewältigung der Herausforderungen in Verbindung mit Big Data, *op. cit.*, S. 11.

¹¹⁰ Tsesis A., „Marketplace of ideas, privacy, and the digital audience“, *Notre Dame Law Review*, 94, 2019, S. 1590, <https://scholarship.law.nd.edu/cgi/viewcontent.cgi?article=4845&context=ndlr>.

¹¹¹ Hartzog W., *Privacy's blueprint*, Harvard University Press, 2018.

¹¹² Europäischer Datenschutzbeauftragter, Stellungnahme 4/2015, Der Weg zu einem neuen digitalen Ethos, 11. September 2015, S. 6, https://edps.europa.eu/sites/edp/files/publication/15-09-11_data_ethics_de.pdf.

¹¹³ Europäischer Datenschutzbeauftragter, Stellungnahme 7/2015, Bewältigung der Herausforderungen in Verbindung mit Big Data, *op. cit.*, S. 15.

Systeme Schlussfolgerungen, die selbst Software-Entwickler nicht unbedingt vorhersehen können. Gerade diese Möglichkeit von Big Data gefährdet die Auslegung der EU-Datenschutzvorschriften. Da die Big-Data-Verarbeitung Ergebnisse liefert, die nicht vollständig vorhersehbar sind, ist es äußerst schwierig, einem Menschen ein detailliertes Bild davon zu vermitteln, wofür seine Informationen verwendet werden.¹¹⁴

2.2.3. China

In China beziehen öffentliche und private Institutionen riesige Datenmengen aus einer Fülle von Quellen, um Personenprofile von höchster Genauigkeit zu erstellen. Das Sammeln und Verarbeiten personenbezogener Daten über die chinesische Bevölkerung ist von entscheidender Bedeutung für Chinas großen Plan zur Umsetzung eines weitreichenden Überwachungs- und Kontrollprogramms, das mit KI-Einsatz Profile von Einzelnen und Gruppen erstellt und deren Verhalten vorhersagt.¹¹⁵ Übergeordnetes Ziel dieses Plans ist der Aufbau eines allgegenwärtigen Sozialkreditsystems – eines KI-basierten Mechanismus, der Informationen aus persönlichen Aufzeichnungen, Smartphones und Massenüberwachungssystemen sammelt und dann ein Ranking der Personen erstellt und ihnen anhand ihres bisherigen Verhaltens Privilegien und Rechte verleiht.¹¹⁶

In China arbeiten öffentliche Institutionen daran, dass das Leben aller Menschen nicht privat ist, sondern transparent. Hierbei kooperieren sie auch mit chinesischen Privatfirmen. So fungiert eine Handvoll großer Technologieunternehmen wie WeChat und Alibaba als digitale Drehscheibe für das Leben der chinesischen Bürgerinnen und Bürger.¹¹⁷ Die Chinesen werden dazu angehalten, dieselbe mobile App für eine Vielzahl von Aktivitäten zu nutzen – von der Reservierung eines Taxis über das Bezahlen im Restaurant und die Pflege privater Kontakte bis zur Interaktion mit Behörden. So wird über alle eine riesige Menge an Informationen gesammelt und zur Profilerstellung an öffentliche Institutionen weitergeleitet.¹¹⁸

¹¹⁴ AGCM, AGCOM und Garante per la protezione dei dati personali, Indagine conoscitiva sui *Big Data*, S. 25–26, <https://www.agcom.it/documents/10179/17633816/Documento+generico+10-02-2020+1581346981452/39c08bbe-1c02-43dc-bb8e-6d1cc9ec0fcf?version=1.0>. In der Studie wird erklärt, wie sich die „dynamische Einwilligung“ als praktikable Option im Datenschutzrahmen der EU durchsetzt. Dieses Konzept versteht die Einwilligung als einen schrittweisen Prozess, in dessen Verlauf die betroffene Person mehr als einmal kontaktiert werden kann, um zu fragen, ob sie in eine bestimmte Verwendung ihrer Informationen einwilligt.

¹¹⁵ State Council, Notice of the State Council Issuing the New Generation of Artificial Intelligence Development Plan, No. 358, Juli 2017, S. 2–5 und 18–21, <https://flia.org/notice-state-council-issuing-new-generation-artificial-intelligence-development-plan>.

¹¹⁶ State Council, Notice concerning Issuance of the Planning Outline for the Construction of a Social Credit System (2014–2020), No. 21, 14. Juni 2014, <https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020>.

¹¹⁷ Pieranni S., *Red Mirror*, Laterza, 2020, S. 22–23.

¹¹⁸ *Ibid*, S. 40 und 115.

2.2.4. Drei verschiedene Ansätze?

Auch wenn es seltsam scheinen mag, haben einige spekuliert, dass ein ähnliches Sozialkreditsystem auch im privaten Sektor der USA bereits existiert.¹¹⁹ Privatunternehmen erstellen Profile ihrer Kundinnen und Kunden nicht nur, damit sie loyal sind. Sie verkaufen die Informationen über sie auch an andere Unternehmen. So werden persönliche Vorlieben und Kaufgewohnheiten abgeglichen, um bessere Profile der Nutzerinnen und Nutzer zu erstellen und ihre Entscheidungen vorausszusehen und zu beeinflussen.¹²⁰ Eine Bank oder Versicherungsgesellschaft kann das finanzielle Risiko einer Person anhand von vielerlei Informationen, von der Ausbildung über die Lebensweise bis zu besuchten Orten und Menschen, genau beurteilen. Eine politische Partei kann die politische Gesinnung einer Person anhand der gesehenen Filme, der bevorzugten Medien oder der Familiengeschichte einschätzen.

Es dürfte nicht sonderlich überraschen, dass der Gesamtansatz der USA zum Datenschutz das negative Potenzial einer solchen privaten Anhäufung personenbezogener Daten übersieht. Hauptanliegen der US-amerikanischen Rechtskultur ist traditionell, die öffentliche Gewalt unter Kontrolle zu halten. Dieser Ansatz lebt unvermindert fort und hält die Aufmerksamkeit der USA auf die Bedrohungen durch die öffentliche Gewalt gerichtet, während Europa der Fähigkeit von Privatunternehmen zu Grundrechtsverletzungen stets größere Aufmerksamkeit geschenkt hat.¹²¹ Das paradoxe Ergebnis ist, dass die USA die globale Drehscheibe für Big-Data-Innovationen sind, aber die Bedrohung der Grundrechte durch Big Data nicht so sehen, wie es Europa zu tun scheint.

Ein so unterschiedlicher Umgang mit der Privatsphäre hat enorme Konsequenzen für den Alltag der Bürgerinnen und Bürger wie auch für Medienunternehmen. Wie sich weiter unten zeigen wird, werden Medienunternehmen durch den Einsatz KI-basierter Technologien zu mehr als bloßen Informationsgebern. Sie können zu Informationssammlern werden und sich an der Erstellung von Personenprofilen beteiligen.

¹¹⁹ Baker L. C., „Next generation law: Data-driven governance and accountability-based regulatory systems in the West, and social credit regimes in China“, *Southern California Interdisciplinary Law Journal*, 28, 2018, S. 170–171, <https://lbackerblog.blogspot.com/2019/05/just-published-next-generation-law-data.html>.

¹²⁰ Das Europäische Parlament hat vor Kurzem die Europäische Kommission aufgefordert, „äußerst zielgruppenspezifische Werbung auf Plattformen zu unterbinden“: Entschließung des Europäischen Parlaments vom 18. Juni 2020 zu der Wettbewerbspolitik – Jahresbericht 2019, https://www.europarl.europa.eu/doceo/document/TA-9-2020-0158_DE.html. According to Morozov E., „Digital socialism?“, *New Left Review*, 116/117, March-June 2019, S. 62, <https://newleftreview.org/issues/1116/articles/evgeny-morozov-digital-socialism>, „Amazon got a patent on ‚anticipatory shipping‘ – allowing it to ship products to us before we even know we want them“.

¹²¹ Zur Wachsamkeit Europas in Bezug auf das Schadpotenzial privater Unternehmen siehe Europäischer Datenschutzbeauftragter, Stellungnahme 8/2016, Stellungnahme des EDSB zur kohärenten Durchsetzung von Grundrechten im Zeitalter von Big Data, 23. September 2016, S. 5, https://edps.europa.eu/sites/edp/files/publication/16-09-23_bigdata_opinion_de.pdf. Siehe auch Pollicino O., „L’„autunno caldo“ della Corte di giustizia in tema di tutela dei diritti fondamentali in rete e le sfide del costituzionalismo alle prese con i nuovi poteri privati in ambito digitale“, *Federalismi*, 15. Oktober 2019, <https://www.federalismi.it/nv14/editoriale.cfm?eid=533>.

2.3. Verzerrung und Diskriminierung durch Big Data

Man könnte meinen, Software könne nicht voreingenommen sein, doch eine der größten Herausforderungen für datengetriebene Technologien liegt in ihrem Diskriminierungspotenzial. Bei der Sammlung, Verarbeitung und Verbreitung von Informationen können Vorurteile einfließen, verankert und verstärkt werden. Das wohl bekannteste Beispiel hierfür ist Microsofts Chatbot Tay. Microsoft legte 2016 ein Facebook-Profil für innovative Software an, die auf der Plattform mit anderen Facebook-Nutzern interagieren konnte, indem sie Informationen aus dem Internet sammelte, Trends erkannte und entsprechende Meinungen austauschte.

Innerhalb von 16 Stunden nach der Eröffnung wurde das Facebook-Konto wieder geschlossen, weil die Ersteller bemerkten, dass es sexistische und rassistische Beiträge postete.¹²² Natürlich hatten die Softwareentwickler ihrem Bot die Vorurteile, die er dann im Web an den Tag legte, nicht selbst mitgegeben. Sie hatten einfach darauf gesetzt, dass der Bot aus dem Web selbst lernt, und der hatte offenbar festgestellt, dass Rassismus und Sexismus dort weit verbreitet sind und viel Aufmerksamkeit bekommen. Tay orientierte sich in Ausdrucksweise und Themenwahl an dem Training, dem er ausgesetzt war. So lernte und übernahm er Vorurteile ganz von allein.

Tays kurzes Leben zeigt, warum Training für eine KI wichtig ist. KI-basierte Systeme benötigen zum Lernen viele Daten. Je mehr Informationen sie sammeln, desto besser können sie Schlüsse ziehen und Entscheidungen treffen. Leider sind große Datensätze zum Trainieren von Algorithmen häufig nicht erhältlich, sodass Programmierer oft nutzen, was im Web bereits verfügbar ist. Dies ist extrem problematisch, weil Menschen den Lernprozess nicht vollständig überwachen können und sich die KI in unvorhergesehene oder sogar unerwünschte Richtungen entwickeln kann. Sie kann in der Gesellschaft vorhandene Vorurteile übernehmen und verinnerlichen und durch die eigene Tätigkeit noch verstärken.¹²³

Uausgewogene Datensätze können unbeabsichtigt zu Benachteiligungen führen, wie der Fall der Gesichtserkennung zeigt. Westliche KI-Systeme zur Gesichtserkennung können Menschen mit nicht-weißer Hautfarbe oft nicht korrekt identifizieren, weil andere ethnische Gruppen im Internet seltener vorkommen als *Weißer*, während in China entwickelte KI-Software unter dem umgekehrten Problem leidet.¹²⁴ Dadurch ist etwa in westlichen Ländern die Wahrscheinlichkeit, verwechselt zu werden, für eine schwarze Person höher als für eine *weiße*. Mediensysteme, die Prozesse auf Basis von Big Data umfassen, stehen daher vor einer gewaltigen Herausforderung, da durch den KI-Einsatz Vorurteile und soziale Ungleichgewichte ins Spiel kommen können.

¹²² „Microsoft ‚deeply sorry‘ for racist and sexist tweets by AI chatbot“, *The Guardian*, 26. März 2016, <https://www.theguardian.com/technology/2016/mar/26/microsoft-deeply-sorry-for-offensive-tweets-by-ai-chatbot>.

¹²³ Stevenson M. T. & Doleac J. L., *Algorithmic Risk Assessment in the Hands of Humans*, Institute of Labor Economics, 1. Dezember 2019, S. 1, <http://ftp.iza.org/dp12853.pdf>; Bruckner M. A., *op. cit.*, S. 25.

¹²⁴ Grother P., Ngan M., Hanaoka K., „Face recognition vendor test (FRVT) Part III. Demographic effects“, National Institute of Standards and Technology Interagency 8280, Dezember 2019, <https://doi.org/10.6028/NIST.IR.8280>.

Der Kampf gegen Diskriminierung ist bei Big Data und neuronalen Netzen wegen der Gefahren der indirekten Diskriminierung („Proxy Discrimination“) sehr schwierig.¹²⁵ Indirekte Diskriminierung liegt vor, wenn private oder öffentliche Stellen Entscheidungen von Erfordernissen oder Faktoren abhängig machen, die scheinbar neutral sind, tatsächlich aber eine diskriminierende Tradition, Praxis oder Überzeugung verkörpern. So können in sozial oder territorial gespaltenen Gesellschaften die Postleitzahl oder der Preis für Wohnraum zu einer indirekten Diskriminierung in Bezug auf Versicherungspolice oder Zoneneinteilung dienen, da diese Kriterien einige Ethnien benachteiligen und andere bevorzugen können. Selbst wenn der KI ausdrücklich verboten wird, bei ihren Schlussfolgerungen die ethnische Herkunft zu berücksichtigen, können andere Faktoren zu einer indirekten Diskriminierung führen.¹²⁶ Innerhalb einer bestimmten Gesellschaft können von Big Data getriebene Marktstrategien, politische Kampagnen oder Wohlfahrtsorganisationen – sogar unfreiwillig – schlechter gestellte Gruppen isolieren und systematisch indirekt diskriminieren.

2.4. Information der Bevölkerung: Medien, Fehlinformationen und illegale Inhalte

KI ist ein mächtiges Medieninstrument. Sie kann Fakten aufdecken, Präferenzen erkennen, Nutzerprofile erstellen und soziale Trends vorhersehen. Kurzum, sie kann den Menschen mehr von dem bieten, was sie haben möchten. Die Individualisierung von Medienangeboten durch Big Data hat aber auch ihren Preis.

KI ist ein sehr gutes Instrument für die Vorauswahl von Inhalten, die für Mediennutzerinnen und -nutzer von Interesse sein könnten. Angesichts des Überangebots an Informationen kann die Fähigkeit der KI, Nutzerprofile zu erstellen, das Interesse einer Nutzerin oder eines Nutzers an einer Information vorhersagen, sodass die Arbeit der Medien effektiver und die Nutzererfahrung angenehmer wird. Der KI-Einsatz kann jedoch dazu führen, dass Mediennutzerinnen und -nutzern nicht bewusst wird, dass sich ihr Horizont verengt, dass also die Informationen, die sie erhalten, möglicherweise nicht die echte Realität darstellen, sondern nur die „Realität“, an der sie nach Meinung der KI interessiert sind.

Wenn Nutzerinnen und Nutzer immer wieder mit Inhalten gefüttert werden, die sie ohnehin schon bevorzugen oder kennen oder für die sie sich interessieren, entstehen leicht soziale Blasen. Big-Data-Technologien können Informationen in Abhängigkeit davon filtern, was einzelne Mediennutzerinnen und -nutzer mutmaßlich mögen oder glauben. Anstatt deren Horizont zu erweitern, kann die KI also dazu führen, dass sie immer selektiver auswählen. Eine nutzerfreundliche Nachrichtenindustrie könnte ihren Zweck aus den Augen

¹²⁵ Prince A. E. R. und Schwarcz D., „Proxy discrimination in the age of artificial intelligence and big data“ *Iowa Law Review* 105, 2020, S. 1260, <https://ilr.law.uiowa.edu/print/volume-105-issue-3/proxy-discrimination-in-the-age-of-artificial-intelligence-and-big-data>.

¹²⁶ Ibid.

verlieren, der Gesellschaft breite Perspektiven, fundierte Nachrichten und herausfordernde Standpunkte zu bieten.

Von Big Data getriebene Medienstrategien können daher ungewollt die Entstehung von Informationsblasen auslösen. Zusätzlich besteht aber auch die Gefahr, dass eine Blase bewusst erzeugt wird. Große Technologieunternehmen können in Abhängigkeit von ihren Marktstrategien oder Agenden Nutzer- und Informationsprofile erstellen, um die Verbreitung bestimmter Informationen zu fördern oder zu behindern.¹²⁷

Big Data fördert auch das Ringen zwischen traditionellen Medien und sozialen Medien. Soziale Medien nutzen den starken Schutz, der normalerweise der Redefreiheit gewährt wird, und leben von ihrer Dauerpräsenz im Internet und ihrer Fähigkeit, das Publikum laufend mit Nachrichten zu versorgen.¹²⁸ Sie bieten daher eine billige und leicht zugängliche Alternative zu professionellen Medienbetreibern und -kanälen. Dieser asymmetrische Wettbewerb hat bei Nachrichtenanbietern einen gefährlichen „Wettkampf nach unten“ ausgelöst.¹²⁹ Um das Publikum nicht zu verlieren, versuchen traditionelle Medien, mit der Geschwindigkeit von nicht-professionellen Diensten wie Blogs Schritt zu halten, oft zulasten der Korrektheit.¹³⁰

Die Blasen auf KI-basierten Medienplattformen beteiligen sich häufig an der Verbreitung von Fake News. Ein Verhängnis der heutigen Nachrichtenindustrie ist, dass einigen Statistiken zufolge Fake News mehr Menschen schneller erreichen können als kuratierte, faktengeprüfte Informationen,¹³¹ sodass „Cyberkaskaden“ entstehen, wie Cass Sunstein es ausdrückt.¹³² Das Blasensystem verschärft den Prozess, weil es Fakten und abweichende Meinungen ausfiltert und dadurch tief verwurzelte Überzeugungen und sogar Vorurteile verstärkt.

Von Big Data getriebene Strategien stellen die Rolle in Frage, die das Mediensystem und die Meinungsfreiheit in demokratischen Regimen seit jeher spielen. Anstatt Horizonte zu erweitern, Standpunkte zu hinterfragen, Vorurteile aufzudecken und die Gesellschaft voranzubringen, laufen heutige Medienplattformen Gefahr, soziale Gruppen voneinander zu isolieren und tief verwurzelte Meinungen zu verstärken. Traditionell schätzt und schützt der liberale Konstitutionalismus die Redefreiheit in hohem Maße, weil unterschiedliche Standpunkte die Gesellschaften durch den freien Meinungsaustausch voranbringen. Im Gegensatz dazu können Big-Data-Technologien „Echokammern“¹³³ schaffen, die abweichende Meinungen verbannen und um unwidersprochene Überzeugungen kreisen.

¹²⁷ Singer H., „How Washington should regulate Facebook“, *Forbes*, 18. Oktober 2017, <https://www.forbes.com/sites/washingtonbytes/2017/10/18/what-to-do-about-facebook>.

¹²⁸ Shefa M. C., „First Amendment 2.0: Revisiting Marsh and the quasi-public forum in the age of social media“, *University of Hawaii Law Review*, 41, 2018, S. 160.

¹²⁹ AGCM, AGCOM und Garante per la protezione dei dati personali, Indagine conoscitiva sui Big Data, *op. cit.*, S. 30.

¹³⁰ Europäischer Datenschutzbeauftragter, Stellungnahme 3/2018, Stellungnahme des EDSB zu Online-Manipulation und personenbezogenen Daten, *op. cit.* S. 15 („Es gibt Hinweise darauf, dass diese Konzentration und Verdrängung des Lokaljournalismus die Verbreitung von Desinformation erleichtert.“)

¹³¹ *Ibid.*

¹³² Sunstein C. R., „*#republic: Divided democracy in the age of Social Media*“, Princeton University Press, 2017, S. 57.

¹³³ Sasahara K. et al., „*On the inevitability of online echo chambers*“, <https://arxiv.org/abs/1905.03919>.

Meinungen, die tief sitzende Weltanschauungen in Frage stellen, werden aus einer Blase ausgeschlossen und finden ihren Platz dann wohl in einer anderen Blase, die praktisch keinen Austausch mit der Außenwelt bietet.¹³⁴ Big Data kann somit Perspektiven einengen und Vorurteile gegen die Vorteile der Meinungsfreiheit immunisieren.

Private und öffentliche Institutionen haben bemerkt, welche Verzerrungen Big Data in den Medien und in der Gesellschaft insgesamt verursachen kann. So hat Twitter vor Kurzem ein umstrittenes Faktenchecker-Tool erstellt, das Fake News oder Tweets, die konkreten Gruppen schaden, aufspüren soll.¹³⁵ Der *EU-Verhaltenskodex zur Bekämpfung von Desinformation*¹³⁶ drängt auf eine umfassende Betrachtung des Phänomens und betont, dass „alle Interessenträger bei der Bekämpfung der Verbreitung von Desinformation mitwirken müssen“. Viele Unterzeichner des Kodex, darunter Facebook, Google, Mozilla, TikTok und Twitter, haben daher versprochen, „die Sichtbarkeit von Desinformation durch Verbesserung der Auffindbarkeit vertrauenswürdiger Inhalte zu verringern“ und „das Auffinden von Inhalten und den Zugang zu unterschiedlichen, alternative Standpunkte vertretenden Informationsquellen zu erleichtern“. Insgesamt fordern viele eine Regulierung des KI-Einsatzes, die die Internet-Diensteanbieter an „traditionelle Standards der Medienverantwortung“ annähern würde.¹³⁷

Die EU-Maßnahmen zielen insbesondere auf terroristische Inhalte, Material über Kindesmissbrauch, Rassismus sowie fremdenfeindliche und hetzerische Äußerungen,¹³⁸ die für die heutigen sozialen Medien meist sehr wichtige Themen sind. Tatsächlich ist es angesichts des massiven Datenaufkommens technisch unrealistisch, Informationen schon vor der Entscheidung, ob sie gehostet werden, zu filtern. Online-Plattformen arbeiten daher in der Regel mit einer Mischung zweier Systeme: Einerseits wenden sie ein „Notice-and-Takedown“-System an – jeder kann sich darüber beschweren, dass eine bestimmte Anzeige von Inhalten gegen das Gesetz verstößt, und eine Bewertung durch die mediale Plattform veranlassen – und andererseits wenden die meisten Plattformen Filtersysteme auf Big-Data-Basis an, die die Inhalte automatisch und durchgängig durchsuchen und entscheiden, was der Öffentlichkeit vorenthalten werden soll.¹³⁹ Die meisten Plattformen verfügen über einen zusätzlichen Schutz gegen solche automatisierten Entscheidungen, der einen Einspruch gegen Software-Entscheidungen zur Entfernung eines Inhalts ermöglicht.¹⁴⁰

In den USA und der EU, die „über einen der umfassendsten Rechtsrahmen zur Verfolgung illegaler Online-Inhalte“¹⁴¹ verfügt, gibt es für Diensteanbieter weitreichende

¹³⁴ Jones R. L., „Can you have too much of a good thing: The modern marketplace of ideas“, *Missouri Law Review*, 83, 2018, S. 987, <https://scholarship.law.missouri.edu/mlr/vol83/iss4/8/>.

¹³⁵ Pham S., „Twitter says it labels tweets to provide ‚context, not fact-checking““, *CNN Business*, <https://edition.cnn.com/2020/06/03/tech/twitter-enforcement-policy/index.html>.

¹³⁶ EU Code of Practice on Disinformation, <https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation>.

¹³⁷ Europäischer Datenschutzbeauftragter, Stellungnahme 3/2018, Stellungnahme des EDSB zu Online-Manipulation und personenbezogenen Daten, *op. cit.*, S. 16.

¹³⁸ Policy Department for Economic, Scientific and Quality of Life Policies, „Online platforms’ moderation of illegal content online“, Juni 2020, S. 9, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652718/IPOL_STU\(2020\)652718_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652718/IPOL_STU(2020)652718_EN.pdf).

¹³⁹ *Ibid*, S. 45 .

¹⁴⁰ *Ibid*, S. 10.

¹⁴¹ *Ibid*, S. 66.

Haftungsbefreiungen. Solche Befreiungen sollen deren positive Rolle beim Verbinden von Menschen und der Verbreitung von Informationen schützen.¹⁴² Das EU-Recht hat diese Regel gestärkt, indem es den Mitgliedstaaten untersagt, Hosting-Plattformen allgemeinen Verpflichtungen zur Überwachung des von ihnen gehosteten Materials zu unterwerfen.¹⁴³ Das Szenario ist jedoch im Wandel begriffen.¹⁴⁴ Bei der Auslegung der Richtlinie über den elektronischen Geschäftsverkehr hat der Gerichtshof der Europäischen Union festgestellt, dass Diensteanbieter, die nicht einfach Materialien passiv anzeigen, mehr tun müssen, als diese erst dann zu überprüfen und gegebenenfalls zu entfernen, wenn sie dazu aufgefordert werden.¹⁴⁵ Tatsächlich, so das Gericht, erstreckt sich eine richterliche Entfernungsverfügung auch auf „Informationen [...], deren Inhalt wegen der verwendeten Worte oder ihrer Kombination im Vergleich zu der Information, deren Inhalt für rechtswidrig erklärt worden ist, zwar leicht unterschiedlich formuliert ist, aber im Wesentlichen die gleiche Aussage vermittelt“.¹⁴⁶ Manche haben dieses vernünftige Prinzip kritisiert, weil es zum „Samariterparadox“ führe: Je mehr eine Plattform verpflichtet ist, die von ihr veröffentlichten Informationen zu kontrollieren, desto stärker wird sie haftbar. Es gibt Bedenken, dass ein solcher juristischer Ansatz die Anbieter dazu ermutigen würde, passiv zu bleiben und ihre Überwachung einzuschränken, um Haftungsrisiken zu vermeiden.¹⁴⁷ Nun muss ausdiskutiert werden, ob die EU ihre Politik revidieren und den US-Ansatz nachahmen sollte, der die Haftungsbefreiung für Plattformen beibehalten hat, da sie dies

¹⁴² Zu den USA siehe Title 47, Section 230 des Communication Decency Act, <https://www.fcc.gov/general/telecommunications-act-1996>; zur EU siehe Richtlinie 2000/31/EG des Europäischen Parlaments und des Rates vom 8. Juni 2000 über bestimmte rechtliche Aspekte der Dienste der Informationsgesellschaft, insbesondere des elektronischen Geschäftsverkehrs, im Binnenmarkt („Richtlinie über den elektronischen Geschäftsverkehr“), <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32000L0031>, Artikel 14: „1. (1) Die Mitgliedstaaten stellen sicher, dass im Fall eines Dienstes der Informationsgesellschaft, der in der Speicherung von durch einen Nutzer eingegebenen Informationen besteht, der Diensteanbieter nicht für die im Auftrag eines Nutzers gespeicherten Informationen verantwortlich ist, sofern folgende Voraussetzungen erfüllt sind: a) Der Anbieter hat keine tatsächliche Kenntnis von der rechtswidrigen Tätigkeit oder Information, und, in Bezug auf Schadenersatzansprüche, ist er sich auch keiner Tatsachen oder Umstände bewusst, aus denen die rechtswidrige Tätigkeit oder Information offensichtlich wird, oder b) der Anbieter wird, sobald er diese Kenntnis oder dieses Bewusstsein erlangt, unverzüglich tätig, um die Information zu entfernen oder den Zugang zu ihr zu sperren. Zum Jugendschutz siehe Richtlinie 2010/13/EU des Europäischen Parlaments und des Rates vom 10. März 2010 zur Koordinierung bestimmter Rechts- und Verwaltungsvorschriften der Mitgliedstaaten über die Bereitstellung audiovisueller Mediendienste (Richtlinie über audiovisuelle Mediendienste (kodifizierte Version) Text von Bedeutung für den EWR). Eine konsolidierte Fassung mit den Änderungen von 2018, <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:02010L0013-20181218>.

¹⁴³ Policy Department for Economic, Scientific and Quality of Life Policies, *op. cit.*, S. 21.

¹⁴⁴ Nunziato D. C., „The marketplace of ideas online“, *Notre Dame Law Review*, 94, 2019, S. 1521, <https://scholarship.law.nd.edu/cgi/viewcontent.cgi?article=4844&context=ndlr>.

¹⁴⁵ C-324/09, *L'Oréal u. a. gegen eBay International AG*, Randnrn. 113–115, <http://curia.europa.eu/juris/document/document.jsf?text=&docid=107261&pageIndex=0&doclang=de>.

¹⁴⁶ C-18/18, *Eva Glawischnig-Piesczek gegen Facebook Ireland Ltd*, Randnr. 41, <http://curia.europa.eu/juris/document/document.jsf?text=&docid=218621&pageIndex=0&doclang=de>.

¹⁴⁷ Policy Department for Economic, Scientific and Quality of Life Policies, *op. cit.*, S. 20; Policy Department Economic and Scientific Policy, „Liability of Online Service Providers for Copyrighted Content – Regulatory Action Needed?“, Januar 2018, S. 10, [https://www.europarl.europa.eu/RegData/etudes/IDAN/2017/614207/IPOL_IDA\(2017\)614207_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/IDAN/2017/614207/IPOL_IDA(2017)614207_EN.pdf).

zu mehr Eigeninitiative ermutigen würde, oder ob dies den Schutz von Personen und Gruppen gefährden würde.¹⁴⁸

Im Zusammenhang mit illegalen Materialien, die auf Online-Plattformen veröffentlicht werden, kann die KI sicherlich eine wichtige Rolle spielen. Angesichts der riesigen Datenmengen, die ausgetauscht werden, und der Tendenz zur Bildung von Blasen, in denen Mediennutzer kaum Informationen oder Standpunkte finden, die ihnen nicht gefallen oder mit denen sie nicht einverstanden sind, kann es sein, dass illegales Material von Menschen lange Zeit nicht entdeckt wird. Die Entwicklung von KI-basierten Systemen, die Inhalte filtern, kann daher ratsam sein oder sogar notwendig werden. KI und Big Data sind nicht nur Teil des Problems – sie können auch Teil der Lösung sein. Natürlich darf KI-basierte Überwachung nicht zu einer automatischen Zensur werden. Anbieter können KI-Systeme einsetzen, um Materialien auszufiltern, die einfach nur kontrovers sind, und so den öffentlichen Raum von Minderheitsmeinungen oder von Informationen freihalten, mit denen viele Menschen Schwierigkeiten hätten. Dieses Risiko muss unter Kontrolle gehalten werden.

2.5. Big-Data-Politik und die politische Blase¹⁴⁹

Demokratien brauchen einen gesunden öffentlichen Raum, um zu überleben und zu gedeihen.¹⁵⁰ Die Existenz und der Austausch alternativer Weltanschauungen und politischer Meinungen sind für sie überlebenswichtig. Allgemeiner gesagt sollten in Demokratien „Menschen mit Materialien konfrontiert werden, die sie sich im Voraus nicht ausgesucht hätten“,¹⁵¹ denn einer der Vorteile, die seit jeher mit Demokratien assoziiert werden, besteht darin, dass „Vorurteile in der großen Republik ausgefiltert werden“.¹⁵²

Soziale Medien haben die heutige Politik überschwemmt. Die Rechtswissenschaft und die Gerichte haben darauf reagiert, indem sie die klassische Idee öffentlicher Foren langsam aber stetig weiterentwickelt haben, sodass nun auch Social-Media-Websites dazugehören, die sich in privatem Eigentum befinden.¹⁵³ Aufgrund ihrer allgegenwärtigen sozialen Rolle und ihrer zentralen Bedeutung für die Versorgung der Öffentlichkeit mit Nachrichten und politischen Meinungen hat der Oberste Gerichtshof der USA Social-Media-Websites als „den modernen öffentlichen Platz“ bezeichnet.¹⁵⁴ Sie seien für das soziale und politische Leben so wichtig, dass sie für die Allgemeinheit zugänglich sein müssen.¹⁵⁵ Seit

¹⁴⁸ Ibid, S. 67.

¹⁴⁹ Für eine andere Sichtweise zum Thema Filterblasen und Echokammern siehe Kapitel 5 dieser Publikation.

¹⁵⁰ Wischmeyer T., „Making social media an instrument of democracy“, *European Law Journal*, 25, 2019, S. 172, <https://onlinelibrary.wiley.com/doi/abs/10.1111/eulj.12312>.

¹⁵¹ Sunstein C. R., *op. cit.* S. 6.

¹⁵² McGinnis J. O., *Accelerating Democracy*, Princeton University Press, 2013, S. 127.

¹⁵³ Nunziato D. C., *op. cit.*, S. 3.

¹⁵⁴ *Packingam gegen North Carolina* 582 U.S. ____ (2017), https://www.supremecourt.gov/opinions/16pdf/15-1194_0811.pdf.

¹⁵⁵ Ibid.

2001 behandeln US-Gerichte auch „Computer und Internetzugang als ‚praktisch unverzichtbar in der modernen Welt der Kommunikation und Informationsbeschaffung“.¹⁵⁶

Soziale Medien sind jedoch keine allgemein zugänglichen Orte, an denen jeder Mensch willkommen ist und Argumente vorbringen kann. Anhand von Big-Data-Analysen können soziale Medien den öffentlichen Raum in selbstreferenzielle Blasen segmentieren.¹⁵⁷ Selbst die Medienplattformen, die Informationen nicht bewusst filtern, passen ihre Nachrichtenfeeds an die Bedürfnisse und Auswahlentscheidungen ihrer Nutzerinnen und Nutzer an und schaffen so Informationsblasen. Solche Blasen können die öffentliche Meinung in undurchdringliche homogene Einflussphasen spalten.¹⁵⁸

Die Schaffung homogener, voreingenommener, nicht dialogorientierter Echokammern ist kein Ersatz für demokratischen Pluralismus¹⁵⁹ und kann ihn sogar bedrohen.¹⁶⁰ Der Skandal um Cambridge Analytica, das mutmaßlich Daten von Facebook-Nutzern ohne deren Zustimmung sammelte, um „Psychogramme“ zu erstellen und dann ausgewählte Personen ins Visier zu nehmen, um deren Wahlverhalten zu beeinflussen,¹⁶¹ ist nur ein Beispiel dafür, wie sich Big Data auf die Politik auswirken kann.¹⁶² Zudem gibt es noch weitere Belege für den Einsatz Big-Data-basierter Bots zur Beeinflussung politischer Agenden.¹⁶³

Der Harvard-Juraprofessor Cass Sunstein hat in seinem gefeierten Band *#Republic* den Einfluss KI-basierter Social-Media-Plattformen auf die Politik untersucht.¹⁶⁴ Sunstein hat überzeugend dargelegt, dass KI Informationscluster erzeugen und die Politik polarisieren kann. Politische Kampagnen können sich gezielt an Nutzerinnen und Nutzer mit klarem Profil richten und sie mit bestimmten Meinungen oder Fakten konfrontieren und gleichzeitig die Aussagen politischer Gegner oder Fakten, die die eigene Plattform und Agenda in Frage stellen würden, unterschlagen oder herunterspielen.¹⁶⁵ KI spaltet damit den öffentlichen Raum in homogene Umgebungen, zwischen denen kaum ein Austausch stattfindet. Erfolgreiche Politikerinnen und Politiker gehen oft bis zum Äußersten, um ihre Anhängerschaft zu mobilisieren und das Blasensystem zu stärken.

¹⁵⁶ Shefa M. C., *op. cit.*, S. 164.

¹⁵⁷ Sunstein C. R., *op. cit.*

¹⁵⁸ Sasahara K. et al., *op. cit.*

¹⁵⁹ Wischmeyer T., *op. cit.*, S. 173–174.

¹⁶⁰ Manheim K. & Kaplan L., *op. cit.*, S. 109.

¹⁶¹ *Ibid.*, S. 139.

¹⁶² Für weitere Beispiele aus verschiedenen Ländern siehe Gurumurthy A. und Bharthur D., „Democracy and the algorithmic turn“, *Sur International Journal of Human Rights*, 27, 2018, S. 43–44, <https://sur.conectas.org/en/democracy-and-the-algorithmic-turn>, und Tenove C., Buffie J., McKay S. und Moscrop D., *Digital threats to democratic elections: how foreign actors use digital techniques to undermine democracy*, Januar 2018, passim, https://democracy2017.sites.olt.ubc.ca/files/2018/01/DigitalThreats_Report-FINAL.pdf.

¹⁶³ Als die Federal Communication Commission 2017 erwog, einige Regeln zur Regulierung des Internet aufzuheben, waren 21 von 22 Millionen Kommentaren, die über ihre Website eingingen, Fake News (Manheim K. & Kaplan L., *op. cit.*, S. 145.)

¹⁶⁴ Sunstein C. R., *op. cit.*

¹⁶⁵ Mor N., „No Longer Private: On Human Rights and the Public Facet of Social Network Sites“, *Hofstra Law Review* 47 (2018), S. 669, <https://www.hofstralawreview.org/wp-content/uploads/2019/04/bb.7.mor.pdf> (6. August 2020).

Big-Data-Politik verwischt oft die Grenze zwischen Person und Amt. Viele politische Persönlichkeiten ziehen es vor, ihre privaten Social-Media-Profile anstelle der amtlichen zu nutzen, auch für die Kommunikation mit der breiten Öffentlichkeit über Angelegenheiten des Amtes. Indem sie ihre privaten Profile verwenden, zwingen sie die Öffentlichkeit – die normalerweise den amtlichen Medienseiten und -profilen folgen würde – in ihren Unterstützerkreis.

Einige Rechtssysteme haben Gegenmaßnahmen ergriffen, um diese Privatisierung des öffentlichen Raums in separate Medien-Echokammern zu bekämpfen. Augenfälligstes Beispiel für diese Entwicklung sind die Erfahrungen in den USA. Viele Persönlichkeiten des öffentlichen Lebens – darunter auch Präsident Donald Trump –, die private Websites für amtliche Zwecke nutzen, haben Personen blockiert, die ihre Posts kritisch kommentieren, und sie damit vom Empfang ihrer öffentlichen Äußerungen abgeschnitten.¹⁶⁶ Einige so aus dem Publikum ausgeschlossene Bürgerinnen und Bürger haben die betreffenden Politiker erfolgreich verklagt. Die Richter prüften, wie die Medienplattformen strukturiert sind und wie Politikerinnen und Politiker sie nutzen, und kamen zu dem Schluss, dass solche Plattformen als öffentliche Orte zu betrachten seien, die für alle offen bleiben sollten. Die Politikerinnen und Politiker könnten ihre Follower zwar immer noch „stumm“ schalten und sie so daran hindern, innerhalb ihres eigenen Profils an einem Dialog teilzunehmen, sie aber nicht „blockieren“, da dies in einigen Fällen verhindert hätte, dass sie Informationen über Angelegenheiten von öffentlichem Interesse erhalten.¹⁶⁷

2.6. Medien als Überwachungsinstanz?

Big-Data-Analysen waren entscheidend für die Entwicklung KI-gestützter Gesichtserkennungsverfahren. Dank KI-Fähigkeiten kann die Software Unmengen an Bildern prüfen und vergleichen, um Übereinstimmungen zu finden. Im Gegensatz zu den altmodischen Überwachungskameras, deren Aufzeichnungen von Menschen auf Übereinstimmungen durchgesehen werden, lassen sich Bilder heute mit Computervision fast augenblicklich verarbeiten. In einer Entscheidung von 2019 befasste sich ein walisisches Gericht mit KI-gestützter Gesichtserkennung.¹⁶⁸ Die Software, die die walisische Polizei bei mehreren öffentlichen Veranstaltungen eingesetzt hatte, war in der Lage, bis zu 40 Gesichter pro Sekunde zu verarbeiten. Dies ergibt eine beeindruckende Gesamtzahl: Bei rund 50 Einsätzen verarbeitete die Software etwa 500 000 Personen – ein Sechstel der gesamten Bevölkerung von Wales. KI kann zu einem mächtigen Instrument der Massenüberwachung werden, wie es bereits in Ländern wie China der Fall ist, wo ein Projekt

¹⁶⁶ *Ibidem*, S. 42 ff.

¹⁶⁷ *Knight First Amendment Inst. at Columbia Univ. gegen Trump* 302 F. Supp. 3d 541 (SDNY 2018), <https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=2780&context=historical> (6. August 2020).

¹⁶⁸ (*Bridges*) *gegen The Chief Constable of South Wales Police et al.*, [2019] EWHC 2341, <https://www.judiciary.uk/wp-content/uploads/2019/09/bridges-swp-judgment-Final03-09-19-1.pdf>.

für ein systematisches Überwachungssystem auf KI-Basis läuft, das mit mehr als einer halben Milliarde Kameras arbeitet.¹⁶⁹

Gesichtserkennung betrifft viele der oben angesprochenen Themen. Erstens berühren Gesichtserkennungsverfahren die Privatsphäre. Sie verarbeiten menschliche Gesichter – nicht nur von Personen in einer Datenbank, sondern von allen. Um einen bestimmten Menschen aus der Gruppe der Personen, die von Interesse sind, auszuschließen, muss eine Software tatsächlich zuerst dessen Gesicht verarbeiten. In der europäischen Rechtskultur erfordert ein solch massiver Eingriff in die Privatsphäre eine ausreichende Begründung. Wie der Europäische Gerichtshof für Menschenrechte wiederholt betont hat, haben öffentliche Interessen keinen Vorrang vor Datenschutzbedenken. Im Gegenteil: Sie erfordern eine vorläufige Kosten-Nutzen-Bewertung, die sicherstellt, dass jeder Einsatz in einem angemessenen Verhältnis zur Aufgabe steht.¹⁷⁰

Zweitens besteht bei Gesichtserkennungsverfahren die Gefahr der Voreingenommenheit. Wie bereits erwähnt, kommen „falsch positive Ergebnisse“ – falsche Übereinstimmungen – häufiger bei ethnischen Gruppen vor, die im Trainingsmaterial unterrepräsentiert sind.¹⁷¹ Falsch positive Ergebnisse haben oft praktische Konsequenzen, da sie rassistische Vorurteile verstärken und öffentliche Institutionen wie Polizeistreifen veranlassen können, ethnische Minderheiten ins Visier zu nehmen, bei denen die Software mehr falsch positive Ergebnisse auswirft.¹⁷²

Drittens kann Gesichtserkennung aus einer Vielzahl von Gründen irreführend sein. Einige Softwareprogramme können die immensen KI-Fähigkeiten ausnutzen, indem sie live übermittelte und aufgezeichnete Bilder aus beliebigen Internetquellen verwenden.¹⁷³ Eine solche Technologie kann die Medienindustrie ausnutzen, um mehr Material zu sammeln und ihre Datenbank zu vergrößern. Aktuell wird über das Für und Wider der Entwicklung oder Einführung einer Software diskutiert, die das Internet durchsucht, um Übereinstimmungen mit Personen zu finden, wie es in den USA in vielen lokalen Polizeibehörden geschah, um Verdächtige ausfindig zu machen. Ein solch riesiger Datensatz stützt sich auf eine Vielzahl von Materialien, die gefälscht sein, Verzerrungen

¹⁶⁹ Carter W. M., „Big Brother facial recognition needs ethical regulations“, *The Conversation*, 22 July 2018, <https://theconversation.com/big-brother-facial-recognition-needs-ethical-regulations-99983>.

¹⁷⁰ *Lopez Ribalda und andere gegen Spanien* (apps. No. 1874/13 and 8567/13: <http://hudoc.echr.coe.int/fre?i=001-197098>); *Gorlov und andere gegen Russland* (app. no. 27057/06; 56443/09; 25147/14: <http://hudoc.echr.coe.int/spa?i=001-194247>); *Antovic und Mirkovic gegen Montenegro* (app. no. 70838/13: <http://hudoc.echr.coe.int/fre?i=001-178904>); *Bărbulescu gegen Rumänien* (app. no. 61496/08: <http://hudoc.echr.coe.int/spa?i=001-177082>).

¹⁷¹ Buolamwini J. & Gebru T., „Gender shades: Intersectional accuracy disparities in commercial gender classification“ *Proceedings of Machine Learning Research* 81, 2018, S. 1 und 15, <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>.

¹⁷² Fung B. und Metz R., „This may be America’s first known wrongful arrest involving facial recognition“, 24 June 2020, *CNN Business*, <https://edition.cnn.com/2020/06/24/tech/aclu-mistaken-facial-recognition/index.html>.

¹⁷³ Hill K., „The secretive company that might end privacy as we know it“, *New York Times*, 18. Januar 2020, <https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html>; Ducklin P., „Clearview AI facial recognition sued again – this time by ACLU“, *Naked Security*, 29. Mai 2020, <https://nakedsecurity.sophos.com/2020/05/29/clearview-ai-facial-recognition-sued-again-this-time-by-aclu>.

enthalten¹⁷⁴ und jedes winzige Element des Soziallebens oder einer Mediensendung in eine Datei verwandeln können.

Es ist keine Überraschung, dass IBM,¹⁷⁵ Microsoft¹⁷⁶ und Amazon¹⁷⁷ kürzlich erklärt haben, dass sie ihre Gesichtserkennungstechnologien nicht mehr der Polizei anbieten werden. Viele US-Bundesstaaten erwägen, KI-gestützte Gesichtserkennung zu verbieten, oder haben bereits Gesetze erlassen, die sie einschränken oder untersagen.¹⁷⁸ Daher besteht in den westlichen Ländern ein wachsender Konsens darüber, dass selbst öffentliche Interessen keine allgegenwärtigen Massenüberwachungssysteme rechtfertigen können, die das Internet nutzen.

2.7. Der Medienmarkt: Big-Data-getriebene Marktstrategien

Big Data hat die Medienwelt revolutioniert. Viele Akteure in der Medienbranche sind heute auf große Technologieunternehmen angewiesen, um den Kontakt zum Publikum besser pflegen zu können.¹⁷⁹ Tatsächlich erfordert das nutzbringende Sammeln und Verarbeiten riesiger Datenmengen Fähigkeiten, die nur wenige besitzen. Der Kreis der Unternehmen, die Big Data abschöpfen können, ist sehr begrenzt, und die meisten Marktteilnehmer arbeiten mit diesem Kreis zusammen, um besser zu verstehen, wer ihre Kundinnen und Kunden sind, auf welche Marktstrategie sie setzen sollen oder wie sie ihre Bekanntheit steigern können. Einige große Technologieunternehmen in diesem Bereich, beispielsweise Amazon, produzieren sogar selbst Medieninhalte. Dank ihrer technologischen Fähigkeiten agieren die großen Technologieunternehmen heute also entweder als Medienmacher oder als Mediatoren zwischen der Medienindustrie und ihren Konsumenten (oder auch in beiden Funktionen).

Der wegweisende Fall Google Spain¹⁸⁰ des Gerichtshofs der Europäischen Union zeigt deutlich, welche überragende Rolle große Technologieunternehmen heute im Nachrichtenbereich spielen und wie sie sich den dort geltenden Gesetzen widersetzen. Als sich eine Person darüber beschwerte, dass eine Google-Suche nach ihrem Namen eine Ergebnisliste lieferte, in der ganz oben ein sehr alter Zeitungsartikel über sie stand, der ihren Ruf immer noch ruinieren könne, bestand die erste Verteidigungslinie von Google darin, dass das Unternehmen keine personenbezogenen Daten verarbeite, sondern nur

¹⁷⁴ Geiger R. S. et al., „Garbage in, garbage out? Do machine learning application papers in social computing report where human-labeled training data comes from?“, <https://arxiv.org/abs/1912.08320>.

¹⁷⁵ Krishna A., „IBM CEO's Letter to Congress on Racial Justice Reform“, 8. Juni 2020, <https://www.ibm.com/blogs/policy/facial-recognition-sunset-racial-justice-reforms/>.

¹⁷⁶ Greene J. Microsoft won't sell police its technology, following similar moves by Amazon and IBM“, *The Washington Post*, 11. Juni 2020, <https://www.washingtonpost.com/technology/2020/06/11/microsoft-facial-recognition/>.

¹⁷⁷ Hao K., „The two-year fight to stop Amazon from selling face recognition to the police“, *MIT Technology Review*, 12. Juni 2020, <https://www.technologyreview.com/2020/06/12/1003482/amazon-stopped-selling-police-face-recognition-fight>. Siehe auch Hartzog W., *op. cit.*, S. 76–77.

¹⁷⁸ Siehe den Illinois Biometric Information Privacy Act, <https://www.termsfeed.com/blog/bipa/>.

¹⁷⁹ Tsesis T., *op. cit.*, S. 1589.

¹⁸⁰ *Google Spain SL u. a. gegen Agencia Española de Protección de Datos*, C-131/12, <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:62012CJ0131>.

Suchanfragen mit Ergebnissen verbinde.¹⁸¹ Google argumentierte also, es sei nicht für das verantwortlich, was es über die Google-Suche zur Verfügung stellt. Der Gerichtshof reagierte mit einem historischen Urteil, das zeigt, dass er sich der einzigartigen Rolle von Google bei der Internetsuche bewusst ist. Er stellte fest, dass Google dafür verantwortlich war, in welche Rangfolge es die Antworten auf eine Anfrage brachte, da es längst vergessene Informationen zutage fördern kann, die der Öffentlichkeit sonst nicht zugänglich gewesen wären.

Große Technologieunternehmen beteiligen sich nicht einfach am Medienmarkt. Sie haben auch großen Einfluss auf seine Dynamik. Ihre einzigartige Fähigkeit zur Erstellung von Marktprofilen hält die Nutzerinnen und Nutzer durch einen Lock-in-Effekt gefangen und erzeugt ein Quasi-Marktmonopol.¹⁸² Sie sind so allgegenwärtig und unverzichtbar, dass diejenigen, die sie nicht nutzen wollen, oft den Markt ganz verlassen müssen. Viele Internetnutzerinnen und -nutzer wissen, dass „der Besuch einer einzigen Website in der Regel zur Offenlegung des Surfverhaltens gegenüber mehr als 100 Dritten führt, die versuchen, ihre eigene rechtliche Haftung durch umfangreiche ‚Datenschutzrichtlinien‘ zu begrenzen, die Hunderte von Seiten umfassen können“, aber sie können nicht umhin, dieselben Websites wieder und wieder zu besuchen.¹⁸³ Die wenigen Unternehmen, die das Potenzial von Big Data ausnutzen, sichern sich ihr Revier womöglich auch noch durch „Killer-Akquisitionen“, bei denen sie innovative Start-ups aufkaufen, um entweder deren gesammelten Daten zu analysieren¹⁸⁴ oder die eigene beherrschende Stellung zu schützen.¹⁸⁵ Mit den Worten von Frank Pasquale: So wie „der Pharao versuchte, Moses als Säugling zu töten“, können große Technologieunternehmen ihren Rivalen die Möglichkeit verwehren, groß zu werden.¹⁸⁶

Die gleichzeitige Präsenz mehrerer Unternehmen, die Big Data nutzen, garantiert nicht, dass in einem Markt Wettbewerb herrscht.¹⁸⁷ Big Data kann die Entwicklung von Marktstrategien, einschließlich der Preisgestaltung, fördern, die den Wettbewerbern nutzen und nicht den Kundinnen und Kunden. Es gibt Hinweise darauf, dass Algorithmen verschiedener Unternehmen durch eine Strategie impliziter Absprachen die Preise maximieren können, indem sie einfach Informationen über den Markt selbst verarbeiten.¹⁸⁸ Ein Algorithmus kann einem Unternehmen vorschlagen, die Preise zu erhöhen, weil er voraussagt, dass die Mitbewerber dasselbe tun werden. Mithilfe von Nutzerprofilen und -clustern können sie zudem „den Markt ... segmentieren“ und von allen Nutzerinnen

¹⁸¹ Ibid., Randnr. 22.

¹⁸² AGCM, AGCOM und Garante per la protezione dei dati personali, Indagine conoscitiva sui Big Data, *op. cit.*, S. 26 und 78.

¹⁸³ Europäischer Datenschutzbeauftragter, Stellungnahme 3/2018, Stellungnahme des EDSB zu Online-Manipulation und personenbezogenen Daten, *op. cit.*, S. 7.

¹⁸⁴ Zuboff S., *The Age of Surveillance Capitalism*, Profile Books, 2019, S. 102–103.

¹⁸⁵ AGCM, AGCOM und Garante per la protezione dei dati personali, „Indagine conoscitiva sui Big Data“, *op. cit.*, S. 81. Siehe auch Hughes C., *op. cit.*

¹⁸⁶ Pasquale F., *The Black Box Society*, Harvard University Press, 2015, S. 67.

¹⁸⁷ European Data Protection Supervisor, Opinion 3/2020 on the European strategy for data, *op. cit.*, S. 8 (wo vor der Schaffung oder Stärkung von „Datenoligopol-Situationen“ gewarnt wird).

¹⁸⁸ Den Boer A. V., „Dynamic pricing and learning: Historical origins, current research, and new directions“, *Surveys in operations research and management science*, 20, 2015, S. 1, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2334429; AGCM, AGCOM und Garante per la protezione dei dati personali, „Indagine conoscitiva sui Big Data“, *op. cit.*

und Nutzern die Preise verlangen, die ihrer jeweiligen Zahlungsbereitschaft entsprechen. Diese Praktiken sorgen für „maximale Erlöse [für Firmen], aber nicht für das Wohl der Konsumenten“.¹⁸⁹ Eine solche datengetriebene Marktstrategie ist in der Regel nicht strafbar, weil sie ohne Absprachen auskommt, bietet aber dieselben Vorteile, die normalerweise mit abgesprochenem Verhalten verbunden sind.¹⁹⁰

2.8. Regulatorische Ansätze für KI-basierte Systeme

Immer wieder werden Forderungen nach neuen Regulierungssystemen laut, die sicherstellen, dass bei der KI-Nutzung Rechtsstaatlichkeit, Grundrechte und ethische Werte gewahrt werden. Große Technologieunternehmen haben sich lange Zeit den öffentlichen Regulierungsbestrebungen widersetzt,¹⁹¹ scheinen sich nun aber damit abgefunden zu haben, dass der KI Grenzen gesetzt werden, wobei sie jedoch auf eine Selbstregulierung der Unternehmen setzen und nicht auf staatliche Vorschriften.

Die meisten Beschränkungen zielen jedoch nicht darauf ab, den KI-Einsatz einzudämmen, sondern sollen vielmehr die Rolle der KI stärken, indem sie sie vertrauenswürdiger und zuverlässiger machen.¹⁹² Es besteht sogar ein breiter Konsens darüber, dass KI „rechtmäßig“ (gesetzeskonform), „ethisch“ (der Einhaltung ethischer Prinzipien und Werte verpflichtet) und „robust“ (technologisch und soziologisch sicher) sein muss, um erfolgreich in menschliche Gesellschaften integrierbar zu sein.¹⁹³

In Debatten wird oft betont, dass Big-Data-Analysen einen neuen Ansatz für die rechtliche Regulierung erfordern. Traditionelle Instrumente reichen möglicherweise nicht aus, um sicherzustellen, dass die Welt von Big Data grundlegende menschliche Werte respektiert. Aufgrund des Blackbox-Charakters und der weitreichenden Auswirkungen der KI können rechtliche Sanktionen Big-Data-basierten Technologien und Strategien kaum Grenzen setzen. Für Klagen kann es bereits zu spät sein, wenn der eigene Ruf oder ein Unternehmen in Trümmern liegt, und eine Haftung lässt sich möglicherweise schwer zuordnen. Die KI muss rechtliche Werte in ihre Datenverarbeitung einbeziehen, um deren Schutz im Betrieb sicherzustellen.

Aufgrund der Fülle an Informationen, die sie sammelt, ihres allgegenwärtigen Einsatzes und ihrer Fähigkeit, menschliche Akteure durch Roboter zu ersetzen, wirft die KI auch ethische Fragen auf. *Digitale Ethik* ist ein neues Feld der KI-Regulierung, das große

¹⁸⁹ Europäischer Datenschutzbeauftragter, Stellungnahme 8/2016, Stellungnahme des EDSB zur kohärenten Durchsetzung von Grundrechten im Zeitalter von Big Data, op. cit., S. 6.

¹⁹⁰ Harrington, J. E. Jr., „Developing competition law for collusion by autonomous artificial agents“, *Journal of Competition Law & Economics*, 14, 2019, S. 349–351, <https://academic.oup.com/jcle/article-abstract/14/3/331/5292366?redirectedFrom=fulltext>.

¹⁹¹ Zuboff S., op. cit., S. 105.

¹⁹² Van Dijk N. & Casiraghi S., „The ethicization of privacy and data protection law in the European Union: The case of artificial intelligence“, *Brussels Privacy Hub*, 6, 22, Mai 2020, S. 5, <https://brusselsprivacyhub.eu/publications/BPH-Working-Paper-VOL6-N22.pdf>.

¹⁹³ High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, S. 2, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>. Siehe N. van Dijk & S. Casiraghi, op. cit., S. 14.

Beachtung findet, vor allem in den USA, Kanada und Europa, wo Ethikkodizes wie Pilze aus dem Boden schießen.¹⁹⁴ Digitale Ethik behandelt in einer Disziplinen und Perspektiven übergreifenden Weise eine Fülle von Themen, darunter „moralische Probleme in Bezug auf *Daten* und *Informationen* ..., *Algorithmen* ... und entsprechende *Praktiken* und *Infrastrukturen*“.¹⁹⁵ Obwohl die Situation äußerst lebhaft ist, entwickelt sie sich im Moment eher wenig, auch weil es so schwierig ist, Grenzen zwischen der rechtlichen und der ethischen Komponente der KI-Regulierung zu ziehen.¹⁹⁶

2.9. Fazit

Big Data ist ein wichtiger Grund für den gesellschaftlichen, wirtschaftlichen und politischen Erfolg der KI. Die Verarbeitung riesiger Datenmengen ist für große Technologieunternehmen von entscheidender Bedeutung. Sie ist jedoch kein reiner Segen, sondern diejenigen, die auf diesem Gebiet arbeiten, müssen dafür sorgen, dass die KI dem Menschen dient.¹⁹⁷ Chris Hughes, Mitbegründer von Facebook, hat davor gewarnt, die Digitalisierung der Wirtschaft könne zu „einem Rückgang des Unternehmertums, einem Stillstand des Produktivitätswachstum sowie höheren Preisen und weniger Wahlmöglichkeiten für Konsumenten“ beitragen.¹⁹⁸ Es steht so viel auf dem Spiel, dass ein Abgeordneter der französischen Nationalversammlung sogar vorschlug, in die Präambel der französischen Verfassung eine *Charta der künstlichen Intelligenz und der Algorithmen* aufzunehmen, um die Menschenrechte besser zu schützen.¹⁹⁹

KI bindet Menschen maximal ein. Ein Schlüsselfaktor für den Erfolg ist, „möglichst vielen Menschen möglichst viele Reaktionen“ zu entlocken, denn dies liefert Feedback und ermöglicht Unternehmen, ihre Geschäftspläne und -modelle in Echtzeit an ihre Kundinnen und Kunden anzupassen.²⁰⁰ Politische Akteure und soziale Influencer nutzen dieses Phänomen aus, indem sie bei ihrem potenziellen Publikum emotionale Reaktionen auslösen. Big-Data-Politik und -Wirtschaft stellen die Medien in den Mittelpunkt, weil sie Nachrichten verbreiten, Informationen sammeln, Emotionen verarbeiten und soziale Räume verbinden.

Big Data mehrt die Bedeutung der Medien für die heutigen Gesellschaften. Unternehmen, Politiker, Influencer und andere politische Figuren nutzen Big Data, um ihre

¹⁹⁴ Jobin A., Ienca M. und Vayena E., „The global landscape of AI ethics guidelines“, *Nature Machine Intelligence*, 1, 2019, S. 393–395, <https://www.nature.com/articles/s42256-019-0088-2>.

¹⁹⁵ Floridi L., *op. cit.*, S. 3.

¹⁹⁶ Siehe z. B. die Empfehlung des Europarats *Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems*, 8. April 2020, https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154, die die Vielfalt der regulatorischen Ebenen aufzeigt, die für die Entwicklung solider KI-basierter Systeme erforderlich sind.

¹⁹⁷ Siehe die Asilomar-Prinzipien, die in Verbindung mit der Asilomar-Konferenz 2017 entwickelt wurden. *Future of Life Institute*, <https://futureoflife.org/ai-principles>.

¹⁹⁸ Siehe auch Hughes C., *op. cit.*

¹⁹⁹ http://www.assemblee-nationale.fr/dyn/15/textes/l15b2585_proposition-loi.

²⁰⁰ Akin Unver H., „Artificial intelligence, authoritarianism and the future of political systems“, Centre for Economics and Foreign Policy Studies, July 2019, S. 3, https://edam.org.tr/wp-content/uploads/2018/07/AKIN-Artificial-Intelligence_Bosch-3.pdf.

Ideen, Agenden und Meinungen zu vermarkten und ihr Publikum zu prägen.²⁰¹ Über Internet-Plattformen können traditionelle Medien ihre Inhalte verbreiten und zwischen traditionellen und neuen Kanälen neuen Wettbewerb erzeugen.

Medienakteure können auch eine negative Rolle spielen. Durch die Erstellung von Profilen der „Denkmuster und der psychischen Verfassung“ können sie ein Publikum bewusst falsch informieren und in die Irre führen.²⁰² Darüber hinaus kann in Ländern, in denen nur wenige Medienakteure tätig sind oder in denen es nur oder fast ausschließlich vom Staat betriebene soziale Medien gibt,²⁰³ ein politisches Regime die Nachrichten und auch die Reaktion der Menschen darauf wirksam kontrollieren, indem es fingiertes positives Feedback verbreitet und negative Kommentare isoliert.²⁰⁴ Innerhalb des Szenarios, das durch Big Data entsteht, können Medien entscheidend dazu beitragen, Demokratie, Gleichheit, Minderheiten und offene Gesellschaften zu schützen – oder zu schwächen.²⁰⁵

Darüber hinaus kann Massenüberwachung einen abschreckenden Effekt auf Kreativität und Innovation haben. Obwohl man früher davon ausging, dass die KI einfach den Erfindungsreichtum fördern würde,²⁰⁶ haben manche eine Tendenz ausgemacht, „Spontaneität, Experimentierfreude oder Abweichungen von der statistischen ‚Norm‘ zu verhindern oder zu ahnden und konformistisches Verhalten zu belohnen“.²⁰⁷

Aufgrund der heutigen weiten Verbreitung der KI muss sich der mediale Raum seiner einzigartigen Rolle bewusst werden. Der Mediensektor muss sich um einen rechtmäßigen, ethischen und robusten KI-Einsatz bemühen. Dank ihrer verbindenden Rolle könnten die Medien den weiteren Kreis KI-basierter Unternehmen dazu anhalten, sich dieselben Werte zu eigen zu machen und rechtmäßig, ethisch und robust zu werden. Insbesondere kann eine ethische Verpflichtung Medienplattformen ermutigen, sich nicht mit einer rein passiven Rolle zu begnügen. Auch wenn viele Vorschriften die rechtliche Haftung von Anbietern für die von ihnen gehosteten Inhalte begrenzen²⁰⁸ und es auch mit härteren Auflagen für die Medien nicht gelungen ist, mehr Kontrollen zu fördern, kann es für Medienplattformen ein lohnendes ethisches Ziel sein, ihre Inhalte zu überwachen.²⁰⁹

²⁰¹ Ibid.

²⁰² Europäischer Datenschutzbeauftragter, Stellungnahme 4/2015. Der Weg zu einem neuen digitalen Ethos, *op. cit.*, S. 7.

²⁰³ Pasquale F., *op. cit.*, S. 10, stellt fest, dass wegen des massiven KI-Einsatzes in strategischen Sektoren von öffentlichem und privatem Interesse „der Unterschied zwischen Staat und Markt schwindet“.

²⁰⁴ Akin Unver H., *op. cit.*, S. 8. Siehe auch Meaker M., „How governments use the Internet to crush online dissent“, *The Correspondent*, 27 November 2019, <https://thecorrespondent.com/142/how-governments-use-the-internet-to-crush-online-dissent/18607103196-db0c0dab>.

²⁰⁵ High-Level Expert Group on Artificial Intelligence, „Ethics guidelines for trustworthy AI, *op. cit.*, S. 11.

²⁰⁶ Perritt, H. H., Jr., *op. cit.*, S. 107.

²⁰⁷ Europäischer Datenschutzbeauftragter, Stellungnahme 4/2015. Der Weg zu einem neuen digitalen Ethos, *op. cit.*, S. 9. Siehe auch Pan S. B., *op. cit.*, S. 257 („The goal of big data is to generalize“) und Pasquale F., *op. cit.*, S. 188.

²⁰⁸ Perritt H. H., Jr., *op. cit.*, S. 149.

²⁰⁹ ERGA2020 Subgroup 1 – Enforcement, ERGA Position Paper on the Digital Services Act, S. 6, https://nellyo.files.wordpress.com/2020/06/erga_sg1_dsa_position-paper_adopted-1.pdf.

Meinungsfreiheit, Vielfalt und Pluralismus

*Ein spezifisches Problem beim KI-Einsatz betrifft den Medienbereich ganz besonders: Vielfalt und Pluralismus. Wer alt genug ist, um sich an die Zeiten zu erinnern, als sich die Fernsehkanäle in einem Land an den Fingern einer Hand abzählen ließen und Zeitungen nur auf Papier erschienen, mag das Problem der Vielfalt (zumindest in quantitativer Hinsicht) allerdings etwas übertrieben finden. Heutzutage gibt es zahllose Fernsehkanäle, und jede Zeitung der Welt ist nur einen Klick entfernt. Man könnte sogar sagen: Das Einzige, was einen Menschen heute daran hindert, sich alle Informationen der Welt zu beschaffen, sind nicht Algorithmen, sondern Bezahlschranken. Doch gerade weil das Informationsangebot so überwältigend groß ist, suchen die Menschen nach Filtern. Und wie schon erwähnt, ist das Filtern eine Aufgabe, die die KI sehr gut beherrscht. Video-on-Demand- oder Nachrichtendienste können diese Personalisierung von Nachrichten für alle Internetnutzerinnen und -nutzer anhand ihres persönlichen Seh- und Leseverlaufs oder anderer Präferenzen übernehmen. Die Kehrseite sind die sogenannten Filterblasen, die laut entstehen, wenn Algorithmen „Fakten und abweichende Meinungen ausfiltern und dadurch tief verwurzelte Überzeugungen und sogar Vorurteile verstärken“.²¹⁰ Über die Existenz und die Auswirkungen solcher Filterblasen besteht allerdings keine allgemeine Einigkeit. **Mira Burri** räumt in ihrem Beitrag zu dieser Publikation zwar ein, dass maßgeschneiderte Medienkost bedenkliche Auswirkungen auf die Vielfalt haben kann und dass die Macht der Plattformen beobachtet werden muss, äußert aber auch Zweifel daran, dass sie in direktem Zusammenhang mit einer Fragmentierung der öffentlichen Debatte und einer Polarisierung der Ansichten stehen.²¹¹ Sogar die Befürworter der Filterblasen-These geben zu, dass sie deren reale Existenz nicht belegen können²¹² und dass es bisher kaum empirische Beweise dafür gibt.²¹³ **Sarah Eskens** stellt in ihrem Beitrag zu dieser Publikation fest: „Die aktuelle Herausforderung für Nachrichtenmedien und Behörden besteht darin, journalistische Ethikkodizes, Selbstregulierungsstandards und möglicherweise auch eine staatliche Regulierung zu entwickeln, um die Risiken der KI für die Meinungsfreiheit einzudämmen und gleichzeitig*

²¹⁰ Siehe Andrea Pins Beitrag zu dieser Veröffentlichung.

²¹¹ Für andere kritische Ansichten hierzu siehe z. B. Bruns A., „It's Not the Technology, Stupid: How the ‚Echo Chamber‘ and ‚Filter Bubble‘ Metaphors Have Failed Us“, <http://snurb.info/node/2526>.

²¹² Zuiderveen Borgesius, F., Trilling, D., Moeller, J., Bodó, B., de Vreese, C. H., & Helberger, N., „Should we worry about filter bubbles?“ Internet Policy Review, 5(1), <https://doi.org/10.14763/2016.1.401>.

²¹³ Helberger N., Eskens S., van Drunen M., Bastian M., Moeller J., Implications of AI-driven tools in the media for freedom of expression, <https://rm.coe.int/cyprus-2020-ai-and-freedom-of-expression/168097fa82>.

dafür zu sorgen, dass die KI zu öffentlichen Debatten, zum Medienpluralismus, zum freien Informationsfluss und zu anderen gesellschaftlichen Zielen beitragen kann.“

3. Auswirkungen des Einsatzes von künstlicher Intelligenz durch Nachrichtenmedien auf die Freiheit der Meinungsäußerung

Sarah Eskens, Universität Amsterdam

3.1. Einführung

Nachrichtenmedien arbeiten in ihrem Geschäft zunehmend mit künstlicher Intelligenz. Das *Reuters Institute for the Study of Journalism* befragte 2018 fast 200 Führungskräfte aus dem journalistischen Bereich. Fast drei Viertel der Befragten gaben an, dass sie in ihrer Organisation bereits KI einsetzen.²¹⁴

Der KI-Einsatz eröffnet den Nachrichtenmedien Möglichkeiten und kann ihnen helfen, ihre demokratische Rolle zu erfüllen. Ein Bericht der Europäischen Rundfunkunion hob auch die Möglichkeiten der KI für den öffentlich-rechtlichen Journalismus hervor.²¹⁵ Der KI-Einsatz durch Nachrichtenmedien kann daher unter den Schutz des Rechts der Medien auf freie Meinungsäußerung fallen. Angesichts der Vorstöße für eine KI-Regulierung in verschiedenen Bereichen ist dies nicht unwichtig. Gleichzeitig kann der KI-Einsatz durch Nachrichtenmedien auch Einfluss darauf haben, inwieweit andere Teilnehmerinnen und Teilnehmer der öffentlichen Debatte ihre Freiheit der Meinungsäußerung wahrnehmen können. Nachrichtenorganisationen etwa können KI einsetzen, um Kommentare auf ihren Websites automatisch zu moderieren. Wenn etwa die automatische Moderation im englischsprachigen Bereich die standardamerikanische Variante bevorzugt, könnten manche Stimmen in der öffentlichen Debatte untergehen.

Da in der Politik über die Notwendigkeit einer KI-Regulierung diskutiert wird, stellt sich die Frage, inwieweit ein KI-Einsatz durch Nachrichtenmedien reguliert werden kann und wie die Medienfreiheit mit anderen Rechten und Interessen in Einklang gebracht werden soll, die mit dem KI-Einsatz durch Nachrichtenmedien verbunden sind. Der KI-Einsatz durch Nachrichtenmedien prägt unser Informationsumfeld und kann erhebliche Auswirkungen auf die offene Debatte, den Medienpluralismus und die Medienvielfalt, den freien Informationsfluss und andere öffentliche Werte haben, die mit der Institution der

²¹⁴ Newman N., „Journalism, media, and technology trends and predictions 2018“. Reuters Institute for the Study of Journalism, S. 29,

<https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-01/RISJ%20Trends%20and%20Predictions%202018%20NN.pdf>.

²¹⁵ European Broadcasting Union, „The next newsroom: Unlocking the power of AI for public service journalism“, <https://www.ebu.ch/publications/news-report-2019>.

Nachrichtenmedien zusammenhängen.²¹⁶ Aufgrund der demokratischen Rolle der Nachrichtenmedien wirft der KI-Einsatz dort ganz spezielle Fragen auf. Das vorliegende Kapitel konzentriert sich daher allein auf den KI-Einsatz bei Nachrichtenmedien und nicht bei anderen Medien, etwa Unterhaltungsmedien, wo eine andere Interessenabwägung gefordert ist.²¹⁷

Dieses Kapitel ist wie folgt aufgebaut: Nach einer Beschreibung des allgemeinen Rahmens für den Schutz der Meinungsfreiheit wird erörtert, inwieweit der KI-Einsatz unter die Medienfreiheit fällt, welche Nachrichtenakteure, die KI einsetzen, von der Medienfreiheit profitieren können und welche Pflichten sich für Nachrichtenmedien aus dem KI-Einsatz ergeben. Anschließend wird beschrieben, wie bestimmte KI-Anwendungen in den Nachrichtenmedien die Freiheit der Meinungsäußerung von anderen an der öffentlichen Debatte Teilnehmenden einschränken können, etwa von Nachrichtennutzerinnen und -nutzern, Bürgerinnen und Bürgern oder Politikerinnen und Politikern, die sich über die Medien Gehör verschaffen. Ein erschöpfender Überblick über alle Risiken des KI-Einsatzes bei Nachrichtenmedien wird nicht angestrebt. Vielmehr geht es darum, die Risiken der KI für die Meinungsfreiheit zu veranschaulichen, um den Wesensgehalt der Menschenrechte verschiedener an der öffentlichen Debatte Teilnehmenden zu diskutieren. Zum Abschluss geht das Kapitel der Frage nach, welche Pflichten Staaten in Bezug auf die Meinungsfreiheit angesichts des KI-Einsatzes bei Nachrichtenmedien haben. Zuvor jedoch wird kurz dargelegt, welche Ziele Nachrichtenmedien mit einem KI-Einsatz überhaupt verfolgen.

