

Beiträge zur empirischen Marketing-
und Vertriebsforschung

RESEARCH

Robin Pade

Moral und künstliche Intelligenz im Marketing

MOREMEDIA



Springer Gabler

Beiträge zur empirischen Marketing- und Vertriebsforschung

Reihe herausgegeben von

Torsten Bornemann, Lehrstuhl für ABWL & Marketing, Universität Stuttgart,
Stuttgart, Deutschland

Martin Klarmann, Karlsruhe Institute of Technology, Karlsruhe, Deutschland

Dirk Totzek, LS für Marketing und Services, Universität Passau, Passau, Bayern,
Deutschland

Auch heute werden Marketing und Vertrieb von vielen als Domäne von „Bauchentscheidungen“ angesehen. Die vorliegende Schriftenreihe umfasst Beiträge, die einen anderen Weg gehen. Wichtige Fragestellungen, zum Beispiel aus den Bereichen Business-to-Business Marketing, Innovationsmarketing, Konsumentenverhalten, Preismanagement und Marketing Analytics, werden mit aktuellen wissenschaftlichen Verfahren empirisch untersucht. Zielsetzung der Beiträge ist es, für akademische und praktische Probleme in Marketing und Vertrieb eine faktenbasierte Grundlage zu schaffen.

Robin Pade

Moral und künstliche Intelligenz im Marketing

 Springer Gabler

Geleitwort

Verfahren der künstlichen Intelligenz haben sich in den letzten Jahren massiv weiterentwickelt. Damit erweitern sich auch die Einsatzmöglichkeiten im Marketing deutlich. Zum Beispiel können immer besser affektive Zustände von Kundinnen und Kunden erkannt werden – wodurch eine Personalisierung der Kommunikation dahingehend möglich wird, dass sie sich sekundengenau an den emotionalen Zustand des Kunden anpasst. Oder die kreativen Elemente der Kommunikation können durch Verfahren der künstlichen Intelligenz selbst generiert – und gleichzeitig optimiert werden.

Es drängt sich die Frage auf, was das für Kundinnen und Kunden bedeutet. Schnell denkt man an die Erfolge von Computern in komplexen Spielen wie Schach und Co. und stellt sich vor, dass derart überlegene Maschinen Menschen vielleicht auch erfolgreich manipulieren könnten. Hier ist die von Herrn Pade vorgelegte Dissertationsschrift angesiedelt. Sie versucht Antworten auf die Frage zu geben, inwieweit Marketing und Moral auch im Zeitalter künstlicher Intelligenz in Einklang zu bringen sind. Herr Pade geht das Thema aus drei empirischen Perspektiven an.

Zentrale Idee der ersten Studie ist die Überlegung, dass die Bewertung einer künstlichen Intelligenz als manipulativ vom Menschenbild der Konsumentinnen und Konsumenten abhängt. Geht man zum Beispiel davon aus, dass Menschen kognitiv limitierte, affektgesteuerte Wesen sind, dann wird man bestimmte Marketingaktivitäten schnell für manipulativ halten. Die gleichen Aktivitäten können aber als harmlos angesehen werden, wenn man davon ausgeht, dass Menschen rational handelnde Wesen sind, die prinzipiell kognitiv hochbegabt sind. Um dies zu untersuchen, entwickelt Herr Pade eine Skala, mit der das Menschenbild von Konsumentinnen und Konsumenten gemessen werden kann. Im Rahmen eines

online durchgeführten Vignetten-Experiments zeigt sich dann, dass die Bewertung des Manipulationsgrades einer Marketing-KI auch (aber nicht nur) vom Menschenbild der Konsumentinnen und Konsumenten abhängt.