3.2. KI-Anwendungen für Nachrichtenmedien

Ähnlich wie in anderen Studien über KI und die Nachrichtenmedien wird künstliche Intelligenz für die Zwecke dieses Kapitels lose definiert als „eine Sammlung von Ideen, Technologien und Verfahren, die sich auf die Fähigkeit eines Computersystems zur Erledigung von Aufgaben beziehen, die normalerweise menschliche Intelligenz erfordern“.²¹⁸ KI ist somit ein Überbegriff für verschiedene digitale Technologien, darunter etwa maschinelles Lernen, Bilderkennung sowie Verarbeitung und Erzeugung natürlicher Sprache. Die anderen Kapitel in dieser Publikation zeigen die Vielfalt von Anwendungen für Medien auf, die als KI gelten.

In diesem Kapitel werden vier Ziele unterschieden, für die Nachrichtenmedien KI einsetzen können: Nachrichtensammlung, Nachrichtenproduktion, Nachrichtenverbreitung

²¹⁶ Council of Europe, Recommendation CM/Rec(2018)1 of the Committee of Ministers to member States on media pluralism and transparency of media ownership, https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680790e13#_ftn1; Council of Europe, Recommendation CM/Rec(2015)6 of the Committee of Ministers to member States on free, transboundary flow of information on the Internet, https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805c3f20.

²¹⁷ Siehe Kapitel 4 dieser Publikation.

²¹⁸ Beckett C., „New powers, new responsibilities: A global survey of journalism and artificial intelligence“, London School of Economics and Political Science, S. 16, <https://blogs.lse.ac.uk/polis/2019/11/18/new-powers-new-responsibilities/>.

und Moderation von Leserkomentaren.²¹⁹ Die ersten drei Ziele betreffen klassische journalistische Prozesse, und das vierte Ziel bezieht sich darauf, dass Online-Nachrichtenmedien manchmal Nutzerkommentare auf ihren Websites zulassen. Wie in einem kürzlich erschienenen Bericht der Parlamentarischen Versammlung festgestellt wurde, konzentrierten sich die meisten Diskussionen über den Einsatz der KI in Online-Kommunikationsprozessen auf die Moderation von Inhalten, während die Art und Weise, in der die KI die Online-Informationsumgebung gestaltet, ebenso wichtig ist.²²⁰

Um den journalistischen Prozess zu beginnen, können die Nachrichtenmedien KI zur Nachrichtensammlung einsetzen. Dies umfasst den KI-Einsatz, um Informationen und berichtenswerte Ereignisse zu finden, Ideen für Storys zu entwickeln und Ereignisse oder Themen zu verfolgen. Sobald die Informationen über potenzielle Storys gesammelt sind, kann die KI zur Nachrichtenproduktion eingesetzt werden. Dazu gehört der KI-Einsatz zum Schreiben von Nachrichtenbeiträgen (manchmal auch als „automatisierter Journalismus“ bezeichnet),²²¹ die Erstellung von Bildern und Videos, die Faktenprüfung oder die Umwidmung von Inhalten für neue Zielgruppen. Im letzten Schritt des journalistischen Prozesses können Nachrichtenmedien KI zur Verbreitung von Nachrichten einsetzen. Hierzu zählen der KI-Einsatz für personalisierte Empfehlungen, das Finden neuer Zielgruppen, die Vermarktung der Nachrichtenmarke und der Verkauf von Abonnements. Personalisierung dient oft einem doppelten Ziel. Zum einen hilft sie den Nachrichtenmedien, ihren Nutzerinnen und Nutzern besser zu dienen und, insbesondere bei öffentlich-rechtlichen Medien, ihren öffentlichen Auftrag zu erfüllen.²²² Zum anderen hilft Personalisierung den Nachrichtenmedien, Abonnenten zu halten, die Nutzereinbindung zu verstärken und dadurch Verkaufs- und Werbeerlöse zu generieren.²²³

Nachdem die Nachrichtenmedien ihre Online-Plattformen für Leserkomentare geöffnet haben, können sie diese durch KI-Einsatz effektiver moderieren. So hat etwa die *New York Times* ein System eingeführt, das maschinelles Lernen verwendet, um Kommentare für die Moderation zu priorisieren und Kommentare automatisch freizugeben.²²⁴ Die *New York Times* moderiert fast 12 000 Kommentare pro Tag, und durch das automatisierte System kann der Kommentarbereich länger offen sein, und Kommentare können schneller freigegeben werden.

Andere Forschungsarbeiten über den KI-Einsatz durch Nachrichtenmedien diskutieren den KI-Einsatz für die Moderation von Kommentaren im Rahmen der

²¹⁹ Beckett C., *Ibid*, p. 20.

²²⁰ Parlamentarische Versammlung, Bericht: „Need for democratic governance of artificial intelligence“ (24. September 2020), Randnr. 18-19, <https://pace.coe.int/en/files/28742/html>.

²²¹ Dörr K.N., „Mapping the field of algorithmic journalism“, *Digital Journalism* 4(6), S. 700–722, <https://doi.org/10.1080/21670811.2015.1096748>.

²²² Van den Bulck H. und H. Moe, „Public service media, universality and personalisation through algorithms: Mapping strategies and exploring dilemmas“, *Media, Culture & Society* 40(6), S. 875–92, <https://doi.org/10.1177/0163443717734407>.

²²³ Bodó B., „Selling news to audiences: A qualitative inquiry into the emerging logics of algorithmic news personalization in European quality news media“, *Digital Journalism* 7(8), S. 1054–75, <https://doi.org/10.1080/21670811.2019.1624185>.

²²⁴ Etim B., „The Times sharply increases articles open for comments, using Google’s technology“, *The New York Times*, <https://www.nytimes.com/2017/06/13/insider/have-a-comment-leave-a-comment.html>.

Nachrichtenverbreitung.²²⁵ Die automatische Moderation von Kommentaren birgt jedoch spezifische Risiken für die Freiheit der Meinungsäußerung derjenigen, die Nachrichtenbeiträge kommentieren und sich an öffentlichen Debatten beteiligen. Beim KI-Einsatz für andere Verbreitungsziele, wie Personalisierung und Marketing, sind diejenigen, die Nachrichten nutzen, zwar Adressaten der Kommunikation, äußern sich aber nicht selbst. In diesem Kapitel wird daher der KI-Einsatz für die Kommentarmoderation auf Nachrichtenplattformen als eigene Kategorie betrachtet. Nach diesem Überblick über die verschiedenen Möglichkeiten des KI-Einsatzes in den Nachrichtenmedien geht es im nächsten Abschnitt darum, inwieweit der KI-Einsatz durch die Medienfreiheit geschützt ist.

3.3. KI-Einsatz durch Nachrichtenmedien als Element der Medienfreiheit

In Europa bildet die Europäische Menschenrechtskonvention (EMRK) die rechtliche Grundlage für das Menschenrecht auf freie Meinungsäußerung. Nach Artikel 10 Absatz 1 EMRK hat jede Person das Recht auf freie Meinungsäußerung. Dieses Recht schließt die Meinungsfreiheit und die Freiheit ein, Informationen und Ideen ohne behördliche Eingriffe zu empfangen und weiterzugeben.

Der Europäische Gerichtshof für Menschenrechte (EGMR) wurde eingerichtet, um sicherzustellen, dass die Staaten ihren Verpflichtungen aus der EMRK nachkommen. Der EGMR hat die Freiheit der Meinungsäußerung in einer Vielzahl von Urteilen interpretiert und weiterentwickelt. In einem seiner ersten Fälle zur Freiheit der Meinungsäußerung bestätigte der EGMR, dass diese eine der wesentlichen Grundlagen einer demokratischen Gesellschaft und eine der grundlegenden Voraussetzungen der Entfaltung jedes Einzelnen ist.²²⁶

3.3.1. Die demokratische Rolle der Nachrichtenmedien

Journalismuswissenschaftler unterscheiden mehrere demokratische Rollen der Nachrichtenmedien.²²⁷ Die Medien dienen als Informationsquelle für demokratische Debatten, indem sie den Bürgerinnen und Bürgern Informationen zu Politik und Zeitgeschehen liefern. Darüber hinaus fungieren sie als „vierte Gewalt“, indem sie die Machtausübung von Staat, Unternehmen und anderen mächtigen Akteuren kritisch hinterfragen. Zudem dienen die Medien als Vermittler zwischen Bevölkerung und Politik, weil sie die Existenz eines öffentlichen Raums ermöglichen, in dem Angehörige beider

²²⁵ Beckett C., *Ibid.*, S. 28.

²²⁶ ECtHR, *Handyside gegen das Vereinigte Königreich* [1976], 5493/72, Randnr. 49, <http://hudoc.echr.coe.int/eng?i=001-57499>.

²²⁷ McNair B., „Journalism and democracy“ in T. Hanitzsch und K. Wahl-Jorgensen (eds.) *The Handbook of Journalism Studies*. Routledge, S. 237–49.

Sphären über Briefe, Gastkommentare, Diskussionssendungen und Beiträge zu Nachrichtenartikeln kommunizieren können.

Der EGMR hat diese verschiedenen demokratischen Rollen der Medien bekräftigt. So hat er die Auffassung vertreten, dass die Nachrichtenmedien Informations- und Kontrollaufgaben zu erfüllen haben.²²⁸ Hierzu hat er festgestellt, dass die Freiheit der Meinungsäußerung sowohl die Sammlung als auch die Veröffentlichung von Informationen²²⁹ und sowohl den Inhalt der Kommunikation als auch die technischen Mittel für die Verbreitung und den Empfang von Informationen schützt.²³⁰ Darüber hinaus hat der EGMR festgestellt, dass es auch zu den Aufgaben der Nachrichtenmedien gehört, Foren für öffentliche Debatten zu schaffen.²³¹

Der KI-Einsatz durch Nachrichtenmedien fällt unter die demokratischen Rollen, die durch das Recht auf freie Meinungsäußerung geschützt sind. Nachrichtenmedien können KI einsetzen, um Informationen zu neuen Themen zu sammeln, z. B. indem sie mithilfe der KI große Datenmengen analysieren. Darüber hinaus können Nachrichtenmedien KI einsetzen, um relevante Informationen in Abhängigkeit von den jeweiligen persönlichen Interessen und Informationsbedürfnissen an verschiedene Bürgerinnen und Bürger zu verbreiten. Nachrichtenmedien können KI auch nutzen, um das Verhalten von Großunternehmen oder die Umsetzung politischer Maßnahmen zu verfolgen und zu überwachen. Eine weitere Einsatzmöglichkeit der KI ist die Verbesserung des Forums für öffentliche Debatten durch die automatische Moderation von Leserkommentaren.

Aufgrund der demokratischen Rolle von Nachrichtenmedien ist deren Freiheit der Meinungsäußerung besonders geschützt. Der EGMR hat erklärt, dass die Freiheit der Meinungsäußerung im Fall der Nachrichtenmedien von besonderer Bedeutung ist.²³² Bei den Nachrichtenmedien spricht er daher von „Pressefreiheit“,²³³ die auch als Medienfreiheit bezeichnet wird.²³⁴ Laut EGMR haben Behörden bei der Entscheidung, ob für einen Eingriff in die Medienfreiheit eine dringende soziale Notwendigkeit besteht,²³⁵ einen geringeren Ermessensspielraum als bei sonstigen Eingriffen in die Freiheit der Meinungsäußerung. Darüber hinaus hat der EGMR erklärt, dass die Medienfreiheit die Nachrichtenmedien vor dem Einfluss mächtiger wirtschaftlicher oder politischer Gruppen in der Gesellschaft schützt und ihre redaktionelle Freiheit gewährleistet.²³⁶ Im Rahmen der Informationssammlung schützt die Medienfreiheit journalistische Quellen,²³⁷ und die

²²⁸ ECtHR, *Barthold gegen Deutschland* [1985], 8734/79, Randnr. 58, <http://hudoc.echr.coe.int/eng?i=001-57432>.

²²⁹ ECtHR, *Dammann gegen Schweiz* [2006], 77551/01, Randnr. 52, <http://hudoc.echr.coe.int/eng?i=001-75174>.

²³⁰ ECtHR, *Autotronic AG gegen Schweiz* [1990], 12726/87, Randnr. 47, <http://hudoc.echr.coe.int/eng?i=001-57630>.

²³¹ ECtHR, *Társaság a Szabadságjogokért gegen Ungarn* [2009], 37374/05, Randnr. 27, <http://hudoc.echr.coe.int/eng?i=001-92171>.

²³² ECtHR, *The Sunday Times gegen Vereinigtes Königreich (Nr. 1)* [1979], 6538/74, Randnr. 65, <http://hudoc.echr.coe.int/eng?i=001-57584>.

²³³ ECtHR, *The Sunday Times gegen Vereinigtes Königreich (Nr. 1)*, Randnr. 66.

²³⁴ Oster J., *Media Freedom as a Fundamental Right*. Cambridge University Press, S. 48.

²³⁵ ECtHR, *Busuioc gegen Moldau* [2004], 61513/00, Randnr. 65, <http://hudoc.echr.coe.int/eng?i=001-67745>.

²³⁶ ECtHR, *Manole und andere gegen Moldau* [2009], 13936/02, Randnr. 98, <http://hudoc.echr.coe.int/eng?i=001-94075>.

²³⁷ ECtHR [GC], *Goodwin gegen das Vereinigte Königreich* [1996], 17488/90, Randnr. 39, <http://hudoc.echr.coe.int/eng?i=001-57974>.

Medien können auch Anspruch auf Zugang zu Informationen der Behörden haben.²³⁸ Soweit der KI-Einsatz unter die Medienfreiheit fällt, sind Behörden also in der Regulierung der KI eingeschränkt. Im nächsten Abschnitt wird erörtert, für welche Akteure die Medienfreiheit gilt.

3.3.2. Für wen gilt die Medienfreiheit?

Heutzutage besteht das Umfeld der Nachrichtenmedien aus einem komplexen Netzwerk verschiedener Akteure, darunter Nachrichtenverlage, Nachrichtennutzerinnen und -nutzer und Online-Vermittler. Traditionelle wie auch digital entstandene Nachrichtenmedien sammeln Originalinformationen und veröffentlichen Nachrichtenartikel über eigene Offline- und Online-Nachrichtenkanäle. Vor dem Internetzeitalter erreichten die Nachrichtenverlage ihr Publikum im Allgemeinen direkt, wenn Menschen eine bestimmte Zeitung kauften oder einen bestimmten Radio- oder Fernsehkanal einschalteten. Im Internet können Nachrichtenverlage ihr Publikum immer noch direkt erreichen, wenn Menschen Nachrichten-Websites aufrufen oder die Apps der Verlage nutzen. Doch immer häufiger finden Menschen Nachrichten über soziale Medien, Suchmaschinen und Nachrichtenaggregatoren.²³⁹ Diese Plattformen fungieren als Vermittlungsinstanz zwischen Nachrichtenverlagen und Nachrichtennutzerinnen und -nutzern. Es liegt auf der Hand, dass die Medienfreiheit sowohl für traditionelle als auch für digital entstandene Nachrichtenmedien gilt. Die Mitwirkung von Online-Vermittlern im Nachrichtenumfeld wirft jedoch die Frage auf, inwieweit auch diese sich beim KI-Einsatz auf die Medienfreiheit berufen können.

Der EGMR hat festgestellt, dass die demokratischen Rollen, die die Medien traditionell wahrnehmen, von verschiedenen Akteuren erfüllt werden können. Er hat darauf hingewiesen, dass ein starkes öffentliches Interesse daran besteht, Interessengruppen die Möglichkeit zu geben, durch die Verbreitung von Informationen über Angelegenheiten von öffentlichem Interesse zu öffentlichen Debatten beizutragen.²⁴⁰ Der EGMR hat daher das Verhalten von Behörden in Bezug auf Interessengruppen vor dem Hintergrund der Medienfreiheit analysiert (aber auch festgestellt, dass Interessengruppen bestimmte Pflichten und Aufgaben zu erfüllen haben, die normalerweise den Medien vorbehalten sind; siehe nächsten Abschnitt). In einem Fall vertrat der EGMR die Auffassung, dass die Schaffung von Foren für öffentliche Debatten nicht auf professionelle Nachrichtenmedien beschränkt ist und dass auch Nichtregierungsorganisationen diese Rolle übernehmen können.²⁴¹ Er hat NGOs daher als „soziale Kontrollinstanzen“ bezeichnet. Bei Aktivitäten zivilgesellschaftlicher Organisationen könne daher ein ähnlicher Schutz im Rahmen der EMRK gerechtfertigt sein, wie er den Nachrichtenmedien gewährt wird.²⁴² Diese Urteile des

²³⁸ Ibid.

²³⁹ Newman N. et al., „Digital News Report 2020“, Reuters Institute for the Study of Journalism, S. 11–12, https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FINAL.pdf.

²⁴⁰ ECtHR, *Steel und Morris gegen das Vereinigte Königreich* [2005], 68416/01, Randnr. 89, <http://hudoc.echr.coe.int/eng?i=001-68224>.

²⁴¹ ECtHR, *Társaság a Szabadságjogokért gegen Ungarn*, Randnr. 27.

²⁴² Ibid.

EGMR könnten eine Grundlage bilden, auf der in Abhängigkeit von ihrer jeweiligen gesellschaftlichen Rolle auch Online-Vermittler als mit den Nachrichtenmedien vergleichbare Akteure anerkannt werden können.

Wie im vorigen Abschnitt erörtert, werden mit Blick auf die Freiheit der Meinungsäußerung drei demokratische Rollen für die Nachrichtenmedien rechtlich anerkannt: die Bereitstellung von Informationen für die Öffentlichkeit, die Schaffung eines Forums für öffentliche Debatten und die Rolle einer Kontrollinstanz. Online-Vermittler können mit KI-Unterstützung insbesondere zwei dieser Rollen erfüllen. Online-Vermittler können die Zugänglichkeit und Auffindbarkeit von Informationen durch personalisierte Nachrichtenfeeds und einfachen Zugang zu verschiedenen Nachrichtenverlagen verbessern. Darüber hinaus können sie Foren für öffentliche Debatten schaffen, indem sie es Nachrichtenverlagen, Politikerinnen und Politikern sowie Bürgerinnen und Bürgern erlauben, auf ihren Plattformen in öffentlichen und privaten Gruppen Inhalte zu veröffentlichen, und verschiedene Reaktionen auf Online-Inhalte ermöglichen, darunter das „Liken“, Kommentieren und Weiterleiten. Online-Vermittler können somit eine ähnliche Rolle spielen wie die Nachrichtenmedien und genau wie diese als Gatekeeper fungieren und Agenden bestimmen. Das Ministerkomitee des Europarats hat zudem angemerkt, dass Online-Vermittler „Formen der Kontrolle ausüben können, die den Online-Informationszugang von Nutzerinnen und Nutzern in vergleichbarer Weise beeinflussen wie die Medien, oder andere Funktionen wahrnehmen können, die denen von Verlagen ähneln“.²⁴³

Anhand dieser Grundsätze für die Freiheit der Meinungsäußerung könnte man argumentieren, dass sich Online-Vermittler auf die Medienfreiheit berufen können, wenn sie KI einsetzen, um ähnliche demokratische Rollen wahrzunehmen wie die Nachrichtenmedien. Zumindest können sich Online-Vermittler auf die „normale“ Freiheit der Meinungsäußerung berufen, wenn sie Nachrichten durch Nachrichtenfeeds und Suchergebnisse leichter zugänglich machen. Hierbei unterscheidet Van Hoboken zwischen der Produktion von „Informationen über Informationen“ durch Suchmaschinen, etwa bei der Veröffentlichung von Suchergebnissen, und dem Verweisen auf Informationen an anderer Stelle. Van Hoboken kommt zu dem Schluss, dass die Veröffentlichung von Suchergebnissen durch eine Suchmaschine nach Artikel 10 EMRK geschützt ist.²⁴⁴ Ebenso könnte man argumentieren, dass es sich bei der KI-gesteuerten Auswahl, Sortierung und Personalisierung von Nachrichtenfeeds durch soziale Medien und Nachrichtenaggregatoren um eine Produktion von „Informationen über Informationen“ handelt, die durch die Freiheit der Meinungsäußerung geschützt sein sollte.

Wenn traditionelle Nachrichtenmedien, digital entstandene Nachrichtenmedien und letztlich auch Online-Vermittler von ihrer Medien- oder Meinungsfreiheit Gebrauch

²⁴³ Council of Europe, Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries, Randnr. 5, https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=0900001680790e14.

²⁴⁴ Van Hoboken J. V. J., „Search engine freedom: On the implications of the right to freedom of expression for the legal governance of Web search engines“, University of Amsterdam, S. 182, <http://hdl.handle.net/11245/1.392066>.

machen, übernehmen sie auch bestimmte Pflichten. Mit diesen Pflichten beschäftigt sich der nächste Abschnitt.

3.3.3. Pflichten und Verantwortung und journalistische Ethikkodizes

Während der erste Absatz von Artikel 10 EMRK das Menschenrecht auf freie Meinungsäußerung garantiert, bestimmt der zweite Absatz, dass die Ausübung der Freiheit der Meinungsäußerung „mit Pflichten und Verantwortung verbunden“ ist und daher Einschränkungen unterworfen werden kann, die gesetzlich vorgesehen und in einer demokratischen Gesellschaft zu einem legitimen Ziel notwendig sind. Mit anderen Worten: Artikel 10 EMRK enthält einen Mechanismus, der sicherstellen soll, dass Personen und Organisationen, die ihre Meinungsfreiheit ausüben, dies in verantwortungsvoller Weise tun. Für die Zwecke dieses Kapitels liegt der Schwerpunkt auf Pflichten und Verantwortung und nicht auf den Bedingungen, unter denen ein Eingriff in das Recht auf freie Meinungsäußerung gerechtfertigt sein kann. Die Frage ist, welche Pflichten und Verantwortlichkeiten mit der Meinungsfreiheit verbunden sind und was diese Pflichten und Verantwortlichkeiten für den KI-Einsatz durch Nachrichtenmedien bedeuten.

Verschiedene Akteure haben Pflichten und Verantwortlichkeiten, wenn sie sich an der Ausübung der Meinungsfreiheit beteiligen oder dazu beitragen. Der EGMR hat festgestellt, dass neben den Sprechern oder Autoren selbst auch Personen oder Organisationen, die anderen Autoren ein Medium oder eine Plattform zur Verfügung stellen, also etwa Verlage²⁴⁵ oder Internet-Nachrichtenportale,²⁴⁶ mit Blick auf die Veröffentlichung und Verbreitung von Inhalten Dritter Aufgaben und Verantwortung übernehmen.

Der EGMR hat festgestellt, dass der Umfang der Pflichten und Verantwortlichkeiten eines Akteurs von verschiedenen Faktoren abhängt. Zunächst einmal hängen die Pflichten und Verantwortlichkeiten einer Person von ihrer Situation und den technischen Mitteln ab, die sie zur Kommunikation einsetzt.²⁴⁷ Da die Nachrichtenmedien eine besondere demokratische Rolle haben, kommt auch ihren Pflichten und Verantwortlichkeiten eine besondere Bedeutung zu. Besonders wichtig sind die Pflichten und Verantwortlichkeiten von Nachrichtenmedien dem EGMR zufolge, wenn ihre Arbeit die Rechte anderer beeinträchtigen könnte.²⁴⁸ Darüber hinaus ist auch der potenzielle Einfluss des Mediums ein Faktor, der den Umfang der Pflichten und Verantwortlichkeiten bestimmt.²⁴⁹ Je

²⁴⁵ ECtHR, *Éditions Plon gegen Frankreich* [2004], 58148/00, Randnr. 50, <http://hudoc.echr.coe.int/eng?i=001-61760>; ECtHR, *Chauvy und andere gegen Frankreich* [2004], 64915/01, Randnr. 79, <http://hudoc.echr.coe.int/eng?i=001-61861>.

²⁴⁶ ECtHR, *Magyar Tartalomszolgáltatók Egyesülete und Index.hu Zrt gegen Ungarn* [2016], 22947/13, Randnr. 62, <http://hudoc.echr.coe.int/eng?i=001-160314>; ECtHR [GC], *Delfi AS gegen Estland* [2015], 64569/09, Randnr. 113, <http://hudoc.echr.coe.int/eng?i=001-155105>.

²⁴⁷ ECtHR, *Handyside gegen das Vereinigte Königreich*, Randnr. 49.

²⁴⁸ ECtHR [GC], *Bladet Tromsø and Stensaas gegen Norwegen* [1999], 21980/93, Randnr. 65, <http://hudoc.echr.coe.int/eng?i=001-58369>.

²⁴⁹ ECtHR [GC], *Jersild gegen Dänemark* [1994], 15890/89, Randnr. 31, <http://hudoc.echr.coe.int/eng?i=001-57891>.

einflussreicher das Medium ist, desto umfangreicher sind die Pflichten und Verantwortlichkeiten. Dennoch hat der EGMR festgestellt, dass sich die Pflichten und Verantwortlichkeiten von Internet-Nachrichtenportalen in Bezug auf Inhalte Dritter ein Stück weit von denen traditioneller Verlage unterscheiden können.²⁵⁰ Ebenso hat das Ministerkomitee des Europarats empfohlen, dass die Pflichten und Verantwortlichkeiten von Online-Vermittlern angesichts der vielfältigen Rollen, die sie spielen, mit Blick auf ihre spezifischen Dienste und Rollen festgelegt werden sollten.²⁵¹

Nach Auffassung des EGMR sollten Nachrichtenmedien aufgrund ihrer Pflichten und Verantwortlichkeiten „in gutem Glauben handeln, um entsprechend der journalistischen Ethik korrekte und zuverlässige Informationen zu liefern“.²⁵² Journalistische Ethikkodizes gab es schon lange bevor der EGMR die Idee einführte, dass die Ausübung der Meinungsfreiheit mit Pflichten und Verantwortung verbunden ist. In den 1920er Jahren nahm die *American Society of Newspaper Editors* die „Canons of Journalism“ an, die als einer der ersten Ethikkodizes für die Nachrichtenmedien gelten.²⁵³ Doch in den Jahren danach wurden aufgrund der Kritik an den Presseunternehmen Forderungen nach einer staatlichen Regulierung der Nachrichtenmedien laut.²⁵⁴ Als Reaktion darauf führten die Nachrichtenmedien Ethikkodizes, Presseräte, Ombudsstellen und andere Formen der Selbstregulierung ein, um eine staatliche Regulierung zu verhindern.²⁵⁵

In Europa gilt die Idee, dass die Nachrichtenmedien nicht reguliert werden sollten, hauptsächlich für die gedruckte Presse. Die europäischen Länder regulieren audiovisuelle Medien auf verschiedene Weise, vor allem durch die EU-Richtlinie über audiovisuelle Mediendienste. Die Interpretation der Begriffe „Pflichten und Verantwortung“ durch den EGMR schafft nun auch eine normative rechtliche Grundlage für eine Theorie der sozialen Verantwortung der gedruckten Presse, wonach diese sich zu Selbstregulierung und Ethikkodizes für ihren Berufsstand verpflichten sollte.

Es stellt sich die Frage, ob die derzeitigen journalistischen Ethikkodizes auch dem KI-Einsatz durch Nachrichtenmedien gewachsen sind. Journalismuswissenschaftler kamen 2001 zu dem Schluss, dass das Internet neue ethische Fragen für Journalisten aufwirft und traditionelle journalistische Ethikkodizes keine ausreichende Orientierungshilfe für das Verhalten von Nachrichtenmedien im Online-Umfeld bieten.²⁵⁶ Mehr als ein Jahrzehnt später ergab die Forschung zu journalistischen Ethikkodizes, dass die meisten Kodizes immer noch keine Regeln für Online-Journalismus und digitale Medien enthalten.²⁵⁷ Helberger und Bastian fordern daher die Entwicklung einer algorithmischen journalistischen Ethik, die den Nachrichtenmedien eine Orientierungshilfe für den KI-Einsatz bei der Produktion,

²⁵⁰ ECtHR [GC], *Delfi AS gegen Estland*, Randnr. 113.

²⁵¹ Recommendation CM/Rec(2018)2 of the Committee of Ministers to Member States on the roles and responsibilities of Internet intermediaries, Randnr. 11.

²⁵² ECtHR [GC], *Bladet Tromsø und Stensaas gegen Norwegen*, Randnr. 65.

²⁵³ Ward S.J.A., *The invention of journalism ethics: The path to objectivity and beyond*. MQUP, S. 236–37.

²⁵⁴ Ward S.J.A., *Ibid*, S. 244.

²⁵⁵ Ward S.J.A., *Ibid*, S. 245.

²⁵⁶ Deuze M. und Yeshua D., „Online journalists face new ethical dilemmas: Lessons from the Netherlands“, *Journal of Mass Media Ethics* 16(4), S. 273–92, https://doi.org/10.1207/S15327728JMME1604_03.

²⁵⁷ Díaz-Campo J. und Segado-Boj F., „Journalism ethics in a digital environment: How journalistic codes of ethics have been adapted to the Internet and ICTs in countries around the world“, *Telematics and Informatics* 32(4), S. 735–44, <https://doi.org/10.1016/j.tele.2015.03.004>.

Publikation und Verbreitung von Nachrichten bietet.²⁵⁸ Ebenso kommen die beiden Kommunikationswissenschaftler Dörr und Hollnbuchner zu dem Schluss, dass Medienorganisationen ethische Kodizes für algorithmischen Journalismus einführen sollten.²⁵⁹

Eine weitere Frage lautet, inwieweit Online-Vermittler, die KI zur Gestaltung des Online-Umfelds für die Meinungsfreiheit einsetzen, Ethikkodizes befolgen sollten. Soziale Medien und Suchmaschinen haben lange versucht, sich mit dem Argument, sie seien keine Medien, der Verantwortung dafür zu entziehen, wie sie Informationen auswählen und priorisieren. In einem Interview betonte der CEO von Facebook, Facebook sei „ein soziales Netzwerk“ und er ziehe dies dem Begriff „soziale Medien“ vor, weil beim Begriff des sozialen Netzwerks mehr der „menschliche Teil“ und weniger der inhaltliche Teil der Plattform im Mittelpunkt stehe.²⁶⁰ Wenn soziale Medien tatsächlich nicht zum Journalismus gehören, dann müssen sie auch keinem journalistischen Ethikkodex folgen.

Somit zeigt sich, dass Artikel 10 EMRK in der Auslegung durch die Rechtsprechung des EGMR eine normative rechtliche Grundlage für Pflichten und Verantwortlichkeiten von Akteuren bietet, die zur Meinungsfreiheit beitragen – ganz gleich, ob es sich um „echte“ Nachrichtenmedienorganisationen handelt oder nicht. Gleichzeitig wird aber auch deutlich, dass sich die Pflichten und Verantwortlichkeiten von Online-Vermittlern von denen traditioneller Nachrichtenmedien unterscheiden können. Dies bedeutet, dass Online-Vermittler nicht zur Befolgung journalistischer Ethikkodizes verpflichtet werden können, obwohl die Grundsätze für die freie Meinungsäußerung deutlich machen, dass Online-Vermittler Verantwortung tragen, wenn sie KI einsetzen, um Meinungsäußerungen auf ihren Plattformen zu regulieren und die eigenen Rechte auf freie Meinungsäußerung wahrzunehmen. Wenn Online-Vermittler keine adäquaten Kodizes oder andere Selbstregulierungsinstrumente entwickeln, steht es dem Staat frei zu regeln, wie sie ihre Meinungsfreiheit beim KI-Einsatz ausüben. Ein ähnliches Denken zeigt sich auch im Umgang der Europäischen Kommission mit Online-Vermittlern. Die Europäische Kommission droht ihnen mit Regulierung, wenn sie sich nicht an Selbstregulierungskodizes für irreführende und falsche Informationen halten. Ein ähnlicher Ansatz könnte auch für den KI-Einsatz bei Online-Vermittlern verfolgt werden, um das Nachrichtenumfeld zu gestalten.

Nach der Analyse der Grundsätze der Meinungsfreiheit beim KI-Einsatz durch Nachrichtenmedien und andere Akteure, die im Online-Nachrichtenumfeld eine Rolle spielen, stellt sich als nächstes die Frage, wie sich der KI-Einsatz durch Nachrichtenmedien auf die Rechte auf freie Meinungsäußerung anderer an öffentlichen Debatten Teilnehmenden auswirkt.

²⁵⁸ Helberger N. und Bastian M., „AI, algorithms and journalistic ethics“, presented at the Future of Journalism conference, Cardiff, 2019.

²⁵⁹ Dörr K. N., „Mapping the field of algorithmic journalism“, *Digital Journalism* 4(6), S. 700–722, <https://doi.org/10.1080/21670811.2015.1096748>.

²⁶⁰ Swisher K., „Zuckerberg: The Recode interview“, *Recode*, <https://www.recode.net/2018/7/18/17575156/mark-zuckerberg-interview-facebook-recode-kara-swisher>.

3.4. Auswirkungen der KI auf das Recht auf freie Meinungsäußerung von Nachrichtennutzerinnen und -nutzern und anderen an öffentlichen Debatten Teilnehmenden

Der KI-Einsatz durch Nachrichtenmedien für die Sammlung, Produktion und Verbreitung von Nachrichten sowie die Moderation von Nutzerkommentaren birgt verschiedene Risiken für das Recht auf freie Meinungsäußerung von Nachrichtennutzerinnen und -nutzern und anderen Interessenträgerinnen und -trägern im Umfeld der Nachrichtenmedien. Zu beachten ist, dass der KI-Einsatz durch Nachrichtenmedien auch das Recht auf freie Meinungsäußerung und das Recht auf den Erhalt von Informationen von Nachrichtennutzenden und anderen an öffentlichen Debatten Teilnehmenden stärken kann. Aus regulatorischer Sicht stellt sich daher die Frage, wie die Risiken eingedämmt werden können, ohne einen positiven Einfluss der KI auf das Umfeld der Nachrichtenmedien und die öffentliche Debatte zu verhindern.

In der Phase der Nachrichtensammlung können Nachrichtenmedien KI einsetzen, um Trends und Fakten in großen Datenmengen zu identifizieren. Hierbei kann die KI in großen Datenmengen Originalstorys erkennen, die einem Menschen nicht auffallen würden. Der KI-Einsatz zur Nachrichtensammlung hängt jedoch von der Verfügbarkeit von (öffentlichen) Datensätzen ab. Ereignisse oder gesellschaftliche Probleme, zu denen es keinen großen Datensatz gibt, können für den Blick des automatisierten Story-Entdeckungssystems unsichtbar bleiben.²⁶¹ Die Stimmen der Menschen, die von Ereignissen und Storys betroffen sind, bei denen keine großen Datenmengen anfallen, kommen unter Umständen gar nicht in den Fokus datengetriebener Nachrichtenmedien und bleiben dadurch in öffentlichen Debatten womöglich ungehört.

In der Phase der Nachrichtenproduktion können Nachrichtenmedien KI einsetzen, um Texte und Bilder zu generieren, Informationen zu verifizieren und einem Faktencheck zu unterziehen, Texte automatisch zu übersetzen, Beiträge für soziale Medien zu schreiben oder massenproduzierte Storys auf bestimmte Zielgruppen zuzuschneiden. Der KI-Einsatz für die automatisierte Produktion von Inhalten kann zu rechtswidrigen Ergebnissen führen, die die Rechte anderer verletzen, wie etwa hetzerische oder verleumderische Inhalte oder Urheberrechtsverletzungen. Es stellt sich dann die Frage, wer für diese rechtswidrigen Inhalte verantwortlich oder haftbar ist: die Nachrichtenorganisation, die den Einsatz des KI-Tools beschlossen hat, der Entwickler des Tools oder die KI selbst?

Juristen haben argumentiert, es sei möglich, KI-Agenten zur Rechenschaft zu ziehen.²⁶² Amerikanische Juristen haben zudem auch argumentiert, der Erste Verfassungszusatz, der das verfassungsmäßige Recht auf freie Meinungsäußerung garantiert, sollte automatisierte Äußerungen schützen.²⁶³ Aus Sicht des positiven Rechts

²⁶¹ Hansen M. et al., „Artificial intelligence: Practice and implications for journalism“, Tow Center for Digital Journalism, S. 17, <https://doi.org/10.7916/D8X92PRD>.

²⁶² Hage J., „Theoretical foundations for the responsibility of autonomous agents“, *Artificial Intelligence and Law* 25(3), S. 255–71, <https://doi.org/10.1007/s10506-017-9208-7>.

²⁶³ Collins R.K.L. und Skover D. M., *Robotica: Speech rights & artificial intelligence*. Cambridge University Press.

kann KI jedoch keine Rechtspersönlichkeit haben und in europäischen Ländern daher auch nicht verantwortlich oder haftbar gemacht werden. Im Jahr 2017 forderte das Europäische Parlament die Europäische Kommission auf, die Schaffung eines speziellen rechtlichen Status („elektronische Person“) für Roboter zu erwägen.²⁶⁴ In ihrer Mitteilung über KI für Europa hat die Kommission eine solche elektronische Persönlichkeit allerdings nicht erwähnt.²⁶⁵ Dies deutet darauf hin, dass die Europäische Kommission bislang nicht vorhat, eine Rechtspersönlichkeit für Roboter oder KI-Systeme nach EU-Recht zu erwägen.²⁶⁶

Am sinnvollsten wäre es wohl, für von einer KI produzierte rechtswidrige Inhalte die Nachrichtenmedienorganisation zur Rechenschaft zu ziehen, die den KI-Einsatz beschließt. Wenn umgekehrt Behörden automatisch generierte Inhalte oder Bots zensieren wollen, die zur öffentlichen Debatte beitragen, dann könnten diese Tools und ihr Output sowohl über das Recht auf freie Meinungsäußerung der Nachrichtenmedienorganisation als auch über das Recht der Nachrichtennutzerinnen und -nutzer, Informationen zu erhalten, geschützt werden.²⁶⁷

In der Phase der Nachrichtenverbreitung können Nachrichtenmedien KI einsetzen, um das Nachrichtenangebot für jeden einzelnen Nachrichtennutzer zu personalisieren, wobei dessen Recht, Informationen zu erhalten, berührt sein kann. Das durch Artikel 10 EMRK geschützte Menschenrecht auf freie Meinungsäußerung schließt das Recht ein, Informationen zu erhalten. Der EGMR hat festgelegt, dass die Medien die Aufgabe haben, Informationen und Ideen zu Themen von öffentlichem Interesse bereitzustellen, und „die Öffentlichkeit auch ein Recht darauf hat, diese zu erhalten“.²⁶⁸ Darüber hinaus hat die Öffentlichkeit ein Recht darauf, angemessen informiert zu werden.²⁶⁹ Aus dem Recht, Informationen zu erhalten, ergibt sich auch, dass die Öffentlichkeit über die Medien Zugang zu vielfältigen Informationen haben sollte.²⁷⁰ Gleichzeitig haben Nachrichtennutzerinnen und -nutzer kein subjektives Recht, Informationen von den Medien zu erhalten.²⁷¹ Dennoch kann der KI-Einsatz durch Nachrichtenmedien die Wahrnehmung des Rechts, Informationen zu erhalten, beeinträchtigen, etwa wenn die Personalisierung die Vielfalt der Informationen verringert, zu denen Menschen Zugang haben. Aus einer durch das Verhalten von Nachrichtenmedien verursachte Einschränkung des Rechts, Information zu erhalten, können sich für Staaten positive Verpflichtungen ergeben (siehe nächsten Abschnitt).

Wenn Online-Vermittler mit Personalisierung arbeiten, können sie sich auch auf das Recht auf freie Meinungsäußerung von Nachrichtenmedien oder Bürgerjournalisten berufen, die Nachrichtenbeiträge auf ihren Plattformen veröffentlichen, um ein größeres

²⁶⁴ Europäisches Parlament, Bericht mit Empfehlungen an die Kommission zu zivilrechtlichen Regelungen im Bereich Robotik, Randnr. 59, https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_DE.html.

²⁶⁵ Europäische Kommission, Mitteilung der Kommission: Künstliche Intelligenz für Europa, <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:52018DC0237>.

²⁶⁶ Für eine eingehende Diskussion zum rechtlichen Status der KI in Bezug auf das Urheberrecht siehe Kapitel 5 dieser Publikation.

²⁶⁷ Kaminski M.E., „Authorship, disrupted: AI authors in copyright and First Amendment law“, *U.C. Davis Law Review* 51(2), S. 589–616.

²⁶⁸ ECtHR, *The Sunday Times gegen Vereinigtes Königreich (Nr. 1)*, Randnr. 65.

²⁶⁹ ECtHR, *The Sunday Times gegen Vereinigtes Königreich (Nr. 1)*, Randnr. 66.

²⁷⁰ ECtHR, *Manole und andere gegen Moldau*, Randnr. 100.

²⁷¹ Eskens S., Helberger N. und Moeller J., „Challenged by news personalisation: Five perspectives on the right to receive information“, *Journal of Media Law* 9(2), S. 259–84, <https://doi.org/10.1080/17577632.2017.1387353>.

Publikum zu erreichen. Als einer der ersten in Europa hat daher der deutsche Staat angesichts der Personalisierung in sozialen Medien und Suchmaschinen rechtliche Garantien für die Medienfreiheit und die Sichtbarkeit von Nachrichtenorganisationen entwickelt. Der neue deutsche Medienstaatsvertrag sieht vor, dass Online-Vermittler „journalistisch-redaktionell gestaltete Angebote, auf deren Wahrnehmbarkeit sie potentiell besonders hohen Einfluss haben, weder mittelbar noch unmittelbar unbillig behindern oder ohne sachlich gerechtfertigten Grund unterschiedlich behandeln“ dürfen.²⁷² Diese Rechtsvorschrift ist ein Novum im europäischen Medienrecht, und auch wenn abzuwarten bleibt, wie sich das Gesetz in der Praxis bewährt, erklären Organisationen für digitale Rechte, die neue deutsche Gesetzgebung habe „hohen symbolischen Wert“ und ihre Kernziele seien „lobenswert“.²⁷³

Darüber hinaus können Nachrichtenmedien KI einsetzen, um Nutzerkommentare auf ihren Websites zu moderieren. Hier kann das Recht auf freie Meinungsäußerung derjenigen berührt sein, die Kommentare posten. Untersuchungen zeigen, dass KI-Systeme in afroamerikanischem Englisch verfasste Posts auf sozialen Medien eher als beleidigend einstufen als Posts in der standardamerikanischen Sprachvariante.²⁷⁴ Wenn eine KI-gesteuerte Moderation von Inhalten bestimmte gesellschaftliche Gruppen benachteiligt, kann dies zu ungleichen Kommunikationschancen führen.

Weitere technische Einschränkungen für die automatisierte Inhaltsanalyse ergeben sich, wenn automatisierte Inhaltsmoderationssysteme Schwierigkeiten haben, den Kontext eines Leserkommentars zu verstehen, wenn Verarbeitungswerkzeuge für natürliche Sprache fehlen, die im jeweiligen Einsatzbereich trainiert wurden, und wenn bestimmte Sprechergruppen in den Trainingsdaten unterrepräsentiert sind.²⁷⁵ Diese technischen Einschränkungen können bei der Prüfung von Kommentaren zu falsch positiven wie auch zu falsch negativen Ergebnissen führen. Wie Llansó und ihre Kollegen anmerken, können falsch positive Ergebnisse die Meinungsfreiheit des Einzelnen beschneiden, während falsch negative Ergebnisse „dazu führen können, dass nicht gegen Hetze, Belästigungen und andere unerwünschte Inhalte vorgegangen wird, die eine abschreckende Wirkung auf die Bereitschaft mancher Personen und Gruppen zur Online-Teilnahme haben können“.²⁷⁶

²⁷² Helberger N., Leerssen P. und van Drunen M., „Germany proposes Europe’s first diversity rules for social media platforms“, *Media@LSE*, <https://blogs.lse.ac.uk/medialse/2019/05/29/germany-proposes-europes-first-diversity-rules-for-social-media-platforms/>.

²⁷³ Nelson, M., „Germany’s new media treaty demands that platforms explain algorithms and stop discriminating. Can it deliver?“, *AlgorithmWatch*, <https://algorithmwatch.org/en/new-media-treaty-germany/>.

²⁷⁴ Sap M. et al., „The risk of racial bias in hate speech detection“, in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2019, S. 1668–78, <https://www.aclweb.org/anthology/P19-1163>.

Die Parlamentarische Versammlung hat hervorgehoben, dass „die Verwendung von verzerrten Datensätzen oder von Datensätzen, die historische Verzerrungen, Vorurteile oder Diskriminierung widerspiegeln, eine Hauptursache für Diskriminierung bei der künstlichen Intelligenz ist“; siehe Parlamentarische Versammlung, Bericht: „Preventing discrimination caused by the use of artificial intelligence“ (29. September 2020), Randnr. 43, <https://pace.coe.int/en/files/28715/html>. Das oben genannte Beispiel zeigt, dass diese Gefahr auch im Bereich der Medien besteht.

²⁷⁵ Llansó E. et al., „Artificial Intelligence, content moderation, and freedom of expression“, Institute for Information Law, S. 7–8, <https://www.ivir.nl/publicaties/download/AI-Llanso-Van-Hoboken-Feb-2020.pdf>.

²⁷⁶ Llansó E. et al., *Ibid* S. 9.

Im Prinzip gibt das Recht auf freie Meinungsäußerung den Menschen kein Recht, sich auf privaten Plattformen zu äußern. Der EGMR hat festgestellt, dass die Freiheit der Meinungsäußerung „keine Forumsfreiheit für die Ausübung dieses Rechts“ beinhaltet.²⁷⁷ Wenn jedoch das Verbot des Zugangs zu Privateigentum jede wirksame Ausübung der Meinungsfreiheit verhindert oder die Essenz der Meinungsfreiheit zerstört, so der EGMR, könnten die Staaten eine positive Verpflichtung haben, die Wahrnehmung des Rechts auf freie Meinungsäußerung zu schützen, indem sie Eigentumsrechte regulieren.²⁷⁸ Hinzu kommt, wie in den vorangegangenen Abschnitten erörtert, dass Nachrichtenmedien oder Online-Vermittler im Grunde das Forum für öffentliche Debatten auch gestalten, wenn sie Kommentare moderieren. Die Schaffung eines Forums für öffentliche Debatten ist eine demokratische Rolle, die mit Pflichten und Verantwortung verbunden ist. Einige Nachrichtenmedien haben diese Verantwortung bei der Moderation von Kommentaren akzeptiert. So hat etwa die *New York Times* als Reaktion auf Studien zu der Frage, wie eine automatische Inhaltsmoderation diskriminierend wirken oder Verzerrungen verstärken kann, die automatische Moderation von Nutzerkommentaren im eigenen Haus untersucht.²⁷⁹

Nach dieser Diskussion über die Grundsätze der Meinungsfreiheit und die damit verbundenen Rechte und Pflichten sowie die Verantwortung verschiedener Interessenträger im Bereich der Nachrichtenmedien stellt sich abschließend die Frage, welche Verpflichtungen Staaten in Bezug auf die Meinungsfreiheit angesichts des KI-Einsatzes durch Nachrichtenmedien haben.

3.5. Verpflichtungen von Staaten in Bezug auf die Medienfreiheit

Die Rechte und Freiheiten in der EMRK sind als negative Rechte formuliert. Die Bestimmungen verbieten es den Behörden, in die Rechte und Freiheiten Einzelner einzugreifen. So heißt es etwa zur Freiheit der Meinungsäußerung in Artikel 10 Absatz 1 EMRK, dass jede Person das Recht auf freie Meinungsäußerung „ohne behördliche Eingriffe“ hat. Die Menschenrechte in der EMRK enthalten somit negative Verpflichtungen für Staaten.

Im Laufe der Jahre hat der EGMR jedoch akzeptiert, dass sich aus den Menschenrechten in der EMRK auch positive Verpflichtungen für Staaten ergeben können. In den 1960er Jahren akzeptierte der EGMR erstmals die Idee, dass Staaten im Rahmen bestimmter Rechte aus der EMRK positive Verpflichtungen haben können.²⁸⁰ Es dauerte einige Jahre, bis der EGMR auch in Bezug auf das Recht auf freie Meinungsäußerung positive Verpflichtungen anerkannte. In den 2000er Jahren stellte er dann aber fest, dass das Recht

²⁷⁷ ECtHR, *Appleby und andere gegen Vereinigtes Königreich* [2003], 44306/98, Randnr. 47, <http://hudoc.echr.coe.int/eng?i=001-61080>.

²⁷⁸ ECtHR, *Appleby und andere gegen Vereinigtes Königreich*, Randnr. 47.

²⁷⁹ Salganik M.J. und Lee R.C., „To apply machine learning responsibly, we use it in moderation“, *NYT Open*, <https://open.nytimes.com/to-apply-machine-learning-responsibly-we-use-it-in-moderation-d001f49e0644>.

²⁸⁰ ECtHR, *Case 'relating to certain aspects of the laws on the use of languages in education in Belgium' v. Belgium* [1968], Randnr. 27, <http://hudoc.echr.coe.int/eng?i=001-57525>.

auf freie Meinungsäußerung positive Verpflichtungen für Staaten enthalten kann, selbst im Bereich der Beziehungen zwischen Einzelpersonen.²⁸¹ So vertrat der EGMR die Auffassung, dass Staaten eine positive Verpflichtung haben, dafür zu sorgen, dass die Öffentlichkeit über Nachrichtenmedien Zugang zu objektiven, korrekten und vielfältigen Informationen hat und dass Journalisten diese Informationen weitergeben können.²⁸² Darüber hinaus stellte er fest, dass die Staaten die positive Verpflichtung haben, einen soliden Rechts- und Verwaltungsrahmen zu verabschieden, um den Pluralismus auf dem Markt für audiovisuelle Medien zu gewährleisten.²⁸³

Im Fall *Dink gegen die Türkei* formulierte der EGMR eine besonders starke positive Verpflichtung für Staaten. Danach müssen Staaten ein günstiges Umfeld für die Teilnahme aller Betroffenen an öffentlichen Debatten schaffen, damit diese ihre Meinungen und Ideen ohne Angst äußern können.²⁸⁴ Im Fall *Khadija Ismayilova* wiederholte der EGMR diese Aussage.²⁸⁵ In beiden Fällen ging es um Angriffe und Schikanen gegen Journalisten. Im Fall *Dink* stellte der EGMR fest, dass der Staat die positive Verpflichtung hat, einen Journalisten vor Angriffen von Personen zu schützen, die sich durch seine Veröffentlichungen beleidigt fühlen. Im Fall *Khadija Ismayilova* vertrat er die Auffassung, dass der Staat eine positive Verpflichtung hat, Eingriffe in das Privatleben einer Journalistin wirksamer zu untersuchen. McGonagle argumentiert, dass der Begriff eines günstigen Umfelds großes Potenzial hat.²⁸⁶ Offen ist jedoch noch, wie weit die positive Verpflichtung der Staaten zur Gewährleistung eines günstigen Umfelds für die Freiheit der Meinungsäußerung reicht. Aus Sicht des positiven Rechts schützt sie derzeit nicht vor den Risiken der KI für die Freiheit der Meinungsäußerung.

Auch wenn sich die zuletzt mehrfach bekräftigte Forderung des EGMR, ein günstiges Umfeld zu gewährleisten, möglicherweise nicht auf den KI-Einsatz durch Nachrichtenmedien erstreckt, zeigt das oben Gesagte, dass Staaten positive Verpflichtungen haben, die für einen KI-Einsatz relevant sein können. Wenn der KI-Einsatz durch Nachrichtenmedien die Vielfalt der Informationen, die Menschen erhalten, schmälert, können die positiven Verpflichtungen der Staaten berührt sein. Konkreter ausgedrückt: Wenn Personalisierung und automatisierte Inhaltsmoderation so allgegenwärtig werden, dass sie die Vielfalt der Nachrichtenmedieninhalte, die Menschen erhalten, verringern, können Staaten die positive Verpflichtung haben, dafür zu sorgen, dass Nachrichtennutzerinnen und -nutzer über KI-gesteuerte Online-Nachrichtenmedien vielfältige Informationen erhalten. Das Ministerkomitee des Europarats betont in seinen Leitlinien zu Medienpluralismus und Transparenz bei Medieneigentum auch, dass Staaten

²⁸¹ ECtHR, *Özgür Gündem gegen Türkei* [2000], 23144/93, Randnr. 43, <http://hudoc.echr.coe.int/eng?i=001-58508>; ECtHR, *Fuentes Bobo gegen Spanien* [2000], 39293/98, Randnr. 38, <http://hudoc.echr.coe.int/eng?i=001-63608>.

²⁸² ECtHR, *Manole und andere gegen Molday*, Randnr. 100.

²⁸³ ECtHR [GC], *Centro Europa 7 S.r.l. und Di Stefano gegen Italien* [2012], 38433/09, Randnr. 134, <http://hudoc.echr.coe.int/eng?i=001-111399>.

²⁸⁴ ECtHR, *Dink gegen Türkei* [2010], 2668/07, 6102/08, 30079/08, 7072/09, 7124/09, Randnr. 137, <http://hudoc.echr.coe.int/eng?i=001-100383>.

²⁸⁵ ECtHR, *Khadija Ismayilova gegen Aserbaidshan* [2019], 65286/13, 57270/14, Randnr. 158, <http://hudoc.echr.coe.int/eng?i=001-188993>.

²⁸⁶ McGonagle T., „Positive obligations concerning freedom of expression: Mere potential or real power?“, in Andreotti O. (ed.) *Journalism at risk: Threats, challenges, and perspectives*, Council of Europe, S. 9–35.

sich bemühen sollten sicherzustellen, dass für jeden Menschen „die größtmögliche Vielfalt an Medieninhalten, auch Inhalten von allgemeinem Interesse,“ zugänglich ist.²⁸⁷ Diese Leitlinien gelten auch für den KI-Einsatz durch Online-Nachrichtenmedien.

Auch auf gesellschaftlicher Ebene kann der KI-Einsatz durch Nachrichtenmedien und andere Akteure in deren Umfeld den Medienpluralismus bedrohen. Da große Nachrichtenmedien und Online-Vermittler auf mehr und bessere Nutzerdaten und leistungsfähigere KI-Technologien zugreifen können, haben sie gegenüber lokalen Nachrichtenmedien einen Wettbewerbsvorteil. Die Einführung von KI in der Nachrichtenindustrie für verschiedene Phasen des journalistischen Prozesses kann diese kleineren und lokalen Nachrichtenmedien aus dem Markt drängen, wodurch eine Abnahme des Medienpluralismus droht. Staaten haben eine positive Verpflichtung, den Medienpluralismus angesichts der steigenden Popularität der KI für das Nachrichtengeschäft zu gewährleisten, etwa indem sie für Online-Nachrichtenmedien beim Einsatz von KI und datengetriebenen Technologien gleiche Wettbewerbsbedingungen schaffen.

3.6. Fazit

In diesem Kapitel wurden die Auswirkungen des KI-Einsatzes durch Nachrichtenmedien auf das Menschenrecht auf freie Meinungsäußerung analysiert, das durch Artikel 10 der Europäischen Menschenrechtskonvention geschützt ist. Dabei wurde auf vier Ziele eingegangen, bei deren Verfolgung die Nachrichtenmedien KI einsetzen können: Nachrichtensammlung, Nachrichtenproduktion, Nachrichtenverbreitung und Moderation von Leserkommentaren. Der KI-Einsatz durch Nachrichtenmedien fällt unter die vom EGMR anerkannten und bestätigten demokratischen Rollen der Medien: die öffentliche Verbreitung von Informationen, die Rolle einer Kontrollinstanz und die Schaffung eines Forums für öffentliche Debatten.

Aufgrund ihrer demokratischen Rolle genießen die Nachrichtenmedien Medienfreiheit. Diese Medienfreiheit schützt den KI-Einsatz zum Sammeln, Veröffentlichen, Verbreiten und Empfangen von Informationen, da die Meinungsfreiheit sowohl den Inhalt der Kommunikation als auch die technischen Mittel schützt.

Online-Vermittler können ähnliche Rollen erfüllen wie die demokratischen Rollen der Medien, wenn sie die Zugänglichkeit von Informationen verbessern und öffentliche Debatten auf ihren Plattformen ermöglichen. Daher kann auch die Auswahl, Sortierung und Priorisierung von Nachrichten durch Online-Vermittler unter die Meinungs- oder sogar die Medienfreiheit fallen.

Die Nachrichtenmedien und andere Akteure, die mit ihrem KI-Einsatz die Freiheit der Meinungsäußerung ausüben oder dazu beitragen, übernehmen auch Pflichten und Verantwortlichkeiten. Diese Pflichten und Verantwortlichkeiten sind für die

²⁸⁷ Council of Europe, Appendix to Recommendation CM/Rec(2018)1, Guidelines on media pluralism and transparency of media ownership.

Nachrichtenmedien in journalistischen Ethikkodizes festgeschrieben und werden durch verschiedene Selbstregulierungsinstrumente durchgesetzt. Die aktuellen journalistischen Ethikkodizes enthalten jedoch keine Leitlinien für den KI-Einsatz. Einige Wissenschaftler fordern daher die Entwicklung einer algorithmischen journalistischen Ethik. Das Konzept der Pflichten und Verantwortlichkeiten bietet auch eine normative rechtliche Grundlage, um Online-Vermittler zu verpflichten, angemessene Selbstregulierungsinstrumente für den KI-Einsatz zu entwickeln, wenn sie zur Ausübung der Meinungsfreiheit beitragen. Wenn journalistische Ethikkodizes und andere Selbstregulierungsinstrumente weiterhin hinter den Realitäten des KI-Einsatzes im Umfeld der Nachrichtenmedien zurückbleiben, haben Staaten einen rechtlichen Grund, diesen Bereich zu regulieren, wenn die KI erhebliche Auswirkungen auf das Recht auf freie Meinungsäußerung von Nachrichtennutzerinnen und -nutzern und anderen an öffentlichen Debatten Teilnehmenden hat.

Der KI-Einsatz durch Nachrichtenmedien birgt für die Meinungsfreiheit unter anderem folgende Risiken: Ereignisse und Nachrichtenstorys, bei denen keine großen Datenmengen anfallen, können von den algorithmischen Augen automatisierter Nachrichtensammelsysteme übersehen werden, und die Stimmen der Menschen, die von diesen Storys betroffen sind, bleiben dadurch möglicherweise ungehört. Automatisierter Journalismus kann rechtswidrige Inhalte produzieren, die die Rechte und die Würde anderer Menschen verletzen. Die personalisierte Verbreitung von Nachrichten kann das Recht der Nachrichtennutzerinnen und -nutzer beeinträchtigen, vielfältige Informationen zu erhalten, das untrennbar zur Meinungsfreiheit gehört. Darüber hinaus kann die automatisierte Moderation von Nutzerkommentaren Minderheiten benachteiligen, was zu ungleichen Chancen auf Kommunikation und Teilnahme an öffentlichen Debatten führen kann.

Staaten können positive Verpflichtungen haben, dafür zu sorgen, dass jeder Mensch sein Recht auf freie Meinungsäußerung ungeachtet des KI-Einsatzes effektiv wahrnehmen kann. Staaten haben eine positive Verpflichtung, dafür zu sorgen, dass Nachrichtennutzerinnen und -nutzer vielfältige Nachrichten erhalten, und ein günstiges Umfeld für die Freiheit der Meinungsäußerung zu schaffen. Soweit der KI-Einsatz durch große Nachrichtenunternehmen und Online-Vermittler die Wettbewerbsfähigkeit kleinerer Akteure bedroht, können Staaten positive Verpflichtungen haben, beim KI-Einsatz durch Nachrichtenmedien gleiche Wettbewerbsbedingungen zu schaffen.

Diese Schlussfolgerungen stehen im Einklang mit den Forderungen von Menschenrechtsorganisationen und Organisationen für digitale Rechte zu den Auswirkungen der KI auf die Freiheit der Meinungsäußerung. Der UN-Sonderberichterstatter für die Förderung und den Schutz des Rechts auf Meinungsfreiheit und freie Meinungsäußerung empfiehlt den Staaten, „ein politisches und rechtliches Umfeld zu schaffen, das einem vielfältigen, pluralistischen Informationsumfeld zuträglich ist“, auch durch „Maßnahmen zur Gewährleistung eines Wettbewerbs im Bereich der künstlichen Intelligenz“.²⁸⁸ Der OSZE-Beauftragte für Medienfreiheit²⁸⁹ sowie die

²⁸⁸ Kaye D., „Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression“, United Nations, Randnr. 64, <https://freedex.org/wp-content/blogs.dir/2015/files/2018/10/Al-and-FOE-GA.pdf>.

²⁸⁹ OSCE Representative on Freedom of the Media, Artificial intelligence & freedom of expression, <https://www.osce.org/representative-on-freedom-of-media/447829?download=true>.

Menschenrechtsorganisationen *Privacy International* und *Article 19*²⁹⁰ weisen ebenfalls auf verschiedene Bedrohungen der Meinungsfreiheit durch KI hin.

Die aktuelle Herausforderung für Nachrichtenmedien und Behörden besteht darin, journalistische Ethikkodizes, Selbstregulierungsstandards und möglicherweise auch eine staatliche Regulierung zu entwickeln, um die Risiken der KI für die Meinungsfreiheit einzudämmen und gleichzeitig dafür zu sorgen, dass die KI zu öffentlichen Debatten, zum Medienpluralismus, zum freien Informationsfluss und zu anderen gesellschaftlichen Zielen beitragen kann. Für die Bewältigung dieser Herausforderung geben die in Artikel 10 EMRK verankerten Grundsätze öffentlichen und privaten Stellen konkrete Leitlinien an die Hand.

²⁹⁰ Privacy International und Article 19, Privacy and freedom of expression in the age of artificial Intelligence, <http://privacyinternational.org/report/1752/privacy-and-freedom-expression-age-artificial-intelligence>.

4. Politik für kulturelle Vielfalt im KI-Zeitalter

Mira Burri, Universität Luzern

4.1. Einführung

Inhaltliche Vielfalt ist eine zentrale Voraussetzung für einen lebendigen öffentlichen Diskurs, für kulturelle und soziale Integration und für den gesellschaftlichen Zusammenhalt. Daher wird kulturelle Vielfalt in der nationalen Medien- und Kulturpolitik seit Langem, vor allem in Europa, als Regulierungsziel definiert, und der Auftrag, sie zu schützen und zu fördern, wurde durch die UNESCO-Konvention zum Schutz der kulturellen Vielfalt von 2005 weiter gestärkt.²⁹¹ Trotz der sehr unterschiedlichen Umsetzung in den einzelnen Ländern ist kulturelle Vielfalt ein zentrales ordnungspolitisches Ziel geblieben, doch während sich das technologische Umfeld tiefgreifend verändert hat, sind einige grundlegende Fragen immer noch nicht beantwortet. In diesem Sinne sind zwei kritische Fragen zu stellen: Erstens inwieweit die Möglichkeiten des digitalen Mediums online Vielfalt ermöglicht, aber auch erschwert haben – sowohl in Bezug auf die Verfügbarkeit als auch auf die tatsächliche Nutzung vielfältiger Inhalte, und zweitens wie gut sich kulturpolitische Instrumentarien, wie sie jetzt angewandt werden, dafür eignen, die Auseinandersetzung mit kulturell vielfältigen Inhalten tatsächlich anzugehen und zu fördern. Dieser Beitrag wird zeigen, dass die Antworten auf diese Fragen nicht einfach sind und dass die Politik möglicherweise komplexe Kompromisse eingehen und bei der Umsetzung ihrer kulturpolitischen Maßnahmen innovativer sein muss, indem sie durch Vermittler und durch Technologien regiert.

Dieser Beitrag befasst sich mit den Möglichkeiten der digitalen Medien und insbesondere der künstlichen Intelligenz (KI) und mit deren Konsequenzen für inhaltliche Richtlinien, er beteiligt sich jedoch nicht an den breiteren Diskussionen über Kreativität im Zeitalter der KI.²⁹² Zudem fragt er weder, wie Vielfalt in der KI verankert wird, um voreingenommene Entscheidungen zu vermeiden,²⁹³ noch wie es um die Vielfalt in der KI-Industrie bestellt ist.²⁹⁴

²⁹¹ Siehe z. B. Burri M., „The UNESCO Convention on Cultural Diversity: An appraisal five years after its entry into force.“ *International Journal of Cultural Property* 20, 4, S. 357–380, November 2013.

²⁹² Siehe in diesem Zusammenhang Kulesz O., „Culture, platforms and machines: The Impact of Artificial Intelligence on the Diversity of Cultural Expressions“, report for UNESCO, DCE/18/12.IGC/INF.4, 2018.

²⁹³ Melendez C., „In AI, Diversity Is a Business Imperative“, *The Forbes*, 14. November 2019.

²⁹⁴ Siehe z. B. Paul K., „Disastrous‘ Lack of Diversity in AI Industry Perpetuates Bias“, *The Guardian*, 17. April 2019.

4.2. Das veränderte Umfeld der Erstellung, Verbreitung, Nutzung und Wiederverwendung von Inhalten

Die Transformationen in der digital vernetzten Umwelt, die die Allgegenwart des Internets in der Gesellschaft versinnbildlicht, sind vielschichtig und ihre Auswirkungen wurden im Laufe der Jahre, wenn auch nicht unumstritten, in einer Vielzahl hervorragender Studien erfasst.²⁹⁵ Durch die zentrale Bedeutung von Daten und die Dominanz mehrseitiger Märkte für Daten sowie den Aufstieg der KI ist eine neue Ebene der Komplexität entstanden, und um angemessene kulturelle Instrumentarien entwerfen zu können, ist es wichtig, die aktuelle Dynamik rund um Inhalte genau zu verstehen. Wir konzentrieren uns auf die spezifischen Entwicklungen, die für die Verfolgung von Zielen der kulturellen Vielfalt in dieser neuen Welt entscheidend sein können, und interessieren uns besonders für die veränderte Art und Weise, wie Inhalte online produziert, verbreitet, abgerufen und konsumiert werden, sowie für die damit verbundenen Veränderungen in den Mustern der Nutzererfahrung und -beteiligung, wann immer diese identifiziert werden können. Als Einstieg in diese komplexe Diskussion nutzt dieses Kapitel die sich wandelnde Rolle von Vermittlungsinstanzen als wichtige Gatekeeper.

4.2.1. Die neuen Vermittler

Es gab – teils auch belegte – Vermutungen, dass das digitale Umfeld eine Fülle, Vielfalt und Nutzerautonomie mit sich bringen werde, die unter den Bedingungen analoger Medien nicht möglich ist.²⁹⁶ Eines der Kernelemente, die diese positiven Aussagen stützen, ist, dass es im Cyberspace keine Vermittlungsinstanzen gibt und man jederzeit jeden Inhalt frei wählen kann. Doch wie die heutige Praxis der digitalen Medien zeigt, entspricht dies nicht unbedingt der Realität. Vielleicht gibt es *tatsächlich* verschiedene Vermittler, die auf unterschiedliche Weise kontrollieren, welche Auswahl wir *treffen* und welche Auswahlmöglichkeiten wir *sehen*. Dabei geht es hier nicht um die physischen Vermittler wie Netzbetreiber oder Internetdiensteanbieter (obwohl diese sehr wichtig sein können²⁹⁷), sondern um die Gatekeeper auf der Anwendungs- und Inhaltsebene – die sogenannten Auswahlvermittler („choice intermediaries“)²⁹⁸ oder neuen Redakteure („new editors“)²⁹⁹, die oft auch KI-Technologien einsetzen.

²⁹⁵ Siehe z. B. Benkler Y., *The wealth of networks: How social production transforms markets and freedom*. New Haven: Yale University Press, 2006); Sunstein C. R., *Republic.com 2.0*. Princeton: Princeton University Press, 2007.

²⁹⁶ Benkler Y., Weinberger D., *Everything is miscellaneous: The power of the new digital disorder*. New York: Henry Holt, 2007; Jenkins H., *Convergence culture: Where old and new media collide*. New York: New York University Press, 2008.

²⁹⁷ Benkler Y. (2006).

²⁹⁸ Helberger N., „Diversity label: Exploring the potential and limits of a transparency approach to media diversity.“ *Journal of Information Policy* 1. S. 337–369, 2011; Helberger N., „Diversity by design“, *Journal of Information Policy*, S. 441–469, 2011.

²⁹⁹ Miel P. und Farris R., *News and information as digital media come of age*, Cambridge: The Berkman Center for Internet and Society, 2008, S. 27.

Um die neue Medienwelt zu verstehen, kann ein Vergleich mit der Funktionsweise der alten Medien hilfreich sein. In der Offline- bzw. Analogwelt waren redaktionelle Aufgaben üblicherweise unter dem Dach einer einzigen Institution zusammengefasst. Redaktionelle Entscheidungen basierten auf einem bestimmten, begrenzten Kreis von Materialien, und die redaktionellen Produkte waren endlich, begrenzt durch die naturgegebenen Beschränkungen des betreffenden Mediums, etwa die Seiten einer gedruckten Zeitung oder die Länge einer Sendung. Zudem wurde das Zielpublikum meist in einem bestimmten Rhythmus angesprochen, der die Breite und Tiefe des Inhalts beeinflusste – zum Beispiel Tageszeitungen oder eine Wochenausgabe. Die redaktionellen Entscheidungen zu Inhalt und Format betrafen das gesamte Publikum einer bestimmten Publikation oder Sendung auf die gleiche Weise – sie waren nicht auf einen bestimmten Nutzer zugeschnitten. Je nach Format bestand auch ein gewisses Gleichgewicht zwischen lokalen, nationalen und internationalen Themen, die eingeordnet und vertrauenswürdig präsentiert wurden. Dies waren die zentralen redaktionellen Funktionen von Rundfunkveranstaltern und anderen traditionellen Medien, die in vielen Ländern auch den konkreten Auftrag hatten, lokale und nationale Inhalte zu bringen. Üblicherweise bestanden Mechanismen zur Überwachung der Erfüllung bestimmter quantitativer und qualitativer inhaltlicher Anforderungen. So gibt es in der Europäischen Union neben der Verpflichtung, auf audiovisuellen Kanälen überwiegend europäische Werke zu übertragen, auch Anforderungen und Verpflichtungen auf nationaler Ebene. Insgesamt hatten diese relativ klar definierten redaktionellen Funktionen weitreichende Auswirkungen auf die Produktion und Verbreitung von Wissen.³⁰⁰ Zudem stützten sie die der Medienpolitik fast aller Länder zugrunde liegende Überzeugung, dass sich die Vielfalt des Angebots in der Vielfalt des Konsums widerspiegeln werde.

Heute bietet sich dagegen ein ganz anderes Bild, weil digitale Medienformen diese analogen Beschränkungen aufheben und zu „fundamentalen Verschiebungen in der Zusammensetzung und im Konsum von Medienprodukten“³⁰¹ führen. Die „neuen Redakteure“ sind zahlreich, desintegriert und verteilt.³⁰² Die „neuen Redakteure“ sind KI-getrieben, und letztlich sind es die Algorithmen,³⁰³ die die neue Medienwelt definieren.

Der erste derartige Redakteur ist die *Aggregation*, der Prozess der maßgeschneiderten Zusammenstellung und ständigen Aktualisierung verschiedenartiger Inhalte. Ein solcher personalisierter Redakteur wird auf verschiedenen Plattformen für verschiedene Arten von Inhalten – ob Nachrichten, Unterhaltung oder Klatsch – angeboten. Scheinbar nahtlos und unentwegt generiert er automatisch Informationen, die auf bestimmte Nutzerprofile und/oder Vorerfahrungen zugeschnitten sind. Der Mechanismus, der diesen Content-Feed steuert, ist in der Regel ein spezifischer Algorithmus der

³⁰⁰ Weinberger D., *Too big to know*, New York: Basic Books, 2012.

³⁰¹ Miel und Farris, S. 27. Siehe auch Kleis Nielsen R., Gorwa R. und de Cock Buning M., *What can be done? Digital media policy options for strengthening European democracy*, Oxford: Reuters Institute Report, 2019.

³⁰² Ibid.; auch Latzer M., Hollnbuchner K., Just N. und Saurwein F., „The economics of algorithmic selection on the Internet“ in Bauer J. M. und Latzer M. (eds.), *Handbook on the economics of the Internet*. Cheltenham: Edward Elgar. S. 395–425, 2016.

³⁰³ Für eine umfassende Definition von Algorithmen siehe Latzer M. und Just N., „Governance by and of algorithms on the Internet“: Impact and consequences“, in *Oxford Research Encyclopedia, Communication*, Oxford: Oxford University Press, 2020.

betreffenden Plattform (z. B. Facebook oder Instagram), der zwischen verschiedenen Arten von Inhalten unterscheiden kann.³⁰⁴

Auch *Social Bookmarking* hat als Mechanismus zum Herausstellen von Inhalten immer mehr an Bedeutung gewonnen. Über verschiedene Ranking- und Bookmarking-Systeme wie Reddit, Technorati oder Del.icio.us fungiert hier die Crowd als Redakteur. Mit der weiten Verbreitung von Twitter und Instagram, insbesondere bei jüngeren Generationen, stieg die Verwendung von Hashtags als eine Art Metadaten-Tag, mit dem Nutzerinnen und Nutzer selbst dynamische Indizes erstellen können. Diese Mechanismen können nicht nur den Medienkonsum maßschneidern, sondern es auch schaffen, die Aufmerksamkeit großer Gruppen zu erregen.³⁰⁵ Dies gilt nicht nur für politische Kampagnen, sondern auch für die Mobilisierung der Aufmerksamkeit der Verbraucher bei Marketingkampagnen.

Darüber hinaus ist die *Suche* als digitale Vermittlungsfunktion heutzutage absolut essenziell. Sie ist der Ausgangspunkt für die meisten Online-Erfahrungen und der wichtigste Traffic-Treiber der meisten Websites. Inhalte, die im Netz nicht indiziert und suchbar sind, existieren im Grunde gar nicht.³⁰⁶ Die Suche wiederum wird meist durch proprietäre Algorithmen gesteuert, und das Geschäft konzentriert sich stark auf sehr wenige Anbieter, wobei Google sich deutlich von seinen Mitbewerbern absetzt.

4.2.2. Implikationen KI-getriebener redaktioneller Akteure

Über all diese verschiedenen Mechanismen fungiert das Netzwerk als Mehrkanal-Redakteur und als wichtige Vermittlungsinstanz in der Content-Wertschöpfungskette – es übernimmt de facto die Rolle der traditionellen Medien als „Vermittler des Allgemeininteresses“.³⁰⁷ Positiv ist angemerkt worden, dass „das vernetzte Medienumfeld als virtueller sozialer Geist [...] etwas produziert, das reichhaltiger, repräsentativer und offener für Ideen ist als das

³⁰⁴ Die Algorithmen kombinieren oft verschiedene Mechanismen und werden von verschiedenen Faktoren gesteuert: (1) Die allgemeine Beliebtheit eines Inhalts bei allen Nutzern ist der einfachste Ansatz, bei dem alle Nutzer dieselbe Empfehlung erhalten, sodass beliebte Inhalte letztlich noch beliebter werden und unbeliebte verschwinden. (2) Eine semantische Filterung empfiehlt Inhalte, die zu einem oder mehreren zuvor vom selben Nutzer verwendeten Inhalten nach einer Reihe vordefinierter Kriterien passen (z. B. Themen, Autor oder Quelle eines Artikels). (3) Kollaborative Filterung oder Social Information Filtering ist ein automatischer „Mundpropaganda“-Empfehlungsgenerator – Inhalte werden einem Nutzer auf der Basis von Werten empfohlen, die ihnen von anderen Personen mit ähnlichem Geschmack zugewiesen wurden. In der Regel kommen Hybridformen dieser Methoden zum Einsatz, darunter auch andere Methoden wie das Gewichten von Inhalten nach Aktualität oder das Puschen von Inhalten mit bestimmten Merkmalen, etwa von kostenpflichtigen Inhalten. Plattformen haben über die Jahre zudem große Mengen wertvoller Daten zum bisherigen Verhalten angesammelt und können den Content-Feed zusätzlich anhand von Nutzerdaten wie Alter oder Ort kalibrieren. Siehe Bozdag E., „Bias in algorithmic filtering and personalization“, in *Ethics and Information Technology* 15. S. 209–227, 2013.

³⁰⁵ Miel und Farris, S. 30.

³⁰⁶ Council of Europe, Draft Recommendation on the Protection of Human Rights with Regard to Search Engines, Strasbourg, 11 March 2010.

³⁰⁷ Sunstein (2007).

Top-down-Massenmedienmodell der Vergangenheit“.³⁰⁸ Wir sollten die Möglichkeiten digitaler Plattformen und die Kommunikations-, Partizipations- und Dialogprozesse, die sie ermöglichen, nicht unterschätzen, doch zumindest bisher bestehen große Unsicherheiten und sogar wachsende Zweifel an der Fähigkeit dieses selbstorganisierenden Mechanismus, maßgebliche Informationen zuverlässig zu identifizieren.³⁰⁹ Zudem gibt es eine gewisse Skepsis hinsichtlich seiner Auswirkungen auf die Vielfalt und Qualität der Inhalte und auf die Fähigkeiten der Nutzerinnen und Nutzer, vielfältige und vertrauenswürdige Inhalte zu finden und darauf zuzugreifen.

Wenn man im Zusammenhang mit der Diskussion in diesem Kapitel über die gesellschaftlichen Funktionen der Medien und das Ziel der kulturellen Vielfalt nachdenkt, könnte es sein, dass dieses komplexe Umfeld gewisse Gefahren einer geringeren Vielfalt und einer Fragmentierung des öffentlichen Diskurses birgt.³¹⁰ Erstens müssen wir die möglichen Eingriffe in die individuelle Autonomie und Wahlfreiheit der Nutzer anerkennen. Das Filtern reduziert zwar die Such- und Informationskosten und erleichtert die soziale Orientierung, wie Latzer et al. argumentieren,³¹¹ kann aber „dadurch kompromittiert werden, dass es soziale Risiken erzeugt, darunter Bedrohungen der Grundrechte und -freiheiten sowie Auswirkungen auf die Vermittlung von Realitäten und die zukünftige Entwicklung von Menschen“.³¹² Die zweite Sorge in diesem Zusammenhang betrifft die Auswirkungen einer maßgeschneiderten Medienproduktion und -nutzung. Bei der Medienproduktion besteht in jüngster Zeit ein Trend zu einer algorithmischen Inhaltsproduktion, bei der Algorithmen die Entscheidungsfindung in Medienunternehmen steuern, indem sie die Konsummuster und Präferenzen des Publikums vorhersagen.³¹³ In einigen Bereichen kann dies positiv gewertet werden, weil das Publikum bekommt, was es sich wünscht, doch in anderen Bereichen, z. B. bei Nachrichten, kann es höchst problematisch sein, wenn lokale Nachrichten und aktuelle Beiträge auf die demografischen, sozialen und politischen Variablen bestimmter Gemeinschaften zugeschnitten werden.³¹⁴ Zudem sollten wir an die sogenannten „Content-Farmen“ denken, die auf der Grundlage von Suchmaschinendaten (wie populären Suchbegriffen, Ad-Word-Verkäufen und tatsächlich verfügbaren Inhalten) schnell und kostengünstig Inhalte produzieren, um diese Nachfrage zu befriedigen. Dabei werden die Inhalte zur reinen Ware, sodass jeder im öffentlichen Interesse liegenden Funktion der Medien Schaden droht, auch im kulturellen Bereich.

Beim Medienkonsum fördert die Personalisierung der Medienkost anhand bestimmter Profile oder Vorerfahrungen „Inhalte, die sowohl geografisch nah als auch

³⁰⁸ Miel und Farris, S. 30.

³⁰⁹ Ibid.

³¹⁰ Siehe z. B. Sunstein C. R. *Going to extremes: How like minds unite and divide*, Oxford: Oxford University Press, 2009; Pariser E., *The filter bubble: What the Internet is hiding from you*, London: Viking, 2011.

³¹¹ Latzer et al.

³¹² Ibid., S. 29–30.

³¹³ Napoli P. M., „On automation in media industries: Integrating algorithmic media production into media industries scholarship“ in *Media Industries Journal* 1. S. 33–38, 2014; sowie Saurwein F., Just N. und Latzer M., „Governance of algorithms: Options and limitations“, *info* 17. S. 35–49, 2015.

³¹⁴ Napoli, *ibid.*, S. 34.

sozial und konzeptionell vertraut sind“³¹⁵ ... „Dies hält die Nutzer in vertrauten Grenzen und befriedigt ihre Neugier mit immer Gleichem. Wenn sie neue Inhalte oder Informationen suchen, verstärkt dies bestehende Meinungen und beseitigt allmählich abweichende Ansichten.“³¹⁶ Natürlich kann man den Standpunkt vertreten, dies sei auch bei den traditionellen Medien der Fall, wo sich die Menschen von Natur aus zu Inhalten hingezogen fühlen, die ihnen in der Vergangenheit gefallen haben. Der entscheidende Unterschied bei den neuen Medien besteht darin, dass die Nutzerinnen und Nutzer *nur diese* Inhalte sehen und ihre aktive Auswahl dadurch eingeschränkt oder manipuliert wird. Hoffman et al. argumentieren, dass soziale Medien diesen Effekt nur verstärken, indem sie zwei Dimensionen der Homophilie kombinieren: die Ähnlichkeit von Gleichgestellten und von Inhalten.³¹⁷ In diesem Zusammenhang ist zu bedenken, dass Social-Networking-Sites trotz eines leichten Rückgangs ihrer Nutzung als Einstiegspunkt für Inhalte und trotz der Unterschiede zwischen den Ländern³¹⁸ immer noch wichtige Gatekeeper sind. Dies verstärkt die Wirkung der Homophilie und verdeutlicht die Macht einiger weniger Akteure und die tiefgreifenden Auswirkungen ihrer Entscheidungen: Als etwa Facebook 2018 seinen Algorithmus änderte und Nachrichten herunterstufte, führte dies automatisch zu einem geringeren Nachrichtenkonsum.³¹⁹

Hinzuweisen ist zudem auf die Kommerzialisierung von Plattformen und die radikale Zunahme kommerziell oder politisch motivierter „Fake News“.³²⁰ Trotz der leichten Verschiebung hin zu zahlungspflichtigen Modellen für Nachrichten ist zu bedenken, dass der größte Teil des Online-Konsums nach wie vor über kostenlose Websites erfolgt, die weitgehend durch Werbung finanziert werden. Auch wenn ein Teil der aggregierten Inhalte aus traditionellen Medien stammt,³²¹ was einen Teil der üblichen Kritik zerstreuen mag, dass Aggregatoren die Wirkung unzuverlässiger nicht-traditioneller Quellen verstärken, so trifft es doch nach wie vor zu, dass Inhalte nicht in größerer Fülle vorhanden sind, sondern lediglich stärker verbreitet werden. In diesem Sinne haben wir nicht mehr und vielfältigere Inhalte, sondern einfach mehr vom Selben. Dennoch lässt sich mit Fug und Recht feststellen, dass die traditionellen Medien auf die technologisch ermöglichte Aggregation reagiert haben und online viel mehr Inhalte anbieten als in ihren Print- oder Rundfunkversionen. Speziell für Nachrichten hat das Reuters Institute for the Study of Journalism festgestellt, dass private Nachrichtenorganisationen viel Geld in soziale Medien investieren und erheblichen Traffic, Offsite-Reichweite und/oder zusätzliche digitale

³¹⁵ Hoffman C. P., Lutz C., Meckel M. und Ranzini G., „Diversity by choice: Applying a social cognitive perspective to the role of public service media in the digital age“, *International Journal of Communication*, 9, 2015, S. 1360–1381.

³¹⁶ Ibid.

³¹⁷ Ibid.

³¹⁸ Für Länderanalysen siehe Reuters Institute for the Study of Journalism, *Digital News Report 2018*, Oxford, 2018.

³¹⁹ Reuters Institute for the Study of Journalism (2018); sowie Tucker J. A. et al. *Social media, political polarization, and political disinformation: A review of the scientific literature*, prepared for the Hewlett Foundation, March, 2018.

³²⁰ Reuters Institute for the Study of Journalism (2018); Europäische Kommission, Bekämpfung von Desinformation im Internet: ein europäisches Konzept, COM(2018) 236 final, 26. April 2018.

³²¹ Reuters Institute for the Study of Journalism (2018). Aggregatoren können durch das Urheberrecht etwas eingeschränkt sein, siehe *Associated Press v. Meltwater U.S. Holdings, Inc.*, 931 F. Supp. 2d 537, 537 (S.D.N.Y. 2013) und neuere Initiativen auf dem Gebiet des EU-Urheberrechts.

Abonnenten melden.³²² Dies mag zwar den Zugang zu einer Vielzahl von Inhalten über mehr Plattformen ermöglichen und damit auch junge Menschen ansprechen, doch zwei Nachteile sind zu beachten: Der erste betrifft die fast vollständige Abhängigkeit der Medienorganisationen von Facebook, was ein gewisses „Plattformrisiko“ mit sich bringt, und der zweite ist die Tatsache, dass der Umgang privatwirtschaftlicher alter Nachrichtenorganisationen mit sozialen Medien stark von pfadabhängigen Geschäftsmodellen geprägt ist, die auf Werbung und Abonnements oder einen Mix aus beidem ausgerichtet sind.³²³ Dies wiederum führt möglicherweise nicht zu einem nachhaltigen Angebot mit vielfältigen lokalen, regionalen und nationalen Inhalten.³²⁴ Insgesamt kann es trotz der gestiegenen Menge an Inhalten weniger lokale, regionale und nationale Inhalte geben, die zudem schwer auffindbar sind, weil sie auf Online-Plattformen marginalisiert sind – oder werden.

Was Suchmaschinen als Vermittler betrifft, so mag es generell im langfristigen Interesse der Suchanbieter liegen, den Bedürfnissen ihrer Nutzerinnen und Nutzer – ob als Konsumentinnen und Konsumenten oder als Bürgerinnen und Bürger – gerecht zu werden. Eine vom britischen Ofcom durchgeführte Untersuchung deutet darauf hin, dass die Nachfrage nach nationalen öffentlich-rechtlichen Inhalten nach wie vor groß ist, sodass es auch weiterhin im Interesse der Suchanbieter liegen sollte, dass ihre Ergebnisse solche Inhalte angemessen herausstellen.³²⁵ Eine kürzlich durchgeführte vergleichende Studie ergab zudem, dass Menschen, die Nachrichten über Suchmaschinen finden, im Durchschnitt mehr Quellen für Online-Nachrichten nutzen, eher sowohl links- als auch rechtsgerichtete Online-Nachrichtenquellen nutzen und über ein ausgewogeneres Nachrichtenrepertoire verfügen.³²⁶ Die Suchergebnisse werden jedoch, wie bereits erwähnt, algorithmisch generiert und weisen Informationen automatisch eine bestimmte Relevanz zu. Die automatisierte Auswahl ist auch anfällig für Manipulationen durch verschiedene Techniken der Suchmaschinenoptimierung, wodurch gesponserte oder andere Inhalte sichtbar werden und mehr Aufmerksamkeit auf sich ziehen.³²⁷

³²² Sie identifizieren drei strategische Hauptziele, die den unterschiedlichen Umgang von Nachrichtenorganisationen mit sozialen Medien prägen: (1) Förderung des Onsite-Traffic durch Verweise, (2) Förderung der Offsite-Reichweite durch native Formate und verteilte Inhalte und (3) Förderung des Verkaufs digitaler Abonnements, oft auch durch Werbung für Inhalte auf Facebook.

³²³ Cornia A., Sehl A., Levy D. A. und Nielsen R. K., *Private sector news, social media distribution, and algorithm change*, Oxford: Reuters Institute for the Study of Journalism, 2018.

³²⁴ Eine Studie von US-Lokalmedien hat beispielsweise gezeigt, dass nur etwa 17 Prozent der Nachrichtenbeiträge, die einer Gemeinschaft angeboten wurden, wirklich lokal waren, also die Kommune oder dortige Ereignisse betrafen; weniger als die Hälfte (43 Prozent) der Nachrichtenbeiträge waren originale, also tatsächlich von dem lokalen Medium selbst produzierte Beiträge. Siehe Napoli P. M., Weber M., McCollough K. und Wang Q., *Assessing local journalism: News deserts, journalism divides, and the determinants of the robustness of local news*. News Measures Research Project, August, 2018.

³²⁵ Ofcom „Ofcom’s Second Public Service Broadcasting Review, Phase Two: Preparing for the Digital Future“, London: Ofcom, 2008, Randnr. 5.60.

³²⁶ Die Autoren verweisen auf das Phänomen der „automatisierten Serendipität“, das Menschen zu Quellen führt, die sie sonst nicht genutzt hätten. Siehe Fletcher R. und Nielsen R. K., „Automated serendipity“ in *Digital Journalism* 6. S. 976–989, 2018.

³²⁷ Siehe z. B. Bradshaw S., „Disinformation optimised: Gaming search engine algorithms to amplify junk news“, in *Internet Policy Review* 8. S. 1–24, 2019.

Zum Abschluss dieses Abschnitts, der nur eine Momentaufnahme der komplexen gegenwärtigen Medienlandschaft bietet, muss betont werden, dass alles im Fluss ist und die Folgen für die Fülle und Vielfalt der Inhalte und die Bedingungen der freien Meinungsäußerung ungewiss sind. Einerseits gibt es in der Literatur einen Diskurs, der unter Etiketten wie „Filterblasen“³²⁸ oder „Echokammern“³²⁹ auf das Risiko hinweist, dass die aktuelle maßgeschneiderte Medienkost zu einer Fragmentierung der öffentlichen Debatte und einer Polarisierung der Ansichten führt.³³⁰ Andererseits sind wir nicht sicher, inwieweit dies zutrifft. So stellte ein länderübergreifender Bericht 2017 fest, dass die Suche zwar eine wichtige Rolle bei der Meinungsbildung spielt, aber im Kontext mehrerer Medien betrachtet werden muss und nicht deterministisch ist.³³¹ Die Untersuchung der „automatisierten Serendipität“ – eines Phänomens, bei dem die Nutzerinnen und Nutzer zu Quellen geleitet werden, die sie sonst nicht konsultiert hätten – verringert auch die Ängste vor „Echokammern“.³³² Im gleichen Zusammenhang ist anzumerken, dass die derzeit angewandten Instrumente zur Verfolgung der Fragmentierung erstaunlich wenig über die Loyalitäten des Publikums und die Verlagerung der öffentlichen Aufmerksamkeit *zwischen* den Medien aussagen.³³³ Webster und Ksiazek etwa finden wenig Belege dafür, dass sich das Publikum aus ergebenen Loyalisten zusammensetzt.³³⁴ „Darüber hinaus können uns auch präziseste Expositionsmessungen nicht sagen, wie sich Inhalte auf Menschen auswirken. Denkbar ist, dass selbst eine kurze Exposition gegenüber Hetz- oder anderweitig obskuren Medien eine starke Wirkung auf Menschen ausübt, die danach suchen“,³³⁵ aber es könnte auch sein, dass die Gesamtwirkung durch andere Bestandteile der Medienkost kompensiert wird.³³⁶ In diesem Sinne sollten wir uns nicht auf Momentaufnahmen konzentrieren, sondern die Dynamik untersuchen und die zeitliche Entwicklung verfolgen.³³⁷

4.3. Handlungsmöglichkeiten: neue Werkzeuge für den Umgang und Dialog mit digitalen Vermittlern

Bei der Beschreibung des Wandels in der Medienlandschaft stellen wir fest, wie komplex die neuen „redaktionellen“ Prozesse sind und wie schwierig es für den Einzelnen ist, sich in diesem potenziell umfangreichen, aber verteilten Inhaltsangebot zurechtzufinden. Wir

³²⁸ Pariser (2011).

³²⁹ Sunstein C. R., *Infotopia: How many minds produce knowledge*. Oxford: Oxford University Press, 2006.

³³⁰ Siehe auch High Level Group on Media Freedom and Pluralism, *A free and pluralistic media to sustain European democracy*. Report prepared for the European Commission, S. 27, Januar 2013.

³³¹ Dutton W. H. et al. „Search and politics: The uses and impacts of search in Britain, France, Germany, Italy, Poland, Spain, and the United States“, Quello Center working paper No 5, S. 1–17, 18. Mai 2017.

³³² Fletcher und Kleis Nielsen (2018).

³³³ Webster J.G. und Ksiazek T. B., „The dynamics of audience fragmentation: Public attention in an age of digital media“, *Journal of Communication* 62, S. 39–56, 2012.

³³⁴ *Ibid.*, S. 40.

³³⁵ *Ibid.*, S. 51.

³³⁶ Sehr interessante Ergebnisse finden sich in Pew Research Center, „Political polarization and media habits: From Fox News to Facebook, how liberals and conservatives keep up with politics“, Washington, DC: Pew Research, 2014.

³³⁷ Zu neueren Trends beim Medienkonsum siehe Reuters Institute for the Study of Journalism (2018).

identifizieren zudem einige potenzielle Risiken des Konsums maßgeschneiderter Inhalte und der Polarisierung von Ansichten in diesem Umfeld, da die neuen Vermittler Angebot und Nachfrage algorithmisch steuern, indem sie Informationselemente auswählen und ihnen eine bestimmte Relevanz zuweisen.³³⁸ Vor diesem Hintergrund sind zwei gangbare Wege zur Einführung von Maßnahmen für kulturelle Vielfalt denkbar: Der erste besteht darin, sich das entstehende Umfeld vorzunehmen und die Algorithmen als kritische Gatekeeper zu regeln, da diese bisher weitgehend unreguliert geblieben sind, vor allem für kulturpolitische Zwecke, und der zweite ist der Einsatz der neuen Vermittler zur Förderung der Exposition gegenüber kultureller Vielfalt im Sinne einer „Regelung durch Vermittler“.

4.3.1. Regulierung von Algorithmen

Zur Regulierung („Governance“) von Algorithmen gibt es eine allgemeinere, nicht unbedingt kulturpolitische Debatte, die mit der Beobachtung zu tun hat, dass Vermittler, insbesondere wenn sie von Algorithmen gesteuert werden, eine kritische Rolle in der Online-Welt erlangt haben und in diesem Sinne diese nun *durch* Algorithmen geregelt wird.³³⁹ Diese Diskussion ist eng mit der Frage verbunden, wie mit dieser neuen Macht am besten umzugehen ist, also mit der Regulierung *von* Algorithmen.³⁴⁰ Besonders wichtig sind in diesem letztgenannten Zusammenhang Fragen nach dem Schutz der Privatsphäre,³⁴¹ aber auch Fragen zur Urheberrechtsdurchsetzung durch Vermittler, wie die jüngste EU-Urheberrechtsreform und die Diskussion um Artikel 17 der Richtlinie über das Urheberrecht im digitalen Binnenmarkt zeigen.³⁴² Latzer et al. nennen neun Risikokategorien, die sich aus der algorithmischen Auswahl ergeben und die möglicherweise angegangen werden müssen: (1) Manipulation, (2) abnehmende Vielfalt, Echokammern und Vorurteile, (3) Einschränkungen der Meinungsfreiheit, (4) Überwachung und Bedrohungen des Datenschutzes und der Privatsphäre, (5) soziale Diskriminierung, (6) Verletzung von Rechten des geistigen Eigentums, (7) Missbrauch von Marktmacht, (8) Auswirkungen auf kognitive Fähigkeiten und (9) wachsende Fremdbestimmung, Verlust der menschlichen Souveränität und Kontrollierbarkeit von Technologie.³⁴³ Besonders interessiert waren wir an den obigen Punkten (2), (3) und (9), die in unmittelbarem Zusammenhang mit den Kernzielen der kulturellen Vielfalt im Medienbereich stehen.

³³⁸ Saurwein et al. (2015), S. 35.

³³⁹ Eine ausgezeichnete Analyse und Rezension der Literatur findet sich in Saurwein et al. (2015).

³⁴⁰ Saurwein et al. (2015).

³⁴¹ Zum Beispiel hinsichtlich des Rechts auf Vergessenwerden, das jetzt in der Datenschutz-Grundverordnung der EU verankert ist.

³⁴² Siehe Richtlinie 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, ABL. L (2019) 130/92, <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32019L0790>. Siehe auch Montagnani M. L. und Trapova A. Y., „Safe harbours in deep waters: A new emerging liability regime for Internet intermediaries in the digital single market“, *International Journal of Law and Information Technology* 26. S. 294–310, 2018; Senftleben M., „Institutionalized algorithmic enforcement – The pros and cons of the EU approach to UGC platform liability“, *Florida International University Law Review* 14, 2020.

³⁴³ Saurwein et al. (2015), S. 37.

Saurwein et al. liefern eine sorgfältige Analyse der verschiedenen Regulierungsmöglichkeiten, mit denen diesen Risiken begegnet werden kann und die neben konventionellen Eingriffen in Form von Anordnungen und Kontrollen auch eine Regulierung durch Markt- und verschiedene Selbst- und Koregulierungslösungen im Zwischenbereich umfassen können.³⁴⁴ Die Autoren merken jedoch auch an, dass es bisher kaum Instrumente gibt, um dem Risiko von Vorurteilen, von Fremdbestimmung und von Auswirkungen auf kognitive Fähigkeiten zu begegnen. Im spezifischen Rahmen unserer Diskussion ist es sogar so, dass die meisten dieser Vermittlungsplattformen nicht in den Geltungsbereich der gegenwärtigen Medienregulierung fallen werden. Napoli hat in diesem Zusammenhang argumentiert, dass wir anfangen sollten, Algorithmen als Medieninstitutionen eigener Art aufzufassen.³⁴⁵ Seiner Ansicht nach sollte für Algorithmen eine restriktive Regulierung gelten – d. h. ein Verbot bestimmter Arten von Aktivitäten der Plattformbetreiber oder von Inhalten auf diesen Plattformen, um die Privatsphäre zu schützen und drastische Gewaltdarstellungen und Hetze zu bekämpfen. Napoli schlägt vor, dass wir angesichts der entscheidenden Rolle dieser neuen Vermittler wie schon bei den traditionellen elektronischen Medien „affirmative Ansätze im öffentlichen Interesse“³⁴⁶ entwickeln sollten, die auf etablierten medienpolitischen Grundsätzen wie Pluralität, Vielfalt und Lokalismus aufbauen – indem wir beispielsweise bestimmte Mengen oder Arten von Inhalten vorschreiben.³⁴⁷ Dies könnte zum einen den Mangel an vielfältigen, vertrauenswürdigen und lokalen Inhalten beheben, die trotz der Verfügbarkeit von mehr Inhalten im Internet knapp zu sein scheinen, und zum anderen die Zugangsmöglichkeiten der Nutzerinnen und Nutzer zu diesen Inhalten verbessern.³⁴⁸

In diesem Zusammenhang ist zu beachten, dass sich die Diskussionen über die Regulierung von Algorithmen, obschon noch relativ jung, rasant entwickelt haben. Im Zeitalter von Big Data, der Gefahren für den Schutz personenbezogener Daten und der Verbreitung von Fake News hat das Thema die Aufmerksamkeit der Politik und der breiten Öffentlichkeit auf sich gezogen, und es wurde erkannt, dass Handlungsbedarf besteht.³⁴⁹ In welcher Form gehandelt werden soll, ist jedoch noch offen. Bisher scheint jedes Handeln in den Bereich der Selbstregulierungs- und Soft-Regulation-Ansätze zu fallen, da harte Eingriffe nicht nur die Innovation von Plattformen und Nutzern behindern, sondern auch das eigentliche Ziel durchkreuzen können, die aktive und passive Redefreiheit zu fördern.

³⁴⁴ Saurwein et al. (2015).

³⁴⁵ Napoli P. M., „The algorithm as institution: Toward a theoretical framework for automated media production and consumption“, Paper presented at the Media in Transition Conference, Massachusetts Institute of Technology, May 2013; Napoli P. M., „Social media and the public interest: Governance of news platforms in the realm of individual and algorithmic gatekeepers. Media + the Public Interest Initiative Working Paper, 2014.

³⁴⁶ Napoli (2014), *ibid.*

³⁴⁷ Napoli P. M. „Social media and the public interest: Governance of news platforms in the realm of individual and algorithmic gatekeepers“, *Telecommunications Policy* 39. S. 751–760, 2015.

³⁴⁸ Siehe Napoli P. M. „Re-evaluating the long tail: Implications for audiovisual diversity on the Internet“, in Alborno L. A. und Garcia Leiva M. T. (eds.), *Audiovisual industries and diversity: Economics and policies in the digital era*. Abingdon: Routledge, 2019, Kapitel 5; Napoli P. M., *Social media and the public interest: Media regulation in the disinformation age*, New York: Columbia University Press, 2019.

³⁴⁹ Siehe z. B. Balkin J., „Free speech in the algorithmic society: Big data, private governance, and new school speech regulation“, *UC Davies Law Review* 51. S. 1149–1210, 2018. Auch in der EU gab es zahlreiche Maßnahmen zu Datenschutz und Fake News – etwa durch die Annahme der DSGVO. Siehe auch Europäische Kommission (2018).

In diesem Zusammenhang muss man bei der Planung geeigneter und zukunftsorientierter Regulierungsinstrumente die bisherigen Erfahrungen bei der Bekämpfung der Online-Desinformation sorgfältig berücksichtigen. Einerseits muss untersucht werden, inwieweit die Unternehmen auf das gestiegene öffentliche Bewusstsein und die Forderungen der Nutzerinnen und Nutzer nach Vertrauen und Qualität reagiert haben und inwieweit verschiedene technische (z. B. Anpassung der Algorithmen) und andere Lösungen (z. B. Zusammenarbeit mit Nutzern und anderen Organisationen) in diesem Fall wirksam zum festgelegten Ziel beigetragen haben – in diesem Fall zur Eindämmung der Menge und Verbreitung von Fake News.³⁵⁰ Die Bemühungen von Facebook nach dem US-Präsidentenwahlkampf 2016 könnten ein gutes Beispiel dafür sein. Als Reaktion auf verschiedene Vorwürfe unterstützte Facebook verschiedene Initiativen: Das Unternehmen verbreitete Lehrmittel zur Informationskompetenz, rief das Facebook Journalism Project ins Leben und schloss sich gemeinsam mit Partnern aus Wissenschaft und Medien der News Integrity Initiative an, die engagierte Gemeinschaften und integrativere Medien fördert und gleichzeitig versucht, Fehlinformationen besser zu verstehen. Konkret ging Facebook zur Eindämmung der Verbreitung von Fake News Partnerschaften mit etwa 40 externen Medienorganisationen ein, darunter Snopes, PolitiFact, Associated Press und FactCheck.org. Sie sind bestrebt, geteilte Nachrichtenbeiträge einem Faktencheck zu unterziehen und mit dem Hinweis „umstritten“ zu versehen, wenn sie den Check nicht bestanden haben. Facebook installierte zudem einen „Mehr Info“-Button, über den Nutzer zusätzlichen Kontext zu Artikeln in ihren Nachrichten-Feeds abrufen können.³⁵¹ Trotz dieser umfangreichen Bemühungen deuten Untersuchungen und Einzelberichte darauf hin, dass das Vorgehen von Facebook möglicherweise immer noch nicht ausreicht, um eine „gesunde“ Medienwelt sicherzustellen.³⁵²

Auch die Rolle von Regierungen, Zivilgesellschaft und anderen Nutzerorganisationen sollte berücksichtigt werden. Ein gutes aktuelles Beispiel für koordinierte Bemühungen und Zusammenarbeit an mehreren Fronten ist der europäische Ansatz für den Umgang mit Fake News.³⁵³ Die Europäische Kommission setzt sich dafür ein, die Transparenz, Vielfalt und Glaubwürdigkeit verbreiteter Informationen zu verbessern und hierfür langfristige integrative Lösungen zu entwickeln. Unter anderem berief die Kommission ein Forum mit verschiedenen Interessenträgern ein, um einen Rahmen für eine effiziente Zusammenarbeit zwischen den relevanten Interessenträgern – darunter Online-Plattformen, die Werbewirtschaft und große Werbetreibende – zu schaffen und ein Bekenntnis zur Koordinierung und Intensivierung der Bemühungen zur Bekämpfung der Desinformation sicherzustellen. Das erste Ergebnis des Forums war ein EU-weiter Verhaltenskodex zur Bekämpfung von Desinformation. Der im September 2018

³⁵⁰ Reuters Institute for the Study of Journalism (2018).

³⁵¹ Mit dem Button „Weitere Informationen“ bei Nachrichtenbeiträgen im Nachrichten-Feed können Nutzerinnen und Nutzer Folgendes abrufen: (1) Hintergrundinformationen von der Wikipedia-Seite über den Post-Ersteller, (2) andere Artikel, die der Post-Ersteller kürzlich gepostet hat, (3) eine Heatmap, die zeigt, wo auf der Welt der Artikel gerade geteilt wird und welche der eigenen Facebook-Freunde ihn geteilt haben.

³⁵² Levin S., „They don't care: Facebook fact-checking in disarray as journalists push to cut ties“, *The Guardian*, 13. Dezember 2018. Für eine vertiefte Analyse siehe Saurwein F. und Spencer-Smith C., „Combating disinformation on social media: Multilevel governance and distributed accountability in Europe“, *Digital Journalism*, 2020.

³⁵³ Europäische Kommission (2018).

verabschiedete Kodex legt Selbstregulierungsstandards zur Bekämpfung von Desinformation fest und soll die Ziele der Kommission erreichen, indem er ein breites Spektrum von Verpflichtungen festlegt, von Transparenz in der politischen Werbung über die Schließung von Fake-Accounts bis hin zur Verringerung der Einnahmen von Desinformationslieferanten. Der Kodex enthält zudem einen Anhang, der bewährte Praktiken nennt, deren Anwendung die Unterzeichner zusagen, um die Verpflichtungen des Kodex umzusetzen.³⁵⁴ Möglicherweise sind jedoch entschlosseneren Schritte in Richtung Rechenschaftspflicht und sogar eine Hinwendung zu Koregulierungsansätzen erforderlich,³⁵⁵ wie das akute Problem der Fake News zur COVID-19-Pandemie zeigt.³⁵⁶

4.3.2. Regulierung durch Algorithmen

Auf dem Weg zu einem zielgerichteteren Instrumentarium für kulturelle Vielfalt könnte man in Erwägung ziehen, sich für neue Formen der redaktionellen Intelligenz auszusprechen,³⁵⁷ als eine Art Vermittlung der digitalen Welt im öffentlichen Interesse, die darauf abzielt, die Sichtbarkeit, Auffindbarkeit und Nutzbarkeit verschiedener Arten von Inhalten zu erhöhen.³⁵⁸ Denkbar wären auch Instrumente, die Anreize für Expositionsvielfalt bieten, d. h. für den tatsächlichen Konsum vielfältiger Inhalte.³⁵⁹

Dies ist kein ganz neues oder exotisches Projekt. Der europäische Medienrahmen gemäß der Richtlinie über audiovisuelle Mediendienste (AVMD-RL)³⁶⁰ enthält den Vorschlag, dass die Förderung europäischer Werke in deren stärkerer „Herausstellung“ bestehen

³⁵⁴ Für alle Dokumente siehe <https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation>. Der jüngste Bericht der Gruppe europäischer Regulierungsstellen für audiovisuelle Medien (ERGA) über Desinformation zeigt einige Grenzen in den Verpflichtungen auf, die die Plattformen im Rahmen des Kodex eingegangen sind; siehe ERGA Report on Disinformation: Assessment of the Implementation of the Code of Practice, 2020, <https://erga-online.eu/?p=732>.

³⁵⁵ Saurwein und Spencer-Smith (2020).

³⁵⁶ Siehe z. B. „Social media firms fail to act on Covid-19 fake news“, *BBC News*, 4. Juni 2020.

³⁵⁷ Wir gehen hier nicht auf Maßnahmen zur Förderung der Medienkompetenz ein, die aus der nutzerzentrierten Perspektive ebenfalls wichtig sein können. Siehe Helberger (2011), S. 357; Burri M., „The global digital divide as impeded access to content“, in Burri M. und Cottier T.(eds.), *Trade governance in the digital age*. Cambridge: Cambridge University Press. S. 396–420, 2012; High Level Group on Media Freedom and Pluralism, „A Free and Pluralistic Media to Sustain European Democracy“, Bericht für die Europäische Kommission, Januar 2013.

³⁵⁸ Miel und Farris, S. 3; sowie Goodman E. P., „Public media 2.0“, in Schejter A. M. (ed.), *And communications for all: A public policy agenda for a new administration*. Lanham, MD: Lexington Books, S. 263–280, 2009; Webster J. G., „User information regimes: How social media shape patterns of consumption“, *Northwestern University Law Review* 104, S. 593–612, 2010.

³⁵⁹ Siehe Helberger N., „Media diversity from the user’s perspective: An introduction“, *Journal of Information Policy* 1, S. 241–245, 2011; Napoli P. M., „Exposure diversity reconsidered“, *Journal of Information Policy* 1, S. 246–259, 2011.

³⁶⁰ Die AVMD-RL wurde zuletzt 2018 überarbeitet, siehe Richtlinie (EU) 2018/1808 des Europäischen Parlaments und des Rates vom 14. November 2018 zur Änderung der Richtlinie 2010/13/EU zur Koordinierung bestimmter Rechts- und Verwaltungsvorschriften der Mitgliedstaaten über die Bereitstellung audiovisueller Mediendienste (Richtlinie über audiovisuelle Mediendienste) im Hinblick auf sich verändernde Marktgegebenheiten, ABL L (2018) 303/69.

kann.³⁶¹ Aus der Konsultation der Regulierungsbehörden im Jahr 2013³⁶² geht hervor, dass sich viele von ihnen für Instrumente der Herausstellung aussprachen, während sie einer Förderung über den Anteil an den Katalogen skeptisch gegenüberstanden. Viele betrachten diese Maßnahme als die effizienteste (auch weil sie sich auf den tatsächlich höheren Konsum europäischer Werke bezieht) und für die Betreiber am wenigsten belastende.³⁶³ Zudem ist es in Europa gängige Praxis, dass öffentlich-rechtliche Rundfunkveranstalter das Privileg haben, die ersten Plätze in den elektronischen Programmführern (EPGs) zu besetzen, und dadurch eine „angemessene Herausstellung“ erhalten.³⁶⁴ Foster und Broughton zeigen, dass EPGs für Konsumentinnen und Konsumenten ein wichtiges Instrument zur Suche und Auswahl von Programmen sind, und es gibt Belege dafür, dass Kanäle mit Plätzen im oberen Bereich eines EPG-Abschnitts bei der Auswahl der Zuschauer einen Vorteil gegenüber denen weiter unten haben.³⁶⁵ „Dieser Ansatz [des ‚Nudging‘, des Beeinflussens der Menschen zu den Entscheidungen, die sie in ihrem eigenen Interesse wie auch im Interesse der Gesellschaft insgesamt treffen sollen] hat bisher recht gut funktioniert.“³⁶⁶ Jüngste Erkenntnisse bestätigen, dass die Positionierung im EPG erheblichen Einfluss auf den Erfolg eines Kanals haben dürfte.³⁶⁷ Ein solches „Nudging“ funktioniert, wenn auch bei kommerziellen Medien, auch mit der Fernbedienung von Smart-TVs, wobei die großen Online-Anbieter wie Netflix, YouTube und Google Play als Buttons erscheinen, die einen direkten Zugriff ermöglichen.

Ebenso kann man einen tieferen Eingriff in Betracht ziehen, der mit einer Art Orientierungshilfe für die Nutzerinnen und Nutzer in Bezug auf „relevante“ und „hochwertige“ lokale, regionale oder nationale Inhalte verbunden ist, damit sie dann die „richtige Mischung“ konsumieren.³⁶⁸ In diesem Zusammenhang – der Kenntnis und der Serendipität – stellen sich zwei entscheidende Fragen: Kennen die Menschen die gesamte Breite des inhaltlichen Angebots, das ihnen online zur Verfügung steht? Und wie oft

³⁶¹ Siehe die neuesten EU-Empfehlungen: Mitteilung der Kommission, Leitlinien gemäß Artikel 13 Absatz 7 der Richtlinie über audiovisuelle Mediendienste für die Berechnung des Anteils europäischer Werke an Abrufkatalogen und für die Definition einer geringen Zuschauerzahl und eines geringen Umsatzes, 2020/C 223/03, ABL C (2020) 223/10.

³⁶² Siehe z. B. Europäische Audiovisuelle Informationsstelle, IRIS Spezial: Videoabrufdienste und die Förderung europäischer Werke. Straßburg: Europäische Audiovisuelle Informationsstelle, 2013.

³⁶³ Europäische Kommission (2014), S. 6.

³⁶⁴ EPGs sind auf EU-Ebene durch die Zugangsrichtlinie geregelt (Richtlinie 2002/19/EG des Europäischen Parlaments und des Rates vom 7. März 2002 über den Zugang zu elektronischen Kommunikationsnetzen und zugehörigen Einrichtungen sowie deren Zusammenschaltung, ABL L 108/7, 24. April 2002, geändert durch die Richtlinie 2009/140/EG des Europäischen Parlaments und des Rates vom 25. November 2009, ABL L 337/37, 18. Dezember 2009). Bei der Umsetzung der Richtlinie gibt es Unterschiede. So erlaubt die britische Regelung eine bevorzugte Behandlung öffentlich-rechtlicher Kanäle, während die EPG-Regulierung in Deutschland auf der Gleichbehandlung von öffentlich-rechtlichen und kommerziellen Sendern in den EPG-Listen beruht.

³⁶⁵ Foster R. und Broughton T., „PSB prominence in a converged media world“, Report commissioned by the BBC, London: Communications Chambers S. 12, 2012. Weitere Faktoren, die die Auswahl beeinflussen, sind eine einprägsame EPG-Kanalnummer und die Nachbarschaft zu einem anderen beliebten Kanal.

³⁶⁶ Foster und Broughton *op cit.*, S. 13–14. Dies wurde durch einen neueren Bericht bestätigt.

³⁶⁷ Ofcom, „EPG prominence: A report on the discoverability of PSB and local TV services“, London: Ofcom, 2018. Die Daten über die Wirkung der Herausstellung auf den VoD-Konsum sind weniger umfassend, deuten aber auf eine ähnliche Korrelation hin.

³⁶⁸ Helberger (2011), S. 346. Zur weiteren Rechtfertigung eines solchen Ansatzes siehe Sunstein C. R., „Television and the public interest“, *California Law Review* 88. S. 499–563.

stolpern sie über Inhalte, die ihnen gefallen, von deren Existenz sie aber nicht wussten?³⁶⁹ Das britische Ofcom hat gezeigt, dass die Hürden für Kenntnis und Serendipität hoch sein können.³⁷⁰

Ein solcher Eingriff könnte etwa über die bestehenden öffentlich-rechtlichen Rundfunksysteme erfolgen.³⁷¹ Zunächst kann man sich eine aktualisierte Variante des EPG als ein Instrument vorstellen, um sowohl den ÖRR-Sender als Marke als auch das lokale, regionale und nationale Inhaltsangebot stärker herauszustellen. Foster und Broughton sehen dieses „Nudging“ als einen zweistufigen Prozess, bei dem Zuschauer von dem ÖRR-Sender oder der ÖRR-Marke angelockt werden und dann verschiedene Techniken eingesetzt werden, um „die Zuschauerinnen und Zuschauer zu einem breiteren Spektrum an Inhalten zu führen, als sie sich selbst vielleicht sonst ausgesucht hätten“.³⁷² Die Autoren haben die Notwendigkeit neuer gesetzlicher Regelungen begründet, die die Herausstellung von ÖRR-Marken oder Marken einzelner Dienste³⁷³ auf Online-Plattformen sicherstellen. Die Anforderungen an die Herausstellung sollten für die Kernelemente jeder Verbraucherschnittstelle gelten, etwa für ein Kanalaraster oder ein Menü für Abrufdienste, und jeder ÖRR-Sender sollte erwarten, dass er mindestens ein Symbol bzw. einen Button auf der ersten Seite eines On-Demand-Führers oder dessen Entsprechung erhält.³⁷⁴ Dieselbe Argumentation lässt sich auch auf europäische Werke anwenden.³⁷⁵

Der zweite Vorschlag (zur Förderung der Serendipität) könnte ebenfalls hilfreich sein – „insbesondere, um die Zuschauer auf Inhalte hinzuweisen, nach denen sie sonst nicht suchen würden, oder um die Ansichten der Nutzer herauszufordern und ihr Wissen ‚zufällig‘ zu erweitern“.³⁷⁶ In diesem Zusammenhang haben Wissenschaftler betont, dass „auf Serendipität beruhende Begegnungen einige Bedenken hinsichtlich restriktiver Gegenstrategien und der Tendenz von Nutzerinnen und Nutzern, sich in ihren ‚Informationskokons‘³⁷⁷ zu verstecken, zerstreuen und Verständnis und Offenheit fördern

³⁶⁹ Ofcom „Ofcom’s Second Public Service Broadcasting Review, Phase Two: Preparing for the Digital Future“, London: Ofcom, 2008, Randnr. 3.95.

³⁷⁰ Ofcom (2008), Randnr. 3.98.

³⁷¹ Für eine umfassende Analyse siehe Burri (2015).

³⁷² Foster und Broughton *op. cit.*, S. 11.

³⁷³ Foster und Broughton *op. cit.* sprechen sich gegen eine Herausstellung einzelner Programme aus, da diese die Nutzererfahrung fragmentieren und der gesamten ÖRR-Marke schaden könne.

³⁷⁴ Foster und Broughton *op. cit.*, S. 4.

³⁷⁵ Die geänderte AVMD-RL enthält in Artikel 13 eine solche Regel: „Die Mitgliedstaaten sorgen dafür, dass die ihrer Rechtshoheit unterworfenen Mediendienstanbieter audiovisueller Mediendienste auf Abruf sicherstellen, dass ihre Kataloge einen Mindestanteil europäischer Werke von 30% enthalten und solche Werke herausgestellt werden.“ Erwägungsgrund 35 erläutert zudem: „... Die Kennzeichnung audiovisueller Inhalte, die als europäische Werke einzustufen sind, in deren Metadaten sollte gefördert werden, damit solche Metadaten Mediendienstanbietern zur Verfügung stehen. Zur Herausstellung gehört, dass europäische Werke durch Erleichterung des Zugangs zu diesen Werken gefördert werden. Eine Herausstellung kann durch verschiedene Mittel gewährleistet werden, beispielsweise durch einen speziellen Bereich für europäische Werke, der von der Hauptseite des Dienstes aus erreichbar ist, durch die Möglichkeit, mit dem als Bestandteil dieses Dienstes verfügbaren Suchwerkzeug nach europäischen Werken zu suchen, durch die Nutzung europäischer Werke in Kampagnen dieses Dienstes oder durch einen Mindestanteil europäischer Werke, für die im Katalog dieses Dienstes zum Beispiel mit Bannern oder ähnlichen Instrumenten geworben wird.“

³⁷⁶ Ofcom (2008a). Randnr. 3.99–3.101.

³⁷⁷ Helberger (2011a). S. 454.

könnten, womit sie auch demokratischen Zielen zugutekämen“.³⁷⁸ Die digitale Welt und die verschiedenen Möglichkeiten, Daten zu analysieren und Inhalte zu aggregieren, ermöglichen die zufällige Bereitstellung verschiedener Arten von Inhalten, die neben den „vom Zuschauer gewählten“ Inhalten oder in speziellen Listen „weniger gesuchter“, „weniger gesehener“ und anderer „weniger populärer“ Inhalte abseits des Mainstreams angezeigt werden können. Da sich Verfügbarkeit und Auffindbarkeit einzelner inhaltlicher Genres (z. B. von Sport- gegenüber Bildungsprogrammen) offenbar stark unterscheiden, könnte es sinnvoll sein, genreübergreifende Verknüpfungen herzustellen, um sowohl auf solche Inhalte hinzuweisen als auch die Chancen eines insgesamt vielfältigeren Konsums zu erhöhen.³⁷⁹

Allerdings ist bei diesen zufälligen Angeboten Vorsicht geboten, denn sie können einfach ignoriert werden oder sogar das Erlebnis des Zuschauers stören. Die Forschung hat gezeigt, dass bei auf Serendipität beruhenden Begegnungen mehr im Spiel sein muss als bloß Zufall. Schönbach erklärt, dass Überraschungen „in Vertrautem eingebettet“ sein müssen, damit sie funktionieren und Nutzern einen Anreiz bieten.³⁸⁰ Helberger führt weiter aus: „Damit in einer zufälligen Informationsexposition ein Sinn erkannt wird, müssen die Informationen an ein gewisses Vorwissen, Interesse oder Erlebnis des Nutzers anknüpfen.“³⁸¹ Ähnlich argumentieren Hoffman et al.: Wir können nur dann von einem „Erlebnis der Vielfalt“ sprechen, wenn Nutzer „diese Inhalte entsprechend ihrer Motivationen, Kenntnisse und Fähigkeiten wahrnehmen und verarbeiten“.³⁸² Die Entwicklung von Instrumenten, die für diesen Zweck gut funktionieren, kann eine schwierige Aufgabe sein, die auch mit dem Thema Medienkompetenz zu tun hat. Solche Instrumente können auch mit bestimmten algorithmischen Designfunktionen verbunden sein, als „Empowerment-Nudges, als Anstöße zur Autonomie, die eine Entscheidungsfindung im Interesse der Bürgerinnen und Bürger nach deren eigenem Ermessen fördern, ohne weitere Regulierung oder Anreize einzuführen oder zu manipulativen Maßnahmen zu greifen“.³⁸³

Insgesamt könnte es Raum geben, über Mechanismen nachzudenken, die als „gute Aggregatoren“ fungieren und die Sichtbarkeit, Verfügbarkeit und Nutzung von hochwertigen und vertrauenswürdigen lokalen, nationalen und regionalen Inhalten auf verschiedenen Plattformen fördern. Im KI-Zeitalter kann davon ausgegangen werden, dass die Entwicklung solcher intelligenten „Redakteure“ machbar ist. Die Abwägung zwischen

³⁷⁸ Ibid., unter Verweis auf Sunstein (2007). S. 27–28.

³⁷⁹ Für ein Experiment zur Förderung von inhaltlicher Vielfalt durch Empfehlungssysteme siehe Möller J. et al., „Do not blame it on the algorithm: An empirical assessment of multiple recommender systems and their impact on content diversity“, *Information, Communication and Society*, 2018.

³⁸⁰ Schönbach K. „The own in the foreign: Reliable surprise – An important function of the Media?“, *Media, Culture and Society* 29. S. 344–353, 2007.

³⁸¹ Helberger (2011a) S. 462.

³⁸² Hoffman et al. (2015) argumentieren, dass Nutzerinnen und Nutzer Vielfalt online nur dann erleben können, wenn sie Vielfalt anstreben, die Voraussetzungen von Vielfalt kennen und den Zugang zu Vielfalt sicherstellen können.

³⁸³ Hansen P. G. und Jespersen A. M. „Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy“, *European Journal of Risk Regulation* 1, S. 3–28, 2013, S. 24. Für eine umfassende Analyse in Bezug auf Expositionsvielfalt siehe Helberger N., Karppinen K. und D’Acunto L., „Exposure diversity as a design principle for recommender systems“, *Information, Communication and Society*, 2017.

der gewünschten Wirkung des Eingriffs und seinen möglichen Nebenwirkungen, die solche paternalistischen Maßnahmen naturgemäß haben, steht noch aus und muss sorgfältig angegangen werden.³⁸⁴

4.4. Schlussbemerkungen

Es liegt auf der Hand, dass sich die Medienlandschaft tiefgreifend verändert hat und immer noch im Wandel begriffen ist. Eine konkrete Veränderung, die für die Produktion, Verbreitung und Nutzung von Medien wirklich disruptiv ist, liegt in den redaktionellen Funktionen der Medien. Digitale Plattformen wie Facebook und Google haben, wenn auch in unterschiedlichem Ausmaß, Schlüsselfunktionen bei der Vermittlung von Inhalten übernommen und spielen daher nun eine entscheidende Rolle bei der Verwirklichung wichtiger öffentlicher Ziele, auch im kulturellen Bereich. Da sie Einfluss auf die Verfügbarkeit von und den Zugang zu lokalen, nationalen und regionalen Inhalten haben, können diese Vermittler auch auf Form und Inhalt des kulturellen Austauschs, auf die demokratische Beteiligung und auf das bürgerschaftliche Engagement einwirken. In den letzten Jahren sind die Risiken einer algorithmischen Filterung und einer maßgeschneiderten Medienkost, die erheblich eingeschränkt und/oder kommerzialisiert werden kann, verstärkt ins Bewusstsein gedrungen. Wissenschaft und Politik sind auf Bezeichnungen wie „Filterblasen“ oder „Echokammern“ aufmerksam geworden. Die tatsächlichen Auswirkungen der Vermittlung durch digitale Plattformen und der kausale Zusammenhang zwischen bestimmten Arten der Medienexposition und kulturellem, politischem und sozialem Engagement sind jedoch bislang ungeklärt. Dies gilt offenbar selbst für eindeutiger „schlechte“ Inhalte wie etwa Fake News. In diesem Sinne lassen sich zwei Lehren für die Politik formulieren:

- (1) Wir brauchen mehr Daten und unabhängige Forschung zur Verfügbarkeit verschiedener Arten von Inhalten, den Konsum und die Auseinandersetzung mit diesen Inhalten, die an diesem Prozess Beteiligten und die Auswirkungen dieser Prozesse auf individuelle und kollektive demokratische und kulturelle Akte.
- (2) Es ist wichtig, den Dialog zwischen Inhaltserstellern, Vermittlern, Nutzern, Werbetreibenden und anderen an der Dynamik der Medienwelt beteiligten Interessenträgern fortzusetzen und in diesem Dialog die entscheidende Bedeutung eines kulturell vielfältigen Medienkonsums zu unterstreichen. Der erhöhte Wert, der der Verfügbarkeit kulturell vielfältiger Auswahlmöglichkeiten beigemessen wird, die Betonung von Vertrauenswürdigkeit und Qualität, die von den Nutzerinnen und Nutzern verstanden und geschätzt wird, kann für Plattformen durchaus einen Anreiz darstellen, solche Optionen anzubieten. Tatsächlich gibt es bereits Schritte in diese Richtung – zum Beispiel in Bezug auf die Meldung oder Entfernung bestimmter

³⁸⁴ Helberger N., „Merely facilitating or actively stimulating diverse media choices? Public service media at the crossroad“, *International Journal of Communication* 9. S. 1324–1340, 2015; Bodo B. et al. „Tackling the algorithmic control crisis – The technical, legal and ethical challenges of research into algorithmic agents“, *Yale Journal of Law and Technology* 19. S. 133–180, 2017.

Arten von Inhalten oder bestimmter Nutzer oder in Bezug auf mehr Transparenz bezüglich der Quellen von Inhalten.

Möglicherweise sind auch noch weitere Maßnahmen erforderlich. Wir skizzieren zwei mögliche Wege, wie der Medienkonsum beeinflusst werden kann: die Regulierung *von* und *durch* Algorithmen. Für die Regulierung durch Algorithmen kann man einige „gute Aggregatoren“ in Betracht ziehen, die die Sichtbarkeit, Verfügbarkeit und Nutzung von hochwertigen und vertrauenswürdigen lokalen, nationalen und regionalen Inhalten fördern. Dies mag nach Interventionismus klingen und nach einem möglichen Eingriff in die Autonomie und Redefreiheit der Nutzerinnen und Nutzer und in die unternehmerische Freiheit der Plattformen, doch vielleicht gibt es Möglichkeiten, ein Design zu haben, das sensibel für Vielfalt ist und nicht im Widerspruch zu autonomen Entscheidungen steht, sondern vielmehr fundiertere Entscheidungen ermöglicht. Die Technologie dürfte viele Variationen des Themas zulassen, und die Politik muss hier möglicherweise offen bleiben und kann damit experimentieren, dass öffentlich-rechtliche Medien Medienerfahrungen kuratieren. Vorsicht ist nach wie vor geboten, und die Verfolgung von Vielfaltszielen passt vielleicht nicht in allen Fällen zu der praktischen Ausgestaltung der Empfehlungssysteme. So kann bei Suchmaschinen wie Google, wo Nutzer aktiv nach Antworten suchen, ein Kompromiss zwischen Richtigkeit und Vielfalt gefordert sein.³⁸⁵

Mit Blick auf die Rolle der Vermittler selbst und die Entschärfung der Risiken eines maßgeschneiderten und potenziell verzerrten Medienkonsums könnte es notwendig sein, im öffentlichen Interesse zu handeln. Wir können jedoch nicht einfach der Plattform oder dem Empfehlungssystem die Schuld geben und alle Maßnahmen auf sie ausrichten. Wie Helberger (2017) et al. anmerken, spielen auch Nutzerinnen und Nutzer „bei der Verwirklichung oder Aushöhlung öffentlicher Werte auf diesen Plattformen eine Rolle“.³⁸⁶ Wir haben es hier in der Tat mit einem „Problem of Many Hands“ – einem Problem gemeinschaftlichen Handelns – zu tun, und daher ist es erforderlich, einen Rahmen mit Beteiligung und unterschiedlichen Verantwortlichkeiten aller Interessenträger – Plattformen, Nutzer, Zivilgesellschaft und Regierungen – zu konzipieren.³⁸⁷ Als Modell hierfür können Multi-Stakeholder-Mechanismen dienen, die sich aus der Verwaltung des Internets ableiten.³⁸⁸ Besonders hilfreiche Erkenntnisse lassen sich aus den Erfahrungen gewinnen, die zuletzt im Bereich der Bekämpfung der Online-Desinformation in Europa gesammelt wurden. Die Realisierung der Verfügbarkeit vielfältiger Inhalte und fundierter und autonomer Auswahlentscheidungen der Nutzerinnen und Nutzer als zentrale öffentliche Werte in der Medienwelt sollte dann das Ergebnis einer dynamischen Interaktion und Beratung zwischen den Interessenträgern sein und kann zu einer Palette von Maßnahmen führen, darunter Verhaltenskodizes, Leitlinien und Prinzipien, staatliche

³⁸⁵ Adomavicius G. und Kwon Y., „Maximizing aggregate recommendation diversity: A graph-theoretic approach“, *Proceedings of the 1st International Workshop on Novelty and Diversity in Recommender Systems*. DiveRS 2011. Chicago, 2011.

³⁸⁶ Helberger N., Pierson J. und Poell T., „Governing online platforms: From contested to cooperative responsibility“, *The Information Society*, 2017, S. 2.

³⁸⁷ Ibid.

³⁸⁸ Siehe z. B. Marda V. und Milan S., „Wisdom of the crowd: Multistakeholder perspective on the fake news debate“, *A Report by the Internet Policy Observatory at the Annenberg School, University of Pennsylvania*, 21 May 2018.

oder nichtstaatliche Aufsichtsgremien, die einen kontinuierlichen und effektiven Dialog gewährleisten, oder bestimmte technologische Lösungen.³⁸⁹

³⁸⁹ Helberger et al. (2017).

Urheberrecht

*Zu den größten Ängsten vor KI zählt, dass Menschen durch Maschinen ersetzt werden. Die Menschen sind zunehmend besorgt, dass Roboter ihnen die Arbeit wegnehmen, und dieses Unbehagen hat auch den audiovisuellen Sektor erreicht. Es gibt immer mehr Beispiele für kreativen KI-Einsatz beim Schreiben von Drehbüchern und beim Komponieren von Musik, um nur zwei Aspekte zu nennen. Diese technobarbarische Invasion in den kreativen Bereich spielt bisher allerdings eine relativ geringe Rolle, sodass die Angst vor der Vernichtung kreativer Arbeitsplätze zumindest vorerst wohl unbegründet ist. Und doch hat die Frage der Urheberrechtsfähigkeit von Werken, die von Maschinen geschaffen wurden, die Wissenschaft überrumpelt. Die Frage liegt ja auf der Hand: Wenn wir uns einig sind, dass Maschinen Schöpfer eines Werks sein können, können sie dann auch das Urheberrecht an dem Werk haben? Oder kann eine natürliche oder juristische Person Inhaberin des Urheberrechts an einem Werk sein, das eine Maschine geschaffen hat? Der Überblick, den **Giancarlo Frosio** in seinem Beitrag zu dieser Publikation gibt, beantwortet verschiedene Rechtsfragen, die sich beim Thema KI und Kreativität neu stellen. KI kann auch gegen die Feinde der Kreativität eingesetzt werden, um Material, das Urheberrechte verletzt, im Internet zu finden und zu entfernen und Piraten und Branchenlecks aufzuspüren. Je nachdem, wie Algorithmen programmiert werden, besteht jedoch immer die Gefahr von falsch positiven Ergebnissen, und dies kann Auswirkungen auf die Freiheit der Meinungsäußerung von Internetnutzerinnen und -nutzern haben.*

5. Urheberrecht – Ist die Maschine eine Urheberin?

Giancarlo Frosio, Center for International Intellectual Property Studies (CEIPI), Universität Straßburg*

5.1. Einführung

Es heißt, künstliche Intelligenz (KI) sei für die Menschheit eine fundamental disruptive Revolution.³⁹⁰ Intelligente Maschinen in verschiedenster Form dienen vielerlei Zwecken und ersetzen den Menschen potenziell überall.³⁹¹ Positive wie auch negative Auswirkungen sind absehbar.³⁹² Offenbar hat KI das Potenzial, sogar Tätigkeiten zu übernehmen, die in der ureigensten Natur des Menschen liegen. Auch wenn die meisten Kreativen bisher nicht befürchten, durch Roboter ersetzt zu werden,³⁹³ ist die Kreativität ein wichtiger Bereich, in dem KI immer besser zu werden scheint. KI schreibt Gedichte, Romane und Nachrichtenartikel, komponiert Musik, bearbeitet Fotos, erstellt Videospiele, malt Bilder und produziert auch sonstige Kunst jeglicher Art. Die meisten Kreativbranchen werden

*Associate Professor, Center for International Intellectual Property Studies (CEIPI), Universität Straßburg; Non-Resident Fellow, Stanford Law School Center for Internet and Society; Faculty Associate, NEXA Center for Internet and Society. Ich danke meiner Forschungsassistentin Varnita Singh für ihre gründlichen Recherchen und ihre hervorragende kritische Unterstützung beim Verfassen dieses Kapitels.

³⁹⁰ Floridi L., *The Forth Revolution: How the Infosphere is Reshaping Human Reality*, OUP, Oxford. Bughin J. et al. (2017), *Artificial Intelligence: The Next Digital Frontier?*, McKinsey Global Institute Discussion Paper, www.mckinsey.com/-/media/mckinsey/industries/advanced%20electronics/our%20insights/how%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/mgi-artificial-intelligence-discussion-paper.ashx. Elsevier, *Artificial Intelligence: How knowledge is created, transferred, and used – Trends in China, Europe, and the United States*, www.elsevier.com/research-intelligence/resource-library/ai-report. Ménière Y., Rudyk I. und Valdes J., *Patents and the Fourth Industrial Revolution: The Inventions behind Digital Transformation*, München, DE: European Patent Office.

³⁹¹ ITU, „Assessing the Economic Impact of Artificial Intelligence“, Issue Paper No.1, International Telecommunications Union, Genf, S. 12-15, www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-ISSUEPAPER-2018-1-PDF-E.pdf. Pricewaterhouse Coopers (PwC), *Sizing the prize: PwC's Global Artificial Intelligence Study: Exploiting the AI Revolution*, PwC, www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html.

³⁹² ITU, *op.cit.*, S. 17-20.

³⁹³ Pfeiffer A., *Pfeiffer Report: Creativity and technology in the in age of AI*, S. 15, 29, <https://www.pfeifferreport.com/wp-content/uploads/2018/10/Creativity-and-technology-in-the-age-of-AI.pdf>.

davon erheblich betroffen sein,³⁹⁴ vom audiovisuellen Sektor³⁹⁵ bis hin zur Musik³⁹⁶ und zum Verlagswesen.³⁹⁷ Künstliche Intelligenz ist längst auch künstlerische Intelligenz – mit weitreichenden Folgen für das Urheberrecht.

In diesem Zusammenhang stößt die Anpassung des Systems des geistigen Eigentums (Intellectual Property, IP) an die Kreativität und Innovationskraft der KI (und die damit verbundenen Herausforderungen) auf zunehmendes Interesse.³⁹⁸ Zurzeit entsteht eine umfangreiche Literatur zum Thema KI und IP.³⁹⁹ Natürlich können bestehende IP-Regelungen, etwa zum Schutz von Urheberrechten, Geschäftsgeheimnissen und Patenten,⁴⁰⁰ die Software schützen, auf der die KI-Technologie basiert.⁴⁰¹ Der Schutz der Software erstreckt sich jedoch nicht auf den Output, den die KI möglicherweise generiert. Ob dieser Schutz verfügbar ist, ist aufgrund der Konstruktion des gegenwärtigen Urheberrechtsrahmens noch eine offene Frage. Zu unterscheiden ist auch zwischen computergestützter Kreativität, die urheberrechtlich geschützt werden kann, solange der Beitrag der Nutzerin oder des Nutzers Originalität aufweist, und der eigentlichen computergenerierten Kreativität, bei der die Nutzerinteraktion mit einem Computer diesen veranlasst, eigene Äußerungen zu generieren.⁴⁰² Ein Bericht der Europäischen Kommission verdeutlicht, welcher Sumpf sich hier abzeichnet:

³⁹⁴ New European Media (2019), *AI in media and creative industries*, <https://arxiv.org/ftp/arxiv/papers/1905/1905.04175.pdf>. Pfeiffer A., *op.cit.*

³⁹⁵ *Artificial Intelligence in the audiovisual industry, Summary of the EAO workshop*, Strasbourg, 17 December 2019, European Audiovisual Observatory, Strasbourg, 2019, <https://rm.coe.int/summary-workshop-2019-bat-2/16809c992a>. Siehe auch Baujard T., Tereskiewicz R., de Swarte A., Tuovinen T., „Entering the new paradigm of artificial intelligence and series“, Studie im Auftrag des Europarates und von Eurimages, Dezember 2019, <https://rm.coe.int/eurimages-entering-the-new-paradigm-051219/1680995331>.

³⁹⁶ Strum B. et al., „Artificial Intelligence and Music: Open Questions of Copyright Law and Engineering Praxis“, Arts 8, S. 115-129. BPI, *Music's smart future: How will AI Impact the Music Industry*, www.musicstank.co.uk/wp-content/uploads/2018/03/bpi-ai-report.pdf.

³⁹⁷ Lovrinovic C. und Volland H., *The future impact of artificial intelligence on the publishing industry*, Gould Finch und Frankfurter Buchmesse, <https://bluesyemre.files.wordpress.com/2019/11/the-future-impact-of-artificial-intelligence-on-the-publishing-industry.pdf>.

³⁹⁸ Cubert J.A. und Bone R.G.A., „The law of intellectual property created by artificial intelligence“ in Barfield W. und Pagallo U. (eds), *Research Handbook on the Law of Artificial Intelligence*, Edward Elgar, Cheltenham, S. 411-427. OECD, *Artificial Intelligence in Society*, OECD Publishing, Paris, S. 104-105, <https://doi.org/10.1787/eedfee77-en>. WIPO (2019a), *Draft Issues Paper on Intellectual Property Policy and Artificial Intelligence*, WIPO, Genf, https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=470053. WIPO (2019b), *WIPO Technology Trends 2019 – Artificial Intelligence*, WIPO, Genf, https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf.

³⁹⁹ Iglesias M., Shamuilia S. und Anderberg A., *Intellectual Property and Artificial Intelligence: A Literature Review*, Publications Office of the European Union, Luxembourg, https://publications.jrc.ec.europa.eu/repository/bitstream/JRC119102/intellectual_property_and_artificial_intelligence_jrc_template_final.pdf.

⁴⁰⁰ Unabhängig davon, ob Software in Abhängigkeit von der Rechtsordnung als computerimplementierte Erfindung oder als Software geschützt wird.

⁴⁰¹ Calvin N. und Leung J., „Who owns artificial intelligence? A preliminary analysis of corporate intellectual property strategies and why they matter“, *Future of Humanity Institute, University of Oxford*, <https://www.fhi.ox.ac.uk/wp-content/uploads/Patents-FHI-Working-Paper-Final-.pdf>.

⁴⁰² *Payer Components South Africa Ltd v Bovic Gaskins* [1995] 33 IPR 407. Clark R. und Smyth S., *Intellectual Property Law in Ireland*, Butterworths, Dublin. Denicola R., „Ex Machina: Copyright Protection for Computer-Generated Works“, *Rutgers University Law Review* 69, S. 269-270.

Der Schutz KI-generierter Werke [...] scheint [...] problematisch zu sein. Angesichts des humanistischen Ansatzes des Urheberrechts ist es fraglich, ob KI-generierte Werke urheberrechtlichen Schutz verdienen. [...] Während einige Wissenschaftlerinnen und Wissenschaftler im Bereich des Urheberrechts klar dafür plädieren, KI-generierte Werke als gemeinfrei zu definieren, haben andere eine Reihe von Vorschlägen unterbreitet, die darauf abzielen, ein gewisses Schutzniveau zu gewährleisten. Von bemerkenswerten Ausnahmen abgesehen, gehen diese Vorschläge [...] nicht immer ausreichend darauf ein, welche Elemente einem solchen Schutz zugrunde liegen könnten.⁴⁰³

Genau in dieser Hinsicht soll dieses Kapitel eine Reihe rechtlicher Fragen beantworten, die sich angesichts des Wirrwarrs um KI-generierte Kreativität stellen. Wie passt KI-generierte Kreativität zur traditionellen Urheberrechtstheorie und den bestehenden Doktrinen? Welches sind insbesondere die Bedingungen, die im Rahmen der wichtigsten Urheberrechtssysteme für den Schutz von Schöpfungen gelten, die von KI und tiefen neuronalen Netzen generiert wurden? Sollte eine Rechtspersönlichkeit für KI in Betracht gezogen werden? Ist KI nach traditionellen Urheberrechtsstandards eine Urheberin? Kann eine Maschine Originalität aufweisen? Diese Fragen – und dies sind nur einige der relevanten Fragen zu KI-generierter Kreativität – lassen sich in der einen Frage zusammenfassen, ob eine KI Urheberstatus haben kann. Darüber hinaus gibt es zwei weitere grundlegende Fragen, die über den Rahmen der vorliegenden Betrachtung hinausgehen. Sie betreffen die KI als (maschinell) Lernende und als Rechtsverletzerin. Hier geht es darum, ob eine KI zum einen durch den maschinellen Lernprozess und das Training, das sie zur Kreativität befähigt, und zum anderen durch die Schaffung von unerlaubtem Output Urheberrechte verletzen kann. Neben echten Herausforderungen im Zusammenhang mit Standards für den Urheberstatus von KI geht dieses Kapitel abschließend auf die weitere Entwicklung ein, indem politische Lösungsmöglichkeiten aus verschiedenen theoretischen Perspektiven betrachtet werden, beispielsweise Persönlichkeitstheorien und utilitaristische bzw. anreizbasierte Theorien des geistigen Eigentums.

5.2. Technologie

Das erste Buch, das je von einem Computer geschrieben wurde, war *The Policeman's Beard is Half Constructed: Computer Prose and Poetry by Racter*.⁴⁰⁴ Dies war 1984, und Racters Prosa war noch recht obskur und ungeschliffen. Doch seitdem hat sich viel getan. Die Qualität von KI-generierter Kreativität hat sich dramatisch verbessert: In Japan setzte sich in den ersten Runden eines Literaturwettbewerbs ein von einer Maschine verfasster Roman gegen Tausende menschlicher Autorinnen und Autoren durch,⁴⁰⁵ und der komplett von einer KI geschriebene Sci-Fi-Film *Sunspring* kam beim jährlichen Filmfestival Sci-Fi London unter

⁴⁰³ Craglia M., *Artificial Intelligence: A European Perspective*, Publications Office of the European Union, Luxembourg, S. 67-68.

⁴⁰⁴ Racter, *The Policeman's Beard is Half Constructed: Computer Prose and Poetry by Racter – The First Book Ever Written by a Computer*, Warner Books, New York.