Die zweite Studie nutzt künstliche Intelligenz um moralisches Verhalten (nicht nur im Marketing) sichtbar zu machen. Konkret identifiziert Herr Pade in dieser Studie mithilfe von Verfahren maschinellen Lernens die impliziten Menschenbilder von Führungskräften amerikanischer Unternehmen im Hinblick auf deren Kunden. Der zugrundeliegende Verdacht ist, dass solche Menschenbilder von CEO's opportunistisch eingesetzt werden. Geht es darum, in großem Stil Daten über das Konsumverhalten zu erheben (z.B. auf Grundlage von Tracking-Daten im Internet) könnte argumentiert werden, dass die Kunden rationale Akteure sind, die sich ihrer Handlungen durchaus bewusst sind. Geht es darum, dass Algorithmen eine Vorauswahl aus Entscheidungsalternativen treffen, könnte dann aber argumentiert werden, dass es die kognitiven Einschränkungen von Konsumenten nötig machen, hier Unterstützung anzubieten. Tatsächlich findet Herr Pade zumindest initiale Evidenz für den opportunistischen Einsatz von Menschenbildern.

Die dritte Studie betrachtet den Zusammenhang zwischen Marketing, Manipulation und moralischem Verhalten dann aus der Perspektive der Gestaltung von KI-Algorithmen. Die zugrundeliegende Idee ist es, einen Algorithmus zum personalisierten Pricing zu entwickeln, der auf Grundlage von verschiedenen Persönlichkeitsmerkmalen etc. Preise adaptiert. Dann geht es darum, dem Algorithmus bestimmte moralische Entscheidungsregeln vorzuschreiben und diese im Hinblick auf kurzfristiges und langfristiges Kundenverhalten zu testen. Angesichts geringer Effekte scheint mir das Kernergebnis dieser Studie aber vor allem zu sein, dass die begrenzte Vorhersagbarkeit menschlichen Verhaltens allgemein eine echte Barriere für die Macht von KI-Algorithmen im Marketing sein kann. Ob das an aktuellen technischen Limitationen oder an der *conditio humana* liegt, mag die Zukunft zeigen.

Robin Pade hat drei beeindruckende Studien durchgeführt. Sie sind aus Forschungssicht hochrelevant und methodisch sehr anspruchsvoll. Aber mehr noch: Sowohl konzeptionell als auch über ihre Ergebnisse sind sie für eine breite Öffentlichkeit von großem Interesse. Es ist deshalb sehr zu wünschen, dass sie sowohl die wissenschaftliche Diskussion um KI im Marketing als auch die öffentliche Debatte zur Regulierung von KI beflügelt.

Martin Klarmann

Danksagung

Die vorliegende Arbeit wurde im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Marketing und Vertrieb von Prof. Dr. Martin Klarman am Karlsruher Institut für Technologie (KIT) erstellt und im November 2023 als Dissertationsschrift angenommen. Ich möchte diese Gelegenheit nutzen, um all jenen meine Dankbarkeit auszusprechen, die mich auf dem Weg zu diesem besonderen Meilenstein begleitet und unterstützt haben.

Zuallererst möchte ich meinem Doktorvater Prof. Dr. Martin Klarman danken. Sein ansteckender Enthusiasmus, seine methodische Akribie und sein solidarischer Führungsstil werden mir stets Vorbild sein. Ich bin dankbar für seine hervorragende und unermüdliche Betreuung meiner Dissertation sowie für die überlassenen Freiheiten innovative und mutige Ansätze zu verfolgen, die einen nachhaltigen Einfluss auf meine Arbeit hatten. Nicht zuletzt danke ich ihm für all die schönen, wohltuend ablenkenden Grill- und Weihnachtsfeste und Fußballdebatten (auch wenn ich dem roten Brustring verbunden bleibe, drücke ich dem BuLi-Dino die Daumen für den Wiederaufstieg). Ebenfalls danke ich meinem Korreferenten Prof. Dr. Martin Ruckes für die Zweitbetreuung meiner Dissertation und den wunderbar anregenden fachlichen Austausch.

Prof. Dr. Sven Feurer – ehemals „gute Lehrstuhlseele“ – möchte ich für all die motivierenden und inhaltlich so wertvollen Gespräche danken. Deine Betreuung meiner Masterarbeit hat den Weg zu dieser Dissertation geebnet und ich freue mich dich auch Jahre später als Wegbegleiter und Freund in meinem Leben zu wissen.

Eine gute Lehrstuhlseele kommt jedoch selten allein. Mein herzlicher Dank gilt daher Juliane Bayer für ihre unermüdliche und umfassende Unterstützung während meiner gesamten Zeit am Lehrstuhl. Ihr Engagement und ihre Nahbarkeit machten nicht nur die Arbeit ein gewaltiges Stück leichter, sondern trugen auch zu einem unbekümmerten und freundschaftlichen Miteinander bei.