⁴⁰⁵ Lewis D., *An AI-Written Novella Almost Won a Literary Prize*, Smithsonian Magazine, Washington, <https://www.smithsonianmag.com/smart-news/ai-written-novella-almost-won-literary-prize-180958577>.

die Top 10.⁴⁰⁶ AIVA betreibt – ebenso wie Amper oder Melodrive – eine KI, die Musik komponiert, die als Begleitung zu audiovisuellen Werken, Werbung oder Videospielen vermarktet wird.⁴⁰⁷ Die japanische Roboterband The Z-Machines verändert bei Auftritten das Tempo der Musik in Abhängigkeit von Aktionen des Publikums und der Personen, die ihre Website besuchen,⁴⁰⁸ und die Flow Machine von Sony kann mit einem musizierenden Menschen interagieren und gemeinsam mit ihm improvisieren.⁴⁰⁹ Der kreative Bereich, in dem die KI am besten abschneidet, scheint jedoch die bildende Kunst zu sein. Ein KI-generiertes „Porträt von Edmond de Belamy“ wurde bei Christie’s für verblüffende USD 432 500 verkauft.⁴¹⁰

Dank massiver Datenverfügbarkeit, verbesserter Rechenressourcen und neuartiger Architekturen auf Basis von Deep Learning hat die KI in den letzten zehn Jahren große Durchbrüche erzielt.⁴¹¹ In engem Zusammenhang mit diesen Fortschritten kam es zu einer grundlegenden Entwicklung von KI-generierter Kreativität durch das Aufkommen von Generative Adversarial Networks (GANs).⁴¹² Diese Entwicklung ist noch recht neu. Im Juni 2014 veröffentlichte Ian Goodfellow ein Paper mit dem Titel „Generative Adversarial Networks“ und stellte den Code unter einer BSD-Lizenz auf GitHub ein.⁴¹³ Das Paper beschreibt einen generativen Prozess, der für maschinelles Lernen ein Modell verwendet, bei dem zwei neuronale Netze in einem Spiel gegeneinander antreten. Anhand eines Trainingsdatensatzes lernt diese Technik, neue Daten zu generieren, die dieselben statistischen Merkmale aufweisen wie der Trainingsdatensatz. Dies wurde zu einer äußerst beliebten Methode, um eine KI mit großen Datensätzen zu trainieren. Die Technologie entwickelte sich weiter zu Creative Adversarial Networks (CANs), die auf GANs aufbauen und „Kunst erzeugen, indem sie Kunst betrachten und etwas über Stil lernen, und kreativ werden, indem sie durch Abweichungen von den erlernten Stilen das Erregungspotenzial der erzeugten Kunst steigern“.⁴¹⁴ Mit GANs und CANs erstellte das Pariser Kunstkollektiv Obvious das „Porträt von Edmond de Belamy“ und eine generative Bilderserie mit dem Titel „La Famille de Belamy“.⁴¹⁵

Ähnlich wie bei Googles DeepMind, das Musik generiert und aufführt oder Kunstwerke erschafft, geschieht dies bei KI durch das Anhören anderer Musik oder die Analyse früherer Kunstwerke im Internet. Pindar Van Arman bringt einer KI seit einiger Zeit bei, kreativ zu sein. Das Projekt cloudpainter.com verdeutlicht die Ähnlichkeiten zwischen menschlichen und maschinellen Lernprozessen zur Schaffung von Kunst.⁴¹⁶ Offenbar würde

⁴⁰⁶ Craig C. und Kerr I., „The Death of the AI Author“, *Osgoode Legal Studies Research Paper*, S. 1-2.

⁴⁰⁷ AIVA, <https://www.aiva.ai>.

⁴⁰⁸ Bakare L., *Meet Z-Machines, Squarepusher's new robot band*, The Guardian, <https://www.theguardian.com/music/2014/apr/04/squarepusher-z-machines-music-for-robots>.

⁴⁰⁹ Deltorn J.M. und Macrez F. (2018), „Authorship in the Age of Machine Learning and Artificial Intelligence“, Center for International Intellectual Property Studies Research Paper No. 2018-10, 22-23, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3261329.

⁴¹⁰ Craig C. und Kerr I., *op.cit.*, 3-4.

⁴¹¹ Goodfellow I., Bengio Y. und Courville A., *Deep Learning*, MIT Press, Cambridge.

⁴¹² Svedman M., „Artificial Creativity: A case against copyright for AI-created visual work“, *IP Theory* 9(1), S. 3-4.

⁴¹³ Goodfellow I. et al., „Generative Adversarial Networks“, *arXiv*, <https://arxiv.org/abs/1406.2661>.

⁴¹⁴ Elgammal A. et al., „CAN: Creative Adversarial Networks Generating ‚Art‘ by Learning About Styles and Deviating from Style Norms“, S. 1-22, <https://arxiv.org/abs/1706.07068>.

⁴¹⁵ Obvious AI & Art, <https://obvious-art.com>.

⁴¹⁶ Cloud Painter, <https://www.cloudpainter.com>.

eine KI in einem mehrstufigen Lernprozess lernen, Kreativität zu generieren – angefangen von technischen Übungen wie dem Verbinden von Punkten zu Bildern über Experimente und Imitation bis zur unabhängigen Kreation.

5.3. Schutz: Ist KI-generierte Kreativität schutzfähig?

KI verändert den kreativen Prozess und hat Auswirkungen auf etablierte urheberrechtliche Konzepte und Doktrinen. Echte Probleme ergeben sich insbesondere in Bezug auf die Schutzfähigkeit von KI-generierter Kreativität im Rahmen des gegenwärtigen Urheberrechtssystems. Beantworten lässt sich die obige Frage, indem man drei Hauptbedingungen für den Schutz von und das Eigentum an urheberrechtlich geschützten Werken betrachtet: (1) Rechtspersönlichkeit, (2) Urheberschaft und (3) Originalität.

5.3.1. Persönlichkeit: Kann eine Maschine eine Rechtspersönlichkeit haben?

Eine erste relevante Frage wäre, ob Maschinen eine Rechtspersönlichkeit besitzen können. Je nach Rechtsordnung, kultureller und religiöser Überzeugung sowie rechtlicher Subjektivität kann die Einführung der Persönlichkeit von Maschinen zu einer politischen Lösungsmöglichkeit werden. Japan hatte aufgrund des Shinto-Glaubens, dass tier- oder menschenähnliche Roboter als beseelt gedacht werden können, schon immer eine besondere Beziehung zu Robotern und Maschinen.⁴¹⁷ Im Oktober 2017 wurde Sophia als erstem Roboter von der saudi-arabischen Regierung die Staatsbürgerschaft verliehen. Dabei handelte es sich natürlich um einen PR-Gag, aber trotzdem ist es ein historischer Schritt in eine mögliche Assimilation von KI und Mensch. So gewährte Tokio einen Monat später, im November 2017, einem Chatbot einen offiziellen Aufenthaltsstatus im Stadtbezirk Shibuya.⁴¹⁸

Auch theoretische Überlegungen stützen die Idee der Rechtspersönlichkeit intelligenter Maschinen. So erklärt etwa Nick Bostrom: „Maschinen, die selbständig die Initiative ergreifen und eigene Pläne machen können, ... sollten vielleicht eher als Personen betrachtet werden statt als Maschinen.“⁴¹⁹ Mehrfach wurde darauf hingewiesen, dass es keine rechtlichen Gründe oder konzeptionellen Motive gibt, um die Persönlichkeit von KI-Robotern zu bestreiten: Das Gesetz sollte das Recht haben, Persönlichkeit aufgrund von rationalen Entscheidungen und empirischen Beweisen zu gewähren statt von Aberglauben

⁴¹⁷ Holland-Minkley D., *God in the Machine: Perceptions and Portrayals of Mechanical Kami in Japanese Anime*, Master's Thesis, University of Pittsburgh.

⁴¹⁸ Cuthbertson, *Tokyo: Artificial Intelligence ‚Boy‘ Shibuya Mirai Becomes World's First AI Bot to Be Granted Residency*, Newsweek, Washington, <https://www.newsweek.com/tokyo-residency-artificial-intelligence-boy-shibuya-mirai-702382>.

⁴¹⁹ Bostrom N., *Superintelligence: Paths, Dangers, Strategies*, OUP, Oxford.

und Privilegien.⁴²⁰ Daher wurden Argumente dafür vorgebracht, eine Persönlichkeit künftigen hypothetischen starken KIs zuzugestehen, die autonom (fähig zu Entscheidungen ohne Input-Aktion), intelligent (fähig zur Selbstprogrammierung und zur Integration von Informationen in einen Rahmen) und selbstbewusst (fähig zu subjektivem Erleben) sind.⁴²¹

Noch bemerkenswerter ist, dass das Europäische Parlament die Möglichkeit erwägt, KI und Roboter zu „elektronischen Personen“ zu erklären. In einer EntschlieÙung zu zivilrechtlichen Regelungen im Bereich Robotik fragt sich das Europäische Parlament, ob die normalen Haftungsregeln ausreichend sind oder ob die KI neue Prinzipien und Regeln erfordert.⁴²² Soll der autonome Charakter von Robotern vor dem Hintergrund der bestehenden rechtlichen Kategorien ausgelegt werden oder soll eine neue Kategorie geschaffen werden?⁴²³ In der EntschlieÙung heiÙt es, „dass Roboter umso weniger als simple Werkzeuge in den Händen anderer Akteure (wie zum Beispiel Hersteller, Betreiber, Eigentümer, Nutzer usw.) angesehen werden können, je autonomer sie sind“.⁴²⁴ Für das Europäische Parlament ist offensichtlich, dass das EU-Recht die außervertragliche Haftung für Schäden durch autonome KI nicht vollständig regeln kann. Traditionelle Regeln würden weiterhin gelten, wenn die Ursache der Handlung oder Unterlassung eines Roboters auf einen konkreten menschlichen Akteur, wie zum Beispiel den Hersteller, den Betreiber oder den Benutzer, zurückverfolgt werden kann. Ebenso gelten die traditionellen Haftungsregeln weiter, wenn es zu einer Fehlfunktion des Roboters gekommen ist oder wenn der menschliche Akteur das schädliche Verhalten des Roboters hätte vorhersehen und verhindern können. Was aber, wenn die Ursache für die Handlung oder Unterlassung des Roboters nicht auf einen konkreten menschlichen Akteur zurückverfolgt werden kann? Was, wenn keine Herstellungsfehler vorliegen? Und die KI keine Fehlfunktion hatte? Und der Geschädigte den tatsächlichen Schaden, den Fehler oder den ursächlichen Zusammenhang zwischen Fehler und Schaden nicht beweisen kann? Was, wenn die KI de facto Schäden verursacht hat, weil sie entsprechend ihrer eigenen Programmierung und Zielsetzung tatsächlich autonom gehandelt hat? In diesem Szenario sollte die Produkthaftungsrichtlinie 85/374/EWG nicht gelten. Die EntschlieÙung unterstreicht, dass die normalen Haftungsregeln hier nicht ausreichen, und fordert neue Regeln, um Klarheit darüber zu schaffen, ob eine Maschine für ihre Handlungen oder Unterlassungen verantwortlich gemacht werden kann.⁴²⁵ Auch wenn die EntschlieÙung darauf hinweist, „dass zumindest im derzeitigen Stadium die Verantwortung bei einem Menschen und nicht bei einem Roboter

⁴²⁰ Solum L.B., „Legal Personhood for Artificial Intelligences“, *North Carolina Law Review* 70, S. 1264.

⁴²¹ Zimmerman E., „Machine Minds: Frontiers in Legal Personhood“, S. 14-21, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2563965. Siehe auch Hubbard F.P., „Do Androids Dream? Personhood and Intelligent Artifacts“, *Temple Law Review* 83, S. 406-474.

⁴²² Vgl. Vladeck D., „Machines without Principals: Liability Rules and Artificial Intelligence“, *Washington Law Review* 89, S. 117-150.

⁴²³ Europäisches Parlament (2017), Zivilrechtliche Regelungen im Bereich Robotik: EntschlieÙung des Europäischen Parlaments vom 16. Februar 2017 mit Empfehlungen an die Kommission zu zivilrechtlichen Regelungen im Bereich Robotik (2015/2103(INL)), https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_DE.pdf.

⁴²⁴ Europäisches Parlament (2017), *op.cit.*

⁴²⁵ Europäisches Parlament (2017), *op.cit.* Siehe auch European Commission, „Liability for Artificial Intelligence and other Emerging Technologies: Report from the Expert Group on Liability and New Technologies – New Technologies Formation“, European Union, Brüssel, <https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608>.

liegen muss,“ fordert sie langfristig: (1) ein obligatorisches Versicherungssystem, das alle potenziellen Verantwortlichkeiten in der Kette berücksichtigt;⁴²⁶ (2) einen speziellen rechtlichen Status für Roboter, „damit zumindest für die ausgeklügeltsten autonomen Roboter ein Status als elektronische Person festgelegt werden könnte, die für den Ausgleich sämtlicher von ihr verursachten Schäden verantwortlich wäre“.⁴²⁷

In zahlreichen Diskussionen kommt die Idee einer KI-Rechtspersönlichkeit auf, doch bestimmt wird die Debatte bisher von widersprüchlichen, hilflosen Versuchen, eine Technologie zu regulieren, deren Entwicklung völlig unvorhersehbar ist. Daher wird der Diskurs über die Zuerkennung der Rechtspersönlichkeit, wie so oft, zu einem politischen Thema ohne rationale Grundlage. Insofern erinnert die Verleihung der saudi-arabischen Staatsbürgerschaft an Sophia daran, wie der römische Kaiser Caligula sein Pferd Incitatus zum Senator machte.⁴²⁸

Ob eine quasi-menschliche oder hyper-menschliche KI nun kommt oder nicht – eine Rechtspersönlichkeit von Maschinen ist im gegenwärtigen Rechtsrahmen gewiss nicht möglich, und dabei dürfte es auf absehbare Zeit auch bleiben. Die Wissenschaft betont immer wieder, dass jede Hypothese über eine volle Rechtspersönlichkeit für KI-Roboter zu verwerfen ist, bis grundlegende technologische Veränderungen eintreten.⁴²⁹ Pagallo hebt unter den normativen Argumenten gegen die Rechtspersönlichkeit das „Missing Something“-Problem hervor, wonach heutigen KI-Robotern die meisten Eigenschaften fehlen, die normalerweise vorausgesetzt werden, um jemandem oder etwas Rechtspersönlichkeit zu verleihen: Solche künstlichen Akteure haben weder ein Selbstbewusstsein oder menschenähnliche Intentionen, noch leiden sie richtig.⁴³⁰ Auch eine statistische Analyse verschiedener Voraussetzungen einer Rechtspersönlichkeit, die etwa in der US-Rechtsprechung aufgestellt wurden, werde die Unvereinbarkeit zwischen Rechtspersönlichkeit und KI-Systemen aufzeigen.⁴³¹ Diese empirische Analyse beweist, dass die Gerichte bei der Zuerkennung der Persönlichkeit berücksichtigen, ob ein Gesetz sie direkt oder indirekt gewährt, ob das KI-System klagen und verklagt werden kann und ob es sich um eine Gesamtheit natürlicher Personen handelt.⁴³²

Diese Überlegungen dienen auch dazu, die KI von Körperschaften abzugrenzen, die als juristische Person behandelt werden. Im Gegensatz zu Körperschaften sind KI-Systeme weder „fiktive“ Gebilde noch Vereinigungen natürlicher Personen.⁴³³ Juristische Personen werden von natürlichen Personen gebildet, die letztlich Rechte wahrnehmen können. Darüber hinaus können juristische Personen zwar Verwertungsrechte besitzen, doch diese Verwertungsrechte bestehen an einem Werk, das von einer natürlichen Person geschaffen

⁴²⁶ Europäisches Parlament (2017), *op.cit.*

⁴²⁷ Europäisches Parlament (2017), *op.cit.*

⁴²⁸ Pagallo U., „Vital, Sophia, and Co. – The Quest for the Legal Personhood of Robots“, *Information 9*(9), S. 239-240.

⁴²⁹ Banteka N., „Artificially Intelligent Persons“, *Houston Law Review 58* (erscheint in Kürze), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3552269. Mik E., „AI as Legal Person?“ in Hilty R. und Liu K-C. (eds), *Artificial Intelligence and Intellectual Property*, Oxford University Press, England, erscheint in Kürze. Pagallo U., *op.cit.*, S. 230-240.

⁴³⁰ Pagallo U., *op.cit.*, S. 237-238.

⁴³¹ Banteka N., *op. cit.*

⁴³² Banteka N., *op. cit.*, S. 50-52.

⁴³³ Banteka N., *op. cit.*, S. 19.

wurde, die Urheberin des Werkes ist, sodass die Voraussetzung der Urheberschaft wie auch der Originalität erfüllt ist. Bei einer KI und einem KI-generierten Werk wäre dies nicht der Fall, wie die nächsten Seiten zeigen werden.

Diese Argumente gegen die Rechtspersönlichkeit der KI mag die Politik bereits verinnerlicht haben, wie die Entschließung des Europäischen Parlaments von 2017 beweisen könnte, die zumindest kurz- und mittelfristig jegliche Rechtspersönlichkeit einer KI ausschließt. Auch konkret in Verbindung mit KI-generierter Kreativität scheint das Europäische Parlament die Rechtspersönlichkeit einer KI nun auszuschließen. Im „Entwurf eines Berichts über die Rechte des geistigen Eigentums bei der Entwicklung von KI-Technologien“ stellte das Europäische Parlament im Rahmen eines Entschließungsantrags kürzlich fest, „dass die Verselbstständigung des kreativen Prozesses Fragen im Zusammenhang mit der Inhaberschaft der Rechte des geistigen Eigentums aufwirft, [war aber] in diesem Zusammenhang der Ansicht, dass es nicht angebracht wäre, KI-Technologien mit einer Rechtspersönlichkeit auszustatten“.⁴³⁴ Die politische Herausforderung bestehe nicht in der Einführung einer Rechtspersönlichkeit für Maschinen, sondern in der korrekten Zuordnung der Rechenschaftspflicht und Haftung für die Aktivitäten von KI-Robotern in Fällen mit komplexer verteilter Verantwortung, etwa durch Verträge und das Wirtschaftsrecht.⁴³⁵

5.3.2. Urheberschaft: Kann eine Maschine Urheberin sein?

Obwohl einer KI mangels Rechtspersönlichkeit keine Urheberschaft zugesprochen werden kann und sie nicht für die Durchsetzung von Rechten an ihrer Kreativität eintreten kann, ist es sinnvoll zu prüfen, ob diese Kreativität im bestehenden Rechtsrahmen schutzfähig ist. Die Beantwortung der weiter gefassten Frage, ob KI-generierte Kreativität urheberrechtlich schutzfähig ist, impliziert die Überlegung, ob eine KI nach traditionellen Urheberrechtsstandards als Urheberin verstanden werden kann. Im Kern geht es darum, ob die Existenz eines Menschen eine zwingende Voraussetzung für die Urheberschaft ist. Kann ein Urheber eine Maschine sein oder muss es sich um einen Menschen handeln?

In internationalen Verträgen findet sich keine Definition, die hierauf eine definitive Antwort geben kann. Es scheint jedoch, dass Verweise auf eine menschliche Schöpfung in der Berner Übereinkunft⁴³⁶ die Möglichkeit ausschließen, unter einem Urheber eine KI zu verstehen. Zum einen scheint die an das Leben des Urhebers gebundene Schutzdauer Maschinen als Urheberin auszuschließen (Berner Übereinkunft, Art. 7). Auch die Bezugnahme auf die Staatsangehörigkeit oder den gewöhnlichen Aufenthalt des Urhebers deutet wohl darauf hin, dass der Begriff der Urheberschaft nur für Menschen gilt (Berner

⁴³⁴ Europäisches Parlament (2020), Entwurf eines Berichts über die Rechte des geistigen Eigentums bei der Entwicklung von KI-Technologien (2020/2015(INI)), 24. April 2020, https://europarl.europa.eu/doceo/document/JURI-PR-650527_DE.html.

⁴³⁵ Europäisches Parlament (2017), *op.cit.* Pagallo U., *op.cit.*, S. 239-240.

⁴³⁶ Berner Übereinkunft zum Schutz von Werken der Literatur und Kunst (Fassung vom 28. September 1979), <https://wipolex.wipo.int/en/treaties/textdetails/12214>. Eine deutsche Fassung ist verfügbar unter: <https://www.admin.ch/opc/de/classified-compilation/19670154/200605050000/0.231.14.pdf>.

Übereinkunft, Art. 3). Insgesamt ist argumentiert worden, „Berns humanistische Prägung“ und Achtung vor idealistischen Persönlichkeitstheorien spreche stark dafür, dass „in der Berner Übereinkunft gegenwärtig ein anthropozentrischer Begriff von Urheberschaft verankert ist“, der eine nicht-menschliche Urheberschaft aus dem Geltungsbereich der Übereinkunft ausschließt.⁴³⁷

5.3.2.1. Europäische Union

Eine genaue Betrachtung des EU-Rechts dürfte zu ähnlichen Schlussfolgerungen führen.⁴³⁸ In keiner gesetzlichen Bestimmung findet sich eine übergreifende Definition des Begriffs der Urheberschaft, doch in Artikel 2 Absatz 1 der Softwarerichtlinie⁴³⁹ wie auch in Artikel 4 Absatz 1 der Datenbankrichtlinie⁴⁴⁰ ist ein Urheber definiert als eine natürliche Person, eine Gruppe natürlicher Personen oder eine juristische Person. Darüber hinaus sieht Artikel 2 Absatz 1 der Schutzdauerrichtlinie⁴⁴¹ vor, dass der Hauptregisseur eines Filmwerks oder eines audiovisuellen Werks als dessen Urheber oder als einer seiner Urheber gilt. Klarer spricht die anthropozentrische Sicht der Urheberschaft aus den *Travaux préparatoires* der Software- und der Datenbankrichtlinie, wo konkret die Rede ist vom „menschlichen Urheber, der das Werk erschafft“, und von der „natürlichen Person, [die] zumindest das unveräußerliche Recht behält, die Vaterschaft ihres Werkes zu beanspruchen“.⁴⁴² Der ursprüngliche Vorschlag für eine Softwarerichtlinie kam zu dem Schluss: „Die menschliche Leistung bei der Schaffung maschinenentwickelter Programme kann relativ gering sein und wird in Zukunft immer bescheidener ausfallen. Ein menschlicher ‚Urheber‘ im weitesten Sinne ist jedoch immer vorhanden, und er muss das Recht haben, die ‚Urheberschaft‘ am Programm geltend zu machen.“⁴⁴³ Denselben Punkt betonte auch Generalanwältin Verica Trstenjak in der Rechtssache Painer des Gerichtshofs der Europäischen Union (EuGH) mit der Feststellung, dass „lediglich das Ergebnis menschlichen Schaffens geschützt [wird],

⁴³⁷ Ginsburg J., „People Not Machines: Authorship and What It Means in the Berne Convention“, *International Review of Intellectual Property and Competition Law* 49, S.131-135. Siehe auch Aplin T. und Pasqualetto G., „Artificial Intelligence and Copyright Protection“, §5.04, in Ballardini R., Kuoppamäki P. und Pitkänen O.(eds.), *Regulating Industrial Internet Through IPR, Data Protection and Competition Law*, Kluwer, Alphen aan den Rijn. Ricketson S., „The Need for Human Authorship – Australian Developments: Telstra Corp Ltd v Phone Directories Co Pty Ltd (Case Comment)“, *E.I.P.R.* 34(1), S. 34. Ricketson S. (1991), „People or machines? The Berne Convention and the changing concept of authorship“, *Columbia VLA Journal of Law & the Arts* 16, S. 34.

⁴³⁸ Deltorn J.M. (2017), „Deep Creations: Intellectual Property and the Automata“, *Frontiers in Digital Humanities*, S. 8, <https://www.frontiersin.org/articles/10.3389/fdigh.2017.00003/full>. Deltorn J.M. und Macrez F. (2018), *op.cit.*, S. 8.

⁴³⁹ Richtlinie 2009/24/EG des Europäischen Parlaments und des Rates vom 23. April 2009 über den Rechtsschutz von Computerprogrammen, ABL L111/16.

⁴⁴⁰ Richtlinie 96/9/EG des Europäischen Parlaments und des Rates vom 11. März 1996 über den rechtlichen Schutz von Datenbanken, ABL L077/20.

⁴⁴¹ Richtlinie 2006/116/EG des Europäischen Parlaments und des Rates vom 12. Dezember 2006 über die Schutzdauer des Urheberrechts und bestimmter verwandter Schutzrechte, ABL L372/12.

⁴⁴² Ramalho A., „Will Robots Rule the (Artistic) World? A Proposed Model for the Legal Status of Creations by Artificial Intelligence Systems“, *Journal of Internet Law* 21, S. 17-18.

⁴⁴³ Europäische Kommission, Begründung zum Vorschlag für eine Richtlinie des Rates über den Rechtsschutz von Computerprogrammen, KOM(88) 816 endg., ABL C91/10, 12. April 1989.

wobei ein solches auch dann vorliegen kann, wenn sich die Person eines technischen Hilfsmittels wie eines Fotoapparats bedient“.⁴⁴⁴

Die nationale Gesetzgebung der EU-Mitgliedstaaten bestätigt diesen Ansatz. So verlangt Art. L.111-1 des französischen Gesetzbuchs über geistiges Eigentum,⁴⁴⁵ dass urheberrechtsfähige Werke die „Schöpfung des Geistes“ sind. Art. 5 des spanischen Urheberrechtsgesetzes bestimmt schlicht: „Als Urheber gilt die natürliche Person, die ein Werk erschafft.“⁴⁴⁶ Und obwohl § 7 des deutschen Urheberrechtsgesetzes die Urheberschaft nicht ausdrücklich auf natürliche Personen beschränkt, bindet § 11 die Urheberschaft an ein persönlichkeitsorientiertes Konzept, indem er „den Urheber in seinen geistigen und persönlichen Beziehungen zum Werk“ schützt.⁴⁴⁷

Darüber hinaus stützt das EU-Recht – ebenso wie mehrere nationale Gesetze (z. B. niederländisches Urheberrechtsgesetz, Art. 4(1); französisches Gesetzbuch über geistiges Eigentum, Art. L113-1; spanisches Urheberrechtsgesetz, Art. 6.1; italienisches Urheberrechtsgesetz, Art. 8)⁴⁴⁸ – einen anthropozentrischen Ansatz mit der Vermutung der Urheberschaft, wonach als Urheber mangels Gegenbeweises die *Person* gilt, deren Name auf dem Werkstück angegeben ist (Richtlinie zur Durchsetzung des geistigen Eigentums, Artikel 5). Einige nationale Gerichte haben sogar klargestellt, dass diese Vermutung nur auf natürliche Personen anwendbar ist, die das Werk erschaffen, nicht aber auf eine juristische Person, die die wirtschaftlichen Rechte erlangt hat.⁴⁴⁹ Theoretisch könnte diese Vermutung der Urheberschaft auch für KI-generierte Werke gelten, sodass die Person(en), deren Name(n) auf dem Werk angegeben ist (sind), als Urheber gilt (gelten). Natürlich bietet diese Lösung keinen wirklichen Schutz vor Rechtsverletzungen, da die Vermutung durch den Nachweis widerlegt werden kann, dass nicht die genannte Person die Urheberin ist, sondern eine KI.

Ein kurzer Überblick über andere wichtige Rechtsordnungen könnte zu ähnlichen Schlussfolgerungen über die Anwendung des Urheberschaftsbegriffs auf KI führen.

⁴⁴⁴ Schlussanträge der GA Trstenjak (12. April 2011), C-145/10 *Eva-Maria Painer gegen Standard VerlagsGmbH u. a.*, ECLI:EU:C:2011:239.

⁴⁴⁵ Code de la propriété intellectuelle [Gesetzbuch über geistiges Eigentum], 1912, <https://www.wipo.int/edocs/lexdocs/laws/fr/fr/fr077fr.pdf> (Frankreich).

⁴⁴⁶ Ley de propiedad intelectual (Gesetz über geistiges Eigentum), 1996, <https://www.wipo.int/edocs/lexdocs/laws/es/es/es177es.pdf> (Spanien).

⁴⁴⁷ Gesetz über Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz, UrhG), 1965, <https://www.gesetze-im-internet.de/urhgf/> (Deutschland).

⁴⁴⁸ Auteurswet (Urheberrechtsgesetz), 1912, <https://wetten.overheid.nl/BWBR0001886/2012-01-01> (Niederlande); Code de la propriété intellectuelle (Gesetz über geistiges Eigentum), 1912, <https://www.wipo.int/edocs/lexdocs/laws/fr/fr/fr077fr.pdf> (Frankreich); Ley de propiedad intelectual (Gesetz über geistiges Eigentum), 1996, <https://www.wipo.int/edocs/lexdocs/laws/es/es/es177es.pdf> (Spanien); Legge n. 633, Protezione del diritto d'autore e di altri diritti connessi al suo esercizio (Gesetz Nr. 633, Schutz des Urheberrechts und verwandter Schutzrechte), 1941, <https://www.wipo.int/edocs/lexdocs/laws/it/it/it158it.pdf> (Italien).

⁴⁴⁹ *Herlitz PBS AG gegen Realister OÜ*, Estnischer Oberster Gerichtshof (7. Februar 2012) AZ 3-2-1-155-11 (Estland). Siehe auch Vasamae E., „Presumption of authorship: only natural persons“, *Kluwer Copyright Blog*, Amsterdam, http://copyrightblog.kluweriplaw.com/2012/03/19/presumption-of-authorship-only-natural-persons/?doing_wp_cron=1594514535.1866068840026855468750.

5.3.2.2. Australien

Das australische Recht enthält eine klare gesetzliche Hürde für nicht-menschliche Urheber: Section 32(1) des australischen Urheberrechtsgesetzes⁴⁵⁰ definiert einen Urheber als „qualifizierte Person“, und diese wiederum ist in Section 32(4) definiert als Person mit australischer Staatsangehörigkeit oder mit Wohnsitz in Australien. Australische Gerichte knüpfen dann den Schutz der Urheberschaft an die Voraussetzung der Originalität. Im Fall *Acohs gegen Ucorp*, in dem es um die Existenz eines Urheberrechts an elektronisch generierten Datenblättern ging, stellte das Gericht klar, ein Werk müsse „den originalen Anstrengungen eines einzelnen menschlichen Urhebers entspringen“.⁴⁵¹ Im Urteil im Fall *Phone Directories* wird dieser Punkt mit der Feststellung untermauert, dass das Urheberrecht „nur besteht, wenn es auf ein Individuum zurückgeht“.⁴⁵² Überdies kam das Gericht im Fall *IceTV gegen Nine Network Australia*, in dem es um das Urheberrecht an computergenerierten TV-Wochenprogrammen ging, zu dem Schluss, dass nur Urheber, also Personen im Sinne der gesetzlichen Definition, Originalität besitzen können.⁴⁵³

5.3.2.3. Vereinigte Staaten

Der Schutz von Produkten einer computerbasierten Kreativität ist in den Vereinigten Staaten nicht neu. Bereits in den späten 1960er Jahren begann die Wissenschaft, die mögliche Schutzfähigkeit von computergenerierter Kreativität zu diskutieren.⁴⁵⁴ Der US-Kongress setzte einen Ausschuss ein, der feststellen sollte, ob Computer oder Computerprogramme Urheber sein können, deren Output urheberrechtlich geschützt werden kann. Im Jahr 1978 erklärte die *Commission on New Technological Uses of Copyrighted Works* (CONTU), Computer seien lediglich „inerte Werkzeuge der Schöpfung“, die noch nicht selbstständig Werke schaffen. Die CONTU befasste sich nicht mit dem urheberrechtlichen Schutz von automatisierten Werken ohne menschliche Urheberschaft, weil dies damals als zu spekulativ galt.⁴⁵⁵ 1986 veröffentlichte das *Office of Technology Assessment* (OTA) des US-Kongresses einen Bericht, in dem es hieß, Computer seien zwar mehr als „inerte Werkzeuge

⁴⁵⁰ Copyright Act 1968, <https://www.legislation.gov.au/Details/C2019C00042> (Australien).

⁴⁵¹ *Acohs Pty Ltd v. Ucorp Pty Ltd* (2010) 86 IPR 492 (AUS).

⁴⁵² *Phone Directories Co Pty v Telstra Corporation Ltd* (2010) 194 FCR 142 (AUS). Siehe auch McCutcheon J., „The Vanishing Author in Computer-Generated Works: A Critical Analysis of Recent Australian Case Law“, *Melbourne University Law Review* 36(3), S. 941-969, https://www.researchgate.net/publication/289409001_The_vanishing_author_in_computer-generated_works_A_critical_analysis_of_recent_Australian_case_law.

⁴⁵³ *IceTV Pty Ltd v. Nine Network Australia Pty Ltd* [2009] HCA 14, 239 CLR 458 (AUS).

⁴⁵⁴ Milde K.F., „Can a Computer Be an „Author“ or an „Inventor?““, *Journal of the Patent Office Society* 51, S. 378-406.

⁴⁵⁵ National Commission on New Technological Uses of Copyrighted Works (CONTU), *Final Report*, United States, p. 44. Siehe auch Bridy A. (2012), „Coding Creativity: Copyright and the Artificially Intelligent Author“, *Stanford Technology Law Review* 5, 22-24, Randnr. 53-60. Miller A.R., „Copyright Protection for Computer Programs, Databases, and Computer Generated Works: Is Anything New Since CONTU?“, *Harvard Law Review* 977, S. 977-1072.

der Schöpfung“, doch die urheberrechtliche Schutzfähigkeit computergenerierter Werke sei unbestimmbar.⁴⁵⁶

Das US-Urheberrechtsgesetz enthält keine ausdrückliche Definition der Urheberschaft, sodass manche Autoren zunächst argumentierten, vom Wortlaut des Gesetzes her sei die Urheberschaft nicht auf menschliche Urheber beschränkt.⁴⁵⁷ Doch sowohl zusätzliche Textverweise als auch die Rechtsprechung schließen offenbar die Möglichkeit aus, unter Urhebern im Sinne des Gesetzes nichtmenschliche Akteure zu verstehen. Insbesondere definiert Section 101 des Urheberrechtsgesetzes⁴⁵⁸ anonyme Werke als Werke, „bei denen keine natürliche Person als Urheber angegeben ist“, was den Schluss zulässt, dass als Urheber nur natürliche Personen in Frage kommen. Zudem gilt seit Langem das Verständnis, dass die Verfassungsgeschichte des Wortes „Copyright“ ausschließlich für Menschen als „Urheber“ spricht.⁴⁵⁹ US-Gerichte haben dieses Verständnis stets unterstützt. Der Oberste Gerichtshof hat klar festgestellt, dass „der Urheber grundsätzlich [...] die *Person* ist, die eine Idee in eine feste, greifbare Äußerung übersetzt, die Anspruch auf urheberrechtlichen Schutz hat“.⁴⁶⁰ In *Feist gegen Rural* erörtert der Oberste Gerichtshof der Vereinigten Staaten die Begriffe Urheberschaft und Urheber ausführlich anhand einer Analyse des Begriffs der Originalität. Dieser beziehe sich auf inhärent menschliche Merkmale wie einen „schöpferischen Funken“ oder „eine geistige Produktion, einen Gedanken und eine Konzeption“.⁴⁶¹ Auch frühere Fälle würden diese Schlussfolgerung stützen. Im Urteil in den sogenannten *Trade-Mark Cases* heißt es, das Urheberrecht schütze nur „die Früchte geistiger Arbeit“, die „auf den schöpferischen Kräften des Geistes gründen“.⁴⁶² Im Fall *Burrow-Giles* erinnerte der Oberste US-Gerichtshof daran, dass das Urheberrecht auf „originale Geistesschöpfungen des Urhebers“ beschränkt sei.⁴⁶³

Einige weitere Klarstellungen in der Sache brachte ein Fall aus jüngster Zeit, in dem der Makak Naruto behilflich war. Im Fall *Naruto gegen Slater* waren zwei weltweit beachtete „Affen-Selfies“ Gegenstand eines Streits darüber, ob Tiere Urheberrechte besitzen können. Die Selbstporträts hatte der siebenjährige Schopfmakak Naruto aufgenommen, als der Tierfotograf David Slater bei einem seiner Besuche in Indonesien seine Kamera unbeaufsichtigt ließ. Kurz darauf veröffentlichte Wikimedia Commons die Bilder auf seiner Website in der Annahme, dass die Affen-Selfies keinen menschlichen Urheber hätten und daher gemeinfrei seien. Wikimedia weigerte sich standhaft, die Bilder zu entfernen. Slater,

⁴⁵⁶ US Office of Technology Assessment, *Intellectual Property Rights in an Age of Electronics and Information*, United States.

⁴⁵⁷ Bridy A. (2012), *op.cit.*, 20, Randnr. 49. Denicola R. (2016), *op.cit.*, S. 275-283. Miller A.R., *op.cit.*, S. 1042-1072; Samuelson P. (1986), „Allocating Ownership Rights in Computer-Generated Works“, *University of Pittsburg Law Review* 47(4), S. 1200-1204.

⁴⁵⁸ The Copyright Act of 1976, <https://www.copyright.gov/title17/title17.pdf> (USA).

⁴⁵⁹ Butler T., „Can a computer be an author? Copyright aspects of artificial intelligence“, (Comm/Ent), *A Journal of Communications and Entertainment Law* 4(4), S. 733-734. Clifford, R.D. (1996), „Intellectual Property in the Era of the Creative Computer Program: Will the True Creator Please Stand up“, *Tulane Law Review* 71, 1682-1686. Kasap A., „Copyright and Creative Artificial Intelligence (AI) Systems: A Twenty-First Century Approach to Authorship of AI-Generated Works in the United States“, *Wake Forest Intellectual Property Law Journal* 19(4), S. 358, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3597792. Milde K.F., *op.cit.*, S. 391-392.

⁴⁶⁰ *Community for Creative Non-Violence v. Reid*, 490 U.S. 730 (1989).

⁴⁶¹ *Feist Publications Rural Telephone Service*, 499 U.S. 340 (1991) (USA).

⁴⁶² *Trade-Mark Cases*, 100 U.S. 82 (1879) (USA).

⁴⁶³ *Burrow-Giles Lithographic Co. gegen Sarony*, 111 U.S. 53 (1884) (USA).

der ursprünglich damit geworben hatte, dass die Affen die Selfies ganz allein, ohne menschliches Zutun, aufgenommen hätten, änderte dann seine Haltung und behauptete, die Selfies seien dadurch entstanden, dass er die Kamera im richtigen Winkel und mit der richtigen Beleuchtung eingerichtet, die Einstellungen optimiert und die Affen einfach dazu verlockt habe, den Kameraknopf zu drücken.

Die Frage, ob die Selfies gemeinfrei sind, wurde letztlich nicht geprüft, doch die Gerichte hatten Gelegenheit zu erwägen, ob Naruto das Urheberrecht an seinem Selfie zugestanden werden kann. 2014 wurden die Affen-Selfies über Blurb Inc. in einem Buch veröffentlicht, in dem Slater und Wildlife Personalities Ltd. als Inhaber der Urheberrechte genannt wurden. 2015 verklagte People for the Ethical Treatment of Animals (PETA) als engste Freunde und im Namen von Naruto Slater, Wildlife Personalities Ltd. und Blurb Inc. vor dem Bezirksgericht Kalifornien wegen Urheberrechtsverletzung. Das Bezirksgericht gab dem Antrag der Beklagten auf Klageabweisung statt, da Naruto es versäumt habe, seine gesetzliche Klagebefugnis nach dem Urheberrechtsgesetz zu begründen, und stellte fest: „Wenn die Menschen, die vorgeben, im Namen des Klägers zu handeln, möchten, dass das Urheberrecht zu den Rechtsgebieten gehört, in denen nicht-menschliche Tiere klagebefugt sind, dann sollten sie diesen zweifelhaften Fall dem Kongress vorlegen und nicht den Bundesgerichten.“⁴⁶⁴ Gegen die Entscheidung wurde Berufung eingelegt, und während die Parteien einem Vergleich zustimmten, lehnte das Berufungsgericht die Abweisung der Berufung ab und bestätigte die Entscheidung der Vorinstanz. Die Mehrheit war der Auffassung, dass Tieren zwar eine verfassungsmäßige Klagebefugnis nach Artikel III der US-Verfassung zukommt, aber keine gesetzliche Klagebefugnis nach dem Urheberrechtsgesetz.⁴⁶⁵ Das Gericht stützte sich auf die Entscheidung des Neunten Gerichtsbezirks im Fall *Cetacean Community gegen Bush*, wonach Tiere nur dann eine gesetzliche Klagebefugnis haben, wenn das Gesetz dies eindeutig vorsieht.⁴⁶⁶ Darüber hinaus ergebe sich aus der Verwendung der Begriffe „Kinder“, „Enkel“, „ehelich“, „Witwe“ und „Witwer“ im Urheberrechtsgesetz notwendigerweise, dass das Gesetz Tiere ausschließt, die „nicht heiraten und keine Erben mit gesetzlichem Eigentumsanspruch haben“.⁴⁶⁷ Die Ergebnisse des Naruto-Urteils lassen sich leicht auf jede nicht-menschliche und KI-generierte Kreativität übertragen.

Die dritte Ausgabe des Kompendiums der Verfahrensweisen des US-Urheberrechtsamts, die im Dezember 2014 nach Beginn des Naruto-Prozesses erschien, enthielt eine nicht bindende fachliche Orientierungshilfe, die eine nicht-menschliche Urheberschaft ausschloss.⁴⁶⁸ Das Kompendium bezieht sich bei der Erörterung der Urheberschaft immer wieder auf Personen oder Menschen. Konkret heißt es in Section 306, „das Erfordernis der menschlichen Urheberschaft“ beschränke die Registrierung auf von einem Menschen geschaffene „originale Geistesschöpfungen des Urhebers“. Nicht urheberrechtsfähig sind gemäß Section 313.2 „Werke ohne menschliche Urheberschaft“, von der Natur, von Tieren oder Pflanzen oder auf ähnliche Weise erzeugte Werke sowie

⁴⁶⁴ *Naruto gegen David Slater*, 15-cv-04324-WHO (N.D. Cal. 2016) (USA) (“Naruto 2016”).

⁴⁶⁵ *Naruto gegen David Slater*, F.3d 418 (9th Cir. 2018) (USA) (“Naruto 2018”).

⁴⁶⁶ *Cetacean Community. gegen Bush*, 386 F.3d 1169 (9th Cir. 2004) (USA).

⁴⁶⁷ *Naruto gegen David Slater*, F.3d 418 (9th Cir. 2018) (USA) (“Naruto 2018”).

⁴⁶⁸ U.S. Copyright Office, *Compendium of U.S. Copyright Office Practices*, 3rd edition, <https://www.copyright.gov/comp3/comp-index.html>

Werke, die von einer Maschine oder durch ein mechanisches Verfahren ohne Zutun eines menschlichen Urhebers geschaffen werden. Unter Verweis auf die *Trade-Mark Cases* und *Burrow-Giles* kam das Urheberrechtsamt zu dem Schluss, dass es die Eintragung eines Anspruchs ablehnen würde, wenn es feststellt, dass das Werk nicht von einem Menschen geschaffen wurde.⁴⁶⁹

5.3.2.4. China

In China haben sich mehrere Gerichte mit den Fragen nach der Urheberschaft von KI und der Urheberrechtsfähigkeit KI-generierter Werke befasst. Die chinesische Position zur Urheberschaft von KI scheint sich mit der anderer Rechtsordnungen zu decken, auch wenn sie gewissen Spielraum für einen möglichen Schutz lässt. Im Fall *Beijing Feilin Law Firm gegen Baidu Corporation* bestätigte das Gericht das Erfordernis menschlicher Urheber und verweigerte den Urheberrechtsschutz für Werke, die allein von Maschinen geschaffen wurden.⁴⁷⁰ Es ging dabei um einen Bericht, den die klagende Pekinger Anwaltskanzlei auf ihrem offiziellen WeChat-Konto veröffentlicht hatte. Nachdem eine nicht identifizierbare Person den Bericht ohne Erlaubnis ins Internet gestellt hatte, erhob die Klägerin vor dem Internetgericht Peking Klage wegen Urheberrechtsverletzung. Der Bericht war anhand von China Law & Reference von Wolters Kluwer erstellt worden, einer Software zur Abfrage juristischer Informationen. Während die Klägerin argumentierte, das Tool sei nur zur Unterstützung eingesetzt worden, behaupteten die Beklagten, die Software habe den gesamten Bericht generiert. Das Gericht gab der Klägerin Recht. Doch obwohl der streitgegenständliche Bericht durch das chinesische Urheberrecht geschützt sei, prüfte das Gericht auch die Schutzfähigkeit des von der Software automatisch generierten Berichts. Zum Schutz von Werken, die allein von einer KI erschaffen wurden, stellte das Gericht fest, der Begriff der Urheberschaft setze voraus, dass das Werk von einer natürlichen Person geschaffen wurde. Das Gericht legte jedoch eine interessante Anreizeanalyse vor, die den Schluss zurückweist, dass das Werk gemeinfrei sein müsse. Dabei vertrat es die Auffassung, dass dem Nutzer – und nicht dem bereits mit dem Urheberrecht an der Software belohnten Softwareentwickler – ein gewisser Schutz gewährt werden sollte, um Anreize zum Kauf der Software sowie zum Generieren und Verbreiten der Werke zu schaffen. Leider klärt das Urteil nicht, welche Form dieser Schutz haben soll.

In einem späteren Urteil im Fall *Shenzhen Tencent gegen Yinxun*, bestätigte das Bezirksgericht Nanshan in Shenzhen im Wesentlichen das Pekinger Urteil.⁴⁷¹ Die beiden Entscheidungen ähneln sich insofern, als die Gerichte den originalen Beiträgen

⁴⁶⁹ U.S. Copyright Office, *op.cit.*, Section 313.2.

⁴⁷⁰ *Beijing Feilin Law Firm gegen Baidu Corporation* (26 April 2019) Beijing Internet Court, (2018) Beijing 0491 Minchu No. 239. He K. (2020b), „Feilin gegen Baidu: Beijing Internet Court tackles protection of AI/software-generated work and holds that copyright only vests in works by human authors“, *The IPKat*, <http://ipkitten.blogspot.com/2019/11/feilin-v-baidu-beijing-internet-court.html>. Chen M., „Beijing Internet Court denies copyright to works created solely by artificial intelligence“, *Journal of Intellectual Property Law & Practice* 14(8), S. 14-18.

⁴⁷¹ *Shenzhen Tencent gegen Yinxun*, Nanshan District People’s Court of Shenzhen, Guangdong Province [2019] No. 14010 (China), <https://mp.weixin.qq.com/s/jjv7aYT5wDBIdTVWXV6rdQ>. He K. (2020a), „Another decision on AI-generated work in China: Is it a Work of Legal Entities?“, *The IPKat*, <http://ipkitten.blogspot.com/2020/01/another-decision-on-ai-generated-work.html>.

menschlicher Akteure Schutz gewährten, nicht aber der ausschließlich KI-generierten Kreativität. Die Klägerin Tencent Technology entwickelte einen KI-Schreibassistenten namens Dreamwriter. Im August 2018 veröffentlichte die Klägerin eines der von der KI geschaffenen Werke auf ihrer Website und teilte den Lesern mit, der Artikel sei von Tencents KI Dreamwriter verfasst worden. Die Beklagte soll den Artikel ohne Zustimmung der Klägerin auf ihrer Website veröffentlicht haben. In ihrer Klage argumentierte die Klägerin, als Urheberin des Artikels habe sie nach dem Urheberrechtsgesetz ausschließliche Rechte. Sie habe die Generierung des Artikels beaufsichtigt, sei für seine Organisation und Erstellung verantwortlich und hafte für ihn. Zugunsten der Klägerin entschied das Gericht, der Artikel erfülle die Voraussetzungen eines literarischen Originalwerks, da der Inhalt ein Produkt der Inputdaten, der Triggerbedingungen und der Anordnung der von einer operativen Gruppe der Klägerin ausgewählten Vorlagen und Ressourcen sei. Da die Formulierung des Artikels auf individuellen Entscheidungen und Anordnungen der Klägerin beruhe, sei der KI-generierte Artikel als Werk juristischer Personen im Sinne von Artikel 11 des Urheberrechtsgesetzes zu betrachten, sodass die Beklagte für die Urheberrechtsverletzung hafte. Doch auch wenn das Gericht das Werk als integrierte Geistesschöpfung betrachtet haben mag, die sowohl auf den Beitrag des menschlichen Teams als auch auf den Einsatz von Dreamwriter zurückgeht, ergibt sich die gewährte Schutzfähigkeit offensichtlich eher aus dem Beitrag des menschlichen Teams als aus irgendeinem KI-Beitrag.

5.3.3. Originalität: Kann eine Maschine Originalität aufweisen?

Neben der Auslegung des Begriffs der Urheberschaft scheint auch der Begriff der Originalität als Voraussetzung für einen Urheberrechtsschutz den Schutz von KI-generierter Kreativität auszuschließen. Textverweise und Rechtsprechung verstehen Originalität im Rahmen eines anthropozentrischen Modells, in dessen Mittelpunkt Selbstbewusstsein steht. Originalität wird anhand eines sogenannten persönlichkeitsorientierten Konzepts definiert, das ein mit Originalität versehenes Werk als Repräsentation der Persönlichkeit des Urhebers beschreibt. Im Englischen steckt diese Bedeutung womöglich schon im Wort für „Urheber“ – „author“ (vgl. Deutsch „Autor“) –, da es sich nach der anerkanntesten Etymologie des Wortes vom altgriechischen *αὐτός* ableiten soll, was „selbst“ bedeutet.⁴⁷² Diese Charakterisierung der Originalität baut auf idealistischen Persönlichkeitstheorien auf, nach denen Geistesschöpfungen Manifestationen oder Erweiterungen der Persönlichkeit ihrer Schöpfer sind.⁴⁷³ Daher wäre Originalität als Repräsentation des „Selbst“ und des Selbstbewusstseins für maschinell erzeugte Kreativität theoretisch unerreichbar. Dieses

⁴⁷² Frosio G., *Reconciling Copyright with Cumulative Creativity: The Third Paradigm*, Edward Elgar, Cheltenham, S. 16.

⁴⁷³ Fichte J., *Beweis der Unrechtmäßigkeit des Büchernachdrucks. Ein Raisonement und eine Parabel*, Berlinische Monatsschrift 21, S. 447. Hegel G.H., *Grundlinien der Philosophie des Rechts*, Ullstein, Frankfurt am Main, § 69. Kant I. (1785), *Von der Unrechtmäßigkeit des Büchernachdrucks*, Berlinische Monatsschrift 5, S. 403–417. Siehe auch Fisher W., „Theories of intellectual property“, in Munzer S. (ed.) *New essays in the legal and political theory of property*, Cambridge University Press, Cambridge, S. 168–200.

Verständnis von Originalität wird von den allermeisten Rechtsordnungen unterstützt. Es hat frühere Ansätze verdrängt, die auf Lockeschen Gerechtigkeitstheorien aufbauten und „Sweat of the Brow“-Lehren („Schweiß des Angesichts“) unterstützten, wonach der Schutz „Skills, Labour and Efforts“ („Sachkenntnis, Arbeitsaufwand und Anstrengungen“) bei der Schaffung eines Geisteswerks belohnt, unabhängig davon, ob das Werk für die Persönlichkeit des Urhebers repräsentativ ist.⁴⁷⁴

In der Europäischen Union haben drei Richtlinien den Begriff der Originalität vertikal harmonisiert. Gemäß Artikel 1 Absatz 3 der Softwarerichtlinie, Artikel 6 der Schutzdauerrichtlinie und Artikel 3 Absatz 1 der Datenbankrichtlinie sind Werke original (bzw. in der deutschen Fassung „individuelle Werke“), wenn sie „das Ergebnis der eigenen geistigen Schöpfung ihres Urhebers“ sind.⁴⁷⁵ Der EuGH dehnte den Begriff der Originalität später „horizontal“ auf alle Urheberrechtsgegenstände aus und präziserte seinen Geltungsbereich. In der Rechtssache *Infopaq* stellte der EuGH fest: „Erst mit Hilfe der Auswahl, der Anordnung und der Kombination dieser Wörter vermag der Urheber seinen schöpferischen Geist in origineller Weise zum Ausdruck zu bringen und zu einem Ergebnis zu gelangen, das eine geistige Schöpfung darstellt.“⁴⁷⁶ Im Urteil in der Rechtssache *Eva-Maria Painer* erklärte er zudem, dass ein Werk – in diesem Fall eine Porträtfotografie – ein Original ist und geschützt werden kann, wenn es (1) die eigene geistige Schöpfung des Urhebers darstellt, in der (2) dessen Persönlichkeit zum Ausdruck kommt und die sich (3) in dessen bei der Herstellung dieser Fotografie getroffenen freien kreativen Entscheidungen ausdrückt.⁴⁷⁷ Mit diesen unterschiedlichen Entscheidungen könne der Urheber einer Porträtfotografie dem geschaffenen Werk seine „persönliche Note“ verleihen.⁴⁷⁸ In der Rechtssache *Football Dataco* schließlic lehnte der EuGH alle noch verbliebenen „Sweat of the Brow“-Doktrinen ab und stellte fest, dass ein bedeutender Arbeitsaufwand und bedeutende Sachkenntnis des Urhebers als solche den Urheberrechtsschutz nicht rechtfertigen können, wenn durch diesen Arbeitsaufwand und diese Sachkenntnis keinerlei Originalität bei der Auswahl oder Anordnung zum Ausdruck kommt.⁴⁷⁹ Werken, die lediglich auf der Grundlage technischer Regeln oder Beschränkungen erstellt werden, fehle die für die Urheberschaft erforderliche kreative Freiheit.

Auch die US-Rechtsprechung unterstützt dieses persönlichkeitsorientierte Originalitätskonzept. Seit frühen Fällen wie *Burrow-Giles gegen Sarony*, in dem es um die Urheberrechtsfähigkeit einer Porträtfotografie von Oscar Wilde ging, hat der Oberste Gerichtshof der USA klargestellt, dass Originalität aus den freien kreativen Entscheidungen des Urhebers resultiert, die dem Werk dessen Persönlichkeit verleihen,⁴⁸⁰ „sodass das Endprodukt seine Vorstellungen und Visionen davon dupliziert,“ wie das Werk aussehen

⁴⁷⁴ Siehe z. B. *International News Service gegen Associated Press*, 248 U.S. 215 (1918) (USA). *Jeweler's Circular Publishing Co. gegen Keystone Publishing Co.*, 281 F. 83 (2nd Cir. 1922) (USA). Siehe auch Rahmatian A., „Originality in UK Copyright Law The Old ‚Skill and Labour‘ Doctrine Under Pressure“, IIC 44, S. 4–34.

⁴⁷⁵ Für eine Diskussion siehe Rosati E., *Originality in EU Copyright – Full Harmonization through Case Law*, Edward Elgar, Cheltenham.

⁴⁷⁶ *Infopaq International A/S gegen Danske Dagblades Forening*, C-5/08 (2009) ECLI:EU:C:2009:465.

⁴⁷⁷ *Eva-Maria Painer gegen Standard VerlagsGmbH u. a.*, C-145/10 (2011) ECLI:EU:C:2011:798.

⁴⁷⁸ *Eva-Maria Painer gegen Standard VerlagsGmbH u. a.*, C-145/10 (2011) ECLI:EU:C:2011:798.

⁴⁷⁹ *Football Dataco Ltd u. a. gegen Yahoo! UK Ltd u. a.*, C-604/10 (2012) ECLI:EU:C:2012:115.

⁴⁸⁰ *Burrow-Giles Lithographic Co. gegen Sarony*, 111 U.S. 53 (1884) (USA).

sollte.⁴⁸¹ Insbesondere hielt das Gericht im Fall *Burrow-Giles* Fotografien für urheberrechtsfähig, weil sie auf die „eigene originale geistige Vorstellung“ des Fotografen zurückgeführt werden könnten.⁴⁸² Später, im Fall *Feist gegen Rural*, stellte der Oberste Gerichtshof der USA klar, dass nur Werken mit einem Mindestmaß an Kreativität, das die Persönlichkeit des Urhebers repräsentiert, Originalität zugesprochen werden könne; Arbeitsaufwand und Anstrengungen bei der Schaffung eines Werkes reichten allein nicht für einen urheberrechtlichen Schutz aus.⁴⁸³ Vor dem Hintergrund dieser systemischen Überlegungen wurde festgestellt, dass Output wie computergestützte Kurzschrift⁴⁸⁴ oder eine mit Hilfe von Softwaresystemen erstellte Auflistung automatisch nummerierter Hardwareteile nicht die Originalität aufweist, die ein urheberrechtlicher Schutz voraussetzt.⁴⁸⁵

Samuelson würde – wie auch andere Autoren – im Grunde argumentieren, dass es in den USA keine gesetzlichen Beschränkungen für die Behandlung einer Maschine als Urheberin gibt, da „der urheberrechtliche Standard der Originalität so niedrig ist, dass computergenerierte Werke, selbst wenn sich herausstellt, dass sie von einer Maschine allein erstellt wurden, schutzfähig erscheinen könnten“.⁴⁸⁶ Ich würde sagen, dass Originalität nach dem Fall *Feist* keine rein quantitative Frage mehr ist. Für KI-generierte Kreativität ist unerheblich, ob der Standard der Originalität niedrig oder hoch ist. Der Standard, den die KI nicht erreicht, ist qualitativ und nicht quantitativ. KI kann kein „Selbst“ ausdrücken. Die Kreativität, die sie generiert, kann die Persönlichkeit des Urhebers nicht zum Ausdruck bringen, weil die KI keine hat. So gesehen stehen der US-Beitritt zur Berner Übereinkunft im Jahr 1988 und der Fall *Feist* von 1991 für die Entstehung einer globalen, stärker harmonisierten Sicht des Urheberrechts. Diese Angleichung der Vereinigten Staaten an das europäische Modell schließt auch ein persönlichkeits-theoretisches Verständnis von Originalität ein.⁴⁸⁷

In jüngerer Zeit haben sich einige wenige noch verbliebene Länder – hauptsächlich aus dem angelsächsischen Rechtskreis – diesem persönlichkeitsorientierten Originalitätskonzept angeschlossen. So haben Australien,⁴⁸⁸ Indien⁴⁸⁹ und das Vereinigte Königreich⁴⁹⁰ ihre früheren „Labour, Skill and Efforts“-Ansätze endlich aufgegeben. Nur noch wenige Länder folgen der „Sweat of the Brow“-Doktrin und lehnen

⁴⁸¹ *Lindsay gegen The Wrecked and Abandoned Vessel R.M.S. Titanic*, 52 U.S.P.Q.2d 1609 (S.D.N.Y. 1999) (USA).

⁴⁸² *Burrow-Giles Lithographic Co. gegen Sarony*, 111 U.S. 53 (1884) (USA).

⁴⁸³ *Feist Publications gegen Rural Telephone Service*, 499 U.S. 340 (1991) (USA).

⁴⁸⁴ *Brief English Systems gegen Owen*, 48 F.2d 555 (2d Cir. 1931) (USA).

⁴⁸⁵ *Southco, Inc. gegen Kanebridge Corporation*, 390 F.3d 276 (3d Cir. 2004) (USA).

⁴⁸⁶ Samuelson P. (1986), *op.cit.*, S. 1199-1200. Siehe auch Brown N., „Artificial Authors: A case for copyright in computer-generated works“, *Columbia Science and Technology Law Review* 9, S. 24-27. Kaminski M., „Authorship, Disrupted: AI Authors in Copyright and First Amendment Law“, *UC Davis Law Review* 51, S. 601. Für eine Gegenmeinung siehe Clifford, R.D. (1996), *op.cit.*, 1694-1695.

⁴⁸⁷ Price M. E. & Pollack M., *The Author in Copyright: Notes for the Literary Critic*, 10 *Cardozo Arts & Entertainment Law Journal* 703 (1992), pp. 717-720, <https://larc.cardozo.yu.edu/faculty-articles/123>.

⁴⁸⁸ *IceTV Pty Ltd gegen Nine Network Australia Pty Ltd* [2009] HCA 14, 239 CLR 458 (Australien).

⁴⁸⁹ *Eastern Book Co. & Ors gegen D.B. Modak & Anr* (2008) 1 SCC 1 (Indien).

⁴⁹⁰ *Temple Island Collections v New English Teas* (No. 2) [2012] EWPC 1. Rahmatian A., *op.cit.*, S. 4-34

persönlichkeitsorientierte Originalitätskonzepte ab, darunter Südafrika⁴⁹¹ und Neuseeland.⁴⁹²

Insgesamt scheint es also ein äußerst konsistentes internationales Verständnis des Begriffs der Originalität zu geben, in dessen Mittelpunkt eine anthropozentrische Vision steht, nach der ein Werk dann Originalität aufweist, wenn es eine Repräsentation des „Selbst“, also der Persönlichkeit des Urhebers ist. Nur wenn diese innere Verbundenheit zwischen Urheber und Werk besteht, wird das Originalitätserfordernis erfüllt und Schutz gewährt. Natürlich wäre nur ein empfindungsfähiges selbstbewusstes Wesen in der Lage, sein „Selbst“ durch ein Werk zu repräsentieren. Umgekehrt kann, selbst wenn jedes denkbare textuelle anthropozentrische Verständnis von Urheberschaft außer Acht gelassen wird, das Originalitätserfordernis, das in der Repräsentation der Persönlichkeit des Urhebers liegt, ohne ein Selbstbewusstsein des Schöpfers niemals erfüllt werden. Solange also nicht behauptet werden kann, dass Maschinen ein Selbstbewusstsein erlangt haben, was bei einer futuristischen, hypothetischen, starken KI möglich wäre, aber heute nicht der Fall ist,⁴⁹³ kann KI-generierte Kreativität das Originalitätserfordernis des bestehenden Rechtsrahmens nicht erfüllen.⁴⁹⁴

Manche argumentieren, dass möglicherweise nur ein neuartiges, vielleicht formelleres und gegenüber dem bestehenden subjektiven auch objektiveres Originalitätskonzept Werke, die von kreativen Robotern geschaffen werden, sowie Kunstwerke, die mit digitalen Werkzeugen erzeugt werden, in den Geltungsbereich des Urheberrechtsschutzes einbeziehen kann.⁴⁹⁵ Aus dieser objektiven Perspektive sollte die Justiz das Endprodukt an sich betrachten und dabei den jeweiligen Kunstbereich, die objektive Meinung der Nutzerinnen und Nutzer und die Ähnlichkeit mit anderen Werken berücksichtigen, die subjektive Intention des Urhebers dagegen außer Acht lassen.⁴⁹⁶ In dieser Hinsicht sollte sich der Standard für Originalität im Urheberrecht stärker am Standard für Neuheit im Patentrecht orientieren, das schutzfähige Gegenstände aus einer sozialen bzw. historischen und nicht aus einer individuellen bzw. subjektiven Perspektive betrachtet.⁴⁹⁷

⁴⁹¹ *Appleton gegen Harnischfeger Corp.* 1995 (2) SA 247 (AD) at 43–44 (Südafrika).

⁴⁹² *Henkel KgaA gegen Holdfast* [2006] NZSC 102, [2007] 1 NZLR 577 (Neuseeland).

⁴⁹³ Zimmerman E., *op.cit.*, S. 14-21.

⁴⁹⁴ Clifford, R.D. (1996), *op.cit.*, 1694-1695. Deltorn J.M. (2017), *op.cit.*, S. 7. Deltorn J.M. und Macrez F. (2018), *op.cit.*, S. 8. Gervais D.J., „The Machine as Author“, *Iowa Law Review Vol. 105*, S. 1-60. Mezei P., „From Leonardo to the Next Rembrandt – The Need for AI-Pessimism in the Age of Algorithms“, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3592187. Ramalho A., *op.cit.*, S. 22-24.

⁴⁹⁵ Yanisky-Ravid S. und Velez-Hernandez L.A., „Copyrightability of Artworks Produced by Creative Robots, Driven by Artificial Intelligence Systems and the Originality Requirement: The Formality-Objective Model“, *Minnesota Journal of Law, Science & Technology* 19(1), S. 40-48.

⁴⁹⁶ Bonadio E. und McDonagh L., „Artificial Intelligence as Producer and Consumer of Copyright Works: Evaluating the Consequences of Algorithmic Creativity“, *Intellectual Property Quarterly* 2, S. 112-137, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3617197.

⁴⁹⁷ Vgl. Boden M., *The Creative Mind: Myths And Mechanisms*, Routledge, London, S. 32.

5.4. Politische Lösungsmöglichkeiten: Sind Anreize notwendig?

Wissenschaftler und Gerichte haben darauf hingewiesen, dass ein Rechtssystem, das KI-generierter Kreativität keinen Schutz gewährt, aus „anreiztheoretischer“ Sicht negative Auswirkungen hätte. Die Anreiztheorie oder der Utilitarismus,⁴⁹⁸ der in den Vereinigten Staaten und anderen Ländern des angelsächsischen Rechtskreises vorherrscht, hebt weniger stark auf die Menschlichkeit des Urhebers ab als Persönlichkeitstheorien, die in Ländern des romanischen Rechtskreises großen Einfluss haben.⁴⁹⁹ Dies bietet Argumenten für eine nicht-menschliche Urheberschaft und eine Schutzfähigkeit von KI-generierter Kreativität mehr Raum. Nach dem anreiztheoretischen Ansatz ist „das Anbieten finanzieller Anreize, um das Wachstum und die Entwicklung der KI-Industrie zu fördern und die Verbreitung KI-generierter Werke sicherzustellen, wohl das oberste Ziel der Zuweisung von Urheberrechten an menschliche Urheber“.⁵⁰⁰ Obwohl ein Computer keinen Anreiz braucht, um Output zu erzeugen, kann der Anreiz für die Person, die mit dem Computer zusammenarbeitet, nützlich sein.⁵⁰¹ Insbesondere argumentieren die Autoren, dass es einen zusätzlichen Anreiz für die Industrie geben sollte, damit sie die nötige Zeit und das nötige Geld investiert, um Maschinen intelligentes Verhalten zu lehren,⁵⁰² oder Nutzer belohnt, die die KI trainieren und anweisen, Inhalte zu generieren.⁵⁰³ Wie einige argumentieren, würden ordnungspolitische Erwägungen es aus utilitaristischer Sicht zwingend erforderlich machen, KI-generierten Output in irgendeiner Form – sei es urheber- oder wettbewerbsrechtlich oder durch einen Schutz sui generis – zu schützen,⁵⁰⁴ da der Prozess der Schöpfung – durch Mensch oder Computer – keinen Einfluss auf dessen Beitrag zum Gemeinwohl hat.⁵⁰⁵

Die meisten Länder des romanischen Rechtskreises sind jedoch möglicherweise nicht so grundlegend von Gemeinwohl- und Anreizerwägungen beeinflusst und legen vielleicht mehr Wert auf ein ausgewogenes System, weshalb sie jede Abweichung vom persönlichkeits-theoretischen Ansatz ablehnen, der den Blick dieser Länder auf das Urheberrecht prägt – und in den allermeisten Ländern auch den Begriff der Originalität. Auch wenn KI-generierte Schöpfungen Anreize zur Förderung von Innovation und Vermarktung rechtfertigen können, ist jedoch die Notwendigkeit solcher Anreize angesichts

⁴⁹⁸ Fisher W., *op.cit.*, S. 177-180.

⁴⁹⁹ Kaminski M., *op.cit.*, S. 599.

⁵⁰⁰ Hristov K., „Artificial Intelligence and the Copyright Dilemma“, *IDEA: The Intellectual Property Law Review* 57, S. 444. Siehe auch Brown N., *op.cit.*, 20-21.

⁵⁰¹ Hristov K., *op.cit.*, 438-439. Miller A.R., *op.cit.*, 1067.

⁵⁰² Bridy A. (2012), *op.cit.*, 1-27. Butler T., *op.cit.*, S. 735. Farr E.H., „Copyrightability of Computer-Created Works“, *Rutgers Computer & Technology Law Journal* 15, S. 73-74. Kasap A., *op.cit.*, S. 361-364; Abbott R., I Think, Therefore I Invent: Creative Computers and the Future of Patent Law“, 57 B.C.L. Rev. 1079 (2016), <http://lawdigitalcommons.bc.edu/bclr/vol57/iss4/2>. Milde K.F., *op.cit.*, p. 390.

⁵⁰³ Brown N., *op.cit.*, S.37. Denicola R., *op.cit.*, S. 283. Ralston W.T., „Copyright in Computer-Composed Music: HAL Meets Handel“, *Journal of the Copyright Society of the USA* 52, S. 303-304. Samuelson S. (1986), *op.cit.*, S. 1224-1228.

⁵⁰⁴ Milde K.F., *op.cit.*, S. 400-403.

⁵⁰⁵ Butler T., *op.cit.*, S. 735. Denicola R., *op.cit.*, S. 273. Kaminski M., *op.cit.*, S. 599.

ihrer möglichen Auswirkungen auf menschliche Schöpfungen fragwürdig.⁵⁰⁶ In Anbetracht der großen Zahl automatisierter Schöpfungen könnte die Gewährung eines Schutzes für diese Werke beispielsweise die menschliche Urheberschaft und die bestehenden Arbeitsplätze in diesem Bereich entwerten und die Kreativität behindern,⁵⁰⁷ da sie dazu führen könnte, dass Künstler ihre Werke aus Angst vor einer Verletzung von geschütztem Material nicht veröffentlichen⁵⁰⁸ oder standardisierter und homogener KI-generierter Output das kreative Ökosystem verstopft, was Auswirkungen auf die kulturelle Vielfalt und die Identitätspolitik hätte.

Die Frage lautet, ob eine Ausweitung des gegenwärtigen Urheberrechtsschutzes auf computergenerierte Werke sinnvoll ist. Der bestehende Rechtsrahmen bietet möglicherweise bereits einen ausreichenden patent- und urheberrechtlichen Schutz für die zugrunde liegende Software, einen Schutz sui generis für Datenbanken oder andere rechtliche Mechanismen wie das Wettbewerbsrecht, um automatisierte Werke zu schützen, ohne das bestehende Urheberrechtssystem auf nicht-menschliche Urheber auszuweiten.⁵⁰⁹ Die Fragen sollten juristisch und ökonomisch geprüft werden, bevor irgendwelche Lösungen bevorzugt werden.⁵¹⁰ Ginsburg und Budiardjo haben diesen Punkt betont: „Wir können uns verschiedenste Szenarien ausdenken, die die Forderung nach einem Schutz sui generis unterstützen oder entkräften, aber ohne empirische Beweise wäre es unklug (und verfrüht) zu versuchen, ein System zu entwickeln, das urheberlosen Output abdeckt.“⁵¹¹

5.4.1. Kein Schutz: Gemeinfreiheit KI-generierter Werke

Wie unsere Erörterung der Schutzvoraussetzungen weiter oben ergeben hat, könnte das Verständnis der Begriffe Rechtspersönlichkeit, Urheberschaft und Originalität im Rahmen des aktuellen Urheberrechtssystems KI-generierte Kreativität vom Urheberrechtsschutz ausschließen.⁵¹² Die Verbannung von KI-generierter Kreativität in die Gemeinfreiheit wäre daher eine mögliche Lösung – die im derzeitigen Rechtsrahmen wohl am ehesten unterstützt würde.

Bei diesem Ansatz hängt das Eigentum am Urheberrecht vom Umfang des menschlichen Zutuns ab. Die bloße Datenauswahl und -klassifizierung durch den Menschen reicht nicht aus, um das Erfordernis der „Originalität“ zu erfüllen; vielmehr setzt der Schutz

⁵⁰⁶ Craglia M., *op.cit.*, S. 67-68.

⁵⁰⁷ Bonadio E. und McDonagh L., *op.cit.*

⁵⁰⁸ Deltorn J.M. (2017), *op.cit.*

⁵⁰⁹ Deltorn J.M. und Macrez F. (2018), *op.cit.*, S. 24.

⁵¹⁰ Craglia M., *op.cit.*, S. 68.

⁵¹¹ Ginsburg J. und Budiardjo L.A., *op.cit.*, S. 448. Siehe auch Ginsburg J., *op.cit.*, S. 131-135.

⁵¹² Aplin T. und Pasqualetto G., *op.cit.*, §5.01-09. Clifford, R.D. (2018), „Creativity Revisited“, *IDEA: The IP Law Review* 59, S. 26-29. Clifford, R.D. (1996), *op.cit.*, 1700-1702. Huson G., „I, Copyright“, *Santa Clara High Technology Law Journal* 35, S. 72-78. Mezei P., *op.cit.* Palace V.M., „What if Artificial Intelligence Wrote This: Artificial Intelligence and Copyright Law“, *Florida Law Review* 71(1), S. 238-241. Saiz Garcia C., „Las obras creadas por sistemas de inteligencia artificial y su protección por el derecho de autor (AI Created Works and Their Protection Under Copyright Law)“ *InDret* 1, S. 38-39, <https://ssrn.com/abstract=3365458>. Gervais D.J., *op.cit.*, S. 1-60. Ramalho A., *op.cit.*, S. 22-24. Svedman M., *op.cit.*, S. 1-22.

einen tatsächlichen substanziellen menschlichen Beitrag zur Anleitung des KI-Systems im Schöpfungsprozess voraus.⁵¹³ Nur wenn ein substanzieller menschlicher Input im Spiel ist und alle kreativen Entscheidungen in den Computercode oder die Nutzeranweisungen eingebettet sind, läge das Urheberrecht beim menschlichen Urheber.⁵¹⁴ In diesem Zusammenhang wurden vier Modelle für die Verortung der Urheberschaft identifiziert: (1) alleinige Urheberschaft für den Nutzer des Tools, wenn dessen Entwickler nicht an dem generierten kreativen Werk mitwirkt; (2) alleinige Urheberschaft für den Entwickler des Tools, wenn der Bediener keine Rolle für den Output spielt und das selbstgenerierende Tool Output auf Basis des vom Entwickler bereitgestellten Trainings und kreativen Rohmaterials erschafft; (3) gemeinsame Urheberschaft für den Nutzer und den Programmierer, wenn der Output die kreativen Beiträge beider widerspiegelt; (4) urheberlose Werke – weder Entwickler noch Nutzer tragen so viel zu der Äußerung bei, dass die Originalität des Werks eine Urheberschaft begründet.⁵¹⁵ Wenn der kreative Output sowohl aus menschlichen als auch maschinellen Entscheidungen resultiert, müssen Materialien, die aus maschinellen Entscheidungen resultieren, auf jeden Fall herausgefiltert werden, wie bei gemeinfreien Materialien üblich.⁵¹⁶ Ein Schutz wird dann nur für selbstständig urheberrechtsfähige menschliche Beiträge in Frage kommen.

Es wurde ein zusätzlicher Test vorgeschlagen, mit dem sich feststellen lässt, ob ein Werk schutzwürdig ist – in Abhängigkeit davon, wie viel von der vom Programmierer intendierten Bedeutung sich auf das endgültige Werk überträgt und ob der Nutzer bzw. Programmierer den Output vorhersagen könnte.⁵¹⁷ Ein solcher Test müsste die Umsetzung traditioneller urheberrechtlicher Anforderungen voraussetzen, die einen Schutz nur dann gewähren, wenn das Werk das Produkt der Fantasie der menschlichen Urheber und eine Konzeption davon ist. Wenn dieser Test eingeführt wird, sollte das Urheberrecht eine Umkehrung der traditionellen Beweislast bewirken, sodass die Person, die sie beansprucht, die menschliche Urheberschaft eines konkreten KI-generierten Outputs nachweisen muss, indem sie belegt, dass der Output vorhersehbar eine Bedeutung oder Botschaft enthält, die der Urheber seinem Publikum vermitteln möchte.⁵¹⁸

5.4.2. Urheberschaft und Rechtsfiktionen: Soll ein Mensch der Urheber sein?

Damit KI-generierte Kreativität nicht gemeinfrei wird, und um notwendige Anreize für am kreativen Prozess der KI beteiligte menschliche Akteure zu schaffen, wurde eine Rechtsfiktion vorgeschlagen (und in Gesetzesform gegossen), wonach die Urheberschaft KI-generierter Werke den Akteuren übertragen wird, die Sachkenntnis, Arbeitsaufwand und

⁵¹³ Selvadurai N. und Matulionyte R., „Reconsidering creativity: copyright protection for works generated using artificial intelligence“, *Journal of Intellectual Property Law & Practice* jpa062, S. 539.

⁵¹⁴ Gervais D.J., *op.cit.*, S. 51-60. Selvadurai N. und Matulionyte R., *op.cit.*, S. 538.

⁵¹⁵ Ginsburg J. und Budiardjo L.A., „Authors and Machines“, *Berkeley Technology Law Journal* 34(2), S. 404-445.

⁵¹⁶ Gervais D.J., *op.cit.*, S. 54.

⁵¹⁷ Boyden B., „Emergent Works“, *Columbia Journal of Law and the Arts* 39, S. 377-394.

⁵¹⁸ Boyden B., *op.cit.*, 393-394.

Anstrengungen aufwenden, um die KI überhaupt erst zu schaffen, zu trainieren oder anzuleiten. Dieser Ansatz wurde auch als Theorie des fiktiven menschlichen Urhebers („Fictional Human Author Theory“) bezeichnet.⁵¹⁹

Dieser politische Ansatz zeichnete sich bereits recht früh ab, als das kreative Potenzial und die Mechanik des maschinellen Lernens und der KI noch völlig unbekannt waren. Im Vereinigten Königreich wurde der urheberrechtliche Schutz einer computergenerierten Sequenz für eine Lotterie schon 1985 im Fall *Express Newspapers gegen Liverpool Daily Post* diskutiert. Der Richter Whitford gewährte der Klägerin Urheberrechtsschutz für den automatisierten Output und lehnte die Vorstellung ab, dass das Urheberrecht an dem Werk bei dem Computer liegen könnte. Der Computer sei ein bloßes Werkzeug für das Schaffen; das Argument, der Computer sei der Urheber, ähnele der Behauptung, bei einem Schriftwerk sei „der Stift der Urheber und nicht die Person, die den Stift führt“.⁵²⁰ Unter dem Aspekt der unmittelbaren Ursache, der Intention und des Willens wurde diese Position kürzlich von Dan Burk prägnant zusammengefasst: „Wenn es einen Urheber gibt, handelt es sich um einen oder mehrere der Menschen, die für die Produktion des Outputs unmittelbar genug ursächlich sind. [...] Aber der Urheber ist niemals die Maschine“.⁵²¹

Das Vereinigte Königreich ist das erste Land, das einen besonderen Schutz für computergenerierte Kreativität eingeführt hat.⁵²² Section 9(3) des Copyright Designs and Patents Act 1988 (CDPA)⁵²³ stellt klar, dass bei computergenerierten Werken die Urheberschaft bei der Person liegt, die die zur Erschaffung des Werks erforderlichen Vorkehrungen trifft. Darüber hinaus sieht Section 178 vor, dass „computergeneriert in Bezug auf ein Werk bedeutet, dass das Werk von einem Computer unter Umständen generiert wird, bei denen es keinen menschlichen Urheber des Werkes gibt“. Nach dieser Regelung würde die Schutzdauer für computergenerierte Werke 50 Jahre ab dem Zeitpunkt der Herstellung des Werkes betragen. Kurz darauf erließen andere Länder des angelsächsischen Rechtskreises, darunter Hongkong, Indien, Irland, Singapur und Neuseeland, ähnliche gesetzliche Regelungen.⁵²⁴

⁵¹⁹ Wu AJ., „From Video Games to Artificial Intelligence: Assigning Copyright Ownership to Works Generated by Increasingly Sophisticated Computer Programs“, *AIPLA Quarterly Journal*, S. 173-174.

⁵²⁰ *Express Newspapers Plc gegen Liverpool Daily Post & Echo Plc* [1985] 1 WLR 1089 (UK).

⁵²¹ Burk D.L., „Thirty-Six Views of Copyright Authorship, By Jackson Pollock“, *Houston Law Review* 58, pp 1-38. Siehe auch Hedrick S.F., „I ‚Think‘, Therefore I Create: Claiming Copyright in the Outputs of Algorithms“, *NYU Journal of Intellectual Property & Entertainment Law* 8(2), S. 324-375, und Grimmelmann J., „There is No Such Thing as a Computer-Authored Work And It is a Good Thing, Too“, *Columbia Journal of Law and the Arts* 39, pp. 403-416.

⁵²² Guadamuz A., „Do Androids Dream of Electric Copyright? Comparative Analysis of Originality in Artificial Intelligence Generated Works“, *Intellectual Property Quarterly*, S. 169-186.

⁵²³ Copyright, Designs and Patents Act 1988, <https://www.legislation.gov.uk/ukpga/1988/48/contents> (UK).

⁵²⁴ Copyright Ordinance cap 528, Section 11(3), <https://www.elegislation.gov.hk/hk/cap528> (Hongkong); Copyright Act 1957, Section 2(d)(vi), <https://www.wipo.int/edocs/lexdocs/laws/en/in/in122en.pdf> (Indien); Copyright and Related Rights Act 2000, Section 21(f), <http://www.irishstatutebook.ie/eli/2000/act/28/enacted/en/html> (Irland); Copyright Act 1987 chapter 63, Section 7A, <https://sso.agc.gov.sg/Act/CA1987#pr27-> (Singapur); Copyright Act 1994, Section 5(2), <http://www.legislation.govt.nz/act/public/1994/0143/latest/DLM345634.html> (Neuseeland).

Bei dieser Regelung gibt es allerdings zwei Probleme. Das erste ist grundsätzlich und systembedingt. Wäre dieser Ansatz in einem Rechtsrahmen tragfähig, der auf dem Begriff der Originalität als Persönlichkeitsäußerung des Urhebers aufbaut, wie er im EU-Recht und in den meisten internationalen Rechtsordnungen übernommen wurde? Natürlich würden Programmierung, Training und die Erteilung von Anweisungen das Kriterium eines originalen Beitrags des menschlichen Gegenübers wohl kaum erfüllen, da letztlich jede „Äußerung“ das Ergebnis des kreativen Prozesses der KI wäre. Solange der gegenwärtige subjektive Standard für Originalität gilt, würde jegliche Theorie des fiktiven menschlichen Urhebers an der fehlenden Originalität von KI-generierter Kreativität scheitern. Dem Werk selbst, dessen fiktive Urheberschaft einem menschlichen Akteur zugeschrieben wird, würde weiterhin die Originalität fehlen, sodass kein Schutz möglich wäre. Zu beachten ist, dass im Vereinigten Königreich und in anderen Ländern des angelsächsischen Rechtskreises Ansätze auf Basis der „Theorie des fiktiven menschlichen Urhebers“ und der „notwendigen Vorkehrungen“ Gesetz wurden, als dort noch die Originalitätskriterien „Sweat of the Brow“ oder „Skill and Labour“ maßgeblich waren. Seitdem hat sich, wie schon erwähnt, vieles verändert, und das Originalitätskriterium Persönlichkeit hat alle alternativen Ansätze ganz oder teilweise ersetzt,⁵²⁵ was die Systemkonformität des Regelungsansatzes von Section 9(3) CDPA in Frage stellt.

Das zweite Problem ist eher praktischer Natur. Mit diesem Ansatz ist die für die notwendigen Vorkehrungen zuständige Person schwer zu ermitteln.⁵²⁶ Gehört das KI-generierte Werk der Person, die das System erstellt hat, also etwa dem Softwareentwickler, dem Hersteller, der Person, die es trainiert hat, oder der Person, die es mit konkreten Inputs gefüttert hat wie ein Nutzer?⁵²⁷ Nach Ansicht von Guadamuz sollte die Ambiguität des Systems jedoch eigentlich als positive Eigenschaft gelten, die die Frage der Dichotomie zwischen Nutzer und Programmierer ins Leere laufen lässt und eine Einzelfallanalyse ermöglicht.⁵²⁸ Auf jeden Fall wurde die Frage, wer für die notwendigen Vorkehrungen zuständig ist, durch Rechtsprechung und Wissenschaft mittlerweile auf vielfältige Weise beantwortet.

5.4.2.1. Soll der Programmierer der Urheber sein?

Eine erste mögliche Antwort auf die Frage, ob der Programmierer der Urheber sein sollte, findet sich im Urteil im Fall *Nova Productions gegen Mazooma Games*, in dem Section 9(3) des CDPA angewandt wurde. Hier ging es um das Urheberrecht an Rahmenbildern, die von einem Computerprogramm anhand von Bitmap-Dateien erzeugt und auf dem Bildschirm angezeigt wurden, wenn die Nutzer ein Snooker-Videospiel spielten. Das Gericht lehnte es ab, die Urheberschaft dem Nutzer zuzusprechen, da dessen Input nicht künstlerischer Natur sei:

⁵²⁵ *Temple Island Collections v New English Teas* (No. 2) [2012] EWPC 1. Guadamuz A., *op.cit.*, S. 178-180; Rahmatian A., *op.cit.*, S. 4-34.

⁵²⁶ Dorotheu E., „Reap the benefits and avoid the legal uncertainty: who owns the creations of artificial intelligence?“, *Computer and Telecommunications Law Review* 21, S. 85.

⁵²⁷ Bonadio E. und McDonagh L., *op.cit.*, S. 117-119. Kasap A., *op.cit.*, S. 364-376.

⁵²⁸ Guadamuz A., *op.cit.* S. 177.

Das Aussehen eines bestimmten Bildschirms hängt bis zu einem gewissen Grad davon ab, wie das Spiel gespielt wird. Wenn zum Beispiel der Drehknopf gedreht wird, dreht sich die Queue um den Spielball. Ebenso wird die Stärke des Stoßes davon beeinflusst, wann genau der Spieler die Spieltaste drückt. Doch bei keinem der künstlerischen Werke, die in den aufeinanderfolgenden Rahmenbildern geschaffen werden, ist der Spieler der Urheber. Sein Input ist nicht künstlerischer Natur, und er hat weder Sachkenntnis noch Arbeitsaufwand künstlerischer Art beigesteuert. Auch hat er keine der zur Erschaffung der Rahmenbilder notwendigen Vorkehrungen getroffen. Er hat nichts weiter getan, als das Spiel zu spielen.⁵²⁹

Stattdessen, so das Gericht, handle es sich bei dem alleinigen Urheber und der Person, die die notwendigen Vorkehrungen getroffen hat, um den Programmierer: „Die für die Erschaffung des Werkes notwendigen Vorkehrungen hat [der Kläger] getroffen, denn er hat das Aussehen der verschiedenen Elemente des Spiels sowie die Regeln und die Logik erdacht, nach denen jeder Frame generiert wird, und das entsprechende Computerprogramm geschrieben.“⁵³⁰ In Wahrheit war das Ergebnis im Fall *Nova Productions* möglicherweise eine direkte Folge der rudimentären Technologie, um die es hier ging, und nach Auffassung von Guadamuz könnte die Urheberschaft je nach den konkreten Umständen des Falles und der betrachteten Technologie auch anders verortet werden.⁵³¹ Trotzdem wurde der Ansatz auch in anderen Ländern wie den USA vorgeschlagen, wo die Gerichte dann den Wortlaut des Urheberrechtsgesetzes verbiegen mussten, sodass in Fällen, in denen weder der Programmierer noch der Nutzer die Voraussetzungen für die Urheberschaft an einem urheberrechtsfähigen Werk erfüllt, das Gericht das Urheberrecht der Person zuweisen sollte, die das Urheberrecht an dem Computerprogramm besitzt.⁵³²

Die Verortung der Urheberschaft beim Programmierer einer KI, die Inhalte generiert, ruft grundlegende Kritik hervor. Die Verortung der Urheberschaft beim Softwareentwickler – oder bei Eigentümern von KI-Technologien wie Unternehmen und Investoren – ist möglicherweise ein eklatanter Irrtum.⁵³³ Tatsächlich scheint es zumindest bei einer Kreativität, die auf neuronalen Netzen neuester Technik (GAN und CAN) basiert, keinen direkten kausalen Zusammenhang zwischen den Softwareentwicklern und dem endgültigen KI-generierten Output zu geben, denn die in diesen Output eingebettete Äußerung ergibt sich aus dem Training der Maschine und den Anweisungen, die zur Erzeugung dieses konkreten Outputs gegeben wurden. Angesichts dessen ließe sich erstens aus systemischer Sicht argumentieren, dass sich bei dieser Regelung ein grundlegender Widerspruch ergibt. Im Grunde ist ja – wie auch das Internetgericht Peking in einem weiter oben erwähnten Fall hervorgehoben hat – der Softwareentwickler bereits mit

⁵²⁹ *Nova Productions Ltd gegen Mazooma Games Ltd & Ors Rev 1* [2006] EWHC 24 (Ch) (20. Januar 2006) (Vereinigtes Königreich). Siehe auch Farr E.H., *op.cit.*, 75-78.

⁵³⁰ *Nova Productions Ltd gegen Mazooma Games Ltd & Ors Rev 1* [2006] EWHC 24 (Ch) (20 January 2006) (UK). Siehe auch Farr E.H., *op.cit.*, 73-74.

⁵³¹ Guadamuz A., *op.cit.*, S. 177.

⁵³² Wu A.J., *op.cit.*, S. 173-174.

⁵³³ Abbott R., „Artificial Intelligence, Big Data and Intellectual Property: protecting computer generated works in the United Kingdom“ in Aplin T. (ed.), *Research Handbook on intellectual property and digital technologies*, Edward Elgar, England, S. 323-324. Svedman M., *op.cit.*, 10-11. Bridy A. (2016), „The Evolution of Authorship: Work made by Code“, *Columbia Journal of Law and the Arts* 39, S. 400-401. Bridy A. (2012), *op.cit.*, 24-25, Randnr. 62. Samuelson P. (1986), *op.cit.*, S. 1207-1212.

Exklusivrechten an der Software belohnt worden, die Werke generiert.⁵³⁴ Angesichts der Tatsache, dass genau diese Rechtsfiktion einen Anreiz zur Erschaffung KI-generierter Werke setzen soll, der wegen deren Gemeinfreiheit sonst wohl fehlen würde, dürfte eine solide wirtschaftliche Analyse zudem eher dagegensprechen, denselben Marktteilnehmer doppelt zu belohnen. Aus eher praktischer Sicht würde diese Lösung Codierern darüber hinaus womöglich einen aggressiven Urheberrechtsschutz für zahllose Kreativitätserzeugnisse einräumen,⁵³⁵ sodass auch jeglicher Anreiz für den ursprünglichen Programmierer schwinden würde, mehr Software zu erstellen.⁵³⁶

Die US-Rechtsprechung scheint auch die Schlussfolgerung zu stützen, dass Software und Output zwei ganz verschiedene Dinge sind und Eigentum an Ersterer keine Rechte an Letzterem impliziert. In zwei Fällen haben die Gerichte entschieden, dass Output, der mit einer unerlaubten Kopie einer Software oder eines Programms erstellt wurde, selbst nicht als unerlaubtes abgeleitetes Werk gilt. Im Fall *Design Data gegen Unigate Enterprises* bestätigte das Berufungsgericht das Urteil des Bezirksgerichts und entschied, dass sich das Urheberrecht der Klägerin an dem Computerprogramm nicht auf den Output des Programms (Zeichnungen und Daten für Stahlbauten) erstreckte.⁵³⁷ Im Fall *Rearden gegen Walt Disney Co.* wurden Filme, die mit einer (unerlaubten) Kopie der Software der Klägerinnen erstellt wurden, nicht als abgeleitete Werke der Software betrachtet, da die Software zwar viel Arbeit übernommen habe, der Löwenanteil der kreativen Äußerung im Film jedoch den Beklagten zuzuschreiben sei.⁵³⁸

5.4.2.2. Soll der Nutzer der Urheber sein?

Eine sinnvollere Lösung könnte sein, die Rechte an KI-generiertem Output der Nutzerin oder dem Nutzer des Generatorprogramms zuzuweisen.⁵³⁹ Zu diesem Schluss gelangte das weiter oben beschriebene Urteil aus Peking.⁵⁴⁰ Samuelson argumentiert, der Grund für die Entstehung des KI-generierten Werkes sei der Nutzer, und daher sei es „unter diesen Umständen nicht unfair, einige Rechte einer Person zu geben, die das Werk zu dem beabsichtigten Zweck nutzt, weitere Werke zu erschaffen“.⁵⁴¹ Diese Lösung wäre nach urheberrechtlichen Standards nicht neu. So wird in den Vereinigten Staaten das Urheberrecht – und die Urheberschaft – den Nutzern als Instrument der Fixierung⁵⁴²

⁵³⁴ *Beijing Feilin Law Firm v Baidu Corporation* (26 April 2019) Beijing Internet Court, (2018) Beijing 0491 Minchu No. 239. He K. (2020b), *op.cit.* Siehe auch Bonadio E. und McDonagh L., *op.cit.*, S. 117. Chen M., *op.cit.*, S. 14-18. Samuelson P. (1986), *op.cit.*, S. 1207-1212.

⁵³⁵ Svedman M., *op.cit.*, S. 14.

⁵³⁶ Huson G., *op.cit.*, S. 74.

⁵³⁷ *Design Data Corp. gegen Unigate Enterprise*, No. 14-16701 (9th Cir. 2017) (USA).

⁵³⁸ *Rearden gegen Walt Disney Co.*, No. 17-cv-04006-JST (N.D. Cal. 2018).

⁵³⁹ CONTU, *op.cit.*, S. 45. Siehe auch Ralston W.T., *op.cit.*, S. 303-304.

⁵⁴⁰ *Beijing Feilin Law Firm v Baidu Corporation* (26 April 2019) Beijing Internet Court, (2018) Beijing 0491 Minchu No. 239. He K. (2020b), *op.cit.*

⁵⁴¹ Samuelson P. (2020), „AI Authorship?“ *Communications of the ACM* 63(7), S. 20-22. Samuelson P. (1986), „Allocating Ownership Rights in Computer-Generated Works“, *University of Pittsburg Law Review* 47(4), S. 1200-1204.

⁵⁴² 17 U.S.C. § 114.

zugestanden, also etwa einer Person, die eine Jazzdarbietung auf Band aufnimmt.⁵⁴³ In diesem Szenario wäre der Nutzer und nicht die Jazzmusiker der Urheber der Tonaufnahme. Ebenso könnte der Nutzer als Urheber der Fixierung eines KI-generierten Werkes angesehen werden. Natürlich müsste dazu eine spezielle Bestimmung wie Section 9(3) des CDPA eingeführt werden. In Ausnahmefällen, etwa wenn der Nutzer die Software lediglich ausführt und sonst keine Kontrolle über sie hat, wäre die Vergabe des Urheberrechts an den Nutzer allerdings wohl eine suboptimale Lösung, die im Widerspruch zur urheberrechtlichen Anreiztheorie stünde.⁵⁴⁴ Ein Ausweg könnte in diesem Fall die gemeinsame Urheberschaft von Nutzern und Programmierern sein⁵⁴⁵ – je nachdem, welche rechtliche Regelung eine Rechtsordnung für die gemeinsame Urheberschaft vorsieht.

Das Europäische Parlament schloss sich dieser Ansicht im „Entwurf eines Berichts über die Rechte des geistigen Eigentums bei der Entwicklung von KI-Technologien“ offenbar ebenfalls an und schlug vor, das Urheberrecht an KI-generierten Werken den KI-Nutzern zu übertragen. Es vertritt darin die „Auffassung, dass der Schutz der durch KI geschaffenen technischen und künstlerischen Schöpfungen in Erwägung gezogen werden muss, um diese Form des Schaffens zu fördern, [und] dass bestimmte durch KI erzeugte Werke mit geistigen Werken vergleichbar sind und daher urheberrechtlich geschützt werden könnten“.⁵⁴⁶ Es empfiehlt daher, „dass der Inhaber der Rechte derjenige sein sollte, der das Werk rechtmäßig erstellt und veröffentlicht, sofern der Entwickler der Technologie sich eine solche Nutzung nicht ausdrücklich vorbehalten hat“.⁵⁴⁷ Offenbar impliziert dieser Vorschlag, dass auch die Urheberrechte auf die Nutzer übertragen werden sollen, denn der Berichtsentwurf nimmt auf den Urheberrechtsschutz Bezug und nicht auf einen Schutz *sui generis*: „Daher wird vorgeschlagen, zu prüfen, ob es zweckmäßig ist, die Urheberrechte an einem solchen „Kunstwerk“ der natürlichen Person zuzuweisen, die es rechtmäßig erstellt und veröffentlicht [...]“.⁵⁴⁸

5.4.2.3. Soll der Arbeitgeber der Urheber sein?

Wissenschaftlerinnen und Wissenschaftler haben die Auftragswerksdoktrin⁵⁴⁹ als rechtlichen Rahmen vorgeschlagen, um die Inhaberschaft an sowie die Schutzfähigkeit und Rechenschaftspflicht von Werken zu gewährleisten, die von KI-Systemen generiert

⁵⁴³ Samuelson P. (1986), *op.cit.*, S. 1200-1204. Die meisten Länder setzen für Tonaufnahmen jedoch auf den Schutz der verwandten Schutzrechte und nicht wie die Vereinigten Staaten auf das Urheberrecht.

⁵⁴⁴ Ralston W.T., *op.cit.*, S. 304-305.

⁵⁴⁵ Z. B. Bonadio E. und McDonagh L., *op.cit.*, 117-118.

⁵⁴⁶ Europäisches Parlament (2020), *op.cit.*

⁵⁴⁷ Europäisches Parlament (2020), *op.cit.*

⁵⁴⁸ Europäisches Parlament (2020), *op.cit.*

⁵⁴⁹ Nach dem US-Urheberrechtsgesetz stellt ein Auftragswerk („work made for hire“) eine Ausnahme von der Regel dar, dass das Urheberrecht an einem schutzfähigen Werk nur der Urheber beanspruchen kann (oder diejenigen, die Rechte vom Urheber ableiten). Nach dieser Doktrin gilt der Arbeitgeber auch dann als Urheber eines Werkes, wenn ein Arbeitnehmer das Werk geschaffen hat. Sie ist in Section 101 des Copyright Act festgelegt. Ein Werk, das von einem Arbeitnehmer im Rahmen seines Arbeitsverhältnisses geschaffen oder speziell für die Nutzung durch den Arbeitgeber bestellt oder in Auftrag gegeben wurde, ist ein Auftragswerk, wenn die Parteien dies in einem unterzeichneten Vertrag vereinbaren (United States Copyright Office, „Works Made for Hire“, <https://copyright.gov/circs/circ09.pdf>).

werden.⁵⁵⁰ Dieses Modell würde auf der Fiktion basieren, dass das KI-System eine kreative Angestellte oder ein unabhängiger Auftragnehmer der Nutzer – natürlicher oder juristischer Personen – ist, die das System nutzen und von ihm profitieren.⁵⁵¹ Samuelson argumentiert: „Wer ein Generatorprogramm kauft oder lizenziert, hat den Computer und seine Programme in gewisser Weise für seine kreativen Bemühungen ‚angestellt‘. Ähnliche Erwägungen wie diejenigen, die der Auftragswerksregel zugrunde liegen, sprechen auch dafür, Rechte an computergenerierten Werken den Nutzern zuzuordnen.“⁵⁵² Damit würden das Urheberrecht und die Haftung für alle von dem Werk ausgehenden Rechtsverletzungen der natürlichen oder juristischen Person auferlegt, die als Arbeit- oder Auftraggeberin des KI-Systems gilt, das das Werk erschafft.

Die Annahme dieses Rechtsrahmens würde auf eine gewisse Kritik stoßen. Erstens wird oft das Argument vorgebracht, Arbeitgeber würden als Urheber von Auftragswerken behandelt, obwohl sie für den Output keine Rolle spielen, und daher könnten ähnliche Regelungen auch für KI-generierte Kreativität entwickelt werden.⁵⁵³ Diese Position scheint allerdings zu übersehen, dass das zugrunde liegende Werk bei der Rechtsfiktion „Auftragswerk“ von einem menschlichen Urheber geschaffen wurde und das Kriterium der Originalität im derzeitigen Rechtsrahmen erfüllt. Dies wäre bei KI-generierter Kreativität nicht der Fall. Zweitens würde diese Regelung mit der Begründung angefochten, dass es sich um einen Missbrauch der Auftragswerksdoktrin handle, da es schwierig sei, ein vertragsrechtliches Arbeits- oder Auftragsverhältnis zwischen Mensch und Maschine zu definieren.⁵⁵⁴ Besonders problematisch wäre die Anwendung der Auftragswerksdoktrin in Ländern wie Frankreich, wo ein Eigentumsübergang auf Arbeit- oder Auftraggeber im Arbeitsvertrag oder Auftrag ausdrücklich vorgesehen sein muss und nicht vorausgesetzt wird. Es liegt auf der Hand, dass die Anwendung der Auftragswerksdoktrin auf KI-generierte Werke in jedem Fall zunächst eine substantielle gesetzliche und rechtswissenschaftliche Rekonstruktion der Begriffe „Arbeitgeber“ und „Arbeitnehmer“ erfordert.⁵⁵⁵

5.4.3. Soll ein Roboter der Urheber sein?

Eine Lösung – aus Sicht der Wissenschaft allerdings eher eine nachrangige⁵⁵⁶ – wäre die Verortung der Urheberschaft bei der KI. Dazu müsste im Gesetz eine Fiktion verankert werden, die der KI Rechtspersönlichkeit verleiht, sodass sie Urheberin eines Werks sein und

⁵⁵⁰ Bridy A. (2016), *op.cit.*, S. 400-401. Bridy A. (2012), *op.cit.*, 26, Randnr. 66. Hristov K., *op.cit.*, 431-454. Kaminski M., *op.cit.* Siehe auch Pearlman R., „Recognizing Artificial Intelligence (AI) as Authors and Inventors under U.S. Intellectual Property Law“, *Richmond Journal of Law & Technology* 24, S. 1-38. Yanisky-Ravid S. (2017), „Generating Rembrandt: Artificial Intelligence, Copyright, and Accountability in the 3A Era, The Human-Like Authors are Already Here: A New Model“, *Michigan State Law Review*, S. 659-726.

⁵⁵¹ Yanisky-Ravid S., *op.cit.*, S. 659-726.

⁵⁵² Samuelson P. (2020), *op.cit.*, S. 20-22; Samuelson P. (1986), *op.cit.*, S. 1200-1204.

⁵⁵³ Brown N., *op.cit.*, S.39. Kaminski M., *op.cit.*, 602.

⁵⁵⁴ Bonadio E. und McDonagh L., *op.cit.*, S. 114-115. Bridy A. (2012), *op.cit.*, S. 27, Randnr. 68. Butler T., *op.cit.*, S. 739-742. Huson G., *op.cit.*, S. 73-75. Ramalho A., *op.cit.*, S. 18-19.

⁵⁵⁵ Hristov K., *op.cit.*, S. 445-447.

⁵⁵⁶ Z. B. Bonadio E. und McDonagh L., *op.cit.*, S. 116. Farr E.H., *op.cit.*, S. 79. Ralston W.T., *op.cit.*, S. 302-303. Samuelson P. (1986), *op.cit.*, S. 1199-1200.

ein Urheberrecht besitzen kann,⁵⁵⁷ oder zumindest müsste das Gesetz so geändert werden, dass eine gemeinsame Urheberschaft von Computer und Mensch möglich wird⁵⁵⁸. Laut Pearlman sollte das Gesetz eine ausreichend kreative KI als Urheberin anerkennen, wenn die KI-Schöpfung Originalität aufweist und unabhängig von menschlichen Anweisungen entwickelt wurde, sodass die KI die Ursache der Kreativität ist und keine bloße Maschine, die Anweisungen eines menschlichen Urhebers abarbeitet.⁵⁵⁹ Sobald die KI zur Urheberin erklärt wird, würden Rechte unmittelbar einer natürlichen oder juristischen Person zugewiesen, etwa dem Schöpfer bzw. Programmierer der KI, dem Nutzer der KI oder mehreren Personen gemeinsam als Gemeinschaftswerk. Das Gesetz müsste dann festlegen, wer zum Empfang der Übertragung und zur Ausübung der Rechte berechtigt ist.

Auch in diesem Szenario könnte jedoch das Kriterium der Originalität für eine Maschine eine unüberwindbare Belastung darstellen. Der Begriff der Originalität müsste wohl so angepasst werden, dass er, wie bereits erwähnt, auch Werke einschließt, die von einer Maschine stammen.⁵⁶⁰ Wenn jedoch einer Maschine eine Rechtspersönlichkeit zuerkannt wird, könnte man auch argumentieren, dass die Maschine, sobald sie als (Rechts-)Person anerkannt ist, nach dem persönlichkeitsorientierten Konzept, das das Urheberrecht bestimmt und Originalität als Persönlichkeitsäußerung des Urhebers versteht, zu originaler Kreativität fähig wäre. In jedem Fall würde die Zulassung der KI als Urheberin wesentliche Änderungen des Rechtsrahmens erfordern. Wie bereits erwähnt, wäre es angesichts des frühen Stands der technologischen Entwicklung für die Politik eine suboptimale Lösung, die Rechtslage zu ändern, bevor wirklich intelligente Maschinen überhaupt entstehen – wobei deren Entstehung und Entwicklung bis heute nur eine hypothetische Spekulation ist.⁵⁶¹

5.4.4. Schutz sui generis für KI-generierte Kreativität

Vorgeschlagen wurde auch die Schaffung eines verwandten Schutzrechts sui generis (ohne Urheberschafts- oder Originalitätserfordernis), das die Investitionen in die Entwicklung und das Training einer KI, die Kreativität generiert, schützen könnte.⁵⁶² So schlägt McCutcheon ein System von Rechten sui generis für KI-generierte Kreativität vor, das den Rechten an Datenbanken ähnelt und so die Investitionen in die Schöpfung schützt, aber weder einen Urheber, noch Urheberschaft, noch Originalität voraussetzt.⁵⁶³ Das australische Copyright Law Review Committee verneinte zwar die Schutzfähigkeit nach dem traditionellen

⁵⁵⁷ Ihalainen J., „Computer creativity: artificial intelligence and copyright“, *Journal of Intellectual Property Law & Practice* 13(9), S. 724–728.

⁵⁵⁸ Abbott R., S. *op.cit.*

⁵⁵⁹ Pearlman R., *op.cit.*, S. 1-38).

⁵⁶⁰ Yanisky-Ravid S. und Velez-Hernandez L.A., *op.cit.*, S. 40-48.

⁵⁶¹ Huson G., *op.cit.*, S. 77-78. Liebesman Y., „The Wisdom of Legislating for Anticipated Technological Advancements“, *J. Marshall Rev. Intell. Prop. L.* 10, S. 172.

⁵⁶² Ciani J., „Learning from Monkeys: Authorship Issues Arising From AI Technology“ in Moura Oliveira P., Novais P., Reis L. (eds), *Progress in Artificial Intelligence EPIA 2019 Lecture Notes in Computer Science* vol. 11804, Springer, Cham.

⁵⁶³ McCutcheon J., *op.cit.*, S. 965-966.

Urheberrechtssystem, stellte aber fest, wenn computergenerierte Kreativität Schutz benötige, solle dieser „eher dem Schutz entsprechen, der für verwandte Schutzrechte gilt, [...] dem Schutz, der ausübenden Künstlern, Herstellern von Tonträgern und Sendeunternehmen gewährt wird“.⁵⁶⁴

Japan erwägt eine neuartige Regelung *sui generis* für nicht von Menschen geschaffenes geistiges Eigentum, die auf einem an das Markenrecht angelehnten Ansatz mit Schwerpunkt auf dem Schutz vor unlauterem Wettbewerb beruht.⁵⁶⁵ Dieser Ansatz soll den Schutz von KI-Werken einschränken, indem er das Schutzniveau flexibilisiert und an die Popularität der KI-generierten Werke als Indikator für deren Wert koppelt.⁵⁶⁶ Dies würde obskure Werke ausschließen, die nur mit Blick auf den Urheberrechtsschutz erschaffen werden. Der Vorschlag würde das Eigentum an dem Werk dem Menschen oder Unternehmen zuweisen, dessen Schöpfung die KI ist.⁵⁶⁷ Zur Begrenzung eines übermäßigen Schutzes von algorithmischer Kreativität schlagen einige Autoren zudem einen engen Geltungsbereich des Rechts *sui generis* vor, verbunden mit starken Schutzmaßnahmen für Fair Use und einer kurzen Schutzdauer von vielleicht etwa drei Jahren.⁵⁶⁸

5.4.5. Verleihung von Rechten an Verleger und Verbreiter

Darüber hinaus regen weitere Vorschläge die Verleihung von Rechten an Verleger und Verbreiter KI-generierter Werke an. Auf der einen Seite könnte auf KI-generierte Werke die Regelung für anonyme bzw. pseudonyme Werke angewandt werden. Mehreren nationalen Regelungen zufolge, wie sie etwa in Spanien, Frankreich, Italien und Schweden⁵⁶⁹ gelten, übt die Person die Rechte aus, die das Werk veröffentlicht. Auf der anderen Seite käme auch eine Lösung in Frage, die einem Verbreiter von KI-generierter Kreativität ein ähnliches Recht einräumt wie die EU-Verlagsrechte an zuvor unveröffentlichten Werken gemäß Artikel 4 der Richtlinie 2006/116/EG.⁵⁷⁰ Dieser Schutz gilt für die erste erlaubte Veröffentlichung bzw. öffentliche Wiedergabe von zuvor unveröffentlichten gemeinfreien Werken. Da auch KI-generierte Werke gemeinfrei wären, sähe die „Verbreiterregelung“ nur eine Belohnung für die Verbreitung KI-generierter Werke vor. Die Schutzdauer hierfür

⁵⁶⁴ Ricketson S. (2012), „The Need for Human Authorship – Australian Developments: Telstra Corp Ltd v Phone Directories Co Pty Ltd (Case Comme nt)“, *E.I.P.R.* 34(1), 54.

⁵⁶⁵ Intellectual Property Strategic Program 2016, S. 10-11, http://www.kantei.go.jp/jp/singi/titeki2/kettei/chizaikeikaku20160509_e.pdf.

⁵⁶⁶ Intellectual Property Strategic Program 2016, *op.cit.*, S. 11.

⁵⁶⁷ Intellectual Property Strategic Program 2016, *op.cit.*

⁵⁶⁸ Bonadio E. und McDonagh L., *op.cit.*, 136-137.

⁵⁶⁹ Ley de propiedad intelectual (Gesetz über geistiges Eigentum), 1996, Artikel 6, <https://www.wipo.int/edocs/lexdocs/laws/es/es/es177es.pdf> (Spanien); Code de la propriété intellectuelle (Gesetz über geistiges Eigentum), 1912, L113-6, <https://www.wipo.int/edocs/lexdocs/laws/fr/fr/fr077fr.pdf> (Frankreich); Legge n. 633, Protezione del diritto d'autore e di altri diritti connessi al suo esercizio (Gesetz Nr. 633, Schutz des Urheberrechts und verwandter Schutzrechte), 1941, Artikel 9, <https://www.wipo.int/edocs/lexdocs/laws/it/it/it158it.pdf> (Italien); Lag om upphovsrätt till litterära och konstnärliga verk (Gesetz über das Urheberrecht an literarischen und künstlerischen Werken), 1960, Artikel 7, <https://www.wipo.int/edocs/lexdocs/laws/sv/se/se006sv.pdf> (Schweden).

⁵⁷⁰ Ramalho A., *op.cit.*, S. 22-24.

könnte beispielsweise, wie in Artikel 4 der Richtlinie 2006/116/EG, auf 25 Jahre begrenzt werden.⁵⁷¹

5.5. Fazit

Anthropozentrismus und Persönlichkeitsorientierung beeinflussen den gegenwärtigen Rechtsrahmen des Urheberrechts stark, insbesondere den Begriff der Originalität, von der Berner Übereinkunft bis hin zum Recht wichtiger Länder. Daher fehlen KI-generierter Kreativität alle Grundvoraussetzungen für einen Urheberrechtsschutz: Rechtspersönlichkeit, Urheberschaft und Originalität.

Utilitaristische bzw. anreizbasierte Ansätze drängen auf die Einführung von Rechtsfiktionen, die aber interne systemische Widersprüche nicht zufriedenstellend angehen und ausräumen. Doch selbst wenn eine Rechtsfiktion die Annahme vorsieht, ein Werk sei von einem Menschen und nicht von einer KI erschaffen worden, fehlt dem Werk selbst Originalität, da Maschinen nach einem persönlichkeitsorientierten Standard von Natur aus nicht zu Originalität fähig sind. Nur eine grundlegende Überarbeitung des Urheberrechtssystems, die sich vom gegenwärtigen anthropozentrischen Ansatz verabschiedet, kann echter KI-generierter Kreativität vollständigen Urheberrechtsschutz bieten, wenn kein menschliches Zutun als originale Äußerung gewertet werden kann. Dies wäre jedoch unüberlegt, zumal die technologische Entwicklung in diesem Bereich noch in den Kinderschuhen steckt. Nachrangig wären auch Sui-generis-Ansätze denkbar, und wenn die Politik beschließt, für KI-generierte Kreativität monopolistische Anreize zu schaffen, wären solche Ansätze wohl vorzuziehen. Da für Programmierung, Vermarktung und Investitionen ohnehin schon Belohnungen winken, die mit Inhalte generierender KI-Software verbunden sind, sollte der Anreiz in diesem Fall an die Nutzerinnen und Nutzer fallen, sofern sie in größerem Umfang Arbeitsaufwand und Anstrengungen zu dem KI-generierten Output beigesteuert haben.

Doch sind Anreize wirklich nötig? Die Notwendigkeit solcher Anreize sollte empirisch nachgewiesen werden, zusammen mit ihren positiven Auswirkungen auf das kreative Ökosystem insgesamt. Tatsächlich ist historisch gut belegt, dass Eigentumsrechte nicht der einzige Anreiz für Kreativität sind.⁵⁷² Verschiedene Erkenntnisse aus Forschung und Marktgeschehen zeigen, dass ein offener und freier Zugang zu kreativen Werken oder alternative Geschäftsmodelle einen stärkeren Anreiz für KI-generierte Kreativität bieten können als auf geistigem Eigentum basierende Schutzmodelle,⁵⁷³ ohne dass es zu den negativen Effekten von Propertisierung und ausschließlichen Rechten kommt. Die kreativen Kapazitäten der KI könnten mit einzigartiger Geschwindigkeit zunehmen und den Kulturmarkt mit einem unüberschaubaren Geflecht von Rechten überziehen, die es zu klären gilt, sodass durch „Copyright-Trolling“ eine Eskalation droht, die geradewegs in ein

⁵⁷¹ Ramalho A., *op.cit.*

⁵⁷² Z. B. Frosio G., *op.cit.*

⁵⁷³ Z. B. Bonadio E. und McDonagh L., *op.cit.*, 122-123. Svedman M., *op.cit.*, S. 13-14.

absolutes Computerchaos führt. Schon jetzt ist die Urheberrechtssuppe ja allzu dick. Endlos tippende KI-Affen⁵⁷⁴ könnten sie sich einverleiben, bis nichts Verwendbares mehr bleibt.

⁵⁷⁴ Borel É., „Mécanique Statistique et Irréversibilité“, J. Phys. (Paris) 5(3), S. 189–196.

Werbung

Befürchtungen in Bezug auf den KI-Einsatz werden nicht zuletzt durch Science-Fiction-Literatur und -Filme genährt. Allerdings nimmt Science-Fiction die Zukunft oft nur vorweg. Bestes Beispiel hierfür ist Steven Spielbergs „Minority Report“.⁵⁷⁵ In einer Szene geht ein Mann durch ein Einkaufszentrum. Dabei werden seine Augen laufend von Kameras mit Augenerkennungssoftware erfasst. Sofort beginnt auf Bildschirmen in den Schaufenstern schrille Werbung zu laufen, die sich konkret an ihn richtet. Ist das wirklich noch Science-Fiction? Im Grunde ist es gar nicht so weit von dem entfernt, was wir schon heute in der Realität erleben. In diesem Zeitalter mit Internet, vernetzten Fernsehern und Second Screens haben die legalen und illegalen Möglichkeiten, an die persönlichen Daten von Mediennutzerinnen und -nutzern zu gelangen, exponentiell zugenommen. Solche Daten sind für Werbetreibende ein überaus wichtiges Gut, denn sie lassen sich nutzen, um auf Onlinediensten und vernetzten Geräten aller Art individualisierte Werbung auszuspielen. Darüber hinaus lässt sich bei einem Onlinedienst mithilfe persönlicher Daten, die über Suchmaschinen, soziale Medien und vernetzte Geräte gewonnen werden, das Nutzungserlebnis verbessern. Justina Raižytė erinnert in ihrem Beitrag zu dieser Publikation daran, dass „der Aufbau nachhaltiger KI-Rahmenbedingungen für Werbung und die Nutzung von Daten und Technologie für gute Zwecke im Interesse der Werbetreibenden selbst liegt“, da sich nur so das Vertrauen der Konsumenten gewinnen lässt – „das in der Werbung ein wahrer Goldstandard ist und immer sein wird“.

⁵⁷⁵ <https://www.imdb.com/title/tt0181689/>.

6. KI in der Werbung: Tor nach *Deadwood* oder sinnvolle Datennutzung?

*Justina Raižytė, European Advertising Standards Alliance*⁵⁷⁶

6.1. Einleitung

Künstliche Intelligenz (KI) hat vor einiger Zeit den direkten Sprung von futuristischen Drehbüchern und fiktionalen Erzählungen in unsere Wohnzimmer, Autos, Taschen und Einkaufsbeutel vollzogen. Es dauerte nicht lange, bis die stets kreativitätsgetriebene und innovationshungrige Werbung das Potenzial automatisierter Technologien erkannte und KI als ihre rechte Hand einsetzte, die Verbraucherinnen und Verbraucher auf ihrem Weg leiten und sie zu Käufern und treuen Kunden machen kann.

Technologien des maschinellen Lernens ermöglichen es Werbetreibenden, Daten zu den Interessen, Vorlieben, Standorten und der Demografie der Menschen schneller zu analysieren und sogar ihre Wünsche vorherzusagen, maßgeschneiderte Zielgruppenpläne zu erstellen und mehr relevante Werbungen zu liefern. Dabei werden die Werbekosten reduziert, während Reaktions- und Effektivitätsraten steigen. Mit personalisierter Werbung, die weniger unangenehme oder lästige Werbebotschaften übermittelt, sollen Verbraucher zudem zufriedener und sogar unterhalten werden.

Allerdings hinterlassen massive Datenströme, die es Vermarktern ermöglichen, Zielgruppen zu segmentieren und Verbraucher gezielt anzusprechen, bei vielen Menschen ein ungutes Gefühl. Angst und das Gefühl, beim Browsen fortwährend verfolgt, durch elektronische Geräte in den Taschen abgehört und durch Angebote mit unlauteren Preisangaben gezielt ins Visier genommen zu werden, lassen Befürchtungen aufkommen, dass die Privatsphäre und Entscheidungsfreiheit den Algorithmen zum Opfer fallen. Aus diesem Grund wird das heutige, technologiegetriebene Werbe-Ökosystem häufig in Farbtönen der klassischen Western dargestellt: mit dunklen Schatten der Gesetzlosigkeit und verblassten Ethikgrundsätzen.

Womit wir bei den Toren von *Deadwood* wären: einer der Heimatstädte der Siedler, die vom „Goldrausch“ gepackt wurden und im Streben nach Reichtum über die Weltmeere jagten. Die Legenden um die Stadt wurden kürzlich mit Filmzauber und tadelloser

⁵⁷⁶ Disclaimer: Die in diesem Artikel geäußerten Ansichten sind die der Autorin und entsprechen nicht notwendigerweise der offiziellen Politik oder Position von sonstigen im Text genannten Organisationen, Unternehmen oder Einzelpersonen.

Erzählkunst, durch die das Wesen des chaotischen Geists des Wilden Westens anschaulich dargestellt wird, auf unseren Bildschirmen zu neuem Leben erweckt, und bieten damit ein interessantes Sinnbild für die Debatte über technologische und Marktentwicklungen, die wir heute führen.

Wo finden wir heutzutage KI in der Werbung? Welche ethischen Grundsätze gelten bei der Nutzung personenbezogener Daten und beim Einsatz automatisierter Technologien zu Marketingzwecken? Worin bestehen Risiken und Chancen der Nutzung der Macht der Algorithmen bei der Erstellung, Schaltung und Überwachung von Werbungen? Diese und andere Fragen werden in diesem Artikel aus der Perspektive bestehender Marktpraktiken und politischer Rahmenbedingungen untersucht, sowie mit Blick auf gezielte, qualitative Interviews mit Expertinnen und Experten ausgelotet, die in den Bereichen KI und Werbung arbeiten (darunter Fachkräfte aus den Werbeagenturen und Technologieunternehmen, unabhängige Analysten sowie Experten für Datenschutzrecht und Werbeselbstkontrolle).

Im ersten Teil werden die wichtigsten Anwendungen von KI in der Werbung beleuchtet und es wird erörtert, wie verschiedene Instrumente des maschinellen Lernens dazu beitragen, kommerzielle Kommunikation besser auf die Zielgruppe abzustimmen, zu optimieren und sogar zu gestalten. Im zweiten Kapitel werden die wesentlichen Bedenken und Probleme, die mit Anwendungen des maschinellen Lernens im Marketing verbunden sind, dargestellt und die Schutzmechanismen ermittelt, die vorhanden sind (oder sein sollten), um sie zu entschärfen. Im dritten Teil wird eruiert, wie Technologien verwendet werden, um die Macht der Daten für sinnvolle Zwecke zu nutzen, etwa den Schutz von Verbrauchern vor schädlicher Werbung, und wie KI möglicherweise in Zukunft den Bereich der Werberegulierung prägen wird.

Der Artikel schließt mit der Beziehung zwischen KI und Werbung, die aus dem Blickwinkel der zuvor eingeführten Wild-West-Metapher betrachtet wird. Im letzten Teil werden die Argumente zusammengefasst, indem Ähnlichkeiten und Unterschiede zwischen der historischen Goldrauschzeit und dem von der Autorin vorgebrachten Datenrauschphänomen diskutiert werden, und es wird eine abschließende Frage gestellt: Sind wir im Begriff, in ein virtuelles *Deadwood* einzutreten?

6.2. KI in der Werbung: Von der Nachverfolgung digitaler Fußabdrücke zur Erstellung von Werbeskripts

„There’s no data like more data“ – dieses bekannte, Robert Mercer zugeschriebene Zitat⁵⁷⁷ wird häufig verwendet, um die wesentliche Rolle von Daten bei der Entwicklung und Anwendung von KI-Algorithmen in einer Vielzahl von Branchen und Bereichen menschlicher Tätigkeit hervorzuheben. Der Einsatz von KI in Werbe-Ökosystemen ist keine Ausnahmeerscheinung. Von der Verarbeitung personenbezogener Daten bei der maßgeschneiderten Schaltung von Werbungen auf der Grundlage von Nutzervorlieben und des Browser-Verlaufs innerhalb von Sekundenbruchteilen bis hin zur Entwicklung von

⁵⁷⁷ Lee, K.-F. (2018), *AI Superpowers: China, Silicon Valley, and the New World Order*, Houghton Mifflin Harcourt: Boston.

Algorithmen, die in riesigen Mengen historischer und kontextueller Daten Muster erkennen (und letztlich neue Inhalte erzeugen), werden KI und Technologien des maschinellen Lernens in der Werbung in breitem Umfang und für vielfältige Zwecke angewendet. Im ersten Kapitel wird der Haupteinsatzbereich von KI im heutigen Werbe-Ökosystem untersucht und erörtert, wie derart hochentwickelte Algorithmen die Werbeindustrie verändert haben.

6.2.1. Programmatische Werbung: die Werbe- und Datenbörse

Laut dem *Interactive Advertising Bureau* (Wirtschaftsverband der Online-Werbebranche – IAB) ist der Großteil des Wachstums von KI im Marketing heutzutage der programmatischen Werbung zuzuschreiben. Dabei handelt es sich um den Einsatz von KI und anderen Technologien des maschinellen Lernens für den Kauf und die Optimierung von Werbung in Echtzeit mit dem Ziel, sowohl für den Werbetreibenden als auch den Publisher die Effizienz und Transparenz zu erhöhen.⁵⁷⁸ In der programmatischen Werbung wird KI im Bereich der Werbeplanung und -platzierung, der automatischen Anpassung, der Optimierung und der Kampagnenkennzahlen angewendet.⁵⁷⁹ Die Schlüsselemente der programmatischen Werbung sind nachstehend im „Glossar zur programmatischen Werbung“ in Tabelle 1 zusammengefasst.

Tabelle 1 **Glossar zur programmatischen Werbung**

Real-Time-Bidding (RTB)	
■	Real-Time-Bidding ist eine Art des Kaufens und Verkaufens von Online-Werbemitteln und erfolgt fallweise durch Echtzeitauktionen, die in der Zeit ablaufen, die eine Website für das Laden im Browser eines Nutzers benötigt (d. h. rund 100 Millisekunden). In dieser Zeit fließt die Information über die Seite, auf der eine Werbung platziert wird, und den Nutzer, der sie lädt, durch ein Ad Exchange, welches die Werbung an den Werbetreibenden versteigert, der gewillt ist, den höchsten Preis dafür zu bezahlen, und die Werbung des erfolgreichen Bieters wird dann in die Website geladen. ⁵⁸⁰ RTB ermöglicht Werbetreibenden eine bessere und schnellere Zielgruppenansprache durch die Freischaltung von Werbungen, welche beispielsweise Werbungen von Marke X nur Nutzern zeigen, die zuvor die Website von Marke X besucht haben, jedoch keinen Kauf getätigt haben. Werbetreibende bieten auf der Grundlage ihrer Interessen und dessen, inwieweit die Daten über den Nutzer ihren Targeting-Parametern entsprechen: je höher die Übereinstimmung, desto höher der Preis. ⁵⁸¹

⁵⁷⁸ Rask, O. (2020), *What is Programmatic Advertising? The Ultimate 2020 Guide*, Match2One, <https://www.match2one.com/blog/what-is-programmatic-advertising/>.

⁵⁷⁹ IAB Data Center of Excellence (2019), *Artificial Intelligence in Marketing Report*, <https://www.iab.com/insights/iab-artificial-intelligence-in-marketing/>.

⁵⁸⁰ Marshall, J. (2014), „WTF is real-time bidding?“, in: *Digiday*, <https://digiday.com/media/what-is-real-time-bidding/>.

⁵⁸¹ SMAATO, *Real-Time Bidding (RTB): The Complete Guide*, <https://www.smaato.com/real-time-bidding/>.

Ad Exchange

- Ein Ad Exchange ist im Wesentlichen eine Börse für programmatische Werbung und der Ort, an dem Werbeinventar (d. h. Werbefläche) versteigert wird. Es ist eine virtuelle Plattform, auf der Publisher Werbetreibende treffen, einander zugeordnet werden und einen Preis für die Werbeeinblendungen vereinbaren. All dies geschieht dank automatisierter Technologie innerhalb von Millisekunden. Im Gegensatz zu Werbenetzwerken, bei denen ursprünglich vorgefertigtes Werbeinventar im Vordergrund stand, steht bei einem Ad Exchange die Reichweitenmessung im Mittelpunkt.⁵⁸² Sie bildet das Herzstück des programmatischen Ökosystems und ist sowohl auf der Seite des Werbetreibenden als auch der des Publishers in die jeweiligen Plattformen integriert.

Demand-Side-Plattform (DSP)

- Eine Demand-Side-Plattform ist ein Tool oder ein Softwareprogramm, das Werbetreibenden den automatischen Kauf von Werbeschaltungen ermöglicht. Werbetreibende melden sich bei einer DSP an, die wiederum mit einem Ad Exchange verbunden ist.⁵⁸³ Wenn ein Nutzer eine Website aufruft, die mit dem Ad Exchange verbunden ist, wird eine Versteigerungsmeldung an die Börse gesendet. Die Börse fragt dann die DSP, ob der Werbetreibende irgendwelche Werbungen hat, die zu der Platzierung passen könnten, und falls dem so ist, wird das Gebot für eine Werbefläche in Echtzeit an die Auktionsbörse zurückgesendet. Die Werbung eines erfolgreichen Bieters wird dem Website-Besucher angezeigt.⁵⁸⁴

Supply-Side-Plattform (SSP)

- Auf der anderen Seite nutzen Publisher eine Supply-Side-Plattform, um ihre Werbefläche zu verwalten. Die SSP verbindet sich mit einem Ad Exchange und meldet, welche Art von Inventar verfügbar ist. Durch RTB wird dieses Inventar automatisch an den Meistbietenden versteigert. Während der Zweck der DSP darin besteht, Fläche für programmatische Werbung von Publishern mit dem erwünschten Inventar so preiswert wie möglich zu erwerben, hat die SSP die gegenteilige Funktion: Werbefläche für den höchstmöglichen Preis zu verkaufen, indem sie sich mit mehreren verschiedenen Ad Exchanges verbindet, um den Publisher möglichst vielen potenziellen Käufern zu präsentieren.⁵⁸⁵

Obgleich programmatische Werbung meist im Zusammenhang mit digitalen Online-Kanälen erwähnt wird, darunter Display-, Handy-, Video- und Social-Media-Werbung, hat der Kauf programmatischer Werbung auch Einzug in die traditionellen Medien gehalten, darunter das Fernsehen⁵⁸⁶, und sogar in digitale Out-of-home-Plakatwände und

⁵⁸² IAB Europe (2017), „The advent of RTB“, <https://iabeurope.eu/blog/laypersons-programmatic/>.

⁵⁸³ Rask, O., op. cit.

⁵⁸⁴ Wang, J., Zhang, W. & Yuan, S. (2017), *Display Advertising with Real-Time Bidding (RTB) and Behavioural Targeting*, ArXiv, <https://arxiv.org/abs/1610.03013>.

⁵⁸⁵ Rask, O., op. cit.

⁵⁸⁶ Obgleich erweiterte, über umfangreiche demografiebasierte Reichweitenmessungen hinausgehende Fragestellungen weiterhin zu den zentralen Herausforderungen bei der Umwandlung der Werbekäufe des traditionellen, linearen Fernsehens in programmatische zählen, beginnen dessen ungeachtet einige Sendergruppen damit, auf Ad Impressions (Sichtkontakt der Zielgruppe) basierende Fernsehwerbungen zu kaufen.

Beschilderungen⁵⁸⁷, bei denen leistungsfähige Algorithmen gepaart mit mobilen Standortdaten und optischen Sensoren Werbeschaltungen auf digitalen Bildflächen an Bushaltestellen ermöglichen, die nunmehr konkrete Ziele anvisieren können – beispielsweise junge Pendler in der morgendlichen Hauptverkehrszeit.

6.2.1.1. Die Rolle von KI in der programmatischen Werbung

Alle in die Wertschöpfungskette der programmatischen Werbung eingebundenen Schlüsselemente unterstreichen die wesentliche Rolle von Anwendungen des maschinellen Lernens in diesem Prozess, die es ermöglichen, Entscheidungen zur Auswahl und Platzierung von Werbungen innerhalb eines Augenblicks zu treffen, und dadurch KI untrennbar mit dem Begriff der programmatischen Werbung selbst verbinden.

Durch diese hochentwickelten, automatisierten Technologien können alle Marktakteure zahlreiche Vorteile genießen, die auf Schnelligkeit, Effizienz und prädiktiver Analyse beruhen. Insgesamt ist dieser breite Nutzen in den folgenden zentralen Zielen der KI in der heutigen programmatischen Werbung zusammengefasst:⁵⁸⁸

- Personalisierung von Werbungen, wobei durch KI und maschinelles Lernen Echtzeit-Verhaltensdaten von Verbrauchern erfasst und auf der Grundlage von Merkmalen wie Alter, Geschlecht, Standort und Millionen anderer Datenpunkte maßgeschneiderte und relevante Werbungen geliefert werden. Der Prozess wird durch leistungsstarke prädiktive Algorithmen angetrieben, mit deren Hilfe ermittelt wird, welche Verbraucher sich voraussichtlich auf die Werbungen der Werbetreibenden einlassen.
- Audience Targeting (Zielgruppenansprache), wobei Vermarkter durch den wirksamen Einsatz von KI Inhalte im Internet durchsuchen und entscheiden können, welche Werbungen für bestimmte Zielgruppen oder Kanäle am besten geeignet sind. Durch die Verwendung von Bilderkennung können diese Systeme zudem die Platzierung von Werbungen unterstützen, die mit den Bildern auf der Seite eines Artikels oder einer Website korrelieren.⁵⁸⁹ Darüber hinaus ist KI bei ordnungsgemäßer Einrichtung in der Lage, die Zielgruppe auf der Grundlage des tatsächlichen Verhaltens fortwährend weiterzuentwickeln und auf andere Segmente auszudehnen, welche möglicherweise dasselbe Kaufverhalten teilen.⁵⁹⁰
- Leistungsoptimierung, wobei Algorithmen des maschinellen Lernens automatisch analysieren können, was Werbungen auf bestimmten Plattformen leisten, und

einzubeziehungen, wobei einige Unternehmen daran arbeiten, die erweiterte Segmentierung der Fernsehzuschauer zu standardisieren und mehr Inventar für Addressable TV (gezieltes Aussteuern von Fernsehwerbung auf SmartTVs) zu erschließen. Siehe Blustein, A. (2020), „The programmatic TV dream is edging closer to reality“, The Drum, Carnyx Group Ltd, <https://www.thedrum.com/news/2020/02/19/the-programmatic-tv-dream-edging-closer-reality>.

⁵⁸⁷ Côté, R. (2020), „What is programmatic DOOH?“, Broadsign, <https://broadsign.com/blog/what-is-programmatic-digital-out-of-home/>.

⁵⁸⁸ IAB (2019), op. cit.

⁵⁸⁹ Schmelzer, R. (2020), „AI Makes A Splash In Advertising“, Forbes, <https://www.forbes.com/sites/cognitiveworld/2020/06/18/ai-makes-a-splash-in-advertising/#24c0287c7682>.

⁵⁹⁰ Rowan, M., „The Impact of Artificial Intelligence in Advertising“, AW360, <https://www.advertisingweek360.com/the-impact-of-artificial-intelligence-in-advertising/>.

Empfehlungen geben. Diese KI-Systeme sind außerdem in der Lage, nicht nur die Kennzahlen der Kampagnen des Werbetreibenden zu verfolgen, sondern auch jene der Mitbewerber des Werbetreibenden. Sie verfügen über ein integriertes „Situationsbewusstsein“, welches es Algorithmen des maschinellen Lernens ermöglicht, sich rasch anzupassen, Werbeausgaben auf alternative Kanäle umzuschichten und Werbebotschaften zu verändern, damit diese den Marktstrukturen und dem Verbraucherverhalten entsprechen.⁵⁹¹

- Media-Mix-Modelling, wobei KI eingesetzt wird, um Verbraucher zu ermitteln, die für die Kampagnen auf verschiedenen Medienkanälen am empfänglichsten sind, und dadurch das Return-on-Investment (ROI) für digitale Werbung zu erhöhen. KI kann fortwährend Empfehlungen zur Weiterentwicklung des Medienmix erteilen; Marken und Agenturen können folglich die Zuordnung in Bezug auf den Marketing-Mix vollständig automatisieren – und wertvolle Zeit und Geld sparen.

Zusammengefasst lässt sich festhalten, dass KI und maschinelles Lernen es den Akteuren im Werbe-Ökosystem ermöglichen, riesige Datenmengen in Echtzeit zu analysieren, Botschaften mittels KI-gestützter Hyperpersonalisierung zuzuschneiden und die besten Zeiten und Kanäle für die Kommunikation zu ermitteln. Daher ist es keine Überraschung, dass über 80 % der befragten Führungskräfte der Werbebranche und frühen Anwender von KI über ein positives ROI bei ihren KI-Initiativen berichteten sowie ein ausgeprägtes Interesse, derartige Investitionen künftig zu erhöhen.⁵⁹²

6.2.2. Algorithmische Kreativität: Wenn KI die Fantasie spielen lässt

„Kreativ sein ist mitunter ungefähr so, als wenn ein Damoklesschwert über dem Kopf schwebt“ – diesen Worten von scheinbar großer Weisheit würden viele Werbedrehbuchschreiber, die mit näher rückenden Deadlines, zu vielen Ideen und zu wenig Zeit kämpfen, wohl zustimmen. Aber der Satz wurde nicht von einer Werbefachperson geäußert, sondern in der Tat von InspiroBot erzeugt – einem KI-Softwareprogramm, „das dazu bestimmt ist, unbegrenzte Mengen einzigartiger inspirierender Zitate zu generieren“ und dem beigebracht wurde, auf der Grundlage unzähliger inspirierender Zeilen aus dem Internet beliebige Sätze hervorzubringen.

Obgleich es sich bei dem durch KI erzeugten Beispiel mit dem „Damoklesschwert“ um ein im Rahmen der Abfassung dieses Artikels zum Vergnügen kreierte handelt, ist die tatsächliche Anwendung von KI bei der Erstellung von Werbung viel weiter fortgeschritten. KI-getriebene Systeme können Werbungen bereits teilweise oder vollständig entwickeln, indem sie Computerlinguistik (CL) und Textgenerierung (TG) wirksam einsetzen, zwei KI-

⁵⁹¹ Schmelzer, R., op. cit.

⁵⁹² Deloitte AI Institute, „State of AI in the Enterprise“, 2nd edition, <https://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/state-of-ai-and-intelligent-automation-in-business-survey.html>.

getriebene Technologien, die zum Verfassen von Werbetexten verwendet werden.⁵⁹³ Darüber hinaus verspricht Computer Vision, gepaart mit Bilderkennungstechnologien, eine Zukunft, in der Nutzer jegliche Videoinhalte oder Bilder „shazamen“⁵⁹⁴ können und sofort auf vorhandene Werbeaktionen abstimmen können.

Einige technologiegetriebene Unternehmen stellen bereits KI-basierte Lösungen bereit, um zu Marketingzwecken Bilder zu generieren. Diese Unternehmen stützen sich in der Regel auf Technologien des maschinellen Lernens, die in die Kategorie der Generative Adversarial Networks (GANs) fallen, und sind in der Lage, „künstliche Models“ zu erzeugen, die immer schwerer von echten Menschen zu unterscheiden sind,⁵⁹⁵ und für die Werbekampagnen eines Vermarkters Gesichter zu erschaffen, die leicht an verschiedene Zielgruppen und Demografien angepasst werden können.

Auch wenn Mausclicks die entscheidende Triebkraft für das digitale Marketing sind, führt KI möglicherweise zudem einen neuen Werbetrend herbei: die Ära der Sprache.⁵⁹⁶ Laut den Daten von 2019 waren 89 % der befragten Marketingexperten überzeugt, dass Sprachassistenten in den kommenden drei bis fünf Jahren ein bedeutender Marketingkanal sein werden, wobei mehr als ein Drittel der Befragten sie als einen „äußerst wichtigen Kanal“ bezeichnete.⁵⁹⁷ Alexa, Siri, Google Home und andere Geräte und Sprachassistenten machten interaktive Sprachschnittstellen populär und bahnten den Weg für dialogorientierte Sprachwerbung. Sie sind KI-getrieben, nutzen Spracherkennung und Sprachverstehen und bieten neue Werbeformate für Vermarkter, Publisher und Verbraucher.

Laut Charlie Cadbury, dem CEO von Say It Now⁵⁹⁸ zählen folgende fünf Entwicklungen im Bereich Sprache zu den interessantesten: verschiedene Dienstleister, die Sprachassistenten nutzen (z. B. Steuerung von Lebensmittelzustellungen); verbessertes Auffinden von Drittanbieterfähigkeiten und Fähigkeitsaufruf ohne Namensnennung bei Sprachassistenten; Voice Commerce; digitale, interaktive Audiobetreuung; und standortabhängige Sprachdienste, die in Fahrzeugen oder über In-Ohr-Kopfhörer erbracht werden.

Auf die Frage nach einer Zukunftsvision für die Sprachanwendungen in der Werbung stellte Cadbury fest, dass „der Schlüssel im ‚Sprachassistenten‘ liegt“, wobei er Kunden

⁵⁹³ Kaput, M. (2020), „AI for Advertising: Everything You Need to Know“, Marketing AI Institute, <https://www.marketingaiinstitute.com/blog/ai-in-advertising>.

⁵⁹⁴ Basiert auf Shazam Entertainment Ltd – deren mobile Anwendung Musik, Filme, Werbung und Fernsehshows anhand eines kurzen, abgespielten Hörbeispiels erkennen kann; von anderen Arten von auf Bilderkennung basierenden Apps wurden in der Branche Prototypen entwickelt, wie etwa Smartify) – Shazam für Kunst und Museen, <https://smartify.org/about-us>.

⁵⁹⁵ Gonfalonieri, A., „Integration of Generative Adversarial Networks in Business Models“, Towards Data Science, Medium, <https://towardsdatascience.com/integration-of-generative-adversarial-networks-in-business-models-47e60263aec4>.

⁵⁹⁶ Tuschinskij, S. (2019), „Voice is the New Click: Why Voice Commands Will Replace the Click as the Standard Measure for Brands“, Instreamatic, Medium, <https://medium.com/@instreamatic/voice-is-the-new-click-why-voice-commands-will-replace-the-click-as-the-standard-measure-for-61225e4e7caa>.

⁵⁹⁷ Kinsella, B. und Mutchler, A. (2019), *The State of Voice Assistants as a Marketing Channel Report*, Voicebot.ai, <https://voicebot.ai/the-state-of-voice-assistants-as-a-marketing-channel-report/>.

⁵⁹⁸ Charlie Cadbury von Say It Now wurde im Rahmen der Abfassung dieses Artikels am 19. Juni 2020 interviewt, <https://www.sayitnow.ai/>.

sehen möchte, die ihre digitalen Assistenten wie eine wirkliche Hilfe in ihrem Leben behandeln, ihrem Rat vertrauen und sinnvoll Aufgaben delegieren.

6.2.3. Von kreativen Spielen zu Gewinnen

Einige der weltweit führenden Werbetreibenden haben bereits in KI-gestützte kreative Prozesse investiert, die bemerkenswerte Ergebnisse und Werbekampagnen liefern, welche die Aufmerksamkeit der Verbraucher auf sich gezogen und für Schlagzeilen in den Medien gesorgt haben. Die in Tabelle 2 aufgeführten Kampagnen und Partnerschaften von Markenagenturen stellen eine Auswahl von Fällen dar, in denen KI-Tools zu kreativen Marketingzwecken eingesetzt wurden und dies zur Erzeugung von zumindest zum Teil durch automatisierte Technologien entwickelten Werbeinhalten führte. Kerry Richardson, einer der Mitgründer von *Tiny Giant*⁵⁹⁹ ist überzeugt, dass Verbraucher bald mehr kreative Anwendungen für KI sehen werden sowie mehr Fälle der Zusammenarbeit zwischen diesen und Menschen bei der Gestaltung von Kampagnen, in denen Kreative einer Agentur eine Idee haben oder sich ein Konzept ausdenken und dann KI-Technologie anwenden, um die Ausführung zu unterstützen.

Tabelle 2 Auf kreative KI-Technologien gestützte Werbe- und Marketingkampagnen

<ul style="list-style-type: none"> ■ Lexus in Zusammenarbeit mit IBM Watson, The&Partnership London und Visual Voice⁶⁰⁰
<ul style="list-style-type: none"> ■ Dieser besondere Lexus-Werbespot ist der erste mit vollständig von KI geschriebenem Drehbuch. Anhand von zahlreichen Datenpunkten als Input – darunter die Ausbeute der mit dem Löwen von Cannes ausgezeichneten Werbungen aus 15 Jahren und der besten Werbungen im Luxussektor aus 10 Jahren sowie Punkte zu „Intuition“ und der Art, wie Menschen Entscheidungen treffen – ermittelte die IBM-Watson-Technologie Elemente, die preiswürdigen Werbespots gemeinsam sind.⁶⁰¹ Ein künstliches Intelligenzsystem gestaltete dann den Skriptablauf und -entwurf. Die Entwickler selbst äußerten sich überrascht darüber, dass sie anstatt des erwarteten „verrückten und seltsamen“ Ergebnisses eine Geschichte mit einer Fußnote zu jeder einzelnen Zeile mit einem Datenpunkt erhielten, die erklärte, warum diese Entscheidung getroffen wurde.⁶⁰² Das Ergebnis war eine eher verblüffende Geschichte eines Lexus-Ingenieurs, der letzte Hand an ein neues Modell anlegt, eines Autos mit Bewusstsein und eines im Fernsehen

⁵⁹⁹ Kerry Richardson von Tiny Giant wurde im Rahmen der Abfassung dieses Artikels am 22. Juni 2020 interviewt, <https://www.tinygiant.io/>.

⁶⁰⁰ Die Werbekampagne kann auf Variety.com eingesehen werden, <https://variety.com/2018/digital/news/lexus-ai-scripted-ad-ibm-watson-kevin-macdonald-1203030693/>.

⁶⁰¹ Spangler, T., „First AI-Scripted Commercial Debuts, Directed by Kevin Macdonald for Lexus“, *Variety*, <https://variety.com/2018/digital/news/lexus-ai-scripted-ad-ibm-watson-kevin-macdonald-1203030693/>.

⁶⁰² Faull, J. (2018), *Lexus reveals ad 'created by AI'. Is it a gimmick? No. Will it win any awards? Probably not*, The Drum, Carnyx Group Ltd, <https://www.thedrum.com/news/2018/11/16/lexus-reveals-ad-created-ai-it-gimmick-no-will-it-win-any-awards-probably-not>.