Selbstverständlich erstreckt sich dieser Dank auf all meine aktuellen und ehemaligen Kolleginnen und Kollegen. So danke ich Dr. Anika Honold für ihre unglaubliche Unterstützung, insbesondere während meiner Anfangszeit, und für ihre ansteckende Unbeschwertheit (an dieser Stelle möchte ich rezitieren: „Cheers Mate!“, Honold 2023, S. VIII). Bei Anne Cordts möchte ich mich für ihr stets wertvolles Feedback und für ihre empathische Art bedanken – ich hätte mir keine bessere Seele vorstellen können, mit der ich die gesamte Zeit am Lehrstuhl teilen darf. Saskia Jacob danke ich für alle inspirierenden Gespräche und gemeinsamen Lacher. Ein herzlicher Dank gebührt Alicia Gerlach für jeden nutzenstiftenden wie motivierenden Zuspruch und für ihr humorvolles Wesen, das auch zu später Überstunde noch positive Stimmung zu verbreiten weiß. Beatrice Weber danke ich für ihre locker progressive Art, die regelmäßig an das wirklich wichtige im Leben erinnert. Ein großes Dankeschön gilt außerdem Dominik Ebinger für seine fachliche und freundschaftliche Expertise während zahlreicher Konferenzen und Seminare. Schließlich möchte ich mich bei Celina Lang bedanken, die eine große Hilfe bei der Organisation des Lehrstuhllalltags darstellte.

Mein Dank gilt weiterhin Dr. Ingo Halbauer, Dr. Ingo Lange und Dr. Martin Moosbrugger, deren Ratschläge und vermittelten Fachkenntnisse an verschiedensten Stellen meiner Zeit am Lehrstuhl von unschätzbarem Wert waren.

Nicht zuletzt gebührt mein Dank Menschen aus meinem engsten Umfeld. Meinen Eltern Silvia und Wolfgang Pade und meinem Bruder Kevin Pade danke ich dafür mir jederzeit ein zuverlässiger Rückhalt zu sein. Das Wissen darum, dass Zuhause stets ein freies Zimmer und ein beratendes Gespräch wartet, war und ist ein großes Geschenk. Meiner Oma Anni danke ich für all die herzlich warmen Gespräche und Mahlzeiten, die ich so gerne noch etwas länger mit meinem Opa Hans geteilt hätte. Mein Dank gebührt außerdem meiner Oma Sophie und meinem Opa Werner als Vorbilder in puncto Resilienz und Vitalität.

Ein großes Dankeschön gilt Christiane und Kurt Willaredt als Inspiration in Sachen Moral und Ausgeglichenheit – euer kostbares pädagogisches Talent geht weit über den Beruf hinaus. Schließlich gilt mein ganz besonderer Dank meiner Freundin Anabel Willaredt. Du warst Tag und Nacht mit absolutem Selbstverständnis für mich da, auch wenn der Abend mal wieder einmal mehr der Arbeit gehörte. Gemeinsam mit dir die herausforderndsten Hürden zu überwinden und große wie kleine Freuden zu teilen, machten diese Reise für mich wahrlich unvergesslich.

Robin Pade

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation und Zielsetzung der Dissertation	1
1.2	Forschungsfragen der Dissertation	6
1.3	Aufbau der Dissertation	9
1.4	Koautorenschaft und Vorleistungen	12
2	Zusammenhang von Menschenbildern und Manipulationswahrnehmungen gegenüber künstlicher Intelligenz im Marketing	
	Intelligenz im Marketing	13
2.1	Abstract	13
2.2	Einleitung	14
2.3	Konzeptionelle Grundlagen	17
2.3.1	KI im Marketing	17
2.3.2	Manipulation und wahrgenommene Manipulationsabsicht	24
2.3.3	Menschenbilder (der Konsumenten)	25
2.4	Skalenenwicklung (Studie 1)	30
2.4.1	Initiale Konzeptualisierung	30
2.4.2	Exploration (Studie 1a)	33
2.4.3	Optimierte Konzeptualisierung	34
2.4.4	Itemreduktion (Studie 1b)	38
2.4.5	Skalensimplifizierung und Güte der finalen Skala (Studie 1c)	39