	übertragenen Crashtests, doch es bewies, dass maschinengeschriebene kreative Texte mehr als eine entfernte Möglichkeit sind.
■	Malaria No More in Zusammenarbeit mit RG/A, Ridley Scott Associates und Synthesia⁶⁰³
■	In der Kampagne „Malaria Must Die“ leiht David Beckham seine Stimme dem Kampf gegen Malaria zugunsten der gemeinnützigen Organisation Malaria No More in ihrem Kampf zur Beseitigung der von Moskitos übertragenen Krankheit. Die neuartige Kampagne zeigt Beckham, wie er scheinbar fließend in neun Sprachen spricht, während er die Zuhörer dazu auffordert, sich zu engagieren. ⁶⁰⁴ Wenn er zwischen den verschiedenen Sprachen wechselt, sind Beckhams unterschiedliche Stimmen tatsächlich jene von Malaria-Überlebenden, deren Merkmale mithilfe KI-getriebener Videosynthesetechnologie digital auf jene des berühmten Sportlers übertragen wurden. ⁶⁰⁵
■	Deutsche Bahn in Zusammenarbeit mit Ogilvy, Frankfurt, Deutschland, Getty Images und Spirable⁶⁰⁶
■	Die Deutsche Bahn startete mithilfe von Fotos deutscher Orte, die weltweiten touristischen Zielen ähneln, eine Kampagne zur Förderung von Inlandsreisen. Die Kampagne „Spar Dir den Flug“ lädt Deutsche dazu ein, die Vorteile günstigerer Zugreisen zu genießen. Sie verwendete KI, um deutsche Orte zu ermitteln, die weltweiten symbolträchtigen Wahrzeichen ähnlich sind. Mit dynamischen Werbevideos, in denen in Echtzeit die mit internationalen Wahrzeichen und ihren deutschen Entsprechungen verbundenen Reisekosten gegenübergestellt wurden, sprach sie dann anhand von Facebook-Daten gezielt Reisebegeisterte und lokale Influencer an. ⁶⁰⁷
■	Cheltenham Science Festival in Zusammenarbeit mit Tiny Giant⁶⁰⁸
■	„AIDA: AI Science Festival Curator“, eine KI-generierte Festival-Kuratorin für das Cheltenham Science Festival verwendete Festivalvorträge aus 10 Jahren als Datenmaterial, wurde als rekurrentes neuronales Netzwerk trainiert und erzeugte rund 800 neue, mögliche Vorträge für das Festival. ⁶⁰⁹ AIDAS Vorschläge wurden der Welt auf Twitter zur Abstimmung vorgelegt, um einen Sieger zu küren. Der Vortrag über „Introvertierte Narwale“ wurde bei der Veranstaltung von AIDA gehalten, deren tatsächliche Lautsprecherstimme mittels Deep Learning geschaffen wurde, um eingegebenen Text in eine nuancierte, menschlich klingende Stimme zu verwandeln. AIDA erwies sich als großer Erfolg und trat später im BBC Radio auf.

⁶⁰³ Die Werbekampagne kann auf TheDrum.com eingesehen werden,

<https://www.thedrum.com/news/2019/04/09/david-beckham-lends-his-voice-malaria-ai-petition>.

⁶⁰⁴ Butcher, M., „The startup behind that deep-fake David Beckham video just raised \$3M“, TechCrunch, <https://techcrunch.com/2019/04/25/>.

⁶⁰⁵ Glenday, J. (2019), „David Beckham lends his voice to Malaria AI petition“, The Drum, Carnyx Group Ltd., <https://www.thedrum.com/news/2019/04/09/david-beckham-lends-his-voice-malaria-ai-petition>.

⁶⁰⁶ Die Werbekampagne kann auf wersm.com eingesehen werden, <https://wersm.com/how-deutsche-bahn-increased-sales-by-24-thanks-to-instagram/>.

⁶⁰⁷ Desreumaux, G. (2019), „How Deutsche Bahn Increased Sales By 24% Thanks To Instagram“, Wersm, <https://wersm.com/how-deutsche-bahn-increased-sales-by-24-thanks-to-instagram/>.

⁶⁰⁸ Die Werbekampagne kann auf TinyGiant.io eingesehen werden, „

⁶⁰⁹ Tiny Giant, „AI Festival curator Cheltenham Science Festival“, <https://www.tinygiant.io/case-study-one-aida>.



Sie wurde außerdem für die beste Anwendung von KI im Jahr 2019 von der Data & Marketing Association mit Gold ausgezeichnet.⁶¹⁰

■ **JPMorgan Chase, laufende Zusammenarbeit mit Persado**

- Im Jahr 2019 kündigte JPMorgan Chase ein auf fünf Jahre angelegtes, unternehmensweites Abkommen mit Persado an, einer der führenden Agenturen bei der Nutzung von KI zur Erzeugung leistungsfähiger Marketingwerbemittel. Das erfolgreiche Pilotprojekt belegte, dass KI-gestützte Marketingtexte hoch effektiv sind und für einen Anstieg auf bis zu 450 % bei den Klickraten der von Persado erstellten Werbungen gegenüber anderen mit einer Reichweite von 50 bis 200 % sorgte.⁶¹¹ Persados firmeneigene Technologien wurden verwendet, um auf der Grundlage einer hochentwickelten Wissensdatenbank zur Marketingsprache mit mehr als einer Million mit Tags und Scoring-Werten versehenen Wörtern und Sätzen Texte und Überschriften zu redigieren. Mithilfe des Tools formulierte JPMorgan Chase Marketingbotschaften im Karten- und Hypothekengeschäft neu, um die überzeugendste Botschaft für Einzelpersonen und bestimmte Kundenzielgruppen zu erzeugen.⁶¹²

6.2.3.1. „Virtualität“ im Influencer-Marketing: Eine Fallstudie

Lil Miquela ist eine Musikerin, deren Musik über Spotify gestreamt werden kann, eine Designerin, die ihre eigene Bekleidungsmarke besitzt, ein Model, das für Luxusmodemarken arbeitet, und ein Social-Media-Star mit mehr als 2,4 Millionen Followern auf Instagram.⁶¹³ Sie präsentiert sich selbst als „Musikerin, Change Seeker und Roboterin“ – eine computergenerierte (CGI) Figur und die erste weltbekannte virtuelle Influencerin.

Ogleich Lil Miquela keine wirkliche KI-Schöpfung ist, hat sie Unternehmen dazu angeregt, massiv in virtuelle Menschen zu investieren und Visionen im Hinblick auf künftige digitale Wesen zu entwickeln, die vollständig KI-getrieben sind und autonom auf Social-Media-Plattformen existieren.⁶¹⁴ Da die Herstellung von hochwertigen 3D-Modellen erschwinglicher wird, ersinnen einige Entwickler bereits virtuelle Menschen, die ohne jegliche menschliche Mitwirkung ihr eigenes „Leben“ führen – vom Posten von Bildern oder Videos und den sie begleitenden Bildtexten bis hin zur Interaktion mit ihren Followern.⁶¹⁵

⁶¹⁰ Data & Marketing Association (2019), „Gold Best Use of AI“, Data & Marketing Association Awards, <https://dma.org.uk/awards/winner/2019-gold-best-use-of-ai>.

⁶¹¹ Persado (2019), „JPMorgan Chase Announces Five-Year Deal with Persado For AI-Powered Marketing Capabilities“, Persado, <https://www.persado.com/press-releases/jpmorgan-chase-announces-five-year-deal-with-persado-for-ai-powered-marketing-capabilities/>.

⁶¹² Business Wire (2019), „JPMorgan Chase Announcement Concerning Preferred Stock“, <https://www.businesswire.com/news/home/20191031005537/en/JPMorgan-Chase-Announcement-Preferred-Stock>.

⁶¹³ Stand vom 1. Juli 2020.

⁶¹⁴ Alexander, J. (2019), „Virtual creators aren't AI – but AI is coming for them“, The Verge, Vox Media, <https://www.theverge.com/2019/1/30/18200509/ai-virtual-creators-lil-miquela-instagram-artificial-intelligence>.

⁶¹⁵ Bradley, S. (2020), „Even better than the real thing? Meet the virtual influencers taking over your feeds“, The Drum, Carnyx Group Ltd, <https://www.thedrum.com/news/2020/03/20/even-better-the-real-thing-meet-the-virtual-influencers-taking-over-your-feeds>.

Während die Idee virtueller Menschen, denen Maschinen zur Handlungsfähigkeit verhelfen, nicht neu ist, haben „die zunehmende Verbreitung von Smartphones und die Beliebtheit von Foto-Sharing-Sites wie Instagram unser Bewusstsein für diese virtuellen Menschen gestärkt und sie in eine einflussreiche Position gebracht“ sagt Scott Guthrie,⁶¹⁶ ein unabhängiger Berater für Influencer-Marketing und Analyst. Die Modeindustrie machte sich als erste die möglichen Vorteile virtueller Models zunutze: Künstliche Menschen benötigen keine Kosmetik, angepasste Konfektionsgrößen oder mehrere Aufnahmen beim Fotoshooting; sie stellen die Kleidungsstücke der auftraggebenden Marke stets bestmöglich zur Schau, erscheinen immer pünktlich und liefern beständig auftragsgemäße Inhalte.⁶¹⁷

6.2.3.1.1. Echte Gefahren in der künstlichen Realität

Trotz ihrer virtuellen Natur werfen digitale Wesen echte Bedenken auf, darunter vor allem: Transparenz. Rupa Shah, Gründerin von *Hashtag Ad Consulting*⁶¹⁸ stellt fest, dass der Einsatz hochentwickelter Technologien bedeutet, dass es zunehmend schwieriger wird, virtuelle Influencer alleine durch die visuelle Wahrnehmung zu erkennen. Durch die Beachtung von Details bei der Übertragung jedes Bildes können sie in jedem Kontext oder an jedem Schauplatz, an jedem Zielort auftreten, um die gewünschte Erzählung der Marke zu realisieren, und durch CL erscheint ihre Kommunikation natürlich und reaktiv.

Die von Organisationen für Werbestandards in ganz Europa durchgesetzten Kodizes zur Werbeselbstkontrolle verlangen bereits, dass jede kommerzielle Kommunikation durch entsprechende Angaben unmittelbar und eindeutig erkennbar sein muss.⁶¹⁹ Allerdings müssen diese Regeln wohl auf virtuelle Influencer ausgedehnt werden und zusätzlich verlangen, dass ihre Inhaber und Entwickler Verbraucher über ihre virtuelle Natur informieren – die der vollständigen Kontrolle einer Marke unterliegt.

Dudley Neville-Spencer, Direktor und Leiter der Datenanalyse bei der *Virtual Influencer Agency*⁶²⁰, ist ebenfalls der Meinung, dass virtuelle Influencer bereits bestehende Streitfragen im Influencer-Marketing verschärfen können, und tritt für ein spezielles Wasserzeichen für virtuelle Wesen ein.⁶²¹ Er arbeitet mit Shah zusammen, um einen Ethikkodex für virtuelle Influencer zu entwickeln, wobei der Schwerpunkt nicht nur auf der Transparenz liegt, sondern auch auf anderen Bereichen der sozialen Verantwortung, wie

⁶¹⁶ Scott Guthrie, ein Unternehmensberater für Influencer-Marketing, Veranstaltungsredner und Blogger, <https://sabguthrie.info/>. Im Rahmen der Abfassung dieses Artikels am 22. Juni 2020 interviewt.

⁶¹⁷ Guthrie, S., „Virtual Influencers: More Human Than Human“, Kap. 15 in: Yesiloglu, S. und Costello, J. (Hg.) (2020), *Influencer Marketing: Building Brand Communities and Engagement*, Routledge, London.

⁶¹⁸ Rupa Shah von Hashtag Ad Consulting, <https://www.hashtagad.co.uk/>. Im Rahmen der Abfassung dieses Artikels am 22. Juni 2020 interviewt.

⁶¹⁹ European Advertising Standards Alliance (2018), *Best Practice Recommendation on Influencer Marketing*, <https://www.easa-alliance.org/products-services/publications/easa-best-practice-guidance>.

⁶²⁰ Äußerungen von Dudley Neville-Spencer, Direktor und Leiter der Datenanalyse bei der Virtual Influencer Agency und Strategy & Insights Director bei Live & Breath, sind einem am 30. Juni 2020 von Persollos veranstalteten Online-live-Webinar entnommen, siehe Persollos, „Ethics, Influencers and Growth“, Persollos Webinar 3, 30. Juni 2020, <https://www.blog.persollos.com/post/persollos-webinar-3-ethics-influencers-and-growth>.

⁶²¹ Persollos, op. cit.

Körperbild und Vielfalt. Viele prominente Stimmen der Branche haben Bedenken geäußert, dass virtuelle Influencer im Zusammenhang mit schädlichen menschlichen Darstellungen und unrealistischen Schönheitsidealen zu Problemen im Hinblick auf das Selbstwertgefühl und die psychische Gesundheit führen könnten.⁶²² Shah fügt hinzu, dass Regulierungsorgane darauf vorbereitet sein sollten und gerüstet, um jegliche Bedenken, die Verbraucher haben könnten, anzugehen.

6.2.3.2. Die Zukunft virtueller Influencer

Insgesamt wächst der Influencer-Markt schnell und soll Prognosen zufolge 2020 auf USD 9,7 Milliarden wachsen.⁶²³ KI wird nicht nur genutzt, um virtuelle Influencer zu entwickeln, sondern sie wird auch in der Auswahl- und Bewertungsphase des Influencer-Marketings eingesetzt und unterstützt Marken dabei, die geeignetsten Social-Media-Influencer für mögliche Kampagnen zu ermitteln und sie zu überprüfen (z. B. im Hinblick auf falsche Follower-Zahlen).

Guthrie ist davon überzeugt, dass weitere technologische Fortschritte den virtuellen Influencern ermöglichen werden, die Fesseln vorgefertigter Animationspfade hinter sich zu lassen und frei zu interagieren sowie von jeder menschlichen Konversation zu lernen. Er sagt außerdem voraus, dass sich die Untergruppe der virtuellen Influencer in der Zukunft in mindestens drei kleinere Kategorien zersplittern wird: virtuelle Markenassistenten (virtuelle Menschen, die von einer Marke entwickelt und betrieben und vollständig von den Markenzielen angetrieben werden); Kundendienstmitarbeiter (virtuelle Wesen, die als Chatbots fungieren, jedoch mit menschenähnlichen Gefühlen und menschenähnlicher Körperform); und virtuelle Influencer (virtuelle menschliche Influencer, die sich entweder im Besitz einer auftraggebenden Marke befinden oder eigenständig agieren).

Schließlich werden die Zielgruppen im Zuge der weiteren Entwicklung der virtuellen Influencer entscheiden, wie schnell sie bereit sind, digitale Avatare in ihre Social-Media-Feeds aufzunehmen und sich auf sie einzulassen. Möglicherweise werden sich die Prognosen einiger Experten, dass es keine Marke ohne irgendeine Art virtueller Repräsentanz geben wird, in der nahen Zukunft bewahrheiten.⁶²⁴

6.2.4. Fazit: KI-gestützte intelligente Werbung

Intelligente Werbung gilt als die dritte Phase und eine neue Dimension der digitalen Werbung, die sich auf interaktives und programmatisches Marketing stützt.⁶²⁵ Sie baut auf der Interaktivität und Automatisierung der vorigen Phasen auf, fügt jedoch neue Attribute

⁶²² Bradley, S., op. cit.

⁶²³ Influencer Marketing Hub (2020), *The State of Influencer Marketing 2020: Benchmark Report*, <https://influencermarketinghub.com/influencer-marketing-benchmark-report-2020/>.

⁶²⁴ Persollo, op. cit.

⁶²⁵ Li, H., „Special Section Introduction: Artificial Intelligence and Advertising“, *Journal of Advertising*, Aug./Sep. 2019, Bd. 48, Ausgabe 4, S. 333-337.

hinzu, wie etwa die Personalisierung, die über die Interessen eines Nutzers hinausgeht, und bewegt sich in Richtung der Vorhersage ihrer Bedürfnisse in verschiedenen Kontexten und Zeitrahmen, wobei sie in Echtzeit und bedarfsgerecht hochgradig individualisierte kommerzielle Inhalte erstellt.⁶²⁶ Sie verbindet Technologien des programmatischen Kaufens und der programmatischen Kreativität und öffnet die Tür zu individuell zugeschnittenen Werbeerlebnissen, die versprechen, noch relevantere und nützlichere Inhalte anzubieten.

Wenn intelligente Werbung allerdings effektiv sein und das Vertrauen der Verbraucher gewinnen soll, kann sie nicht in einer undurchsichtigen Blackbox eingeschlossen sein, die Werbebotschaften erzeugt, die auf widerrechtlich erhobenen Nutzerdaten und unlauteren Algorithmen beruhen. Im nächsten Teil dieses Kapitels werden daher die Herausforderungen im Zusammenhang mit KI in der Werbung untersucht und die notwendigen Schutzmaßnahmen erörtert.

6.3. Bedenken im Hinblick auf Big Data und KI

Die meisten Anwendungen von KI in der Werbung beruhen heute auf Algorithmen und umfangreichen Datensätzen, die Informationen über die Merkmale und persönlichen Vorlieben der Nutzer enthalten und mit maßgeschneiderten, „handverlesenen“ kommerziellen Inhalten aufwarten.⁶²⁷ Natürlich bringt ein derartiger Optimierungsprozess ethische und rechtliche Herausforderungen mit sich, insbesondere jene in Bezug auf den Datenschutz und unausgewogene Algorithmen. Dieses Kapitel wird daher den Blick auf bestehende rechtliche Garantien richten, die Bürger und Verbraucher vor möglichen Schäden im Zusammenhang mit Fragen des Datenschutzes und der automatisierten Entscheidungsfindung schützen. Obgleich in diesem Teil der europäische Kontext im Mittelpunkt steht, sollte angemerkt werden, dass es auch in der restlichen Welt Maßnahmen zu Fragen der Privatsphäre und des Datenschutzes gibt, die Gegenstand einer ähnlichen Analyse sein könnten.⁶²⁸

Laut der Wissenschaft sollte eine gute Datenschutzgesetzgebung im Zeitalter der KI Verbraucher vor möglicher, auf KI beruhender Diskriminierung, fehlender Einwilligung und Datenmissbrauch schützen und mehrere Schlüsselkomponenten beinhalten: ein Transparenzgebot, sodass KI über ein fest verankertes Recht auf die Informationen verfügt, die sie erhebt; eine Opt-out-Möglichkeit für Verbraucher; eine eingebaute Begrenzung der

⁶²⁶ Chen, G., et. al, „Understanding Programmatic Creative: The Role of AI“, *Journal of Advertising*, Aug./Sep. 2019, Bd. 48, Ausgabe 4, S. 347-355.

⁶²⁷ Lee, K.-F., op. cit.

⁶²⁸ Beispielsweise der *California Consumer Privacy Act* (kalifornisches Verbraucherdatenschutzgesetz – CCPA), durch den im Zusammenhang mit der Nutzung und Löschung von sowie dem Zugang zu personenbezogenen Daten durch Unternehmen und mit dem Widerruf der Nutzungserlaubnis neue Rechte für Verbraucher eingeführt werden (siehe Paka, A. (2020), „How Does The CCPA Impact Your AI?“, Forbes Technology Council, Forbes, <https://www.forbes.com/sites/forbestechcouncil/2020/02/20/how-does-the-ccpa-impact-your-ai/#7d27ce7c43c7>), oder das Gesetz zum Schutz personenbezogener Daten in Japan, das 2020 mit der Einführung und Förderung pseudonymisierter Daten im Rahmen der Dateneinspeisung bei KI voraussichtlich weiter geändert wird, siehe Global Legal Insights, *AI, Machine Learning & Big Data Laws and Regulations Japan*, <https://www.globallegalinsights.com/practice-areas/ai-machine-learning-and-big-data-laws-and-regulations/japan>. Siehe auch Kapitel 2 dieser Publikation.

Datenerhebung durch und Zwecke von KI; sowie die Möglichkeit der Datenlöschung auf Verlangen des Verbrauchers.⁶²⁹ Zivilgesellschaftliche Gruppen fordern zudem „klare rote Linien für unzulässige Nutzung, demokratische Kontrolle und einen echten grundrechtsbasierten Ansatz bei der Regulierung von KI“, um ein vertrauenswürdiges KI-System zu schaffen.⁶³⁰ Obgleich der Themenbereich dieses Artikels auf den Einsatz automatisierter Technologien im Marketing begrenzt ist, wäre eine Einbeziehung des menschenrechtsbezogenen Ansatzes, insbesondere in Bezug auf das Recht auf Privatsphäre und Datenschutz (Art. 8, EMRK) und das Diskriminierungsverbot (Art. 14, EMRK), ebenfalls sinnvoll, vor allem im Hinblick auf die Art und Weise, in der personenbezogene Daten letztlich weiterverwendet werden können, und darauf, wie mit absichtlichen oder unbeabsichtigten Verzerrungen entwickelte Algorithmen zur Segmentierung und unterschiedlicher Behandlung von sozialen (oder Verbraucher-)Zielgruppen führen könnten.⁶³¹ Schließlich warnt die Wissenschaft auch vor „asymmetrischer Machtverteilung zwischen jenen, die KI-Technologien entwickeln und anwenden, und jenen, die mit diesen interagieren und diesen ausgesetzt sind“.⁶³² Unter Berücksichtigung dieser Fragen wird sich der folgende Abschnitt mit bestehenden Mechanismen zur Rechenschaftspflicht und Verantwortlichkeit befassen, welche diesen Risiken begegnen und sie minimieren sollen.

6.3.1. Bestehender Rechtsrahmen in Europa

Im Februar 2020 veröffentlichte die Europäische Kommission ein Weißbuch, in dem eine Strategie vorgeschlagen wird, die mittels einer geeigneten Politik und eines entsprechenden Regulierungsrahmens sowie der Schaffung eines „Ökosystems für Exzellenz und Vertrauen“ für eine erfolgreiche Einführung von KI in der Europäischen Union sorgen soll.⁶³³ In dem Weißbuch wird auf den strengen Rechtsrahmen in der EU verwiesen, durch den unter anderem bereits der Verbraucherschutz gewährleistet, gegen unlautere Geschäftspraktiken vorgegangen und die personenbezogenen Daten und die Privatsphäre der Bürgerinnen und Bürger geschützt wird, insbesondere die „Datenschutz-Grundverordnung und andere sektorspezifische Rechtsvorschriften mit Bestimmungen zum Schutz personenbezogener Daten, wie die Richtlinie zum Datenschutz bei der Strafverfolgung“.⁶³⁴ Obgleich die DSGVO nicht ausdrücklich auf KI Bezug nimmt, wurde sie

⁶²⁹ Intel AI (2019), „Rethinking Privacy For The AI Era“, Forbes, Insights Team, <https://www.forbes.com/sites/insights-intelai/2019/03/27/rethinking-privacy-for-the-ai-era/#693cda737f0a>.

⁶³⁰ EDRI (2020), *Can the EU make AI “trustworthy”? No – but they can make it just*, <https://edri.org/can-the-eu-make-ai-trustworthy-no-but-they-can-make-it-just/>.

⁶³¹ Wagner, B. (2017), *Study On The Human Rights Dimensions of Automated Data Processing Techniques (In Particular Algorithms) And Possible Regulatory Implications*, Europarat, Expertenausschuss für Internet-Vermittler (MSI-NET), <https://rm.coe.int/study-hr-dimension-of-automated-data-processing-incl-algorithms/168075b94a>.

⁶³² Yeung, K. (2019), *Responsibility and AI*, Council of Europe study DGI(2019)05, Europarat, Expertenausschuss zur Menschenrechtsdimension der automatisierten Datenverarbeitung und verschiedener Formen der künstlichen Intelligenz (MSI-AUT), <https://rm.coe.int/responsibility-and-ai-en/168097d9c5>.

⁶³³ Europäische Kommission (2020), *Weißbuch Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen*, Brüssel, 19.2.2020, COM(2020) 65, https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en.

⁶³⁴ Europäische Kommission, op. cit.

auf technologieneutrale Weise erstellt, um jeglicher technologischen Veränderung oder Weiterentwicklung zu begegnen, und erfasst daher in vollem Umfang die Verarbeitung personenbezogener Daten durch einen Algorithmus.⁶³⁵ Darüber hinaus können sogar KI-Systeme, die sich auf anonymisierte Daten stützen, weiterhin der DSGVO unterliegen, da bei einigen Anonymisierungstechniken nicht zwangsläufig die Gefahr der Deanonymisierung ausgeschlossen werden kann.⁶³⁶ Des Weiteren ist laut einer Bewertung durch den Europäischen Datenschutzausschuss (EDSA) der derzeitige Rechtsrahmen durch einen „risikobasierten Ansatz, den Grundsatz der Datenminimierung und die Anforderung des eingebauten und standardmäßigen Datenschutzes“ sowie Bestimmungen zur Rechenschaftspflicht, Transparenz und dem Verbot jeglicher Entscheidungsfindung alleine auf der Grundlage automatischer Verarbeitung geeignet, um vielen der möglichen Risiken und Herausforderungen in Verbindung mit der Verarbeitung personenbezogener Daten durch Algorithmen entgegenzutreten.⁶³⁷ Schließlich kam ein im Juni 2020 vom Wissenschaftlichen Dienst des Europäischen Parlaments (EPRS) veröffentlichter Bericht zum Einfluss der DSGVO auf künstliche Intelligenz zu dem Schluss, dass die DSGVO „sinnvolle Datenschutzhinweise im Zusammenhang mit KI-Anwendungen enthält“ und dass „für die Datenverarbeitung Verantwortliche, die sich mit KI-basierter Verarbeitung befassen, die Wertmaßstäbe der DSGVO unterstützen und einen verantwortungsvollen und risikoorientierten Ansatz verfolgen sollten“.⁶³⁸

Geraldine Proust, Direktorin für politische Angelegenheiten bei der Federation of European Direct and Interactive Marketing (europäischer Vermarkterverband – FEDMA)⁶³⁹, die ihr Fachwissen zu Datenschutzgesetzen in der EU zur Verfügung gestellt hat, pflichtet dem bei und sagt, dass die DSGVO in der Tat Grundsätze enthält, durch die mittels KI-gestützter Technologien verarbeitete personenbezogene Daten geschützt werden, und dass sie verlangt, dass Organisationen Rechenschaft ablegen und dass sie diese Grundsätze nicht nur einhalten, sondern auch in der Lage sind, nachzuweisen, dass sie sie einhalten. Proust weist auch darauf hin, dass „Organisationen nur notwendige, zweckentsprechende und sachlich richtige personenbezogenen Daten verarbeiten dürfen“. Sie fügt hinzu: „Überdies muss die Verarbeitung nach Treu und Glauben, transparent, auf rechtmäßige Weise und zu einem legitimen Zweck erfolgen. Wenn der Zweck ohne personenbezogene Daten erreicht werden kann, muss ein alternativer Ansatz gewählt werden, beispielsweise können die Daten anonymisiert werden“. Dies bedeutet, dass IT-Experten, Manager und Datenschutzbeauftragte im kreativen Marketingprozess so früh wie möglich zusammenarbeiten müssen, um ein angemessenes Gleichgewicht zwischen Innovation und Ethik herzustellen. Der EDSA erklärte im Zusammenhang mit Algorithmen des maschinellen

⁶³⁵ Europäischer Datenschutzausschuss (2020), *EDPB Response to the MEP Sophie in't Veld's letter on unfair algorithms*, Brüssel, 29. Januar 2020, OUT2020-0004, https://edpb.europa.eu/our-work-tools/our-documents/letters/edpb-response-mep-sophie-int-velds-letter-unfair-algorithms_de.

⁶³⁶ Zaccaria, E. (2020), „Artificial Intelligence: addressing the risks to data privacy and beyond“, PrivSec Report, <https://gdpr.report/news/2020/06/01/artificial-intelligence-addressing-the-risks-to-data-privacy-and-beyond/>.

⁶³⁷ Europäischer Datenschutzausschuss, op. cit.

⁶³⁸ Wissenschaftlicher Dienst des Europäischen Parlaments (2020), *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, Studie, Wissenschaftlicher Dienst des Europäischen Parlaments, Referat Bewertung wissenschaftlicher und technologischer Optionen (STOA), PE 641.530, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU\(2020\)641530_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf).

⁶³⁹ Geraldine Proust von der FEDMA,+: <https://www.fedma.org/>. Im Rahmen der Abfassung dieses Artikels am 7. Juli 2020 interviewt.

Lernens außerdem, dass in Anbetracht des bereits vorhandenen, umfangreichen Rechtsrahmens das Augenmerk auf der Weiterentwicklung bestehender Normen, der Rechenschaftspflicht und Datenschutzfolgeabschätzungen (DSFA) liegen sollte. Proust unterstützt einen derartigen Ansatz und ist der Ansicht, dass die derzeitige Gesetzgebung zunächst ordnungsgemäß umgesetzt und bewertet werden sollte und dass jegliche neuen Gesetze „ausgewogen sein müssen, um Hemmnisse für die Entwicklung und Anwendung dieser Technologie zu vermeiden, und mit dem Datenschutzrecht in Einklang gebracht werden müssen, um Widersprüche zu vermeiden“.

Von zuständigen Behörden herausgegebene Leitlinien erscheinen als geeigneter Weg, um gegebenenfalls mehr Klarheit zu schaffen und Hinweise zur Anwendung konkreter rechtlicher Anforderungen zu geben. In einem aktuellen Bericht des EPRS wird die Notwendigkeit derartiger Leitlinien unterstrichen und erklärt, dass für die Datenverarbeitung Verantwortliche und Datensubjekte „nicht alleine gelassen werden sollten“ und ihnen „Leitlinien zur Frage, wie KI im Einklang mit der DSGVO auf personenbezogene Daten angewendet werden kann, zur Verfügung gestellt“ werden sollten.⁶⁴⁰ Die Verfasser des EPRS-Berichts fordern einen Mehrebenenansatz und die aktive Beteiligung der Institutionen an breiten gesellschaftlichen Debatten mit allen Akteuren, einschließlich für die Datenverarbeitung Verantwortlichen und Auftragsverarbeitern sowie der Zivilgesellschaft, um angemessene Lösungen und hochwertige Hinweise „auf der Grundlage gemeinsamer Werte und wirksamer Technologien“ zu entwickeln.⁶⁴¹ Zu diesem Zweck haben die europäischen Institutionen, darunter die Europäische Kommission, der Europarat und der EDSA⁶⁴² sowie Interessengruppen der Industrie⁶⁴³ bereits mehrere hilfreiche Leitfäden veröffentlicht. In den kommenden Jahren werden wahrscheinlich viele weitere präsentiert und zusammen mit den Regulierungs- und Selbstkontrollinstrumenten werden sie hoffentlich nützliche Hilfsmittel darstellen, die zum allgemeinen Erfolg der KI-Einführung, zum Verbraucherschutz und zum Vertrauen in Technologien beitragen.

⁶⁴⁰ Wissenschaftlicher Dienst des Europäischen Parlaments, op. cit.

⁶⁴¹ Wissenschaftlicher Dienst des Europäischen Parlaments, op. cit.

⁶⁴² Beispielsweise *Guidelines on Artificial Intelligence and Data Protection* (2019) des Beratenden Ausschusses zum Übereinkommen zum Schutz des Menschen bei der automatischen Verarbeitung personenbezogener Daten (Konvention 108),

<https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8>. Europarat; *Ethik-Leitlinien für eine vertrauenswürdige KI* (2019) der Hochrangigen Expertengruppe für Künstliche Intelligenz, Europäische Kommission, <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>; *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default* (2019) des Europäischen Datenschutzausschusses, https://edpb.europa.eu/our-work-tools/public-consultations-art-704/2019/guidelines-42019-article-25-data-protection-design_en.

⁶⁴³ Beispielsweise *High-Level Principles on Artificial Intelligence* (2020) der European Tech Alliance (EUTA), <http://eutechalliance.eu/wp-content/uploads/2020/02/EUTA-High-Level-Principles-on-AI.pdf>; *The Seven-Step Ad Tech Guide* (2020) der Data & Marketing Association (DMA) und der Incorporated Society of British Advertisers (ISBA) zu Herausforderungen im Bereich des Datenschutzes bei RTB in der programmatischen Werbung, <https://dma.org.uk/uploads/misc/seven-step-ad-tech-guide-v9.pdf>.

6.3.2. Fazit: (Meistens) das Gute, das Schlechte und das Hässliche

Die Antwort auf die bestehenden Hauptbedenken hinsichtlich des Einsatzes von KI im Marketing passt gut in den aus dem berühmten Western von Sergio Leone mit dem englischen Titel „The Good, The Bad and the Ugly“ (dt. Titel: Zwei glorreiche Halunken) übernommenen Rahmen.

Insgesamt gibt es größtenteils gute Nachrichten, nämlich im Hinblick auf die Tatsache, dass die heute vorhandenen Rechtsvorschriften und politischen Leitlinien einen soliden Rahmen bieten, um zukunftsfähige KI-Systeme für die Werbung zu schaffen, welche die Menschenrechte und Verbraucherinteressen schützen, ohne weitere Innovation zu behindern. Darüber hinaus scheinen Unternehmen, von globalen Tech-Konzernen bis hin zu Start-ups, den langfristigen Nutzen des Verbrauchervertrauens einer Neubewertung unterzogen zu haben und fordern einen auf den Menschen ausgerichteten Ansatz bei der Entwicklung und Anwendung von KI, von deren Funktion, Sensorik, kognitiven und Lernfähigkeiten⁶⁴⁴ bis hin zu Erwägungen über ihre Auswirkungen auf Endnutzer (Mikroansatz) sowie die Gesellschaft und das ethische Umfeld im Allgemeinen (Makroansatz). Das niederländische Unternehmen DEUS bezeichnet dies als einen „auf die Menschheit ausgerichteten“ Ansatz im Marketing, wo er besonders relevant ist, da Werbung sich den Bedürfnissen, Wünschen und häufig den „Druckpunkten“ der Kunden widmet, die ihr Kaufverhalten leiten. Da sich die Verbraucher der Nutzung ihrer Daten in der Werbung zunehmend bewusst sind und diesbezüglich sogar noch mehr erfahren wollen,⁶⁴⁵ erscheint es unumgänglich, dass das Thema des ethisch vertretbaren Einsatzes von KI im Marketing von der Zivilgesellschaft, politischen Entscheidungsträgern und verantwortungsbewussten Branchenakteuren, die sich hinsichtlich der Notwendigkeit eines vertrauenswürdigen KI-Rahmens wohl einig sind, weiterhin priorisiert wird.

Die schlechte Nachricht besteht darin, dass alle oben erwähnten Maßnahmen und Strategien nur dann sinnvoll sind, wenn ausreichende Durchsetzungsmechanismen vorhanden sind und wirksame Anforderungen in Bezug auf Rechenschaftspflicht und Transparenz bestehen. Kurz gesagt, muss sich gegen die Desensibilisierung im Hinblick auf all die Daten, die Verbraucher als Gegenleistung für relevantere Informationen und Dienste zur Verfügung stellen, Widerstand regen. Dieser durchschaubare Austausch muss auf der Ebene von Einzelpersonen und der Gesellschaft weiterhin ein Schwerpunkt bleiben.

Schließlich ist die hässliche Wahrheit, dass die Geschwindigkeit der technologischen Entwicklung wahrscheinlich immer die Anstrengungen der Gesetzgeber im Hinblick auf neue Technologien, wie etwa Gesichtserkennung, übertreffen wird und somit neue Herausforderungen erzeugt. Deshalb ist es unerlässlich, ethische Standards in die

⁶⁴⁴ Yamakage, Y. und Okamoto, S., „Toward AI for Human Beings: Human Centric AI Zinrai“, *Fujitsu Scientific & Technical Journal*, Januar 2017, Bd. 53, Nr. 1, S. 38-44.

⁶⁴⁵ Laut einer EDAA-Umfrage zur Wahrnehmung digitaler Werbung durch EU-Bürger seit der Verabschiedung der DSGVO: 97 % der Verbraucher sind sich darüber bewusst, dass Daten für Online-Werbung genutzt werden, 62 % wissen etwas darüber, wie es funktioniert und 72 % würden gerne mehr darüber wissen, wie Informationen über sie im Internet verwendet werden. Siehe European Interactive Digital Advertising Alliance (2019), *Consumer Research – How EU citizens perceive digital advertising since GDPR*, <https://www.edaa.eu/consumer-research-how-eu-citizens-perceive-digital-advertising-since-gdpr/>.

Entwicklung von KI und anderen Technologien einzubeziehen, die im Werbe-Ökosystem und darüber hinaus eingesetzt werden. Wie von allen interviewten Experten festgestellt, ist Technologie zwar wichtig, aber Menschen und Organisationskultur sind wichtiger und sie können die Technologieübernahme in Richtung der gewünschten Ethikgrundsätze lenken.

6.4. Nutzung von KI für intelligente Werberegulierung

Selbstkontrollorganisationen der Werbebranche sind unabhängige Organisationen für Werbestandards, die mittels Selbstkontrolle (SK) und der Durchsetzung von Werbekodizes eine verantwortungsvolle kommerzielle Kommunikation gewährleisten. Gemäß einer von der *European Advertising Standards Alliance* (Europäische Dachorganisation für Selbstkontrollorganisationen der Werbebranche – EASA)⁶⁴⁶ herausgegebenen Empfehlung zu bewährten Praktiken für digitale Marketingkommunikation fällt die gesamte digitale Online-Werbung in den Zuständigkeitsbereich der Selbstkontrollorganisationen, was bedeutet, dass alle neuen oder neu entstehenden digitalen Werbeformate und -praktiken, wie etwa Produktwerbung, die über ein Heim-Sprachgerät oder durch ein Posting eines virtuellen Influencers im Auftrag einer Marke erfolgen, denselben ethischen Normen unterliegen, die verlangen, dass diese rechtmäßig, nicht sittenwidrig, ehrlich und wahrheitsgetreu sein sollten.⁶⁴⁷

Obgleich die Hauptverantwortung der Selbstkontrollorganisationen darin liegt, durch die Bearbeitung von Verbraucher- und Mitbewerberbeschwerden und das Angebot einer unverbindlichen Vorabberatung zur Werbe-Compliance die Ordnungsmäßigkeit von Werbungen sicherzustellen, wenden Organisationen für Werbestandards zunehmend Mittel für die Bereitstellung technologiegestützter Werbeüberwachung auf, insbesondere im Online-Umfeld, das bei den Werbeausgaben jährlich eine Steigerung im zweistelligen Prozentbereich erlebt.⁶⁴⁸ Tatsächlich erhielten 2019 zwei Selbstkontrollorganisationen *Best Practice Awards* der EASA für ihren Technologieeinsatz bei der Online-Überwachung: die Organisation für Werbestandards im Vereinigten Königreich – die *Advertising Standards Authority* (ASA) – gewann eine Platin-Auszeichnung für ihr Avatar-Monitoring-Projekt, während die französische Selbstkontrollorganisationen – *Autorité de Régulation Professionnelle de la Publicité* (ARPP) – eine Gold-Auszeichnung für ihr KI-Programm, das „Compliance as a Service“ bietet, erhielt.⁶⁴⁹

Um den Austausch bewährter Praktiken zu fördern und die Angleichung technologiegestützter SK-Innovationen und von Wissen über KI und Anwendungen des maschinellen Lernens im Umfeld der Werberegulierung auf europäischer Ebene zu unterstützen, richtete die EASA im März 2020 eine Arbeitsgruppe zu „Datengetriebener SK“

⁶⁴⁶ Der EASA gehören 28 derartige SVK in ganz Europa sowie 13 Branchenverbände an, welche die gesamte Werbewertschöpfungskette vertreten (z. B. Marken, Agenturen, Medien, Publisher und digitale Plattformen).

⁶⁴⁷ European Advertising Standards Alliance (2015), *Best Practice Recommendation on Digital Marketing Communications*, <https://www.easa-alliance.org/products-services/publications/easa-best-practice-guidance>.

⁶⁴⁸ IAB Europe (2020), *AdEx Benchmark 2019 Study*, <https://iab europe.eu/knowledge-hub/iab-europe-adex-benchmark-2019-study/>.

⁶⁴⁹ European Advertising Standards Alliance (2019), „Press Release: EASA's Best Practice Awards 2019“, <https://easa-alliance.org/news/easa/press-release-best-practice-awards-2019>.

ein. In den folgenden Abschnitten werden laufende KI-Projekte zur SK untersucht und es wird erörtert, wie diese und andere entsprechende technologische Initiativen zur Förderung einer vertrauenswürdigen Werbung beitragen können.

6.4.1. Sinnvolle Datenerfassung durch Avatare

Im Jahr 2019 nutzte die Organisation für Werbestandards des Vereinigten Königreichs, die ASA, Avatare – automatisierte Datenerfassungsprogramme, welche die Online-Profile bestimmter Altersgruppen replizierten –, um Online-Display-Werbung eingeschränkt zulässiger Produkte, wie etwa Alkohol, Glücksspiel oder Lebensmittel und Erfrischungsgetränke mit hohem Fett-, Salz- oder Zuckergehalt („HFSS-Produkte“), die sich an britische Verbraucher richtet, zu überwachen.

An der Überwachung mittels Avataren waren Avatare beteiligt, die sieben verschiedene Altersprofile repräsentierten und 250 Websites und YouTube-Kanäle sowie verwandte Seiten besuchten, insgesamt 196 000 Seitenaufrufe erzielten und Informationen zu über 95 000 Werbungen aufzeichneten, die innerhalb eines Zeitraums von zwei Wochen dargeboten wurden – ein Umfang, der bei der Verfolgung traditioneller Untersuchungsansätze unmöglich zu erreichen ist.⁶⁵⁰ Bei diesem Versuch wurde das erste Mal Avatar-Technologie auf diese Weise für die regulatorische Überwachung eingesetzt und das Verfahren ermöglichte es der ASA, Informationen zu der Frage zu erheben, ob Werbungen für eingeschränkt zulässige Produkte unangemessen auf bestimmte Zielgruppen gerichtet waren oder im Zusammenhang mit ungeeigneten Medien gezeigt wurden.⁶⁵¹

Obgleich die in diesem erfolgreichen Pilotversuch für automatisierte Überwachungsverfahren genutzte automatisierte Technologie sich nicht auf KI-Systeme stützte, setzt die ASA im Rahmen ihres derzeit vier Projekte umfassenden, langfristigen *Tech4Good*-Programms weiterhin auf die Einbindung von Algorithmen des maschinellen Lernens in die alltäglichen Tätigkeiten der Selbstkontrollorganisationen. Die ASA verwendet auch Brandwatch-Technologie, ein auf maschinellem Lernen beruhendes Social-Intelligence-Softwareprogramm, um rechtswidrige oder nicht normgerechte Werbungen auf Sozialen Medien aufzuspüren. Bei der ersten Überwachungsmaßnahme wurden innerhalb eines Zeitraums von drei Monaten 12 000 Problem-Postings, die Botox-Werbungen⁶⁵² enthielten, entfernt.⁶⁵³

⁶⁵⁰ European Advertising Standards Alliance (2018), *2018 European Trends in Advertising Complaints, Copy Advice and Pre-clearance*, <https://www.easa-alliance.org/products-services/publications/easa-statistics>.

⁶⁵¹ Advertising Standards Authority (2019), *ASA monitoring report on HFSS ads appearing around children's media*, <https://www.asa.org.uk/resource/asa-monitoring-report-on-hfss-ads-appearing-around-children-s-media.html>.

⁶⁵² Botox-Produkte und entsprechende Dienstleistungen dürfen im Vereinigten Königreich nicht öffentlich beworben werden.

⁶⁵³ Advertising Standards Authority, Committee of Advertising Practice (2020), *Annual Report 2019*, <https://www.asa.org.uk/news/using-technology-for-good-our-annual-report.html>.

Guy Parker, Geschäftsführer der ASA⁶⁵⁴, ist der Ansicht, dass KI-Technologien das Potenzial haben, die operative Leistungsfähigkeit von Selbstkontrollorganisationen zu steigern und die Wirksamkeit der Werberegulierung zu verbessern. „Wir sind davon überzeugt, dass Selbstkontrollorganisationen, die diese nicht übernehmen, abgehängt werden“, fügt Parker hinzu. Er glaubt, dass KI-Anwendungen auf dem Gebiet der Werbe-Compliance „die Möglichkeit bieten, der großen Herausforderung der Regulierung von Online-Werbung zu begegnen“, einschließlich der Werbung von „zahlreichen kleinen und mittleren Unternehmen, die sich der Regeln und ihrer Bedeutung nicht immer bewusst sind.“

6.4.2. KI-Entwicklungen für die Werbe-Compliance in Frankreich

Die französische Organisation für Werbestandards, die ARPP, begann 2019 im Rahmen ihres Programms „Compliance as a Service“ (CaaS) mit der Nutzung von KI bei ihren Geschäftstätigkeiten. Im Mittelpunkt der ersten drei Projekte stand die Ermittlung der Branchenvorab-Freigabe von Werben oder Beratungsdienste zur Werbe-Compliance zu ersuchen) und einer Deep-Learning-Anwendung für die Erfassung und Analyse von Alkohol, Einblendungen von Kleingedrucktem und Geschlechterverhältnissen in der Fernsehwerbung.

Anfang 2020 erweiterte die ARPP den Anwendungsbereich ihres Systems für überwachtes Lernen zur Geschlechterdarstellung von der Gesichtserkennung auf die automatische Analyse von Sprache und Altersgruppen und führte eine gründliche Untersuchung von eingegangener Fernsehwerbung durch, deren Vorab-Freigabe die ARPP genehmigen sollte. Mittels dreier Modelle des überwachten Lernens, die sich auf Bild- und Spracherkennungstechnologien und das Tracking von Darstellern in einer Szene stützten, wurden fast 131 Stunden Video-Material untersucht, wodurch die ARPP die Entwicklung bei der Vertretung von Männern und Frauen bewerten konnte.⁶⁵⁵

Mohamed Mansouri, stellvertretender Direktor der ARPP,⁶⁵⁶ erklärt, dass die Hauptgründe für den Einsatz von KI-Technologien durch die Selbstkontrollorganisationen folgende sind: Zeitersparnis dank der Automatisierung durch maschinelles Lernen von Aufgaben mit geringer Wertschöpfung sowie der Wunsch, für alle – Branchenexperten, Behörden und Verbraucher – ein moderne, agile Werbeselbstkontrolle zu schaffen und dabei die Vorteile der neuesten Technologien zu nutzen. Die Konsultation der Branche im

⁶⁵⁴ Guy Parker von der ASA, <https://www.asa.org.uk/>. Im Rahmen der Abfassung dieses Artikels am 23. Juni 2020 interviewt.

⁶⁵⁵ Mansouri, M. (2020), „Intelligence artificielle et représentation féminine/masculine dans la publicité audiovisuelle. Après les visages et le genre : la voix et l'âge !“, ARPP-Blog, ARPP, <https://blog.arpp.org/2020/03/05/intelligence-artificielle-et-representation-feminine-masculine-dans-publicite-audiovisuelle-visages-genre-voix-age/>.

⁶⁵⁶ Mohamed Mansouri von der ARPP, <https://www.arpp.org/>. Im Rahmen der Abfassung dieses Artikels am 24. Juni 2020 interviewt.

Rahmen der Workshops ermöglichte eine genauere Bestimmung der vorrangigen Bereiche bei den Überwachungsleistungen der ARPP und die weitere Entwicklung ihrer KI-Programme.

6.4.2.1. Invenio und darüber hinaus: Eine Fallstudie

Im Jahr 2020 startete Mansouri gemeinsam mit dem Technologiepartner-Unternehmen der ARPP, Sicara,⁶⁵⁷ und anderen Experten der französischen Selbstkontrollorganisationen das aus vier Modulen bestehende Invenio-Projekt:

- ein Webcrawler: ein Programm, das auf einer Reihe von Sites automatisch Werbungen sammelt und derzeit für Display-Werbung eingesetzt wird;
- Ermittlung potenzieller Verstöße durch die Nutzung von auf Computer Vision und Textanalyse gestützter KI;
- Meldung von Verstößen und Validierung durch die Juristen und Experten der ARPP, mit Rückmeldung an das Modell, um dessen Genauigkeit zu verbessern (aktuell bezogen auf irreführende Behauptungen zu Heilverfahren und verbotene Finanzwerbung);
- ein automatisch ausgelöstes Warnsystem, wenn eine Werbung erkannt wird, die auf die definierten verbotenen Sites verweist.

Das lateinische Wort „Invenio“ bedeutet „ich finde“ und ist ein passender Name für das Projekt, mit dessen Hilfe bei der Ermittlung von Non-Compliance-Fällen erhebliche Zeit gespart werde und gegen mehr Verstöße vorgegangen werden könne, so Mansouri. Allerdings wird sein Wert auch auf erzieherischer Ebene gesehen und darin, es verschiedenen Akteuren in der Werbewertschöpfungskette zu ermöglichen, ihren Programmablauf zu prüfen.

Laut Mansouris Ausführungen umfassen die nächsten Phasen der Expansion von Invenio: Ermittlung anderer Arten gängiger Verstöße in Bereichen wie Bewerbung von Kosmetika und elektronischen Zigaretten, fehlende Offenlegung kommerzieller Partnerschaften und andere Formen der Werbung mit Bezug auf Influencer-Marketing sowie Videowerbungen. Er fügt zudem hinzu, dass die Zukunftsvision für Invenio die europäische Dimension einschließt, „was notwendiger denn je ist, um die Geltung der Selbstkontrolle auf europäischer Ebene zu demonstrieren“.

Auf die Frage, ob er die Gefahr sieht, dass KI-Instrumente in der Zukunft zu breit eingesetzt werden und dadurch möglicherweise Werbungen falsch beurteilen und zu stark kontrollieren, beruhigt Mansouri: „Die menschliche Komponente ist in dem System wesentlich“ und nichts, ergänzt er, werde ohne vorherige rechtliche Analyse und Validierung durch Experten entschieden. „KI ermöglicht es, große Datenmengen sehr effizient zu verarbeiten, allerdings bei sehr grundlegenden Dingen. Auch wenn die Modelle ihre Genauigkeit im Zuge der Nutzung verbessern, kann der juristische Vernunftschluss bisher nicht automatisiert werden“.

⁶⁵⁷ Weitere Informationen über Sicara, <https://www.sicara.fr/>.

6.4.3. Nutzung von Technologie zur Steigerung des Vertrauens in den niederländischen Werbemarkt

Eine weitere Vorreiterin, wenn es um die Erschließung von KI und ihres Einsatzes in der Werbeselbstkontrolle geht, ist die niederländische Selbstkontrollorganisation Stichting Reclame Code (SRC), die derzeit verschiedene mögliche Anwendungen automatisierter Technologien untersucht. Das Hauptziel von SRC ist es, „Vertrauen in die Werbebranche aufzubauen“. Ausgehend von seiner beherrschenden Stellung im Bereich der Werbeselbstkontrolle außerhalb des Internets, strebt SRC danach, auch im Online-Umfeld eine Vorbildfunktion einzunehmen. Otto van der Harst, der Direktor von SRC,⁶⁵⁸ ist überzeugt, dass Selbstkontrollorganisationen der Werbebranche „tätig werden und ein tiefes Verständnis für die technisierte Welt der Werbung entwickeln müssen“, während sie gleichzeitig ihre Position als unabhängige Regulierungsorgane behalten. Er ist der Ansicht, dass die niederländische Selbstkontrollorganisation und die Werbebranche gemeinsam „am Beginn eines neuen Jahrzehnts im Online-Handel und in der Online-Werbung stehen“.

Der Technologiepartner von SRC, DEUS, stimmt dem zu und sagt, dass der „Einsatz von Datenwissenschaft und neuen Technologien eine Möglichkeit bietet, damit zu beginnen, die enormen Datenmengen im Online-Werbe-Ökosystem sinnvoll zu nutzen“. Nathalie Post von DEUS⁶⁵⁹ fügt hinzu, dass einer der Schwerpunktbereiche für SRC in den Niederlanden der Schutz besonders gefährdeter Menschen (z. B. Minderjährige, Menschen mit Suchtproblemen, ältere Menschen) vor potenziell schädlichen kommerziellen Inhalten ist (gezielte Werbungen für Alkohol, Glücksspiel, Gewichtsabnahme, plastische Chirurgie und Botoxprodukte sowie übermäßige Anzahlen von HFSS-Werbungen und irreführende Behauptungen zu Heilverfahren). Das DEUS-Team sieht fünf Bereiche, in denen automatisierte Technologien zur Unterstützung von SRC in diesen Betätigungsfeldern besonders hilfreich sein können: Verwendung von Avataren zur Simulation von Profilen bestimmter (gefährdeter) Zielgruppen; Nutzung von Social-Listening-Tools zur Überwachung von Social-Media-Influencern; Einsatz von CL und Computer Vision zur Automatisierung der Registrierung, Kategorisierung und Meldung von Werbungen, die gegen den Werbekodex verstoßen; Anwendung von Computer Vision zur Ermittlung von Werbungen, die bestimmte Verzerrungen enthalten, wie mangelnde Vielfalt; und Erlangung eines besseren Verständnisses von Datenflüssen im RTB-Ökosystem, um möglichen Verstößen gegen den Werbekodex zu begegnen.

Otto van der Harst hat keinen Zweifel daran, dass sich der Traum von KI-gestützten SK-Dienstleistungen bereits am Horizont abzeichnet und dass es in den kommenden Jahren um Experimente und Diskussionen mit Selbstkontrollorganisationen, Plattformen und Agenturen der Werbebranche gehen wird, wobei er hofft, alle relevanten beteiligten Parteien in Einklang zu bringen: „KI und maschinelles Lernen sind bloß Instrumente, die dafür sorgen, dass die Online-Welt ein sicherer Werbeplatz bleibt“.

⁶⁵⁸ Otto van der Harst von SRC, <https://www.reclamecode.nl/>. Im Rahmen der Abfassung dieses Artikels am 20. Juni 2020 interviewt.

⁶⁵⁹ Nathalie Post von DEUS, <https://deus.ai/>. Im Rahmen der Abfassung dieses Artikels am 22. Juni 2020 interviewt.

6.4.4. Technologische Lösungen von einer der Lokomotiven der Werbebranche

Selbstkontrollorganisationen sind nicht die einzigen, die unverantwortliche Werbungen nachverfolgen und versuchen, den Online-Raum durch die Nutzung von KI-Technologien sicherer zu machen. Einige der Giganten der Online-Werbeindustrie, die bei der technologischen Entwicklung führend sind, setzen ebenfalls aktiv KI-Instrumente ein, um ein sicheres und zukunftsfähiges Online-Werbeumfeld zu gewährleisten.

Im April 2020 veröffentlichte Google seinen jährlich erscheinenden *Bad-Ads-Report*, in dem das Unternehmen vermeldet, dass es 2019 2,7 Milliarden schädliche Werbungen (0,4 Milliarden mehr als 2018) gesperrt und entfernt habe, mit anderen Worten 10 Millionen Werbungen pro Tag, 5 000 Werbungen pro Minute und mehr als 100 pro Sekunde.⁶⁶⁰

„Wir sind sehr stolz darauf, ein Hilfsmittel für Menschen auf der ganzen Welt zu sein, die nach wichtigen Informationen suchen“, erklärt das für diesen Artikel interviewte Team von Google⁶⁶¹. „Wir wissen, dass die Nutzer neben dem Vertrauen in die ans Tageslicht beförderten Informationen auch in der Lage sein sollten, den Werbungen, die sie sehen, zu vertrauen“.

Mit dem Wachstum des digitalen Werbe-Ökosystems entstehen neue Bedrohungen, die fortwährende Anpassungen der Unternehmensstrategien erfordern sowie Verbesserungen der Technologie. „Da wir besser darin werden, bei Betrügern und Schwindlern Entwicklungen und Muster zu erkennen, reagieren wir dementsprechend mit neuer Technologie, um diese neu entstehenden Bedrohungen zu stoppen und mehr Maßnahmen auf der Ebene der Kundenkonten zu ergreifen“, so die Google-Experten weiter. Im Jahr 2019 verdoppelten sie ihre Anstrengungen in Bezug auf die Technologie zur Bekämpfung des anhaltenden Missbrauchs durch Phishing und Trick-to-Click-Werbungen. Eine eigene Arbeitsgruppe bei Google widmete sich der Rückverfolgung und Analyse der von den schädlichen Akteuren angewandten Taktiken zur Umgehung der Systeme von Google und zur Sammlung personenbezogener Daten von Nutzern. „Phishing-Werbungen reichen zwar von solchen, die Personen ausbeuten, die sich für Kryptowährung interessieren, oder andere, die Informationen zur Passverlängerung suchen, bis hin zu Werbungen, die Nutzer zu gefälschten Bank-Websites leiten, doch wir fanden gemeinsame Nenner im Hinblick auf die Umgehung unserer Systeme, die uns in die Lage versetzten, unsere Technologie zu verbessern und sie zu stoppen“, erklärt das Google-Team.

Obgleich die Technologie Google dabei hilft, potenzielle Verstöße zu erkennen, ist es die Verbindung von Technologie und Talent, die eine wirksame Durchsetzung ermöglicht. „Wir haben bei Google Tausende Menschen, die sich der Aufgabe verschrieben haben, uns bei der Bekämpfung von schädlichen Werbungen, schädlichen Sites und Betrügern im Internet zu unterstützen“, merkt das Team an. „Wir prüfen fortwährend Werbungen, Sites und Konten, um sicherzustellen, dass unsere Richtlinien eingehalten

⁶⁶⁰ Spencer, S. (2020), „Stopping bad ads to protect users“, Google Ads, Google, <https://blog.google/products/ads/stopping-bad-ads-to-protect-users>.

⁶⁶¹ Ein Expertenteam von Google wurde im Rahmen der Abfassung dieses Artikels am 8. Juli 2020 interviewt und lieferte die Antwort des Unternehmens zu den in diesem Kapitel erörterten Themen, <https://about.google/>.

werden.“ Die Anwendung von KI und anderen Technologien des maschinellen Lernens gestatten es Google, rund um die Uhr Durchsetzungsmaßnahmen zu ergreifen, und erlauben es dem Unternehmen, bei der Entscheidung, eine Werbung zu entfernen, mehrere Variablen in die Beurteilung einzubeziehen.

Als es zu der Tatsache befragt wird, dass jedes Jahr mehr schädliche Werbungen entfernt werden, antwortet das Google-Team, dass die zahlenmäßige Zunahme bezeichnend für eine Reihe von Schlüsselfaktoren sei, darunter die dynamische Natur und der sich kontinuierlich entwickelnde Umfang des digitalen Werbe-Ökosystems sowie laufende Verbesserungen bei der Anpassung an verschiedene Arten, auf die schädliche Akteure versuchen, das System zu manipulieren. „Kein System wird 100 % perfekt sein, doch wir sind wachsam und arbeiten ständig daran, unsere Instrumente zu verbessern“, so die Google-Experten.

6.4.5. Künftige Grenzen für die Werbeselbstkontrolle

Intelligente Werbung wurde bereits als ein mögliches Ziel jener genannt, die KI-Anwendungen für die Werbung entwickeln und einsetzen. Es lohnt sich daher, auch die Frage zu stellen, wie die Zukunft KI-gestützter Werberegulierung aussehen sollte und wie sie intelligent gestaltet werden kann, das heißt nicht bloß reaktiv und strafend, sondern so, dass auftretende Probleme antizipiert und Branchenakteure auf proaktive Weise unterstützt werden.

Die befragten Selbstkontrollorganisationen sind sich alle darin einig, dass Organisationen für Werbestandards auf Werbeinnovationen mit eigenen technologischen Innovationen antworten müssen. Guy Parker von der ASA ist der Ansicht, dass KI und andere Formen automatisierter Technologien Selbstkontrollorganisationen bessere Informationen darüber liefern werden, wo und wie sie eingreifen oder Dienstleistungen bereitstellen sollten. Obgleich die ASA „niemals den Idealzustand erreichen wird“, werde die Technologie bei der Verringerung der Anzahl unverantwortlicher und schädlicher Online-Werbungen künftig im Zentrum stehen, fügt er hinzu. Mohamed Mansouri von der ARPP ergänzt, dass KI Selbstkontrollorganisationen im Wesentlichen dabei unterstützen wird, die Effizienz zu steigern, tiefer und weiter zu gehen und in der digitalen Werbewelt, in der Kanäle, Formate und Inhalte exponentiell zunehmen, schnell zu reagieren. Er fügt hinzu, dass die Zusammenarbeit bei der Entwicklung und Nutzung von KI-Technologien sehr wichtig ist, da sie die Entwicklung kollektiver Intelligenz ermöglicht und eine Antwort auf die Herausforderungen darstellt, die mit dem naturgemäß grenzüberschreitenden digitalen Umfeld verbunden sind. Otto van der Harst von SRC stimmt dem zu und sagt, dass Selbstkontrollorganisationen das Werbe-Ökosystem in den Griff bekommen müssen, wenn sie „in der Zukunft relevant sein wollen“, und technische Hilfe nutzen müssen, um irgendeine Form unabhängiger Kontrolle zu haben und in der Lage zu sein, eine glaubwürdige Antwort auf die derzeitigen Marktentwicklungen zu geben.

Schließlich wird die europäische Zusammenarbeit voraussichtlich eine wichtige Rolle bei der Anpassung bestehender Initiativen, dem Austausch von Fachwissen und dem gegenseitigen Lernen der verschiedenen Akteure spielen. Das von der EASA koordinierte Selbstkontrollorganisations-Netzwerk in Europa hat erlebt, wie neue Formate

kommerzieller Kommunikation die Werbebranche in der Vergangenheit viele Male revolutioniert haben, und sich unter Wahrung der kulturellen Vielfalt in Europa bei der Anpassung an den technologischen und gesellschaftlichen Wandel als agil und flexibel erwiesen. Durch den Austausch bewährter Praktiken und die gemeinsame Erörterung künftiger Herausforderungen inspirieren und unterstützen die Selbstkontrollorganisationen weiterhin die Werbebranche bei der Erzeugung verantwortungsvoller Werbung und der Erhaltung des Verbrauchervertrauens. Die fortlaufende Zusammenarbeit bei der Nutzung von KI und anderer automatisierter Technologien ist eine der spannendsten europaweiten Initiativen, eine, die voraussichtlich ein Inkubator für die vielversprechende Zukunft intelligenter Werbeselbstkontrolle sein wird.

6.5. Fazit: Der große Datenrausch

Viele Fachleute aus der Praxis und Wissenschaftler, die über KI schreiben, beschreiben ihre zunehmende und sich ausbreitende Anwendung in verschiedenen Branchen, darunter der Werbung, unter Verwendung der Analogie einer großen Datenwelle – mit Bezug auf eine tsunamiähnliche Kraft, die gewaltsam das Umfeld umwandelt und neu gestaltet, mitunter bis zur Unkenntlichkeit. Im Unterschied zu Tsunamis findet die Operationalisierung von Daten jedoch nicht als Naturereignis statt. Sie muss organisiert werden.

Daher bin ich der Meinung, dass der Goldrausch des 19. Jahrhunderts, der zu massiven Veränderungen in der Weltwirtschaft und den Handelsstrukturen führte und Migration und eine rasante soziale Mobilität in Gang setzte, eine sinnvollere Analogie zur Beschreibung von KI-Anwendungen im Marketing und in der Werbung ist. Nicht anders als diese neuen Siedler, die zur Verwirklichung des amerikanischen Traums über die Weltmeere segelten, erlebt die Werbebranche heute einen großen Datenrausch, der jenen sogar noch größere Profite verspricht, die Daten erschließen und richtig kanalisieren und so aus Verbrauchern gewinnbringende Kunden machen.

Ganz wie die während der Goldrauschzeit eingeführten Innovationen haben KI-Anwendungen in der Werbung sowohl für die Industrie als auch die Verbraucher bereits zahlreiche Vorteile hervorgebracht. Durch KI kann Werbung leichter in Umlauf gebracht und kostengünstiger vor den relevanten Verbrauchern platziert werden. Durch Technologien des maschinellen Lernens ist es zudem möglich geworden, neue Werbeformate nahtlos in unseren Alltag zu integrieren, von Sprachassistenten, die uns auf der Grundlage unserer bisherigen Kaufhistorie und anderer Einflussfaktoren maßgeschneiderte Produkte und Dienstleistungen anbieten, bis hin zu virtuellen Influencern, die auf digitalen Plattformen mit uns natürlich interagieren, mit Zielgruppen Reisetipps austauschen und erstrebenswerte Lebensstile bewerben.

Auf der anderen Seite wurde der Goldrausch häufig auch mit dem gesetzlosen Streben nach Reichtum in Verbindung gebracht, was in dem Begriff „Wilder Westen“ zusammengefasst ist. Dies weist auf die revolutionäre Natur der neuen KI-gestützten Werbeformen und -formate hin. Obgleich das heutige Wettrennen der Märkte bei der Entwicklung und Umsetzung der effizientesten KI-Instrumente für die Werbung nicht mit den in Western dargestellten, dramatischen Showdowns verglichen werden sollte,

bestehen nach wie vor Herausforderungen und Bedenken im Hinblick auf die möglicherweise fehlende Kontrolle von KI-Technologien.

Die Zivilgesellschaft und politische Entscheidungsträger sind beunruhigt über die zunehmende Sammlung personenbezogener Daten, Automatisierung von Entscheidungen und Marktasymmetrie. Rechenschaftspflicht und Transparenz sind von allergrößter Bedeutung, um sicherzustellen, dass Menschen durch KI nicht abgeschreckt werden, sondern diese stattdessen bereitwillig annehmen und in vollem Umfang von den Technologien profitieren können, die ihnen helfen, relevante Informationen schneller zu finden und rascher Antworten auf ihre Rückfragen zu erhalten – und die sie sogar bei ihren kreativen Projekten unterstützen können.

Ist – abschließend gefragt – KI ein Tor zu einem *Deadwood*-ähnlichen Werbe-Ökosystem, das ohne Einhaltung bestehender Regeln funktioniert, Verbraucher ausbeutet und dazu gedacht ist, nur jenen Nutzen zu bringen, die das technologische Know-how entwickeln und die Goldgrube der Daten an sich reißen? Nein und in Wirklichkeit sind eine derart düstere Zukunft und in der Tat das Wildwestszenario sehr unwahrscheinlich. Wie in diesem Artikel diskutiert, rückt die ethische Nutzung von KI durch den konstanten Druck der Zivilgesellschaft sowie die detaillierten Anforderungen der Regulierungsrahmen und Selbstkontrollinitiativen in den Mittelpunkt anhaltender öffentlicher Debatten. Die jüngsten politischen Entwicklungen und die Ansichten der interviewten Experten lassen ein starkes Engagement für eine auf den Menschen ausgerichtete KI und entsprechende Anwendung in der Werbung erkennen. Allerdings müssen Strategien und Leitlinien kontinuierlich in die Praxis umgesetzt werden und Schutzmaßnahmen funktionieren nur dann, wenn der Wille besteht, diese zu veranlassen und durchzusetzen – weshalb Verbraucher ebenfalls die Verantwortung haben, ihre Rechte zu kennen, aktiv zu sein und wachsam zu bleiben, wenn es darum geht, von Marktakteuren angemessenen Schutz und Rechenschaft zu fordern.

Schließlich, und dies ist vielleicht am allerwichtigsten, ist die Schaffung eines zukunftsfähigen KI-Rahmens für die Werbung und die sinnvolle Nutzung von Daten und Technologie im Interesse der Werbegemeinschaft selbst. Nur durch die Einbettung von Ethikgrundsätzen in die Algorithmen des maschinellen Lernens, die achtsame Erhebung und Nutzung personenbezogener Daten, welche hohe Anforderungen an Transparenz und Verantwortung widerspiegeln, und die Förderung von KI-Anwendungen im Marketing, die dazu beitragen, schädliche Akteure zu ermitteln und zur Rechenschaft zu ziehen, kann die Branche damit rechnen, nachhaltiges Verbrauchervertrauen zu gewinnen. Letzten Endes ist Vertrauen ein echter Goldstandard in der Werbung und wird es auch immer bleiben.

6.6. Danksagung

Genau wie KI-Algorithmen zum Lernen einen guten Daten-Input brauchen, hätte dieser Artikel nicht ohne die Erkenntnisse der interviewten Expertinnen und Experten entstehen können, die sich freundlicherweise bereit erklärt haben, ihr Fachwissen, ihre Kenntnisse und Zukunftsprognosen im Zusammenhang mit der Nutzung von KI in der Werbung zu teilen. Ich möchte Kerry Richardson von Tiny Giants, Charlie Cadbury von Say It Now und Nathalie Post von DEUS für die Weitergabe ihres Fachwissens zu den Schnittpunkten

zwischen KI und Kreativität, technologiegestützten neuen Werbeformaten und der Entwicklung eines auf den Menschen ausgerichteten KI-Ansatzes danken. Ein großes Dankeschön an den unabhängigen Berater für Influencer-Marketing, Scott Guthrie, und an Rupa Shah von Hashtag Ad Consulting für die Mitteilung ihrer Ansichten zur Entwicklung virtueller Influencer, Ethik und der möglichen Zukunft von virtuellen Menschen als Marketingkanälen. Ich möchte auch Geraldine Proust von der FEDMA danken, die mich freundlicherweise durch die Rechtsrahmen zum Einsatz von KI in der Werbung geleitete und mich an ihren Erkenntnissen zu den wichtigsten Herausforderungen für Anwendungen des maschinellen Lernens in der Zukunft teilhaben ließ. Vielen Dank an das Google-Team, das Erkenntnisse und nähere Erklärungen zu den Anstrengungen des Unternehmens lieferte, hochentwickelte Technologien und menschliches Fachwissen zu nutzen, um schädliche Werbungen zu entfernen und für die Sicherheit ihrer Plattformen zu sorgen. Zu guter Letzt bin ich meinen Kollegen von den Selbstkontrollorganisationen der Werbebranche – Guy Parker von der ASA, Otto van der Harst von SRC und Mohamed Mansouri von der ARPP – sehr dankbar dafür, dass sie herausfordernde Schritte unternommen haben, um KI und andere Technologien des maschinellen Lernens zur weiteren Gewährleistung hoher ethischer Standards in der Werbung einzusetzen, und dass sie es mir ermöglicht haben, sie auf dieser spannenden Reise zu begleiten, unterwegs zu lernen und gemeinsam künftige Entwicklungen zu erörtern und dabei nach innovativen Lösungen mit dem Schwerpunkt auf einer starken europäischen Zusammenarbeit zu suchen.

KI in der Werbung gleicht in mancher Hinsicht einem Labyrinth, bei dem man unmögliche jeden Winkel aufsuchen kann. Unter Anleitung all der interviewten Expertinnen und Experten hoffe ich jedoch, verschiedene mögliche Wege hindurch gefunden zu haben. Diese in diesem Artikel analysierten Wege stellen meine eigene Sicht auf KI in der Werbung dar und entsprechen nicht den Sichtweisen irgendeiner bestimmten, im Text genannten Organisation oder Einzelperson. Schließlich ist die Europäische Audiovisuelle Informationsstelle nicht für in diesem Artikel dargebotene Zitate und Meinungen, einschließlich jener der interviewten Expertinnen und Experten und der Autorin selbst, verantwortlich zu machen.

6.7. Interviewverzeichnis

- Charlie Cadbury von Say It Now, <https://www.sayitnow.ai/>. Interviewt am 19. Juni 2020.
- Expertenteam von Google, <https://about.google/>. Im Rahmen der Abfassung dieses Artikels am 8. Juli 2020 interviewt, lieferte die Antwort des Unternehmens.
- Scott Guthrie, ein unabhängiger Unternehmensberater für Influencer-Marketing, Veranstaltungsredner und Blogger, <https://sabguthrie.info/>. Interviewt am 22. Juni 2020.
- Otto van der Harst von SRC, <https://www.reclamecode.nl/>. Interviewt am 20. Juni 2020.
- Mohamed Mansouri von der ARPP, <https://www.arpp.org/>. Interviewt am 24. Juni 2020.
- Guy Parker von der ASA, <https://www.asa.org.uk>. Interviewt am 23. Juni 2020.
- Geraldine Proust von der FEDMA, <https://www.fedma.org/>. Interviewt am 7. Juli 2020.
- Nathalie Post von DEUS, <https://deus.ai/>. Interviewt am 22. Juni 2020.
- Kerry Richardson von Tiny Giant, <https://www.tinygiant.io/>. Interviewt am 22. Juni 2020.
- Rupa Shah von Hashtag Ad Consulting, <https://www.hashtagad.co.uk/>. Interviewt am 22. Juni 2020.