2.5	Einfluss des Menschenbildes auf Manipulationswahrnehmungen gegenüber Marketing-KI (Studie 2)	41
2.5.1	Hypothesenentwicklung	41
2.5.2	Studie 2: Zusammenhang zwischen Menschenbild und wahrgenommener Manipulationsabsicht	44
2.6	Übergreifende Diskussion	57
2.6.1	Beiträge und Implikationen für die Forschung	57
2.6.2	Implikationen für Manager und politische Entscheidungsträger	59
2.6.3	Limitationen und künftige Forschung	61
3	Kontextabhängiges Framing von Menschenbildern als Indikator für opportunistisches Unternehmensverhalten	63
3.1	Abstract	63
3.2	Einleitung	64
3.3	Theoretisch-konzeptionelle Grundlagen	69
3.3.1	Kontextabhängiges Framing der Konsumentennatur	69
3.3.2	Außenkommunikation von Managern	71
3.4	Extrahierung von Darstellungen der Konsumentennatur aus Textdaten	73
3.4.1	Datenvorbereitung	73
3.4.2	Generierung von Trainings- und Validierungsdaten	75
3.4.3	Modelltraining und -selektion	77
3.4.4	Textklassifikation	79
3.4.5	Teilüberwachtes Topic Modeling	82
3.4.6	Metrik für Glauben an ein Topic	90
3.5	Deskriptive Analyse von Darstellungen der Konsumentennatur	94
3.5.1	Darstellungen von Managern und US-Kongressabgeordneten	94
3.5.2	Varianz innerhalb der Darstellungen von Managern	97
3.6	Analyse der Varianz in Darstellungen der Konsumentennatur als Indikator für opportunistisches Verhalten	99
3.6.1	Diskrepanzen zwischen Wirtschaftssektoren	100
3.6.2	Mimikry gesellschaftlicher Darstellungen	102
3.6.3	Zusammenhang von Varianz in Managerdarstellungen und Insiderhandel	108

3.7	Übergreifende Diskussion	112
3.7.1	Beiträge und Implikationen für die Forschung	112
3.7.2	Praxisimplikationen	116
3.7.3	Limitationen und künftige Forschung	117
4	Künstliche Intelligenz mit moralischer Kapazität? Faires personalisiertes Pricing	119
4.1	Abstract	119
4.2	Einleitung	120
4.3	Konzeptionelle Grundlagen	124
4.3.1	Preisfairnessnormen	124
4.3.2	Personalisiertes Pricing im Ultimatumspiel	125
4.3.3	Artificial (Moral) Agents zur Personalisierung von Preisen	129
4.4	Hypothesenentwicklung	131
4.5	Studie 1: Training neuronaler Netze für (fares) algorithmisches Pricing	134
4.5.1	Methode	134
4.5.2	Modelltraining und Analyse	137
4.5.3	Berechnungslogik algorithmisch personalisierter Preise	143
4.6	Studie 2: Wahrnehmungen, Verhaltensintentionen und Profitabilität von fairem algorithmischem Pricing	144
4.6.1	Methode	144
4.6.2	Analyse und Ergebnisse: Proband als Rezipient (Szenario 1)	147
4.6.3	Analyse und Ergebnisse: Proband als Anbieter (Szenario 2)	156
4.7	Übergreifende Diskussion	157
4.7.1	Beiträge und Implikationen für die Forschung	157
4.7.2	Praxisimplikationen	161
4.7.3	Limitationen und künftige Forschung	163
5	Schlussbetrachtung	167
	Literaturverzeichnis	177