Persönlichkeitsrechte

*KI kann so kreativ sein, dass sie weit mehr kann, als beim Drehbuchschreiben zu helfen. Ein Drehbuch ist schließlich nur der Anfang des kreativen Prozesses. Die Geschichte und die Ideen in einem Drehbuch müssen in Bilder übersetzt werden. In den meisten Fällen geht es in diesen Geschichten um Menschen. Menschen, die von Schauspielerinnen und Schauspielern gespielt werden. KI kann nicht nur das Drehbuch schreiben und die Musik spielen, sondern auch die Schauspieler stellen. Oder zumindest jeden Schauspieler in den Schauspieler verwandeln, mit dem man schon immer einen Film drehen wollte. Dazu kann sie ihn beispielsweise jünger oder älter machen. Dafür gibt es ganz aktuelle Beispiele: In „Gemini Man“ muss die von Will Smith gespielte Figur gegen einen jüngeren Klon seiner selbst kämpfen. Ein ähnliche digitale Verjüngung kam bei den Hauptfiguren in Martin Scorseses „The Irishman“ zum Einsatz. In „Star Wars: Rogue One“ wirkt Carrie Fisher jünger als je zuvor, und auch Peter Cushing, der 1994 starb, hat seine postmortale Sternstunde. Doch auf den Hype folgt meist die Hysterie. Wer heute Zeitung liest, kann auf Schlagzeilen stoßen wie: „Könnten virtuelle Schauspieler im Zeitalter der Deepfakes Menschen aus dem Geschäft drängen?“⁶⁶² Man stelle sich etwa einen Gangsterfilm mit einem digitalen Marlon Brando vor, aber ohne dessen berüchtigtes Benehmen hinter den Kulissen. Für die Regie wäre das ein Traum. Man braucht dafür nur die entsprechende Hard- und Software ... und einen „Ghost Actor“ – also einen Schauspieler, dessen Gesicht durch das des berühmteren ersetzt wird. Billiger und im Fall Marlon Brandos wohl auch mit besserem Benehmen. Darüber hinaus lassen sich mit KI-Hilfe auch wesentlich leichter digitale Komparsen erschaffen. Wie man sich denken kann, werfen diese – technologischen wie auch künstlerischen – Entwicklungen persönlichkeitsrechtliche Fragen auf. Diese Rechtsfragen sind gesetzlich geregelt und werden vertraglich geklärt. Doch für die Seite mit geringerer Verhandlungsmacht kann ein Vertrag unfair sein. Etwa für unbekannte Schauspielerinnen und Schauspieler. Oder für Tote. Und die Frage hat auch eine dunklere Seite: Deepfakes. Sie können auf unterschiedliche Weise eingesetzt werden und Schaden anrichten. Zunächst einmal die widerrechtliche kommerzielle Nutzung: Deepfakes können für gefakte Produktempfehlungen eingesetzt werden. Ein weiteres Thema ist Identitätsmissbrauch (meist in Pornofilmen). **Kelsey Farish** schreibt in ihrem Beitrag zu dieser Publikation, dass „Persönlichkeitsrechte speziell bei neuartigen Technologien wie Deepfakes und Ghost Acting [...] eine sorgfältige Berücksichtigung des situativen Kontexts erfordern“.*

⁶⁶² Kemp L., „In the age of deepfakes, could virtual actors put humans out of business?“, *The Guardian*, <https://www.theguardian.com/film/2019/jul/03/in-the-age-of-deepfakes-could-virtual-actors-put-humans-out-of-business>.

7. Persönlichkeitsrechte: Von Hollywood bis Deepfakes

Kelsey Farish, DAC Beachcroft LLP

„Der Körper ist keine Sache, er ist eine Situation: Er ist das Instrument, mit dem wir die Welt begreifen“

Simone de Beauvoir

7.1. Einleitung

Juristen pflegen gewöhnlich, wenn sie sich mit komplizierten beruflichen Fragen befassen, keine Hollywood-Schauspieler zu bemühen. Wenn es jedoch um neue Technologien geht, die dem öffentlichen Ansehen einer Person schaden können, ist Scarlett Johansson eine rühmliche Ausnahme. In einem Interview mit der *Washington Post* erklärte die Schauspielerin, die eine der renommiertesten und bestbezahlten Filmstars weltweit ist: „Nichts auf der Welt kann einen Menschen daran hindern, mein Bild oder das eines anderen Menschen irgendwo herauszukopieren und auf einen anderen Körper zu montieren und das Ganze geradezu unheimlich realistisch aussehen zu lassen“.⁶⁶³ Nun sind zwar computergenerierte Spezialeffekte keine neue Erfindung. Aber Scarlett Johansson bezog sich in ihrem Interview auf eine Technik, die zum Zeitpunkt des Interviews (2018) noch relativ unbekannt war, das „Face-Swapping“. Dabei werden die Gesichter von Prominenten in Pornovideos kopiert, und zwar mit Hilfe einer hoch entwickelten Technologie der künstlichen Intelligenz. Heute wird diese Technik als „Deepfake“ bezeichnet, unabhängig davon, ob es sich dabei um Pornos handelt oder um andere Arten von Videos.

Auf die Frage, ob sie gerichtlich gegen Deepfakes vorgehen wolle, erklärte die Schauspielerin, sie habe das Gefühl, dass dies „sinnlos“ sei, weil „jedes Land das Recht am eigenen Bild anders definiert“. Sie fügte hinzu: „Vielleicht schaffen Sie es ja, Seiten in den Vereinigten Staaten sperren zu lassen, die Ihr Gesicht nutzen. Aber in Deutschland kann das

⁶⁶³ Harwell, D., *Scarlett Johansson on fake AI-generated sex videos: 'Nothing can stop someone from cutting and pasting my image'*, The Washington Post, www.washingtonpost.com/technology/2018/12/31/scarlett-johansson-fake-ai-generated-sex-videos-nothing-can-stop-someone-cutting-pasting-my-image/.

ganz anders aussehen.“⁶⁶⁴ Diese Vorschriften, die das Erscheinungsbild einer Person in Veröffentlichungen und Videos schützen sollen, werden allgemein als Persönlichkeitsrecht bezeichnet. Das Persönlichkeitsrecht besteht jedoch aus zahlreichen unterschiedlichen Gesetzen mit komplizierten Ausnahmen und Nuancen. Verallgemeinerungen sind äußerst schwierig, der Ausgang von Rechtsstreitigkeiten ist häufig unvorhersehbar, und, wie Scarlett Johansson ganz richtig festgestellt hat, unterscheiden sich die Gesetze von Land zu Land erheblich.

Dieses Kapitel versucht daher, einige der Missverständnisse im Zusammenhang mit dem Persönlichkeitsrecht aufzuklären und praxisnahe Kommentare in Bezug auf die Risiken und Chancen zu geben, die Deepfakes und Ghost Acting für die Film-, Fernseh- und Unterhaltungsindustrie darstellen. Der Abschnitt „Künstliche Intelligenz macht es möglich“ gibt einen kurzen Einblick in die Technologie, um die es hier geht, und der folgende Abschnitt über „Persönlichkeitsrechte und die Auswirkungen“ beleuchtet den Rechtsrahmen unter vier unterschiedlichen Aspekten mit echten Beispielen aus der Praxis. Der Abschnitt „Gesetze in ausgewählten Rechtssystemen“ baut auf dem Rechtsrahmen auf und gibt einen Überblick über die Art und Weise, wie Persönlichkeitsrechte in Deutschland, Frankreich, Schweden, Guernsey, dem Vereinigten Königreich⁶⁶⁵ und Kalifornien definiert werden. Der letzte Abschnitt „Wie geht es weiter mit dem audiovisuellen Sektor in Europa?“ weist auf einige der gravierenden Lücken hin, die in den Gesetzen bestehen, aber auch auf potenzielle Trends.

7.2. Künstliche Intelligenz macht es möglich: Deepfakes und Ghost Actors

Die unendliche Vielfalt menschlicher Gesichter ist das Ergebnis der Evolution, das lehrt uns die Forschung. Weil Menschen sich hauptsächlich am Gesicht erkennen, hat der Selektionsdruck zur Entwicklung einer ungeheuren Vielfalt von Gesichtszügen geführt, um das gegenseitige Erkennen zu erleichtern.⁶⁶⁶ Und weil wir besonders gut darin sind, Gesichter voneinander zu unterscheiden, können wir auch schnell feststellen, wenn Gesichter unheimlich oder unnatürlich aussehen,⁶⁶⁷ und Wissenschaftler und Visual-Effects-Spezialisten haben lange vergeblich versucht, den Gesichtsausdruck von Menschen in Animationen so nachzuahmen, dass er täuschend echt wirkt. Dies änderte sich erst, als 2014

⁶⁶⁴ Harwell D., *op.cit.*

⁶⁶⁵ Auf die Gesetze des Vereinigten Königreichs wird über das gesamte Kapitel hinweg Bezug genommen. Aber es gibt unterschiedliche Ausprägungen des Persönlichkeitsrechts in den einzelnen Landesteilen des Vereinigten Königreichs, also Schottland, England und Wales und in Nordirland. Sofern nicht anders angegeben, beziehen sich die Referenzen auf Gesetze des Vereinigten Königreichs, wie sie von den Gerichten von England und Wales typisiert werden, oder auf „englisches Recht“.

⁶⁶⁶ Sanders R., *Human faces are so variable because we evolved to look unique*, Berkeley News, University of California Berkeley, <https://news.berkeley.edu/2014/09/16/human-faces-are-so-variable-because-we-evolved-to-look-unique/>.

⁶⁶⁷ Mori M., *The Uncanny Valley: The Original Essay by Masahiro Mori*, IEEE Spectrum, übersetzt von K. F. MacDorman und N. Kageki [Japanisches Original . 不気味の谷 (Bukimi No Tani) 1970]. Englische Übersetzung: <https://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley>.

die künstliche Intelligenz erhebliche Fortschritte in diesem Bereich möglich machte, mit der Erfindung des Deep machine learning.⁶⁶⁸

7.2.1. Deepfakes

Der Begriff „Deepfakes“ tauchte erstmals 2017 auf einem Forum der Reddit-Plattform in den USA auf, als Nutzer hyperrealistische Videos teilten, die Schauspielerinnen wie Scarlett Johansson, Gal Gadot und Emma Watson zeigten.⁶⁶⁹ Heute sind Deepfakes nicht mehr auf die Welt von Porno-Videos beschränkt und können von jedem Internetnutzer zu Unterhaltungszwecken hergestellt oder genutzt werden. Die Software zur Herstellung von Deepfakes ist für jedermann frei zugänglich, sie kann von Seiten wie BitTorrent und GitHub heruntergeladen werden, YouTube-Tutorials bieten Step-by-step-Anleitungen, und Freelance-Entwickler verkaufen maßgeschneiderte Deepfakes für ganze EUR 5 pro Video auf Marktplätzen wie Fivver. Mit Hilfe mobiler Apps wie *ZAO*, *Doublicat* und *AvengeThem* lassen sich Face-Swaps auf der Grundlage eines einzigen Selfies herstellen,⁶⁷⁰ und selbst die Mainstream-Apps *Instagram* und *Snapchat* verfügen über „Filter“, die dasselbe können. Auch wenn man von einer raschen Entwicklung dieser Technik ausgehen muss, so lassen sich die meisten Deepfakes, die nicht professionell hergestellt wurden, leicht als Fälschung erkennen, wenn man sich die Mühe macht, genauer hinzusehen. Trotzdem bleiben sie ein beliebter Freizeitspaß, weil keine speziellen technischen Kenntnisse erforderlich sind. Und für Menschen, die diese Technik einfach aus Neugier oder zum Spaß nutzen, sind kleinere Unstimmigkeiten oder Pannen unerheblich.

7.2.2. Ghost Actors

Die Kunst, jemanden älter, jünger oder ganz anders aussehen zu lassen ist so alt wie die Kunst des performativen Geschichtenerzählens, und seit den 1970er Jahren werden Computer für die Animation des menschlichen Gesichts genutzt.⁶⁷¹ Aber filmtaugliche Gesichtsveränderungen sind erst seit Beginn des Jahrtausends möglich, zum Beispiel in Filmen wie „Der Herr der Ringe“ (2001) und „The Matrix Reloaded“ (2003). Beim Einsatz von Visual Special-Effects-Techniken (VFX-Techniken) werden Gesicht und Körper eines Schauspielers millimetergenau gescannt, um ein virtuelles Modell zu schaffen, ein so genanntes „digitales Double“. Das digitale Double wird anschließend verändert und auf den

⁶⁶⁸ Han J., *Ian Goodfellow: Invented a Way for Neural Networks to Get Better by Working Together*, MIT Technology Review, www.technologyreview.com/lists/innovators-under-35/2017/inventor/ian-goodfellow.

⁶⁶⁹ Cole S., *AI-Assisted Fake Porn Is Here and We're All Fucked*, Motherboard Tech by VICE, www.vice.com/en_us/article/gyddm/gal-gadot-fake-ai-porn.

⁶⁷⁰ Beebom staff, *8 Best Deepfake Apps and Websites You Can Try for Fun*, Beebom, www.beebom.com/best-deepfake-apps-websites.

⁶⁷¹ Parke F. I., „Computer generated animation of faces“, Association for Computing Machinery '72 Proceedings of the ACM Annual Conference, 1: 451–457.

Körper eines Lichtdoubles montiert.⁶⁷² Dabei kann es sich um die Person handeln, deren Gesicht ursprünglich gescannt wurde, und häufig bildet der letzte Filmschnitt einen Schauspieler entweder als eine jüngere oder ältere Version seiner selbst ab. Das digitale Double kann auch als Ersatz für einen Schauspieler genutzt werden, der während der Dreharbeiten verstorben ist.

Diese Praktiken sind als „Ghost acting“ oder „Hologram-Acting“ bekannt. Sie können genutzt werden, um sicherzustellen, dass die narrative Kontinuität eines Films nicht unterbrochen wird und die Produktion einheitlich erscheint. VFX-Spezialisten nutzen allerdings eher Archivaufnahmen des verstorbenen Schauspielers, um ein digitales Double zu erstellen, das anschließend verändert und auf den Körper eines Lichtdoubles montiert wird. So hat zum Beispiel *Lucasfilms* nach dem plötzlichen Tod von Carrie Fischer im Jahr 2016 Archivaufnahmen der verstorbenen Schauspielerin genutzt, um sie als Prinzessin Leia in dem Film „Star Wars: The Rise of Skywalker“ aus dem Jahr 2019 wiederauferstehen zu lassen. Das Filmstudio hielt es jedoch für wichtig, darauf hinzuweisen, dass es nicht die Absicht habe, „mit diesem Film einen Trend der Wiederbelebung verstorbener Schauspieler auszulösen.“⁶⁷³

7.3. Persönlichkeitsrechte und ihre Auswirkungen

Zweck des Persönlichkeitsrechts ist es, jeden Menschen in die Lage zu versetzen, sein Erscheinungsbild oder bestimmte Merkmale dieses Erscheinungsbilds in der Öffentlichkeit zu kontrollieren. Geschützt sind zum Beispiel Fotos und Abbildungen von Gesicht und Körper, aber auch der Name und der Stil eines Menschen, seine Signatur, die Stimme oder sogar bestimmte Eigenheiten. Aber auch Accessoires, die untrennbar mit einer Person verbunden sind, können geschützt sein, wie zum Beispiel der Hut oder die Brille eines berühmten Sängers. So wurde zum Beispiel bereits vor Jahren die Verwendung dieses Markenzeichens eines italienischen Sängers für einen Werbespot ohne die Einwilligung des Sängers von italienischen Gerichten als Verletzung seines Persönlichkeitsrechts gewertet.⁶⁷⁴ Zusammengenommen können diese Attribute als „Persona“ oder das „Bild“ eines Menschen bezeichnet werden. Der Begriff „Bild“ bezieht sich in diesem Fall nicht auf eine Abbildung, sondern auf das Erscheinungsbild einer Person in der Öffentlichkeit. Aus diesem Grund werden die Persönlichkeitsrechte in einigen Ländern wie Frankreich und im Vereinigten Königreich auch als „droits d’image“ oder „image rights“ bezeichnet, also als „Imagerechte“. Aber gleichgültig, wie man diese Rechte nennen mag – für den audiovisuellen Sektor können sich daraus einige seltsame rechtliche Komplikationen ergeben.

Die Technik und die Gesellschaft haben sich vor allem in den letzten Jahren so rasant verändert, dass das Recht dieser Entwicklung hoffnungslos hinterherhinkt. Daher

⁶⁷² Cosker D., Eisert P, und Helzle V., *Facial Capture and Animation in Visual Effects*, S. 311-321 in *Digital Representations of the Real World: How to Capture, Model, and Render Visual Reality*, University of Bath, Department of Computer Science, http://cs.bath.ac.uk/~dpc/papers/VFX_2015.pdf.

⁶⁷³ Robinson J., *Star Wars: The Last Jedi – What Happened to Leia?*, Vanity Fair, www.vanityfair.com/hollywood/2017/12/star-wars-the-last-jedi-does-leia-die-carrie-fisher-in-episode-ix.

⁶⁷⁴ *Dalla v Autovox SpA* Pret. Di Roma, 18 apr. 1984, Foro It. 1984, I, 2030, Giur. It. 1985, I, 2, 453.

müssen Juristen bei rechtlichen Auseinandersetzungen, die sich auf die Verwendung des Bildnisses einer Person beziehen, in der Regel auf ältere, stärker etablierte Doktrinen zurückgreifen. Das Persönlichkeitsrecht wird häufig in zwei unterschiedliche Formen der Ausprägung aufgespalten: in ein „Right of publicity“ und ein „Right of privacy“. Das „Right of publicity“ bezieht sich auf die kommerzielle Nutzung des Bilds einer Person, und das „Right of privacy“, das Recht auf Privatsphäre, schützt vor unberechtigtem Eindringen in die Privat- oder Intimsphäre eines Menschen.

An dieser Stelle sollten wir vielleicht kurz innehalten und uns bewusst machen, dass Recht kein statischer Begriff ist, sondern eine überaus differenzierte und nuancenreiche Sprache, deren obskure Regeln und Vorläufer gewisse mythologische Eigenschaften haben, in denen sich die sozialen Normen der Zeit widerspiegeln, aus der sie stammen. Wie Persönlichkeitsrechte genutzt werden, hängt sowohl vom Umfeld als auch der Kultur ab, und selbst innerhalb der Europäischen Union gibt es keinen einheitlichen Ansatz. Was wir hier aus einer rechtlichen Perspektive beobachten, ist ungefähr so, als würden wir einen Blick in einen Nachtclub mit schummeriger Beleuchtung werfen, in dem das Recht am geistigen Eigentum, Markenrecht und Werbung, das Recht der Kontrolle des guten Rufs, das Recht auf freie Meinungsäußerung, das Recht auf Schutz der Würde und der Verbraucherschutz kunterbunt zusammen tanzen – aber nicht unbedingt harmonisch. Im Folgenden soll daher versucht werden, die Tanzfläche auf irgendeine Weise zu organisieren, indem die Persönlichkeitsrechte unter vier verschiedenen Aspekten untersucht werden.

7.3.1. Erster Aspekt: Publicity als (geistiges) Eigentum

Filmstars, Profisportler und andere Prominente können auch über ihre berufliche Arbeit im Studio oder im Stadion hinaus sehr viel Geld verdienen. Die lukrative Verwertung der *Persona* eines Prominenten, ihres Bekanntheitsgrads, kann aus bezahlter Werbung, offiziellem Merchandising oder der Zusammenarbeit mit Modehäusern bestehen. Aber dabei geht es nicht nur um die finanziellen Interessen der prominenten Person selbst: Es geht auch um die Interessen des Managementteams, von Sponsoren, Filmstudios, Plattenlabels und vielem mehr, die alle um den monetären Wert der „*Persona*“ eines Prominenten bemüht sind. Bei diesem ersten Aspekt geht es also um den Publicity-Aspekt der Persönlichkeitsrechte als eine Form des wirtschaftlichen Schutzes.

Geht man von der Arbeitstheorie des Eigentums aus, hat eine Person das Recht auf die Früchte ihrer eigenen Arbeit.⁶⁷⁵ Beim geistigen Eigentum handelt es sich entsprechend um das Recht auf die Früchte geistiger und schöpferischer Arbeit. Eine Form des Rechts auf geistiges Eigentum ist das Urheberrecht, das die Anstrengungen und Investitionen anerkennt, die mit der Produktion schöpferischer Werke verbunden sind, von Spielfilmen bis hin zu Tiktok-Clips und allem dazwischen. Das Urheberrecht soll verhindern, dass solche schöpferischen Werke von anderen kopiert, verändert oder genutzt werden können, ohne

⁶⁷⁵ Hughes J., *The Philosophy of Intellectual Property*, Georgetown University Law Center and Georgetown Law Journal, 77 Geo. L.J. 287.

dass der Schöpfer des Werkes zugestimmt hat, und im Online-Ökosystem sind Klagen gegen Urheberrechtsverletzungen ein häufig genutztes Mittel für schöpferisch Tätige und Studios, die Kontrolle über ihre Inhalte sicherzustellen.

Leider gibt es ein paar praktische Probleme, wenn man versucht, sich dem Persönlichkeitsrecht unter dem Aspekt der Urheberrechtsverletzung zu nähern. Der Versuch, das Urheberrecht von Seiten der abgebildeten Person geltend zu machen, kann zu geradezu paradoxen Komplikationen führen. Erinnern wir uns: Auch die Arbeit des Photographen oder des Videofilmers ist durch das Urheberrecht geschützt, und diese Rechte werden häufig direkt vom Urheber eines Werkes an das Studio oder das Unternehmen abgetreten, das den Film produziert. Daher kann jemand nicht einfach einen Urheberrechtsstreit beginnen, weil er in dem audiovisuellen Inhalt abgebildet wird. Auch wenn es vorkommt, dass Rechteinhaber sich auf einen Urheberrechtsstreit von Seiten der abgebildeten Person einlassen, so ist das nicht der Normalfall, und vor allem nicht, wenn es um einen Rechtsstreit zwischen dem Rechteinhaber und dem Darsteller geht.

Es kann nämlich sein, dass Hunderte oder Tausende von Bildern für die Schaffung eines digitalen Doubles genutzt wurden. Sobald diese Bilder vermischt werden, kann es schwierig werden, die Inhaber des Urheberrechts an den Originalbildern festzustellen – wenn nicht unmöglich. Außerdem gibt es Fälle, in denen die Verwendung von urheberrechtsgeschütztem Material rechtmäßig ist, auch ohne die Genehmigung des Rechteinhabers. Diese Ausnahmen sind zum Beispiel Satire und Parodie und in den Vereinigten Staaten die Umwandlung einer schöpferischen Arbeit, um neue Ausdrucksformen oder Bedeutungen zu schaffen.

7.3.2. Zweiter Aspekt: Publicity und Schutz eines persönlichen Markenzeichens

Auch der zweite Aspekt des Persönlichkeitsrechts bezieht sich auf die kommerzielle Nutzung der Bekanntheit einer Person. Allerdings liegt in diesem Fall der Schwerpunkt nicht auf der Betrachtung des Images einer Person als Eigentum, sondern auf dem Grundsatz der ungerechtfertigten Bereicherung. Nach diesem Grundsatz ist es rechtlich nicht zulässig, von dem Firmenwert oder dem Ansehen einer anderen Person finanziell zu profitieren. So wie zum Beispiel bestimmte Produktsymbole und Namen durch Handelsmarken geschützt sind, um Dienstleistungen und Produkte einer Marke von anderen Wettbewerbern zu unterscheiden, kann auch das Image einer Person geschützt werden, wenn dieses Image dazu genutzt wird, eine bestimmte Qualität, Herkunft oder Zugehörigkeit zu unterscheiden. Aber anders als der markenrechtliche Schutz eines bestimmten Symbols oder eines Satzes ist es unmöglich, jeden Aspekt der *Persona* eines Prominenten zu schützen. Erstens können zum Beispiel Videoaufnahmen nicht markenrechtlich geschützt werden. Und zweitens, selbst wenn eine bestimmte Aufnahme oder ein bestimmtes Bild geschützt ist, kann nicht verhindert werden, dass andere Fotos für Deepfakes oder Ghost Acting genutzt werden. Um diese Lücke im Persönlichkeitsrecht zu schließen, haben einige Länder einen Anspruch auf Wiedergutmachung des finanziellen Schadens eingeführt, der aus dem Missbrauch des Images einer Person entstanden ist.

Dieser Faktor spielt eine wichtige Rolle in Fällen, in denen Produkte mit einer falschen oder irreführenden Einwilligung vermarktet werden. Ein Beispiel hierfür ist der Fall des britischen Modekonzerns Topshop, der ein T-Shirt mit dem Bild der Grammy-Gewinnerin Rihanna verkauft hat. Topshop hatte eine ordnungsgemäße Copyright-Lizenz für die Nutzung des Fotos der Sängerin, aber da Rihanna glaubhaft nachweisen konnte, dass das T-Shirt ihren Ruf in der Modewelt und ihre persönliche Marke geschädigt hat, gewann sie den Rechtsstreit mit Topshop wegen Missbrauchs ihrer Publicity.⁶⁷⁶ In einem ähnlichen Rechtsstreit in Deutschland, bei dem es ebenfalls um das Foto eines Popstars auf T-Shirts ging, sprach der Bundesgerichtshof der Sängerin Nena („99 Luftballons“) Schadensersatz zu, weil ihr Markenzeichen ohne Entgelt genutzt worden war.⁶⁷⁷ In einigen Ländern sind Klagen wegen unbefugter Nutzung einer persönlichen Marke jedoch nur erfolgreich, wenn die betroffene Person nachweisen kann, dass ihr tatsächlich ein finanzieller Schaden entstanden ist. Eine solche Regelung ist mehr als unbefriedigend für Personen, die weniger bekannt sind oder für Personen, die den monetären Wert ihrer *Persona* nicht nachweisen können.

7.3.3. Dritter Aspekt: Schutz der Privatsphäre

Die beiden bisher untersuchten Aspekte haben sich mit der Frage befasst, wie Publicity-Gesetze die Verwendung des Images einer Person für Werbung und andere kommerzielle Zwecke schützen können. Der dritte Aspekt befasst sich mit der Art und Weise, wie Gesetze zum Schutz der Persönlichkeitsrechte unbefugtes Eindringen in die Privatsphäre oder in das Familienleben verhindern können und wie sich diese Gesetze auf Deepfakes und Ghost Acting anwenden lassen. Während Eigentumsrechte auf dem Eigentum an materiellen oder immateriellen Vermögenswerten beruhen, leitet sich das *Right to privacy*, das Recht auf Schutz der Privatsphäre, automatisch aus dem Menschsein ab und gilt daher als Menschenrecht.

Die Europäische Konvention zum Schutz der Menschenrechte (EMRK)⁶⁷⁸ ist das wichtigste Gesetz zum Schutz der Menschenrechte und der politischen Freiheiten in Europa. Sie gilt für alle Mitgliedstaaten der Europäischen Union und für 20 weitere Länder. Nach Artikel 8 EMRK hat jede Person ein grundlegendes Recht auf Achtung seines Privat- und Familienlebens. Dies schließt auch den Schutz vor unbefugtem Eindringen in seinen persönlichen Raum ein. Wer Dokumentarfilme oder Reality-TV kennt, weiß, dass visuelle Bilder dem Zuschauer einen voyeuristischen Blick auf das Leben der anderen ermöglichen.⁶⁷⁹ Fotos und Filmaufnahmen werden daher juristisch mit Vorsicht behandelt und von den Gerichten besonders sorgfältig geprüft.

⁶⁷⁶ Robyn Rihanna Fenty v Arcadia Group Brands Ltd (t/a Topshop), Court of Appeal (Civil Division) - [2015] EWCA Civ 38.

⁶⁷⁷ BGH, Urt. v. 14. Oktober 1986 - VI ZR 10/86.

⁶⁷⁸ Europäische Konvention zum Schutz der Menschenrechte und Grundfreiheiten, www.echr.coe.int/Documents/Convention_DEU.pdf.

⁶⁷⁹ Michael Douglas, Catherine Zeta-Jones & Ors v Hello! Ltd. & Ors [2005] EWCA Civ 595.

Trotzdem bieten Gesetze zum Schutz der Privatsphäre keinen absoluten Schutz für das Image eines Menschen. In den meisten Ländern wird eine Interessenabwägung vorgenommen zwischen dem Recht auf Schutz der Privatsphäre und anderen konkurrierenden Rechten, und häufig wird zugunsten letzterer entschieden – so haben zum Beispiel Nachrichtenmedien das Recht und in einigen Fällen sogar die Pflicht, Fotos von bekannten Persönlichkeiten, die normalerweise als „privat“ gelten, zu veröffentlichen, wenn dies im öffentlichen Interesse ist. Der Schutz der Privatsphäre ist zwar ein grundlegendes Recht nach der EMRK, aber das Recht auf freie Meinungsäußerung nach Artikel 10 ist dies ebenso. Die meisten Länder, in denen die Konvention nicht gilt, haben ähnliche Vorschriften: In den Vereinigten Staaten ist dies in erster Linie der Erste Zusatz zur Verfassung.⁶⁸⁰

In den meisten Fällen, bei denen es um den Schutz der Privatsphäre geht, konzentriert sich die Analyse daher auf die Frage: Wurde das Bild in der Öffentlichkeit aufgenommen oder in einer Umgebung, in der die betreffende Person aus gutem Grund annehmen konnte, dass sie sich in einem privaten Raum befindet? Fotos von Paparazzi, die mit einer Teleskopkamera aus großer Entfernung aufgenommen wurden und eine Prominente zeigen, die sich oben ohne am Pool ihrer abgeschirmten Villa sonnt, werden naturgemäß völlig anders bewertet als Aufnahmen, die von derselben Person in voller Bekleidung auf dem roten Teppich eines Filmfestivals gemacht wurden. Ein Problem, dem wir unter dem Privacy-Aspekt begegnen, ist, dass in der Praxis die Millionen Fotos, die jeden Tag gemacht und veröffentlicht werden, in der Regel irgendwo zwischen diesen beiden Extremen liegen. Es ist nicht leicht, zu sagen, welche Fotos als Eindringen in das Privatleben einer Person angesehen werden können oder als Missbrauch vertraulicher Informationen. Außerdem wird in der Öffentlichkeit häufig die Meinung vertreten, dass es Personen, die finanziell von ihrer Bekanntheit profitieren, durchaus zuzumuten ist, den einen oder anderen Aspekt ihrer Privatsphäre aufzugeben. Ob das fair ist, ist eine andere Frage.

Wenn audiovisuelle Inhalte vertrauliche Einzelheiten enthüllen, zum Beispiel, wenn eine Person in einem Gespräch private Fakten preisgibt, kann es sich um eine Verletzung der Privatsphäre handeln. Allerdings sind Gesetze zum Schutz der Privatsphäre wahrscheinlich kein geeignetes Instrument, um unerwünschte Deepfakes oder Ghost Acting zu verhindern. In erster Linie, weil Bilder und Videos häufig vermischt werden, um ein digitales Double zu schaffen, und festzustellen, welches dieser Bilder ursprünglich privat war, könnte nahezu unmöglich sein. Zweitens, und das ist wahrscheinlich der gewichtigere Grund, bilden Deepfakes etwas ab, was nicht real ist, und Phantasieszenarien können keine Verletzung der Privatsphäre darstellen. Und wenn Aufnahmen genutzt werden, die einen verstorbenen Schauspieler zeigen, würden die Erben oder der Nachlassverwalter eines Schauspielers alles daransetzen, um post mortem die Persönlichkeitsrechte des Schauspielers durchzusetzen. Der Europäische Gerichtshof für Menschenrechte war bei der Anerkennung von Persönlichkeitsrechten Verstorbener bisher stets zurückhaltend, außer wenn ihre Privatsphäre mit noch lebenden Personen verbunden ist.⁶⁸¹ Von einigen Ausnahmen abgesehen hat ein Verstorbener keine nennenswerte Privatsphäre, die man

⁶⁸⁰ First Amendment to the Constitution of the United States, www.constitution.congress.gov/constitution/amendment-1/.

⁶⁸¹ Buitelaar, J.C., *Post-mortem privacy and informational self-determination*, Ethics and Information Technology, 19(2), pp.129–142, <https://link.springer.com/article/10.1007/s10676-017-9421-9>.

verletzen könnte, und weil Persönlichkeitsrechte unveräußerlich sind, können sie weder vererbt noch übertragen werden.

7.3.4. Vierter Aspekt: Persönliche Würde und verwandte Schutzrechte

Die oben diskutierten Aspekte sollten einige hilfreiche Kommentare zu den Publicity- und Privacy-Aspekten des Persönlichkeitsrechts liefern. Obwohl die genannten Gesetze in einigen Fällen angemessenen Schutz bieten, können sie nicht ausschließen, dass diese Rechte missbraucht werden. Um das zu verhindern, können Persönlichkeitsrechte auf das emotionale und seelische Wohlbefinden einer Person ausgedehnt werden, unabhängig von finanziellen Auswirkungen oder Auswirkungen auf die Privatsphäre. Dieser vierte Aspekt befasst sich mit der Vorstellung, dass jeder ein grundlegendes Recht auf Achtung der Menschenwürde und auf persönliche Integrität hat. Es ist fraglich, ob Gesetze zum Schutz der Integrität aufgezeichneter Darstellungen auch auf virtuelle Darbietungen anwendbar sind.

Die Genfer Konvention zum Schutz der Zivilbevölkerung in Kriegszeiten erkennt an, dass „die Achtung der Persönlichkeit und Würde eines Menschen ein allgemein gültiger Grundsatz ist, der auch ohne vertragliche Vereinbarungen rechtlich verbindlich ist“.⁶⁸² Diese ausdrückliche Anerkennung der Würde des Menschen ist auch in der Verfassung der Bundesrepublik Deutschland verankert, ebenso wie in der Rechtsprechung und den Gesetzen der meisten europäischen Länder. Man kann daher aus gutem Grund davon ausgehen, dass diese Gesetze von dem Wunsch inspiriert wurden, die Integrität, Individualität und Selbstbestimmung eines Menschen zu schützen.⁶⁸³ Ein Beispiel für die Anerkennung dieses Grundsatzes ist die vielzitierte Rechtssache *von Hannover gegen Deutschland Nr. 2*,⁶⁸⁴ in der es um Prinzessin Caroline von Hannover geht, die älteste Tochter von Fürst Rainier III. von Monaco. Prinzessin Caroline hatte lange Jahre versucht, gegen die Presse vorzugehen, um die Veröffentlichung von Fotografien aus ihrem Privatleben zu verhindern. Als erneut Fotos von ihr und ihrer Familie ohne ihre Zustimmung veröffentlicht wurden, landete der Rechtsstreit schließlich vor dem Europäischen Gerichtshof für Menschenrechte. In seinem zweiten Urteil im Jahr 2012 erklärte das Straßburger Gericht, dass „das Bild des Einzelnen eines der Hauptmerkmale seiner Persönlichkeit ist, weil es die Besonderheit einer Person zum Ausdruck bringt und ihr ermöglicht, sich von anderen Mitmenschen zu unterscheiden. Das Recht der Person auf Schutz des eigenen Bildes stellt eine der wesentlichen Bedingungen für ihre persönliche Entfaltung dar.“

⁶⁸² Präambel zum Genfer Abkommen (IV) über den Schutz von Zivilpersonen in Kriegszeiten, Genf, 12. August 1949,

<https://ihl->

[databases.icrc.org/ihl/385ec082b509e76c41256739003e636d/6756482d86146898c125641e004aa3c5#:~:text=\(1\)%20Persons%20taking%20no%20active,on%20race%2C%20colour%2C%20religion%20or.](https://ihl-databases.icrc.org/ihl/385ec082b509e76c41256739003e636d/6756482d86146898c125641e004aa3c5#:~:text=(1)%20Persons%20taking%20no%20active,on%20race%2C%20colour%2C%20religion%20or.)

⁶⁸³ Abraham, K. and White E., “The Puzzle of the Dignitary Torts”, *Cornell Law Review* 104 (2), pp.317–380, www.core.ac.uk/download/pdf/228302795.pdf.

⁶⁸⁴ *Von Hannover gegen Deutschland* (nr. 2) 40660/08 [2012] Europäischer Gerichtshof für Menschenrechte 228.

Das Internet macht es sehr leicht, audiovisuelle Inhalte zu veröffentlichen, die Menschen verletzen oder herabwürdigen können. In den sozialen Medien werden solche Inhalte mit exponentieller Geschwindigkeit verbreitet und verspottet. Solche Videos können nicht nur der Selbstachtung und geistigen Gesundheit einer Person schaden, sie können auch erhebliche negative Folgen für Film- oder Unternehmenspartner haben. Obwohl ein relatives Novum in der Rechtswissenschaft, zu dem Thema gehen Argumente für juristische Lösungen im Zusammenhang mit der Menschenwürde heute zunehmend von einem stärker ganzheitlichen Verständnis von der Identität des Menschen aus. In absehbarer Zeit dürfte die juristische Anerkennung dieses Aspekts der Persönlichkeit zu einer Modernisierung des Persönlichkeitsrechts führen und auf diese Weise die Menschen wirksamer vor Schäden durch den Missbrauch der Technik schützen.⁶⁸⁵

Vor Gericht jedoch wird das Konzept der Menschenwürde in Fällen, wo es um die persönliche Ehre des Menschen und seinen Seelenfrieden geht, häufig als bloßer Platzhalter für ein abstraktes Ideal verwandt, und das ist ziemlich frustrierend. Obwohl der Gesetzestext eigentlich klar ist und die Bedeutung der Würde eines Menschen betont, sehen wir häufig, wie Rechtswissenschaftler und Rechtsanwälte krampfhaft versuchen, kohärente Regeln zu definieren, wie ein solches Recht in der Praxis funktionieren soll. In vielen Fällen wird bei der Beilegung solcher Streitfälle, auch im Fall *von Hannover*, letztlich auf Gesetze zum Schutz des Persönlichkeitsrechts oder auf Verleumdungsgesetze zurückgegriffen. Wie oben erläutert, muss die geschädigte Person in vielen Fällen gegen das Recht der Presse auf freie Meinungsäußerung ankämpfen.

Für darstellende Künstler gibt es jedoch so etwas wie einen Hoffnungsschimmer. Während Urheberrechte die Interessen des Urhebers eines Werkes schützen, wird die Leistung von Musikern, Tänzern oder Schauspielern in einer aufgezeichneten Aufführung durch die verwandten Schutzrechte geschützt. Diese Rechte werden als „verwandte“ Schutzrechte bezeichnet, weil sie dem Konzept des Urheberrechts „verwandt“ oder benachbart sind (eine andere Bezeichnung ist auch „Rechte der Darsteller“, um eine Verwechslung mit dem Begriff „Rechte der Autoren“ zu vermeiden). Verwandte Schutzrechte sollen in der Regel sicherstellen, dass die Leistung der ausübenden Künstler vergütet wird. Ihre Aufgabe ist aber auch, Änderungen an einer Aufzeichnung zu verhindern, die den intellektuellen oder persönlichen Interessen eines ausübenden Künstlers schaden könnten. Im Wesentlichen befassen sich verwandte Schutzrechte mit Problemen, die im Zusammenhang mit Raubkopien oder Ausstrahlungen von Shows entstehen, die die Leistung der beteiligten Darsteller nicht angemessen vergüten.

Obwohl die Bedeutung verwandter Schutzrechte für Darstellungen, die über ein audiovisuelles Medium verbreitet werden, offensichtlich ist, werden sie doch nur selten im Zusammenhang mit Persönlichkeitsrechten diskutiert. In der Regel befassen sich Rechtsstreitigkeiten im Zusammenhang mit Persönlichkeitsrechten eher mit dem Missbrauch statischer Bilder, die leicht verändert werden können, und weniger mit Aufzeichnungen von Darbietungen. Aber wie wir gesehen haben, machen es Fortschritte

⁶⁸⁵ Dunn S., *Identity Manipulation: Responding to advances in artificial intelligence and robotics*. Dieser Beitrag war auf der We Robot-Konferenz vom April 2020 als Arbeitspapier vorgestellt und der Autorin Kelsey Farish über private Korrespondenz im Juli 2020 zur Verfügung gestellt worden, mit freundlicher Genehmigung von Miss Dunn, aus dem Dokument zitieren zu dürfen.

bei der Entwicklung der künstlichen Intelligenz und von VFX möglich, Videos so zu manipulieren, dass sie von echten kaum zu unterscheiden sind. Diese neue Ära in der Schaffung audiovisueller Werke, die echte Darbietungen täuschend echt nachahmen, rechtfertigt möglicherweise eine Aktualisierung der Art und Weise, wie in den verwandten Schutzrechten der Begriff „Darstellung“ bisher definiert wurde.

In Kalifornien hat es bereits Fälle gegeben, in denen Gerichte Schadensersatz für die digitale „Versklavung“ von Künstlern in Videospiele gefordert haben. Die Musiker der Band *No Doubt* haben zum Beispiel den Hersteller des Videospiele *Band Hero* verklagt, weil es Spielern ermöglicht, einen äußerst realistischen, digitalen Avatar auszuwählen, um einen Auftritt in einer Rockband zu simulieren.⁶⁸⁶ Wenn Deepfakes- oder Ghost acting-Darstellungen vor Gericht als „Darstellung“ anerkannt werden, können diese Rechte das Image einer Person in Fällen schützen, in denen sie gegen ihren Willen zu einer digitalisierten Darstellung gezwungen wurde. Im Interesse der Vollständigkeit sollten daher verwandte Schutzrechte zusammen mit anderen Rechten zum Schutz der Publicity oder der Privatsphäre als Persönlichkeitsrechte angesehen werden.

7.4. Gesetze in ausgewählten Ländern

Wie in Abschnitt 2 erläutert, werden Persönlichkeitsrechte häufig in zwei umfassende Konzepte aufgespalten. Das erste ist die kommerzielle Verwertung der Bekanntheit und des Ansehens einer Person, das, was als das „Right of publicity“ bezeichnet wird. Publicity-Rechte können wir unter verschiedenen Aspekten untersuchen, die jedoch alle miteinander in Zusammenhang stehen, also Recht des geistigen Eigentums (erster Aspekt, s. oben), und Markenschutzrechte (zweiter Aspekt, s. oben). Getrennt von der kommerziellen „publicity“ versucht das zweite Konzept, das Privacy-Recht, das persönliche Leben eines Menschen zu schützen, indem es die Veröffentlichung bestimmter Bilder durch „privacy laws“, durch Gesetze zum Schutz der Privatsphäre, verbietet (dritter Aspekt, s. oben). Der menschlichen Würde als ein Aspekt der Persönlichkeit wurde in der Diskussion über Persönlichkeitsrechte weniger Aufmerksamkeit gewidmet und den verwandten Schutzrechten noch weniger (vierter Aspekt, s. oben). Ob diese Rechte über den Tod des Inhabers hinaus Geltung haben, wird in den einzelnen Ländern unterschiedlich gehandhabt, aber postmortale Rechte werden in der Regel als eine Form des Eigentums angesehen, das an die Erben weitergegeben werden kann.

Wie künftige Rechtsstreitigkeiten in Bezug auf Deepfakes und Ghost acting geregelt werden, bleibt abzuwarten. In Bezug auf die Gesetze über den Markenschutz und den Schutz des guten Rufs eines Menschen sind zumindest drei der ausgewählten Länder besonders relevant. Das Recht auf Schutz der Menschenwürde und verwandte Schutzrechte dürften ebenfalls eine Rolle spielen und könnten möglicherweise einen aktuelleren Ansatz in Bezug auf den Schutz der *Persona* eines Menschen bieten. Auf der anderen Seite sind das Recht am geistigen Eigentum und die Rechte zum Schutz der Privatsphäre vielleicht weniger wichtig, wenn es darum geht, manipulierte Bilder von Personen, die in audiovisuellen

⁶⁸⁶ *No Doubt gegen Activision Publishing, Inc.*, 122 Cal.Rptr.3d 397 (Cal. Ct. App. 2001).

Inhalten erscheinen, rechtlich einzuordnen. Auf jeden Fall müssen Aspekte wie Publicity, Privacy und die Würde eines Menschen von Fall zu Fall untersucht werden, und jedes Mal muss eine Abwägung mit den konkurrierenden Rechten anderer vorgenommen werden, insbesondere mit dem Recht der freien Meinungsäußerung.

7.4.1. Deutschland

Mit rund EUR 23 Milliarden Umsatz im Jahr 2018⁶⁸⁷ und mehr als 520 000 Beschäftigten⁶⁸⁸ ist Deutschland der größte audiovisuelle Markt in Europa. Die Anerkennung der Persönlichkeitsrechte spielt in Deutschland eine wichtige Rolle und ist in der deutschen Verfassung verankert. Die Tatsache, dass der Schutz der Menschenwürde an erster Stelle im deutschen Grundgesetz steht, ist ein fast perfektes Beispiel für den Schutz der menschlichen Würde unter Aspekt 4: „Die Würde des Menschen ist unantastbar“, und „Jeder hat das Recht auf die freie Entfaltung seiner Persönlichkeit“ heißt es in Artikel 1 und 2 des deutschen Grundgesetzes.⁶⁸⁹ Dieser fundamentale Schutz wird ergänzt durch die Vorschriften des Bürgerlichen Gesetzbuchs. Darin wird festgelegt: „Wer vorsätzlich oder fahrlässig das Leben, den Körper, die Gesundheit, die Freiheit, das Eigentum oder ein sonstiges Recht eines anderen widerrechtlich verletzt, ist dem anderen zum Ersatz des daraus entstehenden Schadens verpflichtet (§ 823 BGB).⁶⁹⁰ In Deutschland kann die Schutzwirkung der Menschenwürde sogar über den Tod eines Menschen hinausgehen. Daher können Angehörige auch eine verstorbene Person gegen Erniedrigung oder diffamierende Darstellung schützen.⁶⁹¹

Alle Aspekte der Identität einer Person fallen unter das allgemeine deutsche Persönlichkeitsrecht. Für Abbildungen und Fotos von Personen gilt jedoch ein erhöhter Schutz. Das deutsche Urheberrechtsgesetz sieht vor, dass Abbildungen einer erkennbaren Person „nur mit Einwilligung des Abgebildeten verbreitet oder öffentlich zur Schau gestellt werden dürfen“ (§ 19a and § 22 UrhG).⁶⁹² Erkennbarkeit umfasst nicht nur Abbildungen des Gesichts, sondern auch anderer Merkmale sowie Cartoons und sogar *Doppelgänger* oder

⁶⁸⁷ Europäische Audiovisuelle Informationsstelle, Jahrbuch 2019/2020 Schlüsseltrends – Fernsehen, Film, Video und audiovisuelle Abrufdienste – Die gesamteuropäische Lage. Europäische Audiovisuelle Informationsstelle, <http://yearbook.obs.coe.int/s/document/key-trends/2019>.

⁶⁸⁸ Weidenbach B., *Beschäftigte in der Medienbranche in Deutschland 2017*, Statista.

⁶⁸⁹ Grundgesetz, formal Grundgesetz für die Bundesrepublik Deutschland, <https://www.gesetze-im-internet.de/gg/>.

⁶⁹⁰ BGB, formal Bürgerliches Gesetzbuch in der Fassung der Bekanntmachung vom 2. Januar 2002 (Bundesgesetzblatt Seite 42, 2909; 2003 Seite 738) zuletzt geändert durch Artikel 4 Absatz 5 des Gesetzes vom 1. Oktober 2013, <http://www.gesetze-im-internet.de/bgb/>.

⁶⁹¹ Seyfert C., *Landgericht Frankfurt am Main: Postmortales Persönlichkeitsrecht einer Holocaust-Überlebenden setzt sich gegen Veröffentlichungen einer britischen Geschichtspräsidentin durch* (Az. 2-03 O 306/19), Zeller & Seyfert Partnerschaft von Rechtsanwälten, <http://www.zellerseyfert.com/de/litigationblog-detail/items/regional-court-frankfurt-am-main-postmortem-personality-right-of-a-holocaust-survivor-prevails-over-publications-by-a-british-hi.html>.

⁶⁹² UrhG, formal Gesetz über Urheberrecht und verwandte Schutzrechte vom 9. September 1965 (Bundesgesetzblatt I, S. 1273), zuletzt geändert durch Artikel 1 des Gesetzes vom 28. November 2018 (Bundesgesetzblatt I, S. 2014), <https://www.gesetze-im-internet.de/urhg/>.

Personen, die der betreffenden Person ähnlich sehen. Wichtig ist dabei die Tatsache, dass der Begriff „Verbreitung“ sehr umfassend interpretiert wird. Das bedeutet, dass sogar ein Hobbyfotograf die Einwilligung der betreffenden Person einholen muss, bevor er ein Bild einer anderen Person in Social Media postet oder mit Freunden über Messenger Apps teilt. Ebenso muss die Einwilligung eines ausübenden Künstlers eingeholt werden, bevor seine aufgezeichnete Darbietung ausgestrahlt oder auf irgendeine andere Weise öffentlich zugänglich gemacht werden darf. (Abschnitt 3 UrhG).

Für Bildnisse aus dem Bereich der Zeitgeschichte ist jedoch eine solche Einwilligung nicht erforderlich, auch nicht für Bilder, auf denen Personen lediglich zufällig erscheinen (als „Beiwerk“), und auch nicht für Bildnisse, deren Verbreitung „einem höheren Interesse der Kunst dient“ (§ 23 KunstUrhG).⁶⁹³ In den letzten Jahren hat sich diese Abwägung jedoch mehr und mehr in Richtung des Schutzes des Inhabers der Bildrechte verschoben.⁶⁹⁴ Was besonders intime Fotos betrifft, zum Beispiel Fotos, die Personen in ihrem eigenen Zuhause zeigen, sieht das Strafgesetzbuch eine Freiheitsstrafe bis zu zwei Jahren oder eine Geldstrafe für Personen vor, die eine solche Aufnahme unbefugt herstellen oder übertragen und dadurch den höchstpersönlichen Lebensbereich der abgebildeten Person verletzen (§ 201a StGB).⁶⁹⁵

Was den Publicity-Aspekt der Persönlichkeitsrechte betrifft, so sind deutsche Gerichte zunehmend bereit, Personen gegen die unerwünschte kommerzielle Verwertung ihres Bildnisses zu verteidigen.⁶⁹⁶ Anders als im Vereinigten Königreich ist die frühere Vermarktung der *persona* keine ausdrückliche Voraussetzung für ein schützenswertes Right of publicity. In Fällen, in denen eine verstorbene Person gezeigt wird, ist die Einwilligung der Angehörigen des Verstorbenen erforderlich, und zwar bis zum Ablauf von 10 Jahren nach dem Tod. In einem Urteil zu Fotos des Filmstars Marlene Dietrich (die mehrere Jahre zuvor verstorben war) entschied der Bundesgerichtshof, dass gegenüber einer unbefugten kommerziellen Verwertung von Bildern dem Träger des Persönlichkeitsrechts ein Schadensersatzanspruch zusteht, unabhängig von der Schwere des Eingriffs.⁶⁹⁷

7.4.2. Frankreich

Frankreich gilt allgemein als das Geburtsland des Films. Das Land hat in vielerlei Hinsicht die produktivste Filmindustrie in Europa: Seit 2015 wurden in Frankreich jedes Jahr rund

⁶⁹³ KunstUrhG, formal Gesetz betreffend das Urheberrecht an Werken der bildenden Künste und der Photographie, Gesetz vom 9. Januar 1907 (RGBl. I S. 7) zuletzt geändert durch das Gesetz vom 16. Februar 2001 (BGBl. I, p. 266), 1 August 2001, www.gesetze-im-internet.de/kunsturhg/_23.html.

⁶⁹⁴ Coors C., „Image Rights of Celebrities vs. Public Interest – Striking the Right Balance Under German Law“, *Journal of Intellectual Property Law & Practice* (2014) 9 (10): 835-840, www.ssrn.com/abstract=2738514.

⁶⁹⁵ StGB, formal Strafgesetzbuch in der Fassung vom 13. November 1998 (Bundesgesetzblatt I S. 3322) zuletzt geändert durch Artikel 3 des Gesetzes vom 2. Oktober 2009 (Bundesgesetzblatt I S. 3214), <https://www.gesetze-im-internet.de/stgb/>.

⁶⁹⁶ Peters M., *The Media and Entertainment Law Review – Germany*, *The Law Reviews*, www.thelawreviews.co.uk/edition/the-media-and-entertainment-law-review-edition-1/1211744/germany.

⁶⁹⁷ BGH, Urteil v. 1. Dezember 1999 - 1 ZR 49/97 (Marlene Dietrich).

300 Filme gedreht.⁶⁹⁸ Die französischen Persönlichkeitsrechte, die sich wörtlich als das „Recht am eigenen Bild“ übersetzen lassen, umfassen Privacy-Gesetze, die eine Person vor ungewolltem Eingriff in die Privatsphäre schützen, sowie kommerzielle Rechte, die eine Verwertung solcher Bilder als vermarktungsfähige Vermögenswerte ermöglichen.⁶⁹⁹ Grundsätzlich gilt, dass die Einwilligung der abgebildeten Person eingeholt werden muss, bevor ihr Bild der Öffentlichkeit zugänglich gemacht wird. Wie in anderen Ländern wird auch in Frankreich der Begriff „image“ umfassend definiert als das Bildnis einer Person, die Stimme, Photographie, Portrait oder Vervielfältigung von Videos. Französische Gerichte haben bestätigt, dass die Unkenntlichmachung des Gesichts eines Models allein nicht ausreicht, um die Verletzung des Rechts am eigenen Bild zu beenden, wenn andere Teile des Körpers sichtbar bleiben.⁷⁰⁰

Diese Philosophie ist tief in der französischen Tradition verwurzelt, wo der Schutz der Privatsphäre und des Familienlebens eine wichtige Rolle spielt. Der *Code Civil*, das französische Zivilgesetzbuch, legt fest, dass jeder das Recht auf Achtung seines Privatlebens hat, und ermächtigt französische Gerichte, alle Maßnahmen anzuwenden, die geeignet sind, um ein unbefugtes Eindringen in die Privatsphäre zu verhindern oder zu beenden (Artikel 9 Code civil).⁷⁰¹ Die französischen Gerichte haben über die Jahre in ihren Urteilen systematisch die Persönlichkeitsrechte für Kläger gestärkt.⁷⁰² Ausschlaggebend für die Durchsetzung der Persönlichkeitsrechte vor Gericht ist die Art der Einwilligung, und die ist nicht immer eindeutig festzustellen. In einem Urteil des Pariser *Tribunal de Grande Instance* (TGI) haben die Richter entschieden, dass nicht davon ausgegangen werden kann, dass die Tatsache, dass ein Model an einem Fotoshooting für Werbezwecke teilgenommen hat, bedeutet, dass sie in alle Formen der kommerziellen Verwertung ihrer Fotos eingewilligt hat.⁷⁰³

Ausnahmen von diesem strengen Schutz gibt es auch in Frankreich. Eine Einwilligung ist nicht erforderlich, wenn die Aufnahmen in der Öffentlichkeit gemacht werden oder wenn Bilder bekannte Persönlichkeiten zeigen, die offizielle Aufgaben wahrnehmen oder Aktivitäten durchführen, die mit ihrer Bekanntheit in Zusammenhang stehen, sofern die Aufnahmen nicht gegen die Menschenwürde der betreffenden Person verstoßen. Parodien und die ausschließlich private Verwendung zählen ebenfalls zu den Ausnahmen. Französische Gerichte sprechen in der Regel auch höheren Schadensersatz zu, wenn Prominente in der Vergangenheit darauf verzichtet haben, ihr Bild kommerziell zu verwerten. Für die Verletzung der Privatsphäre sieht das französische Strafbuch eine Freiheitsstrafe bis zu einem Jahr und eine Geldstrafe in Höhe von EUR 45 000 vor (Artikel

⁶⁹⁸ Lemerrier F., *301 feature films produced by France in 2019*, Cineuropa - the best of European cinema, <https://cineuropa.org/en/newsdetail/387425/>.

⁶⁹⁹ Logeais E. und Schroeder J-B., „The French Right of Image: An Ambiguous Concept Protecting the Human Persona“, *18 Loyola University Entertainment Law Review* 511, <https://digitalcommons.lmu.edu/cgi/viewcontent.cgi?article=1366&context=elr&httpsredir=1&referer=>.

⁷⁰⁰ Mr X v Umanlife, TGI de Paris, Urteil vom 16. November 2018.

⁷⁰¹ Code civil [Bürgerliches Gesetzbuch], www.legifrance.gouv.fr/.

⁷⁰² Sullivan C. L. und Stalla-Bourdillon S., *Digital Identity and French Personality Rights – A Way Forward in Recognizing and Protecting an Individual's Rights in His/Her Digital Identity*, *Computer Law & Security Review* (2015), www.ssrn.com/abstract=2584427.

⁷⁰³ Mrs X v SARL Denim, TGI de Paris, 17. Kammer, Urteil vom 21. November 2018.

226-1 Code pénal).⁷⁰⁴ Über den Bereich des *Privacy law* hinaus sieht das französische Gesetz zum Schutz geistigen Eigentums (*Code de la propriété intellectuelle* – CPI) vor, dass ein ausübender Künstler das Recht auf Achtung seines Namens, seiner Qualität und Darbietung hat (Artikel 212-2 CPI).⁷⁰⁵ Da es sich hierbei um ein ewiges Recht handelt, das mit der betreffenden Person verbunden ist, kann es vererbt werden, und die Erben können die Darbietung und die Erinnerung schützen.

7.4.3. Schweden

Gemessen an seiner Einwohnerzahl hat Schweden eine bemerkenswerte Zahl internationaler Film- und Fernsehstars hervorgebracht, etwa Greta Garbo, Ingrid Bergman, Max von Sydow, Stellan Skarsgård und Alicia Vikander, um nur die bekanntesten zu nennen. Außerdem ist das Land ein Hotspot für das Nordic Noir-Krimigenre. Aus Schweden kommt eine Reihe von Krimiserien, die auch in anderen Ländern großen Erfolg haben, etwa „Wallander“, „The Bridge“ (deutscher Titel: „Die Brücke – Transit in den Tod“) und „Girl With The Dragon Tattoo“ (deutscher Titel: Verblendung). Dass Schweden kein eigenständiges Persönlichkeitsrecht hat, ist daher mehr als überraschend.

Im Gegensatz zu Deutschland und Frankreich kennt das schwedische Recht kein allgemeines Persönlichkeitsrecht.⁷⁰⁶ In Fällen, in denen Medien das Bild eines Menschen widerrechtlich genutzt haben, findet das Gesetz über die Pressefreiheit Anwendung. Aber Schweden tendiert generell dazu, das Recht auf freie Meinungsäußerung über das Recht auf Schutz der Privatsphäre zu stellen. Diese Praxis ist im Vergleich zu vielen anderen Ländern ungewöhnlich. Vielleicht liegt es daran, dass im schwedischen Recht Transparenz und Offenheit eine wichtigere Rolle spielen als der Schutz der Persönlichkeit. Dies gilt auch für den öffentlichen Zugang zu Informationen. Ein Beispiel hierfür ist das umstrittene *Utgivningsbevis*, ein freiwilliges Veröffentlichungszertifikat, das es Unternehmen erlaubt, personenbezogene Daten ihrer Mitarbeiter zu veröffentlichen.⁷⁰⁷ Auch das steuerpflichtige Einkommen der Schweden ist kein Geheimnis – jeder Bürger kann es einsehen.⁷⁰⁸ Seltsamerweise bezieht sich der „Publicity-Grundsatz“ in Schweden nicht auf die Publicity-Aspekte des Persönlichkeitsrechts, sondern darauf, dass Regierungsdokumente zu einem großen Teil nicht geheim sind und der öffentlichen Kontrolle unterliegen.

Trotz dieses mehr als legeren Umgangs mit dem Schutz der Persönlichkeit kann auch in Schweden das Bild eines Menschen nicht ohne Einwilligung genutzt werden. Das

⁷⁰⁴ Code penal [Französisches Strafgesetzbuch], www.legifrance.gouv.fr/.

⁷⁰⁵ CPI, Code de la propriété intellectuelle [Gesetz zum Schutz des geistigen Eigentums], www.legifrance.gouv.fr/.

⁷⁰⁶ Ondreasova E., „Personality Rights in Different European Legal Systems: Privacy, Dignity, Honour and Reputation“, *The Legal Protection of Personality Rights*, pp.24–70. Brill | Nijhoff, https://brill.com/view/book/edcoll/9789004351714/B9789004351714_004.xml.

⁷⁰⁷ Herlin-Karnell E., *Corona and the Absence of a Real Constitutional Debate in Sweden*, Verfassungsblog on Matters Constitutional, www.verfassungsblog.de/corona-and-the-absence-of-a-real-constitutional-debate-in-sweden/.

⁷⁰⁸ Marçal, K., *Sweden shows that pay transparency works*, *The Financial Times*, www.ft.com/content/2a9274be-72aa-11e7-93ff-99f383b09ff9.

schwedische Gesetz über Bilder- und Namensnutzung in der Werbung schreibt vor, dass die Einwilligung einer Person eingeholt werden muss, bevor ihr Name oder Bild oder eine Darstellung, die eine eindeutige Identifizierung dieser Person ermöglicht, für kommerzielle Zwecke genutzt werden können (*Lag om namn och bild i reklam* (1978:800)).⁷⁰⁹ Nach dem schwedischen Schadensersatzgesetz muss für immaterielle Schäden ein Schadensersatz gezahlt werden, aber nur, wenn die Verletzung eine Straftat darstellt oder die Gesundheit und das Leben einer Person in Gefahr bringt [*Skadeståndslag* (1972:207), Kapitel 2].⁷¹⁰ Verleumdung stellt in Schweden nach dem Strafgesetzbuch (*Brottsbalken*)⁷¹¹ zwar eine Straftat dar. Geahndet wird sie jedoch nur, wenn jemand einen anderen fälschlich oder „ohne vernünftigen Grund“ beschuldigt, „ein Krimineller zu sein oder einen verwerflichen Lebenswandel zu führen oder Informationen liefert, um die betreffende Person der Missachtung durch andere auszusetzen“ (*Brottsbalken*, Kapitel 5 §1). Das Kriminalgesetzbuch ermöglicht es auch Familienmitgliedern oder dem Staatsanwalt, ein Verfahren wegen „Störung der Ruhe“ eines Verstorbenen einzuleiten, wenn dies im öffentlichen Interesse ist (*Brottsbalken* Kapitel 5 § 4). Sieht man von diesen wenigen strengeren Schutzbestimmungen ab, sind die Lücken im schwedischen Persönlichkeitsrecht nur allzu offensichtlich, und die schwedischen Richter haben bisher wenig Interesse gezeigt, sie in ihren Urteilen zu füllen.⁷¹²

7.4.4. Guernsey

Die *Bailiwick of Guernsey* („Vogtei Guernsey“) ist eine Insel im Ärmelkanal vor der französischen Nordseeküste. Die Insel ist zwar britischer Kronbesitz, aber sie hat ein Rechtssystem, das sich von dem des Vereinigten Königreichs erheblich unterscheidet. Die Insel hat zwar nur 63 000 Einwohner, aber ihr unternehmensfreundliches Steuersystem, die Schönheit der Landschaft und die englischsprachige Bevölkerung haben dazu beigetragen, dass sich auf Guernsey eine blühende Filmproduktion und ein erfolgreicher Filmfinanzsektor entwickeln konnten.⁷¹³ Was im Zusammenhang mit dem Schutz des Persönlichkeitsrechts von Interesse ist, ist die Tatsache, dass Guernsey 2012 mit seinem *Image Rights of Bailiwick Guernsey Ordinance* ein Gesetz zur Registrierung von Persönlichkeitsrechten eingeführt hat, das weltweit einmalig ist.⁷¹⁴ Vereinfacht gesagt funktioniert dieses Gesetz ähnlich wie ein Markenschutzgesetz.

Geschützt werden Fotos und Bilder von Personen, aber auch Filmaufnahmen, der Name einer Person, Stimme, Signatur, Abbildungen, Eigenheiten und persönliche Attribute,

⁷⁰⁹ *Lag om namn och bild i reklam* (1978:800) [Gesetz über die Bilder- und Namensnutzung in der Werbung], englische Übersetzung, www.kb.se/Dokument/Bibliotek/biblioteksjuridik/names_pictures.pdf.

⁷¹⁰ *Skadeståndslag* (1972:207) [Schwedisches Schadensersatzgesetz], www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/skadestandslag-1972207_sfs-1972-207.

⁷¹¹ *Brottsbalken*, SFS 1962:700 [Schwedisches Kriminalgesetz], englische Übersetzung, www.legislationline.org/download/id/1700/file/4c405aed10fb48cc256dd3732d76.pdf.

⁷¹² Ondreasova E., *op.cit.*

⁷¹³ Tustin B., *Guernsey Film: More Than You Might Think*, Mondaq, www.mondaq.com/guernsey/film-television/171220/guernsey-film-more-than-you-might-think.

⁷¹⁴ *Image Rights (Bailiwick of Guernsey) Ordinance*, 2012.

wie zum Beispiel eine Sporttrikotnummer. Nicht nur lebende Personen können sich registrieren lassen, sondern auch Personen, die in einem Zeitraum bis 100 Jahre vor der Anmeldung verstorben sind, aber auch Gruppen und Mannschaften oder sogar fiktive Charaktere. Die Antragsteller müssen Inhaber der betreffenden Persönlichkeit sein. Dabei kann es sich um die Person selbst oder ihren autorisierten Vertreter oder Erben handeln.

Ein Verstoß gegen dieses Gesetz liegt dann vor, wenn das geschützte Bild (oder ein Bild, das diesem ähnlich sieht) ohne Einwilligung des Inhabers „für kommerzielle Zwecke oder finanzielle oder wirtschaftliche Vorteile“ genutzt wird und diese Nutzung entweder zu einer Verwechslung mit der registrierten Person führt oder den guten Ruf der abgebildeten Person verletzt. Bei der Bewertung des Schadens prüft das Gericht alle relevanten Faktoren, um sämtliche wirtschaftliche Folgen und Einnahmenverluste zu berücksichtigen, aber auch immaterielle Schäden, die dem Opfer entstanden sind. Die Registrierung bezieht sich auf einen Zeitraum von zehn Jahren, sie kann um weitere zehn 10 Jahre verlängert werden. Der Antragsteller muss zwar nicht in Guernsey seinen Wohn- oder Firmensitz haben. Verstöße gegen das Gesetz werden jedoch nur geahndet, wenn sie in Guernsey erfolgt sind. Allerdings wurde das Gesetz mit modernen Medien gestaltet und mit grenzüberschreitenden digitalen Diensten.⁷¹⁵ Daher wäre es klug, eine sorgfältige Recherche durchzuführen, bevor registrierte Bilder oder Aufnahmen weiterverbreitet werden.

7.4.5. Vereinigtes Königreich

Zum Zeitpunkt der Abfassung dieses Beitrags ist die Regierung des Vereinigten Königreichs damit beschäftigt, das Land von dem politischen und wirtschaftlichen System der Europäischen Union abzukoppeln. Trotz des Brexit werden britische Fernsehshows und Filme auch in Zukunft auf dem audiovisuellen Markt in Europa eine wichtige Rolle spielen und von Zuschauern in ganz Europa gesehen werden. Das Vereinigte Königreich kennt genauso wie Schweden kein formales Persönlichkeitsrecht. Und anders als in vielen anderen europäischen Ländern haben die Gerichte in England bisher jeder Versuchung widerstanden, solche Rechte in ihrer Rechtsprechung anzuerkennen.

Um aus dem Urteil in dem bereits genannten Fall von Rihanna zu zitieren: „In England gibt es heute kein eigenständiges allgemeines Recht für Prominente (oder auch andere Personen), das es ihnen ermöglicht, die Vervielfältigung ihres Bilds zu kontrollieren.“ Richter im Vereinigten Königreich verlassen sich lieber auf traditionelle Rechtsinstrumente anstatt in ihren Urteilen auch nur ansatzweise so etwas wie Persönlichkeitsrechte zu gestalten. Trotzdem gibt es mehrere Möglichkeiten, wie Personen ihr Bild schützen können. Diese konzentrieren sich normalerweise auf die Enthüllung privater Daten und auf Schädigung der Geschäftsinteressen oder des guten Rufs einer Person.

⁷¹⁵ Shires S (2015), *Guernsey image rights exposed*, Lexology, www.lexology.com/Library/detail.aspx?g=9829178f-e93d-4ead-94d0-bfbfaa81f8b7.

Der Schutz der Privatsphäre wird im Vereinigten Königreich nicht als ein eigenständiges Recht angesehen. Die Gründe hierfür liegen in der Art und Weise, wie sich das *Common law* im Vereinigten Königreich entwickelt hat, aber auch in der Tatsache, dass es keine geschriebene Verfassung gibt. Allerdings erkennen die Gerichte die Tatsache an, dass die Privatsphäre ein wichtiger sozialer Wert ist. Dies hat dazu geführt, dass besondere Gesetze gegen den Vertrauensbruch oder den Missbrauch besonders sensibler Informationen entwickelt wurden. 1998 wurde das Human Rights-Gesetz verabschiedet, mit dem die Europäische Konvention zum Schutz der Menschenrechte in britisches Recht umgesetzt wurde. Seit dem Ende der 1990er Jahre haben die Richter in ihren Urteilen daher erstmals dem Schutz der Privatsphäre größere Bedeutung beigemessen. Allerdings haben die Richter in den seltenen Fällen, in denen Fragen des Schutzes der Privatsphäre überhaupt vor englischen Gerichten verhandelt wurden, dazu tendiert, den Schutz der Meinungsfreiheit und andere Pressefreiheiten über den Schutz der Privatsphäre zu stellen. Ausnahmen von diesem Grundsatz betreffen Bilder, die vertrauliche Informationen darstellen oder Kinder oder besonders intime Szenen zeigen.

Wenn es um den Schutz der Publicity-Rechte als kommerzielle Werte geht, wird in der Regel auf die markenrechtliche Kennzeichenverletzung (tort of „passing off“) zurückgegriffen. Der Erfolg hängt jedoch zunächst davon ab, ob die prominente Person „einen signifikanten guten Ruf oder Goodwill“ hat.⁷¹⁶ Unter dem Aspekt der Würde und der Reputation kann die unautorisierte Nutzung des Bilds einer Person nach dem Gewohnheitsrecht als böswillige Verleumdung gewertet werden und stellt dann eine Straftat dar - aber nur, wenn sie falsche Behauptungen enthält, die zu quantifizierbaren finanziellen Verlusten führen.⁷¹⁷ Falschaussagen, die dem guten Ruf einer Person schaden, können eine Verleumdung darstellen, aber nach der Reform des *Defamation Act 2013* (Verleumdungsgesetz)⁷¹⁸ nur, wenn sie der abgebildeten Person „ernsthaften Schaden“ zufügen.

Diese Feststellungen sollten auch im Zusammenhang mit dem *Copyright, Designs and Patents Act* (CDPA – dem britischen Urheberrechtsgesetz)⁷¹⁹ gesehen werden. Dieses Gesetz erlaubt es einem ausübenden Künstler unter bestimmten Umständen, „gegen die nachteilige Bearbeitung seiner Darbietung Einspruch zu erheben“ und sicherzustellen, dass seine Darbietung „... ohne Verzerrung oder andere Veränderungen, die dem Ruf des darstellenden Künstlers schaden“ gezeigt wird (CDPA, s. 205). Bevor Aufzeichnungen von Darbietungen genutzt werden dürfen, muss die Einwilligung des Schauspielers, Musikers, Tänzers oder anderer darstellender Künstler eingeholt werden (CDPA, s. 182). Diese Regelung geht auf einen konkreten Fall zurück, und zwar auf Clips, die den Schauspieler Peter Sellers zeigen. Diese wurden nach seinem Tod genutzt, um einen neuen Pink Panther-Film zu produzieren. Seine persönlichen Vertreter setzten sich mit Erfolg für die postmortalen Rechte des Schauspielers ein, um die unautorisierte Nutzung seiner Darbietung zu verbieten.⁷²⁰

⁷¹⁶ Edmund ‘Eddie’ Irvine v Talksport [2002] EWHC 367 (Ch).

⁷¹⁷ Marathon Mutual Ltd v Waters [2009] EWHC 1931 (QB).

⁷¹⁸ Defamation Act 2013, www.legislation.gov.uk/ukpga/2013/26/contents/enacted.

⁷¹⁹ Copyright, Designs and Patents Act 1988, www.legislation.gov.uk/ukpga/1988/48/contents.

⁷²⁰ *Rickless gegen United Artists Corporation* [1988] QB 40.

7.4.6. Kalifornien

Es gibt in den Vereinigten Staaten von Amerika auf Bundesebene kein Recht auf Schutz der Privatsphäre: Der Schutz der Privatsphäre wird unter Bezugnahme auf bestimmte Sektoren oder Themen geregelt, etwa Finanz- oder Gesundheitsdaten. Es gibt zwar auf Bundesebene Gesetze zum Schutz des geistigen Eigentums und zum Schutz vor unlauterem Wettbewerb und vor unlauterer Werbung. Aber der Schutz von Publicity- und Privacy-Rechten fällt größtenteils in den Kompetenzbereich der einzelnen Bundesstaaten. Kalifornien ist die Heimat von Hollywood und Silicon Valley und daher wie kein anderer amerikanischer Staat prädestiniert, den Zeitgeist von Prominenz und Technologie zu repräsentieren.