Abbildungsverzeichnis

Abbildung 2.1	Studie 2: k-means Clusteranalyse	54
Abbildung 3.1	Sensitivitätsanalyse des FinBERT Transformer-Modell Trainings	79
Abbildung 3.2	Divergenz in Abhängigkeit der Anzahl zusätzlicher, freier Topics	84
Abbildung 3.3	Darstellungen der Konsumentennatur $m_{i,t}$ ausgewählter Manager i für alle Topics t	96
Abbildung 3.4	Standardabweichungen des geäußerten Glaubens von Managern an Topics	98
Abbildung 3.5	Darstellungen der Konsumentennatur nach Wirtschaftssektor und Topic	101
Abbildung 3.6	Zeitreihen der Darstellungen der Konsumentennatur (y-Achse: Quartalsmittelwerte $m_{i,t}$ über alle Referenten i)	103
Abbildung 4.1	Paarweiser Vergleich von Mittelwerten (post-hoc Tukey Tests) der unterbreiteten Angebotshöhen	148
Abbildung 4.2	Paarweiser Vergleich von Mittelwerten (post-hoc Tukey Tests) der Wahrnehmungen der Probanden	150
Abbildung 4.3	Paarweiser Vergleich von Mittelwerten (post-hoc Tukey Tests) und Anteilen (Bonferroni-Korrektur) der Verhaltensintentionen der Probanden	151
Abbildung 4.4	Paarweiser Vergleich von Anteilen angenommener Angebote (Bonferroni-Korrektur) und Vergleich mittlerer Pro-Kopf-Profitte	153

Abbildung 4.5	Vergleich von langfristig erwarteten mittleren Pro-Kopf-Profiten in Euro	155
Abbildung 4.6	Paarweiser Vergleich von Mittelwerten (post-hoc Tukey Tests) der unterbreiteten Angebotshöhen (Szenario 2)	157

Tabellenverzeichnis

Tabelle 2.1	Literaturübersicht über Konsumentenreaktionen nach Begegnung mit KI-Algorithmen (oder verwandten künstlichen Entitäten)	19
Tabelle 2.2	Exemplarische Menschenbilder verschiedener Disziplinen (Synthese von Menschenbildern präsentiert von Glock (1964), Hjelle/Ziegler (1992) und Müller (2018), ergänzt um weitere Menschenbilder)	26
Tabelle 2.3	Überblick über den Prozess der Skalenentwicklung, -validierung, und –anwendung	31
Tabelle 2.4	Übersicht der Dimensionen von Menschenbildern in der vorliegenden Forschungsarbeit (abgeleitet von Menschenbilderkollektionen präsentiert von Glock (1964), Hjelle/Ziegler (1992) und Müller (2018) und eigens ergänzten Menschenbildern)	36
Tabelle 2.5	Studie 1c: Korrelationsmatrix der latenten Konstrukte (Pearson), DEV	42
Tabelle 2.6	Studie 2: Vignetten	46
Tabelle 2.7	Studie 2: Korrelationsmatrix der latenten Konstrukte (Pearson), DEV	50
Tabelle 2.8	Studie 2: Multiple Regressionsanalysen zur Vorhersage der wahrgenommenen Manipulationsabsicht	52
Tabelle 2.9	Studie 2: Statistische Unterschiede zwischen Mittelwerten der Clustervariablen	55

Tabelle 2.10	Studie 2: Multiple Regressionsanalyse zur Vorhersage der wahrgenommenen Manipulationsabsicht	56
Tabelle 3.1	Methodische Herangehensweise	68
Tabelle 3.2	Interrater-Reliabilitäten	77
Tabelle 3.3	Vergleich trainierter Transformer-Modelle	78
Tabelle 3.4	Exemplarische Ergebnisse der Textklassifikation	80
Tabelle 3.5	Ergebnisse des teilüberwachten Topic Modelings	86
Tabelle 3.6	Anzahl eindeutiger Satz-Topic Zuordnungen nach Menschenbild-Topic und Datenquelle	90
Tabelle 3.7	Exemplarische Sätze für (Un-)Glauben an das fokale Topic (GPT3.5 Bewertungen)	92
Tabelle 3.8	Mittelwertvergleich des geäußerten Glaubens an ein Topic von S&P-900 Managern und US-Kongressabgeordneten	95
Tabelle 3.9	Vector Auto Regression: Zeitlicher Zusammenhang gesellschaftlicher und unternehmerischer Darstellungen der Konsumentennatur	106
Tabelle 3.10	Durchschnittliche Profitabilität der Insider Trades nach Quintil	110
Tabelle 3.11	Multilevel Regressionsanalyse zur Vorhersage des Anteils an Quintil-5 Insidern	111
Tabelle 4.1	Psychografische Determinanten der Rezipientenentscheidung (Ultimatums spiel)	127
Tabelle 4.2	Verwendete Artificial (Moral) Agents	131
Tabelle 4.3	Basisszenario für Studie 1 und 2	136
Tabelle 4.4	Vergleich der selektierten neuronalen Netze mit Regressionsmodellen	139
Tabelle 4.5	Ergänzungen der Szenarien in Studie 2 im Vergleich zu Studie 1	146