Die Publicity-Rechte von Prominenten werden in Kalifornien sowohl auf der Grundlage des Zivilrechts als auch auf der Grundlage des *Common law* geschützt, dank seiner weltberühmten Unterhaltungsindustrie und der kalifornischen Neigung, Emotionen zuzulassen und zu respektieren. Der *California Civil Code* (das Bürgerliche Gesetzbuch Kaliforniens, CIV)⁷²¹ verbietet die unbefugte Nutzung des Namens, der Signatur, der Stimme, der Fotografie oder des Bildnisses einer anderen Person für Werbezwecke ohne die Einwilligung der betreffenden Person (CIV § 3344). Im Falle der Verurteilung hat die geschädigte Person Anspruch auf Schadensersatz, entweder USD 750 (ungefähr EUR 635 im August 2020) oder die Erstattung des tatsächlich erlittenen Schadens – je nachdem, welche Summe höher ist, außerdem hat der Geschädigte Anspruch auf Gewinnherausgabe. Mit dem *Fred Astaire Celebrity Image Protection Act* wurde nachträglich § 3344 CIV geändert, um die postmortale kommerzielle Nutzung des Namens eines Verstorbenen, seines Bilds oder seiner Stimme für einen Zeitraum von 70 Jahren nach seinem Tod zu schützen.⁷²² Allerdings beschränkt sich dieser Artikel ausschließlich auf Merchandising, Werbung und Waren. Unterhaltung (Fiction und Nonfiction), Bühnenwerke, literarische und musikalische Werke sind von der Haftung ausgeschlossen. Das Publicity-Recht auf der Grundlage des *Common law* gilt für jede Art von Missbrauch der Identität einer Person. Das ist sehr viel umfassender als die besonderen Merkmale, die in dem Gesetz, „zum Vorteil des Klägers, kommerzieller oder anderer Art“, aufgeführt werden⁷²³

Kalifornische Gerichte erkennen an, dass das Recht einer Person, eine unautorisierte und unbefugte Nutzung ihrer Publicity-Rechte zu verbieten, ein Aspekt des Privacy-Rechts ist.⁷²⁴ Es gibt mehrere Gesetze zum Schutz der Privatsphäre, die auf den Missbrauch des Bilds einer Person Anwendung finden, und jedes davon ist anders, je nachdem, ob es sich um einen wirtschaftlichen Schaden handelt oder um eine Verletzung der Würde des Menschen. Falsche Angaben, die den guten Ruf einer Person verletzen, fallen unter den Straftatbestand der Verleumdung, und eine Berichterstattung, die zwar sachlich nicht falsch ist, aber trotzdem dem seelischen oder emotionalen Wohlbefinden des Opfers Schaden zufügt, können ein „*False Light*-Delikt“ darstellen, durch das das Opfer in der Öffentlichkeit

⁷²¹ CIV (California Civil Code), www.codes.findlaw.com/ca/civil-code/.

⁷²² Fred Astaire Celebrity Image Protection Act, dazu siehe § 3344.1 des California Civil Code.

⁷²³ *Clint Eastwood gegen National Enquirer*, Superior Court, 149 Cal.App.3d 409 (Cal. Ct. App. 1983).

⁷²⁴ *Fairfield gegen American Photocopy Equipment Co.*, (1955) 138 Cal. App. 2d 82, 291 P.2d 194

in ein „falsches Licht“ gerückt wird.⁷²⁵ Da in Kalifornien eine sehr große Zahl von Prominenten lebt, die häufig im Rampenlicht stehen (sowohl buchstäblich als auch im übertragenen Sinn), hat der Staat eine der klägerfreundlichsten Rechtsprechungen, wenn es um die Verletzung des Persönlichkeitsrechts geht.

Allerdings darf man nicht außer Acht lassen, dass der konkurrierende Schutz der freien Meinungsäußerung, der durch den Ersten Zusatz zur amerikanischen Verfassung eingeführt wurde, auch in Kalifornien ein sehr starkes Recht darstellt, auch für Bilder und Videos, die von mächtigen Filmkonzernen veröffentlicht werden. Ein Beispiel hierfür ist die Klage von Dustin Hoffman gegen eine Zeitschrift, die ein digital manipuliertes Foto in einer Modestory veröffentlicht hatte, das den Schauspieler in Designer-Mode zeigte. Die Gerichte wiesen die Klage zurück mit der Begründung, dass die Zeitschrift nicht die Absicht gehabt habe, dem Schauspieler zu schaden, außerdem sei das Foto nicht ausschließlich aus kommerziellen Gründen veröffentlicht worden. Aus diesen Gründen entschieden die Richter, dass in diesem Fall das Recht auf freie Meinungsäußerung das Schutzinteresse des Betroffenen überwiegt.⁷²⁶

Prominente, einschließlich Personen, die bei Ereignissen von öffentlichem Interesse eine Rolle spielen, haben alles in allem einen geringeren Anspruch auf Schutz ihrer Privatsphäre als normale Bürger. Wichtig für alle diejenigen, die versuchen, das Aussehen eines Schauspielers durch Ghost Acting oder ähnliche Techniken zu verändern: Kalifornien nutzt einen „transformative work test“, um festzustellen, ob die Nutzung des Bilds einer Person durch den Ersten Verfassungszusatz geschützt ist. Nach diesem Test ist es so, dass je mehr Originalaufnahmen verändert werden, um eine neue Bedeutung zu schaffen oder eine andere Botschaft zu liefern, desto wahrscheinlicher ist es, dass der Urheberrechtsschutz in diesem Fall nicht greift (Copyright Code, § 107).⁷²⁷ Insbesondere in Bezug auf das Bild von Personen prüfen kalifornische Gerichte, ob „das Aussehen eines Prominenten so verändert wird, dass es in erster Linie der eigene Ausdruck des Angeklagten ist und nicht das Abbild der prominenten Person.“⁷²⁸

⁷²⁵ In Common-law-Rechtssystemen kann eine Veröffentlichung, die falsch ist und den Ruf einer Person schädigt, eine Verleumdung darstellen. In einigen US-Bundesstaaten einschließlich Kalifornien gibt es jedoch ein eigenständiges Delikt (legal harm – rechtliche Schäden), das für Äußerungen gilt, die zwar nicht sachlich falsch sind, aber trotzdem irreführend sein können. Gerichte in Kalifornien erkennen an, dass falsche Angaben (*implications*) dazu führen können, dass von einer Person ein falscher und letztendlich schädlicher Eindruck in der Öffentlichkeit entsteht. Diese Unterscheidung ist sehr subtil und nuanciert, aber allgemein kann man sagen, dass der Tatbestand der Verleumdung versucht, den Schaden am Ruf einer Person wiedergutzumachen. Wenn die Äußerung jedoch der Wahrheit entspricht, kann sie nicht als Verleumdung gewertet werden. Umgekehrt kann eine sachlich richtige Behauptung oder ein Foto „jemanden in ein falsches Licht (false light) rücken“. Kaliforniens „False light“-Delikt versucht daher, Schadenersatz für die Verletzung von Gefühlen festzusetzen – unabhängig davon, ob eine Behauptung wahr ist. CACI (Judicial Council of California Civil Jury Instructions) 2017 edition. No. 1802. False Light, <https://www.justia.com/trials-litigation/docs/caci/1800/1802/>.

⁷²⁶ *Dustin Hoffman gegen Capital Cities/ABC, Inc.*, 255 F. 3d 1180 (9th Cir. 2001).

⁷²⁷ Copyright Code (Urheberrechtsgesetz), formal U.S. Code, Title 17 Copyrights, § 107 Limitations on exclusive rights: Fair use, www.law.cornell.edu/uscode/text/17/107.

⁷²⁸ *Comedy III Prods., Inc. gegen Gary Saderup, Inc.* - 25 Cal. 4th 387, 106 Cal. Rptr. 2d 126, 21 P.3d 797 (2001).

7.5. Wie geht es weiter mit dem audiovisuellen Sektor in Europa?

Dieses Kapitel hat versucht, einen Einblick in die Vielschichtigkeit des Persönlichkeitsrechts zu geben, das charakteristischerweise unterschiedliche Publicity- und Privacy-Rechte umfasst. In diesen Rahmen fügen sich auch das Recht auf Menschenwürde und das Recht auf persönliche Integrität ein, zusammen mit den verwandten Schutzrechten von ausübenden Künstlern. Aber unabhängig davon, wie diese Gesetze theoretisch eingeordnet werden, es gibt einige praktische Tipps und Beobachtungen, die dabei helfen könnten, die Risiken im Zusammenhang mit Deepfakes und Ghost Acting abzumildern.

Einige Länder haben bereits spezielle Gesetze erlassen, die die Verwendung von Deepfake-Technologie verbieten, um sexuellen Missbrauch auf der Grundlage von Bildern (image-based sexual abuse) oder die Manipulation von Wahlen zu verhindern. Der Staat Virginia war der erste, der seine so genannten „revenge porn“-Gesetze aktualisiert hat (so genannte Rache-Pornos). In Virginia wird es als ein Vergehen gewertet, wenn manipulierte Fotos oder Videos, die eine Person nackt darstellen oder ihre Genitalien zeigen, ohne die Einwilligung der betreffenden Person veröffentlicht werden.⁷²⁹ Texas hat als erster Staat die Verbreitung von Deepfake-Videos von Kandidaten im Wahlkampf verboten, wenn diese den Kandidaten beleidigen oder die Wahlen zu beeinflussen versuchen.⁷³⁰ Natürlich können Deepfakes auch Schaden anrichten, ohne dass sie sexuellen oder politischen Charakter haben.

Technologiekonzerne gehen zunehmend dazu über, Deepfakes in ihrem Zuständigkeitsbereich zu reglementieren. *Facebook* und sein Schwesterunternehmen *Instagram* haben Deepfakes offiziell ‚verbannt‘, das gilt auch für *Twitter*, *Reddit* und *Pornhub*. *TikTok* hatte zwar anfänglich den Deepfakes-Trend übernommen, aber die Technik ebenfalls ‚verbannt‘ nach Übernahmediskussionen mit US-Konzernen.⁷³¹ Dass „verbannt“ in Anführungszeichen gesetzt wurde, ist kein Zufall: Unabhängig von offiziellen Statements oder Maßnahmen ist die Entfernung oder Veränderung von Deepfakes sehr schwierig. Das heißt, das Verbot von Deepfakes besteht häufig nur formal.

Zunächst einmal ist es sehr schwierig, ein Deepfake-Video von einem echten Video zu unterscheiden, wie die *Deepfake Detection Challenge* unter der Führung von *Facebook*, *Microsoft* und einer kleinen Armee von Experten der künstlichen Intelligenz feststellen musste.⁷³² So wird zum Beispiel versucht, Quellendateien mit Metadaten anzureichern und durch digitale Wasserzeichen zu markieren, aber bisher ist noch keine praktikable Lösung in Sicht. Zweitens, selbst wenn ein Deepfake aufgedeckt wird, ist es praktisch unmöglich,

⁷²⁹ Virginia House Bill 2678, a Bill to amend and reenact § 18.2-386.2 of the Code of Virginia, relating to unlawful dissemination or sale of images of another; falsely created videographic or still image, <https://law.lis.virginia.gov/vacode/title18.2/chapter8/section18.2-386.2/#:-:text=Any%20person%20who%2C%20with%20the,or%20female%20breast%2C%20where%20such..>

⁷³⁰ Texas Senate Bill 751, a Bill Relating to the creation of a criminal offense for fabricating a deceptive video with intent to influence the outcome of an election, www.legiscan.com/TX/text/SB751/2019.

⁷³¹ Statt N., *TikTok is banning deepfakes to better protect against misinformation*, The Verge, www.theverge.com/2020/8/5/21354829/tiktok-deepfakes-ban-misinformation-us-2020-election-interference.

⁷³² Wiggers K., *Facebook, Microsoft, and others launch Deepfake Detection Challenge*. VentureBeat.

den Urheber eines Deepfakes festzustellen, da es sehr leicht ist, im Internet anonym zu bleiben. Neben diesen technischen Problemen muss man auch die Probleme berücksichtigen, die den Kontext und die Absicht umgeben. Der Begriff „Deepfake“ bezieht sich auf die Technologie, nicht auf die Substanz: Medizin und Bildung sind zum Beispiel zwei Bereiche, in denen diese Technologie überaus segensreich eingesetzt werden kann.⁷³³ Kurz gesagt, es gibt nicht immer eine klare Trennlinie zwischen einem unerwünschten Deepfake und einem akzeptablen Deepfake.

Unabhängig von der Frage, wie „schlechte“ Deepfakes sich von „guten“ trennen lassen, verbreiten sich neuartige und unterhaltsame Inhalte mit rasender Geschwindigkeit im Netz, auch wenn sie gefälscht oder irreführend sind. Es mag zynisch klingen – aber Plattformen, die von Werbeklicks und Seitenaufrufen leben, können es für gerechtfertigt halten, aufregende Inhalte nicht zu löschen, wenn sie ihren Geschäftsinteressen dienen. Außerdem kann sich selbst dann, wenn ein Deepfake eigentlich entfernt werden müsste, die Entfernung negativ auf die Meinungsfreiheit auswirken. Und nicht zuletzt könnte ein Video, selbst wenn es von einer Plattform entfernt wurde, immer noch irgendwo anders auftauchen. Mehr noch, der so genannte Streisand-Effekt⁷³⁴ hat gezeigt, dass wenn man versucht, eine Information zu unterdrücken, das Gegenteil erreicht wird: Sie verbreitet sich umso stärker.

Wenn audiovisuelle Inhalte online gestellt werden, kann es schwierig werden, zu verhindern, dass sie auf eine Art und Weise genutzt werden, die nicht beabsichtigt war. Trotzdem ist es nach wie vor für private Parteien möglich, sich vertraglich darauf zu einigen, wie die Persönlichkeitsrechte verwertet oder geschützt werden sollen. Ausübende Künstler schließen häufig für jedes getrennte Projekt einen neuen Vertrag ab, und wenn von Anfang an klare Bedingungen festgelegt werden, kann dies helfen, spätere Auseinandersetzungen zu vermeiden. Diese Vereinbarungen müssen noch nicht einmal in juristischer Fachsprache abgefasst werden, aber intelligenterweise sollte schriftlich festgelegt werden, auf welche Art und Weise das Bild einer Person genutzt, verbreitet und an andere weitergegeben werden darf.

Beide Parteien täten gut daran, sorgfältig zu überlegen, wie der Inhalt über die Originalproduktion hinaus genutzt werden soll, ebenso wie diese Nutzung zusätzlich vergütet werden sollte. Wenn ein Filmstudio frühere Aufnahmen nutzt, um sie für neue Zwecke zu verwerten, ist es unwahrscheinlich, dass der ausübende Künstler eine zusätzliche Vergütung dafür erhält, wenn dies nicht ausdrücklich vertraglich geregelt wurde. Was den guten Ruf eines Darstellers betrifft, so kann es auch sinnvoll sein, Klauseln zur Verhinderung einer Herabwürdigung oder so genannte Moralklauseln („morality

⁷³³ Kalmykov M., *Positive Applications for Deepfake Technology*, Hackernoon.com, www.hackernoon.com/the-light-side-of-deepfakes-how-the-technology-can-be-used-for-good-4hr32pp.

⁷³⁴ Der Streisand-Effekt ist benannt nach der amerikanischen Schauspielerin und Sängerin Barbara Streisand. Sie hatte 2003 gegen den Fotografen Kenneth Adelman geklagt, um die Veröffentlichung von Fotos zu verhindern, die der Fotograf von ihrem Haus am Strand von Malibu gemacht hatte. Die Fotos von Adelman waren jedoch für das *California Coastal Records Project* bestimmt, ein wissenschaftliches Projekt zur Erforschung der Küstenerosion, das Fotos von der kalifornischen Küste für wissenschaftliche Zwecke liefert. Die Klage von Barbara Streisand wurde abgewiesen, und als Ergebnis der Publicity um den Rechtsstreit wuchs das Interesse an Fotos ihres Hauses. Siehe Cacciottolo M., *The Streisand Effect: When censorship backfires*, BBC News, <https://www.bbc.co.uk/news/uk-18458567>.

clauses“) in den Vertrag aufzunehmen, um zu verhindern, dass eine der Parteien den guten Ruf der anderen schädigt. Für den Schauspieler wird eine solche vertragliche Festlegung vor allem in Ländern wie Schweden, dem Vereinigten Königreich und den Vereinigten Staaten wichtig sein, denn in diesen Ländern ist es sehr schwierig, den Tatbestand der Verleumdung oder der Verletzung der menschlichen Würde festzustellen.

Aber auch unter rein finanziellen Gesichtspunkten können sich Probleme ergeben. So kann eine nicht gewollte Veränderung des Erscheinungsbilds dazu führen, dass die Zusammenarbeit zwischen einem Schauspieler und denjenigen, die hinter der Kamera arbeiten gestört oder sogar unmöglich gemacht wird. Es ist kein Geheimnis, dass kreative Differenzen und Forderungen, seien sie gerechtfertigt oder nicht, dazu führen, dass eine Filmproduktion aus dem Ruder läuft oder gar abgebrochen werden muss. Der Ausstieg eines Filmstars mitten in den Dreharbeiten oder die Notwendigkeit von Neuaufnahmen kann für ein Filmstudio sehr teuer werden. Substanzielle Veränderungen über die künstlerische Notwendigkeit hinaus sollten daher vorab geklärt werden, vor allem, wenn solche Änderungen die Gefühle oder Interessen des Schauspielers verletzen könnten.

Abschließend sollte noch darauf hingewiesen werden, dass die Schaffung eines digitalen Doubles oder Ghost Acting – so aufregend und innovativ diese Technik auch sein mag – unvermeidliche Risiken mit sich bringt. Wenn wir uns in aller Ruhe einen Film oder eine Fernsehshow ansehen wollen, dann erwarten wir Illusion, um Ingmar Bergman zu zitieren. Wir schalten Willen und Verstand aus, um uns mental auf die Entfaltung einer fiktiven Story einzustellen.⁷³⁵ Aber die Echtheit digital geschaffener Gesichter ist etwas ganz anderes, denn sie ist in der Lage, die Autonomie und Selbstbestimmung eines Schauspielers außer Kraft zu setzen. Sie kann auch die bewertenden Prozesse oder Entscheidungsfindungsprozesse der Zuschauer manipulieren. Wenn Personen in nicht-fiktionalen Videos falsch abgebildet werden, auch in Dokumentarfilmen, Interviews und Werbespots, ist das Risiko noch sehr viel greifbarer, dass dem Schauspieler ein finanzieller Schaden zugefügt wird, sein Ruf und sein gesellschaftliches Ansehen beschädigt wird. Es besteht auch die Gefahr, dass mit der zunehmenden Weiterentwicklung der künstlichen Intelligenz die Unterscheidung zwischen einer echten Darbietung und einer Darbietung, die digital generiert wurde, nicht mehr möglich sein wird. Wenn das menschliche Auge nicht mehr in der Lage sein wird, den Unterschied zu erkennen, wird vielleicht auch das Recht aufhören, diese beiden Formen zu unterscheiden und die Persönlichkeitsrechte werden sich auch auf die virtuelle Abbildung ausdehnen.

Wenn man nur eine einzige Sache aus diesem Kapitel im Gedächtnis behalten sollte, dann die Tatsache, dass Persönlichkeitsrechte sorgfältig in ihrem jeweiligen Kontext betrachtet werden müssen. Kommt es zu einem Rechtsstreit, dann interpretieren Gerichte in der Regel eher bestehendes Recht, sie schaffen kein neues Recht. Das heißt, sie bestätigen lediglich die Anwendbarkeit bestehender Vorschriften, die sich aus den Gewohnheiten der betreffenden Gesellschaft entwickelt haben. Unabhängig von der juristischen Ausbildung oder Autorität ist es keine einfache Aufgabe, sich mit Fragen von Kunst, Wahrheit, Ausdruck und Identität zu befassen – von denen jede einzelne den Kern des Menschseins betrifft. Fragt man Juristen nach der Verwertung und dem Schutz von

⁷³⁵ Bergman I., *Four Screenplays*, Secker & Warburg, London. Aus dem Schwedischen übersetzt von L. Malmstrom and D. Kushner.



Persönlichkeitsrechten, vor allem im Zusammenhang mit neuen Technologien wie Deepfakes und Ghost Acting, dann lautet die Antwort in der Regel: „Kommt darauf an“.

Regulierung von KI

Die von Kelsey Farish in dieser Veröffentlichung erwähnte „*Berücksichtigung des situativen Kontexts*“ wäre wohl bei allen in dieser Publikation behandelten Rechtsfragen sinnvoll. Zudem ist eine weitere wichtige Tatsache zu bedenken (die bereits im Vorwort dieser Publikation erwähnt wurde): Computer sind eben Computer – dumme Maschinen, die nur den Unterschied zwischen Nullen und Einsen kennen – und daher werden die Ergebnisse ihrer seelenlosen Rechnerei immer von menschlichen Eingriffen und Kontrollen abhängen bzw. darauf angewiesen sein. Und sehr oft bedeuten menschliche Eingriffe Regulierung. **Atte Jääskeläinen** skizziert in seinem Beitrag zu dieser Publikation einige Prinzipien, die aus seiner Sicht für die KI-Regulierung gelten sollten. Wie bereits ausgeführt, ist das grundlegendste Prinzip hier die Transparenz, denn sie „*dient dem menschlichen Bedürfnis, die Funktionsweise der Systeme zu verstehen und Verantwortung bei den richtigen Personen anzusiedeln*“. Jääskeläinen vermutet jedoch, dass wir akzeptieren müssen, dass „*unbekannte Risiken möglicherweise nicht regulierbar sind, zumindest wenn sich die Regulierung an der Technologie orientiert und nicht an Zielen*“. Darüber hinaus sollte KI-Regulierung „*das Risiko für die Öffentlichkeit verringern, ohne Kreativität und Innovation zu zerstören*“, und „*unnötige Hindernisse beseitigen, die es erschweren, Daten zu nutzen, um Wohlstand zu schaffen und Gutes zu tun*“.

8. Ansätze für einen nachhaltigen Regulierungsrahmen für die audiovisuelle Industrie in Europa

Atte Jääskeläinen, LUT, Finnland, und LSE, London⁷³⁶

8.1. Einführung

Ein Softwareprogramm erkennt im Bild eines Fußballspiels den Ball und die Spieler und zoomt auf das Geschehen. Da die Kameras extrem scharf sind, kann der „virtuelle Regisseur“ – ein Computer, der sich auf maschinell lernende Algorithmen stützt – ohne menschliches Zutun die Übertragung eines kompletten Spiels mit mehreren Kameraperspektiven produzieren. Der niederländische Fußballverband KNVB und das niederländische Medienunternehmen Talpa haben sich zum Ziel gesetzt, ohne jeden Eingriff von Menschen jährlich 80 000 Amateurfußballspiele live zu übertragen.

Ein Softwareprogramm analysiert Ihr Verhalten in sozialen Medien und kann Ihre Persönlichkeit in dem weit verbreiteten psychologischen Modell Big 5 erfassen. Dabei erreicht es mit Informationen über etwa 200 Facebook-Likes die gleiche Genauigkeit wie mit dem Ausfüllen eines Fragebogens. Das Ergebnis: ein System, das zur Beeinflussung von Wahlen genutzt werden könnte, indem politische Botschaften gezielt an diejenigen gerichtet werden, die sich durch einen bestimmten Stil von personalisierter Werbung am ehesten beeinflussen lassen.

Ein vom schwedischen Regionalverlag Mittmedia entwickeltes Softwareprogramm verwendet Personalisierung in Verbindung mit journalistischer Zielformulierung mit einem klaren Ansatz zur Kontextualisierung von Daten und der Nutzung maschinellen Lernens. Ich denke, wir werden in naher Zukunft eine vollständige Personalisierung und Automatisierung des Verlagswesens erleben, sagt der Chief Technology Officer des Unternehmens.

⁷³⁶ Atte Jääskeläinen ist ordentlicher Professor an der Technischen Universität Lappeenranta (LUT), Finnland, und Gastprofessor an der London School of Economics (LSE). Er war 2006–2017 Direktor für Nachrichten und Aktuelles bei Finnlands öffentlich-rechtlicher Rundfunkanstalt Yle und 2004–2006 CEO der finnischen Nachrichtenagentur STT. Zusammen mit Maike Olij verfasste er den „EBU News Report 2019: The Next Newsroom. Unlocking the Power of AI for Public Service Journalism“.

Ein von der Schweizer Firma Largo entwickeltes Softwareprogramm sagt anhand einer Drehbuch-Rohfassung voraus, wie die Reaktion des Publikums in einzelnen Ländern ausfallen wird. Anhand dieser Analyse empfehlen Datenwissenschaftler Änderungen bestimmter Drehbuchteile, damit ein Film für den Markt attraktiver wird und maximale Einnahmen erzielt.

Diese realen Beispiele veranschaulichen, wie künstliche Intelligenz die audiovisuelle Industrie oder ihr Umfeld bereits strategisch beeinflusst: Die audiovisuelle Produktion kann weitgehend automatisiert werden, das Zuschauerverhalten kann wie nie zuvor analysiert werden, und diese Informationen können sowohl dazu verwendet werden, den Nutzen zu optimieren, den die Kunden aus dem Markt ziehen, als auch dazu, die kreative und journalistische Arbeit zu lenken und in demokratischen Prozessen zu nutzen – manchmal auch so, wie Demokratie nicht funktionieren sollte.

Einige haben vorausgesagt, dass diese Entwicklung ohne die richtige Regulierung zum Untergang der Menschheit führt. Wie Amu Webb vom Future Today Institute 2019 schrieb: „Mangelnde Nuancierung ist ein Teil des Entstehungsproblems der KI: Einige überschätzen die Anwendbarkeit der KI an ihrem Arbeitsplatz dramatisch, während sie für andere zu einer unaufhaltsamen Waffe werden wird.“⁷³⁷

Die Regulierung der KI, vor allem in Bereichen, die für die audiovisuelle Industrie relevant sind, ist eine komplexe Frage, auf die es keine klaren Antworten gibt. Diese Komplexität wirkt sich auch auf praktische Entscheidungen aus. Sowohl für Behörden als auch für die Politik ist das Gebiet der KI schwer in den Griff zu bekommen, da das gesamte Konzept unklar ist und ein „bewegliches Ziel“ abgibt: Gerade wenn man glaubt, es zu „erwischen“, sind bereits neue Formen entstanden. Regulierung ist ein reaktiver Prozess, und in einem sich schnell verändernden Bereich ist sie oft schon überflüssig, bevor sie zur Anwendung kommt.⁷³⁸

Einige der KI-Beispiele vermitteln den Eindruck, dass Computer etwas fast Magisches erschaffen. Meist beruht diese Illusion einfach auf dem Einsatz intelligenter Mathematik und riesiger Rechenressourcen, die durch Cloud Computing und schnelle Telekommunikation zur Verfügung stehen. Was als Kreativität erscheint, ist nur die Freiheit von Vorurteilen, die das menschliche Denken einschränken, kombiniert mit der Fähigkeit, alle Alternativen zu prüfen und zu produzieren, auch solche, auf die Menschen nicht kommen würden, selbst wenn sie eingefahrene Denkmuster ausdrücklich aufbrechen sollen.

Die Autoren eines kürzlich erschienenen Buchs über die „einfache Ökonomie“ der prädiktiven KI erklären: „Alle menschlichen Aktivitäten können durch fünf allgemeine Komponenten beschrieben werden: Daten, Vorhersagen, Urteile, Handlungen und Ergebnisse. Wenn die maschinelle Intelligenz besser wird, werden menschliche Prognosefähigkeiten an Wert verlieren, weil die maschinelle Vorhersage einen billigeren

⁷³⁷ Webb A., „The big nine before it's too late“, WMG Weekly, <https://www.wmgweekly.com/post/2019/06/08/the-big-ninebefore-it-s-too-late>.

⁷³⁸ Petit N., „Law and regulation of artificial intelligence and robots – Conceptual framework and normative implications“, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2931339.

und besseren Ersatz bietet. Der Wert der menschlichen Urteilsfähigkeit wird jedoch steigen.“⁷³⁹

Die OECD hat geschätzt, dass fast die Hälfte aller Berufe innerhalb von 15–20 Jahren aufgrund von Automatisierung und neuen selbstlernenden Technologien entweder verschwinden oder sich grundlegend verändern werden.⁷⁴⁰ Aus praktischer Sicht besteht die Herausforderung wohl darin, Prozesse so auszurichten, dass die Stärken von Mensch und Maschine bestmöglich genutzt werden können.

Technologien existieren nicht isoliert von Kultur und Werten. Einigen Wissenschaftlerinnen und Wissenschaftlern zufolge führt verstärkte Automatisierung zu einer wesentlichen Neudefinition von Werten, die sich enger an die Informatik anlehnt.⁷⁴¹ Andere argumentieren, Kommunikation könne nicht mehr als rein zwischenmenschliches Phänomen definiert werden.⁷⁴² Eine grundlegende Beobachtung ist, dass wir am Übergang in ein Zeitalter der Kommunikation zwischen vielen Beteiligten stehen, in dem Milliarden und Abermilliarden von Verbindungen zwischen Menschen bestehen, die durch Automatisierung angetrieben werden und unser Verständnis der vernetzten Welt bestimmen, in der wir leben.⁷⁴³

In diesem Kapitel sollen einige der ethischen und moralischen Fragen geklärt werden, die bei der Entwicklung einer wirksamen und ethisch vertretbaren Regulierung für diesen faszinierenden, für manche aber auch erschreckenden Technologiebereich berücksichtigt werden sollten.

8.1.1. Grundbegriffe der KI

Der Begriff „künstliche Intelligenz“ ist nicht klar. Einer gängigen Definition zufolge beschreibt KI maschinelle Prozesse, für die Menschen Intelligenz benötigen würden. Der Begriff wurde ursprünglich von John McCarthy geprägt, der in den 1950er Jahren mit der Erforschung der KI begann. Er ging davon aus, dass menschliches Lernen und menschliche Intelligenz durch eine Maschine simuliert werden könnten. In ihrer grundlegendsten Form ist künstliche Intelligenz jedoch „ein System, das autonome Entscheidungen trifft, ein Zweig der Informatik, in dem Computer so programmiert werden, dass sie Dinge tun, die

⁷³⁹ Agrawal A., Gans J. und Goldfarb A., „Prediction machines: The simple economics of artificial intelligence“, Harvard Business Press.

⁷⁴⁰ OECD, The Future of Work. OECD Employment Outlook 2019. Highlights, https://www.oecd-ilibrary.org/employment/oecd-employment-outlook-2019_9ee00155-en.

⁷⁴¹ Coddington M. „Clarifying journalism’s quantitative turn“, Digital Journalism, Routledge, 3(3), S. 331–348. Milosavljević M. und Vobič I., „Our task is to demystify fears: Analysing newsroom management of automation in journalism“, Journalism, SAGE Publications, <https://journals.sagepub.com/doi/abs/10.1177/1464884919861598?journalCode=joua>.

⁷⁴² Lewis S. C., Guzman A. L. und Schmidt T. R., „Automation, journalism, and human–machine communication: Rethinking roles and relationships of humans and machines in news“ Digital Journalism, Routledge, 7(4), S. 409–427.

⁷⁴³ Jääskeläinen A. und Olij M., „The next newsroom: Unlocking the power of AI for public service journalism“, European Broadcasting Union, <https://www.ebu.ch/publications/news-report-2019>.

normalerweise menschliche Intelligenz erfordern“.⁷⁴⁴ Einer der Pioniere der künstlichen Intelligenz, Marvin Minsky, beschrieb sie als „Kofferbegriff“, in dem viele Konzepte zusammengepackt sind.

Die meisten Anwendungen der künstlichen Intelligenz nutzen heute Technologien, die in den Bereich des **maschinellen Lernens** fallen. Beim maschinellen Lernen Computer aus Daten, ohne dafür explizit programmiert zu sein.

- Bei der traditionellen Programmierung werden die Daten auf dem Computer ausgeführt, der den Output produziert. Programmierer müssen daher die Regeln kennen, wenn der gewünschte Output erzielt werden soll.
- Beim maschinellen Lernen werden Daten und Beispiele für den gewünschten Output zunächst auf dem Computer ausgeführt, und dieser lernt dann aus ihnen, eigene Regeln zu erstellen. Diese werden dann zur Produktion des Outputs verwendet.

Eine sehr leistungsfähige und aufwändige Art des maschinellen Lernens, das **Deep Learning**, verwendet Algorithmen, die als „künstliche neuronale Netze“ bezeichnet werden und der Funktionsweise neuronaler Netze im menschlichen Gehirn nachempfunden sind. Fortschritte im Deep Learning haben Sprachtechnologien und Bilderkennung stark vorangebracht und damit beträchtliche Möglichkeiten im audiovisuellen Sektor eröffnet.

Der Hauptvorteil beim Deep Learning besteht darin, dass riesige Datenmengen berücksichtigt werden können. Daher lassen sich durch maschinelles Lernen Aufgaben erfüllen, die früher unlösbar erschienen. Andererseits erfordert es aber auch riesige Mengen an Trainingsdaten und teure Rechenressourcen, sodass sein Einsatz in großem Maßstab erst seit Kurzem möglich – und immer noch begrenzt – ist.

Die verschiedenen Arten des maschinellen Lernens unterscheiden sich darin, wie die Maschinen Daten nutzen, um Regeln zu lernen. Die älteste Art des maschinellen Lernens wird als **überwachtes Lernen** bezeichnet. Hierbei wird der Computer anhand gewünschter Ergebnisse trainiert. Der Algorithmus stellt dann Regeln auf, nach denen der Computer Ergebnisse erzeugen kann, die denen des Trainingsdatensatzes ähneln. Beim **unüberwachten Lernen** wird der Computer verwendet, um einen riesigen Datensatz sinnvoll zu gruppieren. Dieser Ansatz erfordert keine Trainingsdaten, da der Datensatz in der Regel durch seine interne Logik strukturiert ist.

Eines der neueren Modelle ist das **bestärkende Lernen**, bei dem das System spontan aus dem Feedback lernt, das es aus seiner Umgebung erhält. Dieser Bereich entwickelt sich schnell und ist recht nützlich, da auf bestärkendem Lernen basierende Systeme sich schnell an neue Situationen anpassen und aus dem Nutzerverhalten lernen können – das sich aus irgendeinem Grund plötzlich ändern kann. So nutzen die fortschrittlichsten Empfehlungssysteme bestärkendes Lernen als Möglichkeit, das Nutzerverhalten zu beobachten und sich daran anzupassen.

Alle drei Lernmodelle können in einem komplexen System kombiniert werden. Ein fortschrittliches System verfügt in der Regel über mehrere Algorithmen und kann

⁷⁴⁴ <https://futuretodayinstitute.com/trend/artificial-intelligence/>.

verschiedene Ansätze miteinander verbinden, um den gewünschten Output für eine bestimmte Situation zu erzielen. Für ein solches System ist es wichtig, das Problem, den Kontext und die in den Daten enthaltenen Informationen zu verstehen.

Verwirrend an den verschiedenen Bereichen der künstlichen Intelligenz ist, dass die Dimensionen der KI oft verwechselt werden. So taucht etwa maschinelles Lernen in Listen wichtiger KI-Anwendungsbereiche häufig neben der Bilderkennung und der Verarbeitung natürlicher Sprache auf. Doch maschinelles Lernen ist bei fast allen modernen Anwendungen der künstlichen Intelligenz im Spiel, etwa bei medizinischen Diagnosen, selbstfahrenden Autos, Vorhersagesystemen und automatischer Klassifikation.

Mit Blick auf die Notwendigkeit einer KI-Regulierung ist eine sehr wichtige Kategorisierung die enge oder schwache KI, bei der das System eine einzige Aufgabe im Zusammenhang mit einem spezifischen Problem ausführt. Ein weitaus schwierigerer Regulierungsbereich ist die sogenannte künstliche allgemeine Intelligenz, die komplexe Probleme in jedem Kontext lösen und ihre Ziele autonom definieren kann. Es gibt kühne Initiativen, um eine derartige Autonomie von Computern zu erreichen, wie etwa DeepMind von Google und OpenAI, doch diese Art der KI ist bisher nicht realisiert worden. Die meisten Forscher glauben, dass wir mindestens noch Jahrzehnte davon entfernt sind, ein solches KI-Niveau zu erreichen.⁷⁴⁵

8.2. Wie wird KI in der audiovisuellen Industrie eingesetzt?

Unter der audiovisuellen Industrie verstehen wir traditionell die Produktion und Vermarktung von Filmen, Fernsehprogrammen und audiovisuellen Internet-Inhalten. Da es jedoch schwierig ist, für die KI klare Grenzen zu definieren, ist es nicht sinnvoll, unsere Überlegungen nur auf die Besonderheiten dieses speziellen Sektors zu beschränken. Die Übermittlung von Botschaften erfolgt heutzutage oft mit gemischten Methoden – und diese Methoden können sogar autonom, anhand individueller Präferenzen, von datengetriebenen KI-Systemen ausgewählt werden. Daher hängen die Herausforderungen der KI-Regulierung im audiovisuellen Sektor eng mit anderen Sektoren zusammen, die mit diesem verbunden sind. Die Schlüsselfrage lautet, was wir mit einer Regulierung erreichen wollen und wie sie in den für diesen Sektor relevanten Bereichen angewandt werden soll. Das Spektrum möglicher Anwendungen ist bereits vielfältig und entwickelt sich schnell in Bereiche hinein, die wir uns noch gar nicht vorstellen können. Es wäre jedoch weder sinnvoll noch effektiv, die gleichen Regeln auf automatische Übersetzungen, selbstfahrende Autos, sensible personenbezogene Daten und moderne Kamertechnologie anzuwenden, um nur einige Anwendungsbereiche zu nennen.

Wie lässt sich diese Frage dann in den Griff bekommen? Erstens sollten wir versuchen, die Folgen unterschiedlicher KI-Verwendungen in der audiovisuellen Industrie

⁷⁴⁵ Siehe z. B. Joshi N., et al. „How far are we from achieving artificial general intelligence?“, <https://www.forbes.com/sites/cognitiveworld/2019/06/10/how-far-are-we-from-achieving-artificial-general-intelligence/#5ebe24f06dc4> und Fjelland R., „Why general artificial intelligence will not be realized“, <https://www.nature.com/articles/s41599-020-0494-4>.

besser zu verstehen. Da sie sehr unterschiedlich sind, hilft selbst eine grobe Kategorisierung, die Werte zu verstehen, die durch Regulierung geschützt werden sollten.

Da es keine bessere und differenziertere gibt, kann hier die Kategorisierung aus dem Bericht „The Next Newsroom“⁷⁴⁶ für die Europäische Rundfunkunion über den KI-Einsatz im öffentlich-rechtlichen Journalismus verwendet werden, die als grundlegender Rahmen dienen könnte, um die strategische Bedeutung verschiedener Arten von KI-Technologien für die audiovisuelle Industrie zu verstehen.

Erstens kann künstliche Intelligenz als eine wachsende Sammlung praktischer Instrumente betrachtet werden. Dabei handelt es sich um eine KI auf rein operativer Ebene, die in erster Linie darauf abzielt, sich wiederholende Aufgaben zu automatisieren und Kosten zu senken.

So gibt es etwa zahlreiche Lösungen mit KI-Systemen für Aufgaben im Redaktionswesen und Medienmanagement, die sich immer schneller durchsetzen. Werkzeuge, die für die Transkription und Übersetzung von Sprachen und für die Erkennung von spezifischem Material in Archiven verwendet werden, erleichtern und beschleunigen die Wiederverwendung des Materials und verändern die Wertschöpfungslogik in der Produktion. So setzt die Deutsche Welle beispielsweise KI-basierte Sprachverarbeitung ein, nur um die Nachrichtenredaktion in ihrem mehrsprachigen Nachrichtenbetrieb in Echtzeit über das Geschehen auf dem Laufenden zu halten. Der Schweizer Rundfunk ist eines der Unternehmen, die fortschrittliche Systeme zur Erkennung von Personen und Orten in archiviertem Videomaterial entwickelt haben.⁷⁴⁷

Der irische Rundfunkveranstalter RTE und Al Jazeera haben sogar gemeinsam ein System geschaffen, das im Wahlkampf die Sendezeit von Politikern misst und Inhalte kennzeichnet, die ein regulatorisches Problem bergen könnten. Das System basiert auf einer fortschrittlichen Methode, die nicht nur Gegenstände und Personen auf den Bildern erkennt, sondern auch deren Kontext.⁷⁴⁸

Die Frage ist berechtigt, ob an diesen Werkzeugen etwas Besonderes ist, das einer besonderen Regulierung bedarf, oder ob es einfach ausreicht, dass sie verantwortungsvoll genutzt werden. Schließlich handelt es sich dabei nur um Werkzeuge und Technologien, die helfen sollen, die Arbeit zu erledigen.

Zweitens ermöglicht die KI die Schaffung einer datentechnisch kompetenten Kultur, die darauf beruht, eigene Ziele festzulegen und zu kennen und Mess- und Optimierungsmöglichkeiten zu finden, die auf ihnen basieren. Dies ermöglicht auch einen sehr strategischen KI-Einsatz: die gezielte Vermittlung von Botschaften und die Optimierung des Nutzens für einzelne Kunden auf Basis von Informationen über deren Präferenzen und Verhalten.

Eine solche Optimierungs- und Personalisierungs-KI kommt auch zum Einsatz, um die finanziellen Ergebnisse wirtschaftlicher Aktivitäten im audiovisuellen Bereich zu

⁷⁴⁶ Jääskeläinen A. und Olij M., *op.cit.*

⁷⁴⁷ Rezzonico P., „Artificial intelligence at the service of the RTS audiovisual archives“, FIAT/IFTA, <http://fiatifta.org/index.php/artificial-intelligence-at-the-service-of-the-rts-audiovisual-archives/>.

⁷⁴⁸ TM Forum, „AI indexing for regulatory practise“, <https://www.tmforum.org/ai-indexing-regulatory-practise>.

optimieren, etwa durch die Schaffung von mehr Effizienz bei Marketingkampagnen, indem die Wirksamkeit verschiedener Botschaften getestet wird oder interessante Marktcluster und Möglichkeiten identifiziert werden.

Eine der grundlegenden Herausforderungen im Zeitalter des Überflusses, den die digitalen Technologien geschaffen haben, liegt darin, in einem ungeordneten Überangebot gute Inhalte zu finden. Es besteht also die Notwendigkeit, Inhalte an das Publikum heranzutragen, das sich für sie interessiert. Diese Art der Optimierung ist nicht auf Online-Angebote beschränkt. So gibt es beim spanischen RTVE ein vielversprechendes Forschungsprojekt zur Planung des Fernsehprogramms mit Hilfe von KI-Algorithmen. Die Fragestellung lautet, welche Fernsehprogramme den Geschmack des Publikums zu einer bestimmten Tageszeit oder an einem bestimmten Wochentag treffen.⁷⁴⁹

Technisch ist es bereits möglich, alle Geräte einzelner Nutzerinnen und Nutzer, die Informationen über ihre musikalischen Vorlieben, gesehenen Filme und Fernsehgewohnheiten übermitteln, zu verbinden und anhand ihrer persönlichen Gesundheitsgeräte sogar ihre Stimmung zu erkennen und sie dann zu den besten Inhalten zu leiten. Nimmt man noch Informationen über das Wetter, den Arbeitskalender und den lokalen Verkehr hinzu, hat man einen leistungsstarken persönlichen Assistenten, der Nutzerinnen und Nutzern immer das serviert, was sie genau hier und jetzt interessieren könnte. All diese Dienste gibt es bereits, und die meisten von ihnen basieren auf audiovisuellen Inhalten. Tatsächlich bieten einige der globalen Technologieriesen bereits Ähnliches an. Denken wir nur an die Möglichkeiten des Google Assistant oder des Portfolios an Apple-Services. All diese Funktionen sind im Betriebssystem unseres Mobiltelefons vereint – das von denselben Unternehmen kontrolliert wird.

Drittens kann KI in speziellen Prozessen eingesetzt werden, die darauf abzielen, bessere und markantere Inhalte zu schaffen. Dabei geht es nicht nur um die Optimierung und Wiederholung dessen, was bereits getan wurde, sondern um die Schaffung völlig neuer Ansätze ohne die Einschränkungen und Voreingenommenheiten des menschlichen Gehirns. Dieser Bereich bringt interessante Herausforderungen mit sich, denn manchmal wirkt ein gutes Ergebnis nicht wirklich besser, wenn es von einer Maschine statt von einem Menschen kommt, oder vielleicht ist es nach unserem ethischen Kodex auch nicht akzeptabel. Ein weiterer interessanter Aspekt ist die Frage, wie sich Kreativität und künstlerische Motivation und Identität erhalten lassen, wenn Computer am Prozess beteiligt sind und auf eine Art und Weise arbeiten, die man – zumindest nach manchen Definitionen – als kreativ bezeichnen könnte.

8.3. Ist KI anders als frühere Technologien?

Wie kann verhindert werden, dass KI Schaden anrichtet? Wie können Risiken entschärft werden, die mit der Schaffung solcher KI-Systeme verbunden sind? Ist die Welt sicher, wenn

⁷⁴⁹ Cibrián E. et al. "Artificial intelligence and machine learning for commercial analysis in the audiovisual sector: A case study of designing TV schedules", <http://www.kr.inf.uc3m.es/artificial-intelligence-and-machine-learning-for-commercial-analysis-in-the-audiovisual-sector-a-case-study-of-designing-tv-schedules/>.

Autos autonom auf den Straßen fahren? Oder ist die Welt gerecht, wenn über die Verbrechensaufklärung entscheidet, wo man wohnt oder wie man in die Kamera schaut?

Bevor wir uns in die KI-Regulierung stürzen, sollten wir uns fragen: Was an ihr ist eigentlich anders? Was ist an der KI so speziell, dass sie der besonderen Aufmerksamkeit der Regulierer bedarf? Sind die Probleme aus menschlicher und gesellschaftlicher Sicht wirklich neu? Oder geht es um dieselben Grundprobleme, die das Gesetz vielleicht schon regelt, nur in neuem Gewand?

8.3.1. Wer ist verantwortlich, wenn die KI Schaden anrichtet?

Das Schlüsselkonzept, das sich aus der Frage nach der Verantwortung für Schäden der KI ergibt, ist die unterstellte Fähigkeit der KI, ohne menschenähnliche Verantwortung und Kontrolle menschenähnlich zu handeln. Mit anderen Worten: Wenn KI potentielle Autonomie besitzt, muss sie dann kontrolliert und reguliert werden?

Eine weitere Besonderheit ist die Tatsache, dass das Handeln der KI bei der Entwicklung von Systemen schwer vorhersehbar ist. Dies gilt insbesondere dann, wenn Systeme „kreativ“ sein sollen, wie man es im audiovisuellen Sektor erwarten würde. Was ist, wenn die Systeme so unerwartete Ergebnisse erzeugen, dass man nicht sagen kann, dass der Entwickler des Systems sie hätte vorhersehen müssen und daher auch verantwortlich für sie ist?

Bisher gehen wir in der Regel nicht davon aus, dass Maschinen für ihr Handeln Verantwortung tragen. Verantwortlich ist derjenige, der die Maschine benutzt oder gebaut hat – ohne ausreichende Sorgfalt⁷⁵⁰.

Theoretisch ist es vernünftig, sich eine Welt vorzustellen, in der Maschinen Verantwortung tragen und für Schäden aufkommen müssen, die sie verursachen. Wir haben Konstrukte wie Gesellschaften mit Rechtspersönlichkeit, die Pflichten haben und Rechnungen einfordern können.

In den meisten Fällen halten wir jedoch den Menschen für am besten geeignet, moralische Verantwortung zu tragen und eine ethische Werteabwägung vorzunehmen. Bei allen High-End-Technologien ist es allerdings schwierig, den maßgeblichen Akteur zu bestimmen, der für Ergebnisse verantwortlich ist, die möglicherweise erst später und in Zusammenhängen auftreten, an die der ursprüngliche Entwickler nie gedacht hat.

Wie die Europäische Gruppe für Ethik der Naturwissenschaften und der Neuen Technologien (EGE) in ihrem Bericht über künstliche Intelligenz, Robotik und „autonome“ Systeme feststellt, ist Autonomie ein wichtiger Aspekt der Menschenwürde, der nicht relativiert und auf die Technik verlagert werden darf. Maschinen können weder die moralischen Positionen von Menschen einnehmen noch Menschenwürde erben. Alle unsere moralischen und juristischen Institutionen basieren auf der Idee, dass Menschen moralische

⁷⁵⁰ Petit N., *op.cit.*

Verantwortung übernehmen, indem sie etwa Rechenschaft ablegen, haften oder Pflichten übernehmen.⁷⁵¹

8.3.2. Es geht nicht nur um die Wirtschaft

Es ist unbestreitbar, dass Personalisierung für Einzelne wertvoll ist, aber gerade in einer Branche mit erheblicher kultureller und gesellschaftlicher Relevanz ist es von entscheidender Bedeutung, dass diese Systeme über die Interessen Einzelner hinausgehen, also den Interessen des Publikums oder der Gesellschaft insgesamt dienen. Es ist wichtig, einen akzeptablen Ausgleich zwischen diesen potenziell konkurrierenden Interessen zu finden. In jüngster Zeit sind einige der grundlegendsten Herausforderungen für Demokratien durch „Targeting-Maschinen“ zur gezielten Beeinflussung entstanden, die in den falschen Händen lagen: Fake News in gewaltigem Umfang, Deepfakes oder die Fähigkeit, eine Filterblase zu schaffen, in der man leben kann.

Die audiovisuelle Industrie hat in den demokratischen Prozessen unserer Gesellschaften eine besondere Bedeutung, sowohl auf nationaler als auch auf europäischer Ebene. Und wir leben in einer globalen Arena, besonders wenn es um KI geht. Daten- und Softwaredienste werden zunehmend über globale Systeme angeboten, oft auf einem Markt, der von Unternehmen aus den USA oder China kontrolliert wird. Wie eine nachhaltige Kreativbranche und audiovisuelle Wirtschaft gefördert werden kann, wenn ein strategischer KI-Einsatz in der Praxis immer wichtiger wird, ist auch eine Frage der europäischen Wettbewerbsfähigkeit und wird darüber entscheiden, was für eine Rolle Europa künftig als Soft Power spielen wird.

Der audiovisuelle Sektor als Ganzes ist eng mit den kulturellen und politischen Bedürfnissen einer Gesellschaft verbunden. Doch Kultur hat nicht nur höhere gesellschaftliche Werte und Relevanz, sondern auch einen wirtschaftlichen Wert auf dem Markt und in der ökonomischen Realität. In ihrem Verständnis der wirtschaftlichen Auswirkungen von Kultur unterscheiden einige zwischen einem „institutionellen Wert“, der mit den makroökonomischen Auswirkungen auf nationaler oder internationaler Ebene für Volkswirtschaften verbunden ist, und einem „mikroökonomischen Wert“, der die Grundlage für einzelne Transaktionen bildet, bei denen Menschen für kulturelle Erfahrungen bezahlen. Der institutionelle Wert ist die theoretische Grundlage für die Subventionierung der Kulturwirtschaft.⁷⁵²

Das Konzept der AI4SG (AI for Social Good – KI für soziales Wohl) liefert einige Anregungen für die Gestaltung von Systemen, die den gesellschaftlichen Wert fördern. Dazu

⁷⁵¹ European Group on Ethics in Science and New Technologies, „Statement on artificial intelligence, robotics and ‚autonomous‘ systems“, European Commission, Directorate-General for Research and Innovation, http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf.

⁷⁵² La Torre M., „Defining the audiovisual industry“, in La Torre M. (ed.) The economics of the audiovisual industry: Financing TV, film and web. Palgrave Macmillan UK, S. 16–34.

gehören Vertrauenswürdigkeit als Grundlage der Gesellschaft und die Nutzung aller Technologien mit dem Ziel, das Leben positiv zu beeinflussen.⁷⁵³

Weitere Empfehlungen lauten, dass Systeme schrittweise entwickelt und vor dem Einsatz in der Produktion im Labor erprobt werden und Manipulationen durch Eingriffe von außen verhindert werden sollten. Die Förderung des sozialen Wohls ist oft erfolgreich, wenn Nutzerinnen und Nutzer in die Gestaltung der Systeme einbezogen werden, ihre Autonomie respektiert wird und in die Systemgestaltung einfließt und wenn ihnen die Möglichkeit gegeben wird, den Systemen und ihrem Output Sinn und Bedeutung zu verleihen.

8.4. Wir sind moralisch verpflichtet, mit der KI Gutes zu tun

Bei der Überlegung, wie eine Branche am besten reguliert werden kann, geht es immer um die Abwägung zwischen den Kosten und dem Nutzen der Regulierung. Wenn wir nur Risiken sehen und versuchen, Schaden abzuwenden und mit der Regulierung eine Rechenschaftspflicht einzuführen, schmälert dies möglicherweise den Nutzen, da eine Regulierung in der Regel Innovationen bremst und Kosten verursacht.⁷⁵⁴

Bei einer strengen Regulierung besteht das Problem, dass sie in der Regel die Risikobereitschaft, Kreativität und Wertschöpfung einschränkt. Deshalb müssen wir uns fragen, welches Risiko akzeptabel ist und wer das Risiko tragen soll, wenn etwas schiefgeht.

Bei der Gestaltung der Regulierung ist es also wichtig, beide Seiten des Phänomens zu betrachten. Auf der philosophischen Ebene stellt sich die Frage, wie man Innovation und Kreativität anregen und wie man vermeiden kann, dass das System die Risikobereitschaft dämpft. Risiko ist eine Triebfeder für Wirtschaft und Wohlstand. Auch in der audiovisuellen Industrie gehört zu Kreativität und Wertschöpfung das Eingehen von Risiken.

Daher sind sowohl im KI-Bereich im Allgemeinen als auch im Bereich der audiovisuellen Industrie vor der Regulierung einer spezifischen – oder in diesem Fall unspezifischen – Gruppe von Technologien Fragen zu beantworten: Besteht überhaupt ein Problem, das gelöst werden muss? Oder gibt es sogar eine Regulierung, die abgeschafft werden sollte, um mit der KI Gutes tun zu können? Oder sollte eine Regulierung eingeführt werden, um einen leichteren Zugang zu Daten oder Rechten zu ermöglichen, die auch in der audiovisuellen Industrie für den Übergang in ein neues Industriezeitalter benötigt werden?

Die Diskussion über die KI-Regulierung wurzelt im Bereich grundlegender menschlicher Instinkte und Gefühle, die uns befürchten lassen, dass Maschinen eines Tages die Kontrolle über uns übernehmen werden. Dies gilt insbesondere für die imaginierte künstliche allgemeine Intelligenz mit eigenem Willen. Diese Fähigkeit, unabhängig von

⁷⁵³ Taddeo M., „Trusting digital technologies correctly“, *Minds and Machines*, 27(4), 565–568. Taddeo M., Floridi, L., „The case for e-trust“, *Ethics and Information Technology*, 13(1), 1-3.

⁷⁵⁴ Gurkaynak G., Yilmaz I. und Haksever G., „Stifling artificial intelligence: Human perils“, *Computer Law & Security Review*, 32(5), S. 749–758.

Menschen zu handeln, und das teilweise fiktive Konstrukt einer menschenähnlichen Intelligenz ist die Grundlage von Weltuntergangsprognosen und untermauert Forderungen nach einer spezifischen KI-Regulierung.

Oft gilt eine Regulierung als notwendig, um Risiken zu bewältigen und Schaden abzuwenden, doch im Fall der audiovisuellen Industrie ließe sie sich auch aus einem anderen ethischen Blickwinkel betrachten: Aufgrund unserer moralischen und ethischen Pflicht, Gutes zu tun, stellt sich die Frage, wie wir den KI-Einsatz zum Guten verantwortungsvoll fördern können und ob Regulierung für ein Umfeld sorgen könnte, das Kreativität und Innovation fördert. Mit anderen Worten also für ein reicheres europäisches Leben und wirtschaftlichen Wohlstand.

Eine Regulierung wird meist dann eingeführt, wenn ein Problem zu lösen ist. Aber wir können nur verlieren, was wir vorher gewonnen haben.

8.5. Regulierung sollte den Menschen in den Mittelpunkt stellen und sich an Zielen orientieren

Erinnern wir uns daran, dass in einer Welt voller Risiken einige Risiken wünschenswert sind, weil Kreativität und Wohlstand darauf beruhen, dass jemand sie eingeht. Grundsätzlich geht es bei der Frage der KI-Regulierung in der audiovisuellen Industrie darum, welche Risiken so unerwünscht sind, dass sie es verdienen, reguliert zu werden.⁷⁵⁵ Wenn wir wollen, dass ein KI-System kreativ ist, sollten wir dann zulassen, dass es unvorhergesehene Ergebnisse hervorbringt? Was sind die tatsächlichen Konsequenzen dieser Risiken, wenn sie eingegangen werden?⁷⁵⁶

Die mechanische Anwendung ethischer Regeln kann manchmal heikel sein. Die Forderung nach Transparenz etwa kann manchmal zu lächerlichen Situationen führen. Wer hat bei kreativer Arbeit im audiovisuellen Sektor jemals gewusst, wie die Künstlerin oder der Künstler zu einem bestimmten künstlerischen Ergebnis gekommen ist? Wollen wir es überhaupt wissen? Gehört zur Faszination nicht auch das Geheimnis? Warum sollten wir von den Maschinen, die im kreativen Teil der Branche eingesetzt werden, Transparenz fordern, wenn wir sie nicht auch von Menschen verlangen? In diesem Fall besteht kein öffentliches oder privates Interesse daran, jemanden mit Hilfe von Technologie vor Kreativität zu schützen.

Andererseits können die Vorteile eines KI-Einsatzes in der audiovisuellen Industrie speziell durch eine wirksame Regulierung gefördert werden, da sie die mit fehlender Klarheit verbundenen Risiken minimiert. Gibt es im gegenwärtigen Regulierungssystem Hindernisse, die es zu beseitigen gilt? Sollten wir regulierte Vereinbarungen über den

⁷⁵⁵ Buiten M. C., „Towards Intelligent Regulation of Artificial Intelligence“. *European Journal of Risk Regulation*. 10 (1): 41–59, <https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/towards-intelligent-regulation-of-artificial-intelligence/AF1AD1940B70DB88D2B24202EE933F1B>.

⁷⁵⁶ Scherer M. U., „Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies“. *Harvard Journal of Law & Technology*, Vol. 29, No. 2, Spring 2016, p. 364, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2609777.

Datenaustausch fördern, damit es leichter wird, neue Unternehmen zu gründen und den Nutzen für Kundinnen und Kunden, Bürgerinnen und Bürger zu mehren?

Aus meiner Sicht leidet Europa daran, dass die Einstellung zur Regulierung allzu sehr auf Risiken fixiert ist. Während wir Risiken prüfen und über das Risikomanagement nachdenken, sind die Akteure aus den USA und China bereits weltweit aktiv und errichten Verteidigungslinien für ihre schon jetzt beherrschende Rolle in der Branche.

8.5.1. Größere Risiken sollten angegangen werden

Unsere Technologie ermöglicht es uns bereits, gemeinsam die Natur zu zerstören und sogar die gesamte Menschheit zu vernichten. Laut Huber handelt es sich dabei um „öffentliche Risiken“, die so breit gestreut sind und das Verständnis und die Kontrolle des einzelnen Risikoträgers so weit übersteigen, dass sie eine Bedrohung für die menschliche Gesundheit und Sicherheit darstellen.⁷⁵⁷ Im audiovisuellen Sektor können solche Risiken tatsächlich den größten Schaden anrichten. Sie zerstören unbeabsichtigt die Grundlagen der Gesellschaft, weil Bürgerinnen und Bürger zu kommerziellen und politischen Zwecken manipuliert und beeinflusst werden. In gewisser Weise ist es ein Beispiel für die „Tragik der Allmende“,⁷⁵⁸ ein noch immer diskutiertes Konzept, wonach Menschen, die im eigenen Interesse handeln, dem Gemeinwohl schaden.

Ein möglicher Umgang mit solchen Risiken besteht darin, die Menschenrechte als Linse und politisches Instrument einzusetzen. Unser Rechtssystem verpflichtet zur Vermeidung von Menschenrechtsverletzungen. Doch wer ist dann verantwortlich, wenn die Demokratie wegen der Zunahme von gezielter Beeinflussung in Gefahr ist? Und können, selbst wenn das Verantwortungskonzept formuliert werden könnte, Richter und nationale Rechtssysteme die verantwortlichen Parteien so schnell identifizieren, dass die Regulierung wirkt?

Maschinell lernenden Systemen ist wiederholt – und zu Recht, wie sich herausgestellt hat – vorgeworfen worden, voreingenommene Entscheidungen oder Vorhersagen zu treffen.⁷⁵⁹ Es ist tatsächlich möglich, dass das System voreingenommen sein könnte, weil es so konzipiert wurde. Häufiger stammt die Voreingenommenheit jedoch aus den Daten, mit denen das System trainiert wurde. Außerdem ist es oft so, dass diese Daten lediglich die Voreingenommenheit der gegenwärtigen Realität und der von Menschen getroffenen Entscheidungen widerspiegeln und sichtbar machen. Sie spiegeln daher unsere aktuellen gesellschaftlichen Werte und Entscheidungen sowie unsere menschlichen und kulturellen Vorurteile wider. Dies wiederum spiegelt den gesellschaftlichen Charakter von KI-Systemen wider, insbesondere in der audiovisuellen Industrie. Wir müssen uns davor

⁷⁵⁷ Huber P., „Safety and the second best: The hazards of public risk management in the courts“, *Columbia Law Review*, 85(2), S. 277–337.

⁷⁵⁸ Feeny D. et al., „The tragedy of the commons: Twenty-two years later“, *Human ecology*, 18(1), S. 1–19; Hardin G., „The tragedy of the commons“, *Journal of Natural Resources Policy Research*; Taylor & Francis, 1(3), S. 243–253; Stavins R. N., „The problem of the commons: Still unsettled after 100 years“, *The American economic review*, 101(1), S. 81–108.

⁷⁵⁹ Siehe Kapitel 1, 2 und 3 dieser Publikation.

hüten, KI-Systemen etwas zur Last zu legen, was gesellschaftlich sogar von Nutzen sein könnte: nämlich dass sie aufdecken, wie voreingenommen eigentlich die Entscheidungen sind, die bisher von Menschen getroffen werden. Stattdessen sollten wir die Systeme begrüßen, weil sie uns unsere Werte bewusster machen – wobei wir natürlich darauf achten müssen, dass unethische Entscheidungen nicht praktisch umgesetzt werden, bevor die Ergebnisse getestet und analysiert sind.

8.5.2. Verantwortlich ist der Mensch

Wir kommen nicht um den Schluss herum, dass ein gut funktionierendes Regulierungssystem in dieser schnelllebigen Welt nicht versuchen darf, die Technologie zu regulieren, sondern flexibel sein muss und auf Grundprinzipien basieren sollte. Es gibt bereits Ansätze, die gut funktionieren können. Sie orientieren sich an Zielen und stellen den Menschen in den Mittelpunkt, nicht die Technologie.

Eine grundlegende Frage ist, wo die rechtliche Verantwortung angesiedelt wird. In jeder Gesellschaft gibt es ein System, das in Fällen, in denen anderen Schaden zugefügt wird, die Verantwortung, die Haftung und teilweise auch die Rechenschaftspflicht regelt. In unseren gegenwärtigen Systemen kann diese Verantwortung nicht bei Technologien oder Maschinen liegen, sondern nur bei Menschen, Gesellschaften und anderen Rechtssubjekten. In einigen Branchen gelten die Risiken als so groß, dass Organisationen auch ohne Fahrlässigkeit eine größere Verantwortung für die Ergebnisse tragen. KI wird von Menschen geschaffen, die gut ausgebildet sind und sich oft der möglichen Risiken bewusst sind, die ihre Technologie für Personen oder Sachen mit sich bringen kann. Eine Verantwortung für ihr Handeln und in einigen Bereichen sogar eine verschuldensunabhängige Gefährdungshaftung ist für sie – oder in der Praxis für die Institutionen, die sie vertreten – nicht untragbar.

Wenn es bei der Regulierung scheinbar um Technologie geht, geht es in Wirklichkeit bislang um die Personen, die diese Technologie geschaffen oder genutzt haben, und die Organisationen, in denen sie beschäftigt sind. Und auch wenn KI-basierte Technologie eigene Fehler haben mag, sind Menschen oft noch fehleranfälliger. Das Problem ist, dass wir uns an menschliche Fehler gewöhnt haben, es aber weniger akzeptabel finden, wenn die Technologie die gleichen Fehler macht.⁷⁶⁰ Eine Kernfrage lautet, ob durch KI wirklich ein neues, vielleicht sogar noch unbekanntes Risiko entsteht.⁷⁶¹

8.5.3. Transparenz als Zwischenlösung?

Ohne Transparenz stehen Bürger und Verbraucher vor Entscheidungen, die sie nicht verstehen und über die sie keine Kontrolle haben. Um beurteilen zu können, ob es eine

⁷⁶⁰ Petit N., *op.cit.*

⁷⁶¹ Scherer M. U., *op.cit.*, p. 364.

Haftung für KI-basierte Entscheidungen geben sollte, müssen auch die Gerichte verstehen, wie die KI zu ihrer Entscheidung gelangt ist. Die Forderung nach Transparenz und Erklärbarkeit könnte also eine geeignete Lösung für einige der ethischen und rechtlichen Probleme sein und wurde auch schon als Regulierungsinstrument empfohlen.⁷⁶² Tatsächlich ist Transparenz in der Diskussion um eine nachhaltige KI-Regulierung ein zentraler Begriff, zumindest als Zwischenlösung.

Bei Daten bedeutet Transparenz die Verpflichtung, Nutzer, Kunden oder Klienten über die Verwendung ihrer Daten auf dem Laufenden zu halten. In der Welt der Algorithmen heißt Transparenz, dass ihre Funktionsweise erklärt werden muss, soweit dies für die Nutzerinnen und Nutzer verständlich ist. Erklärbarkeit – ein ähnliches Konzept wie Transparenz – bedeutet letztlich sogar, dass die Werte, auf die die Systeme ihr Verhalten gründen, nachvollziehbar sein sollten.

Allerdings gilt es, ein Gleichgewicht zwischen Nutzen und Kosten dieser Transparenz zu finden. Manchmal kann die Forderung nach Transparenz in der Praxis dazu führen, dass die Implementierung des Systems technisch nicht durchführbar ist.⁷⁶³ Die Forderung nach mehr Transparenz könnte uns sogar zwingen zu akzeptieren, dass Systeme nicht so korrekt sind, wie sie technisch sein könnten. Eigentlich eine interessante Wahl: Wenn ein Diagnosesystem als unerklärliche Blackbox korrekter ist, würden man damit trotzdem ein Leben retten – ohne zu wissen, wie? Eine Schwarz-Weiß-Regulierung, die bei jedem KI-Einsatz Transparenz verlangt, ist also auch nicht praktikabel.

Transparenz ist aber auch mit weiteren Problemen verbunden. Es kann Geschäftsgeheimnisse geben, die nicht preisgegeben werden dürfen, oder die Kosten für die Wahrung der Transparenz können zu einer unerwünschten Branchenkonzentration führen.⁷⁶⁴ Die Forderung nach kostspieliger Transparenz kann sich negativ auf die Innovation auswirken.⁷⁶⁵ Manchmal lässt sich die Logik von Maschinen einfach nicht in einer für Menschen verständlichen Sprache ausdrücken. Und wie lässt sich eine Rechenschaftspflicht in einem Technologiebereich nachvollziehen, der weitgehend darauf beruht, dass Ressourcen wie Codeschnipsel und KI-Algorithmusmodelle weltweit geteilt werden?⁷⁶⁶

8.6. Im Mittelpunkt der Mensch, nicht die Technologie

Die Diskussion über die KI-Regulierung folgt vor allem zwei Linien: Die eine basiert auf dem Standpunkt des Rechtssystems, die andere auf der Vorstellung, von den Technologien auszugehen und dann von unten her aus spezifischen KI-Anwendungen Regulierungsbedarf

⁷⁶² Siehe Kapitel 1 dieser Publikation.

⁷⁶³ Buiten M. C., *op.cit.*

⁷⁶⁴ Scherer M. U., *op.cit.*

⁷⁶⁵ Buiten M. C., *op.cit.*

⁷⁶⁶ Leonelli S., „Locating ethics in data science: Responsibility and accountability in global and distributed knowledge production systems“, *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 374(2083), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5124067/>.

abzuleiten.⁷⁶⁷ Ein gutes, funktionierendes Regulierungssystem beruht jedoch nicht auf diesen Konzepten, sondern auf den Bedürfnissen von Menschen und Gesellschaften.

All diese Diskussionen über die besonderen Merkmale der KI sind zwar wichtig, doch im Grunde ist KI nur eine Technologie – eine Computersoftware, deren eigentliche Domäne mathematische Berechnungen sind – und es ist fraglich, ob sie so grundlegend neuartig ist, wie sie oft dargestellt wird.⁷⁶⁸

Bei der KI-Regulierung den Menschen in den Mittelpunkt zu stellen, kann also Folgendes bedeuten:

- Transparenz dient dem menschlichen Bedürfnis, zu verstehen, wie die Systeme funktionieren, und Verantwortung bei den richtigen Personen anzusiedeln.
- Es ist zu akzeptieren, dass unbekannte Risiken möglicherweise nicht regulierbar sind, zumindest wenn sich die Regulierung an der Technologie orientiert und nicht an Zielen.
- Wir sollten versuchen, das öffentliche Risiko zu verringern, ohne Kreativität und Innovation zu zerstören.
- Wir sollten uns das aktuelle rechtliche Umfeld, insbesondere in Europa, ansehen und unnötige Hindernisse beseitigen, die es erschweren, Daten zu nutzen, um Wohlstand zu schaffen und Gutes zu tun.
- Menschen sollten mehr echte Kontrolle darüber haben, wie ihre Daten verwendet werden, doch der gegenwärtige Rahmen der DSGVO nutzt hierbei wenig,⁷⁶⁹ da die Zustimmung in Wahrheit erteilt wird, ohne sie wirklich zu verstehen, und die Nutzung personenbezogener Daten größtenteils durch Dritte erfolgt. Die Europäische Union sollte ihre Politik im Bereich KI und Daten kritisch überprüfen und vor allem verstärkt ermöglichen, Gutes zu tun, ohne zu vergessen, dass große Risiken, insbesondere für Demokratien, angegangen werden sollten.

⁷⁶⁷ Petit N., *op.cit.*

⁷⁶⁸ Edelman R. D., „Here’s how to regulate artificial intelligence properly“, <https://www.post-gazette.com/opinion/Op-Ed/2020/01/14/R-David-Edelman-Here-s-how-to-regulate-artificial-intelligence-properly/stories/202001140013>.

⁷⁶⁹ Siehe Kapitel 2 dieser Publikation.

Schlussbemerkungen

Es lässt sich feststellen, dass zum Thema KI perfekt die Devise passt: „*Aus großer Kraft folgt große Verantwortung*“ (die durch die Spider-Man-Comics berühmt wurde, aber mindestens auf die Zeit der Französischen Revolution zurückgeht).⁷⁷⁰ Die KI hat enormes Potenzial, zum Guten wie zum Bösen. Deshalb finden wir sie faszinierend und beängstigend zugleich, und für manche werden die Sorgen im Vordergrund stehen, während andere den Akzent darauf legen werden, welche wunderbaren Dinge sich mit dieser bahnbrechenden technologischen Entwicklung erreichen lassen. Tatsächlich zeigt die Lektüre der verschiedenen Beiträge in diesem Bericht, dass es nicht die eine Vision gibt, wie KI reguliert werden sollte, wohl aber bestimmte Prinzipien, die alle Autoren und Autorinnen (auf die eine oder andere Weise) im Kopf zu haben scheinen: Erklärbarkeit, Vertrauen, Privatsphäre, Pluralismus, aber auch Meinungsfreiheit, Kreativität und Innovation. Wenn es uns gelingt, all diese Ziele zu vereinen, kann KI für die Menschheit in vielerlei Hinsicht ein Segen sein.

Natürlich nur, wenn nicht eines Tages Elon Musks schlimmste Alpträume wahr werden und Maschinen die Welt übernehmen. Doch eine solche dystopische Zukunft ist nicht in Sicht.

Zumindest noch nicht.

⁷⁷⁰ <https://quoteinvestigator.com/2015/07/23/great-power/>.

Eine Publikation
der Europäischen Audiovisuellen Informationsstelle