Einleitung

1

1.1 Motivation und Zielsetzung der Dissertation

Künstliche Intelligenz (KI, engl. *AI*) gilt als eine der weltweit vielversprechendsten Technologien und befindet sich auf einem globalen Siegeszug (vgl. PWC 2017a). Bereits 2017 prognostizierte PWC einen KI-bedingten Anstieg des globalen Bruttoinlandproduktes als Wert aller für den Endkonsum produzierten Güter und erbrachten Dienstleistungen um 14 % bis zum Jahr 2030. Laut dieser Prognose entfallen 45 % des Zuwachses auf erhöhte Nachfrage infolge zunehmender Personalisierung, Attraktivität, Qualität und Erschwinglichkeit von Gütern oder Dienstleistungen durch die Einbindung von KI (vgl. PWC 2017a). Die aktuellen Errungenschaften um generative KI mit der Fähigkeit zur eigenständigen Erzeugung multimodaler Artefakte (kombinierte Text-, Bild-, Audio- oder Videoformate) bestätigen und befeuern diesen Trend (vgl. Deveau/Griffin/Reis 2023). Nach einer Umfrage von McKinsey unter global agierenden Business-to-Consumer (B2C) und Business-to-Business (B2B) Marketing- und Vertriebsorganisationen, streben 90 % der Führungskräfte an, (generative) KI und Machine Learning Algorithmen künftig „oft“ oder „fast immer“ verwenden zu wollen (vgl. Deveau/Griffin/Reis 2023). Oracle's Befragung (vgl. 2022) von 853 Marketingmanagern aus den USA, Kanada, Indien und Großbritannien zeigt dabei, dass KI in der Marketingpraxis am häufigsten für zielgruppengenaue Werbung (47 %) und zur Personalisierung von Inhalten und Preisen (42 %) entlang der Online Journey der Konsumenten eingesetzt wird.

KI umfasst Programme, Algorithmen, Systeme und Maschinen, die intelligentes menschliches Verhalten imitieren (vgl. Huang/Rust 2018, Kopalle, et al. 2022, Shankar 2018). Dafür bedient sich KI hauptsächlich innovativer Methoden des

Machine Learnings (z. B. Neuronale Netze) und des Natural Language Processings (NLP), um autonom verstehen, lernen und handeln zu können (vgl. Kopal, et al. 2022). Neben KI in Form von physischen Entitäten, wie Service Robotern, beschäftigen sich Marketingforscher vor allem mit KI in algorithmischer Form, der Konsumenten während ihrer Onlinesuche nach Produkten oder Dienstleistungen inflationär und oftmals unbewusst begegnen (vgl. Davenport, et al. 2020). Ma und Sun (vgl. 2020) fassen diese latente Omnipräsenz algorithmischer KI in ihrer Forschungsagenda „Machine learning and AI in marketing – Connecting computing power to human insights“ wie folgt zusammen:

„Unbeknownst to the consumer, much of this journey is guided by automated systems. The search results are generated by a sophisticated Google ranking system (...). The content at the websites is customized based on her profile through website morphing (...). The reviews she reads are placed prominently as they are deemed helpful by an evaluation algorithm, and the ads she repeatedly sees are delivered through retargeting algorithms via real-time bidding. The coupon which offers her the personalized price is generated by the firm’s pricing engine at just the right time. Finally, her posts on social media are collected by social listening engines and analyzed for sentiment and feedback. These automated systems that make split-second context-dependent decisions are known as Artificial Intelligence (...), generally implemented using state-of-the-art machine learning algorithms“ (Ma/Sun 2020, S. 482).

Diese Verkettung unbewusster Interventionen birgt das Potenzial rationale Entscheidungsprozesse zu torpedieren. Für zahlreiche Philosophen im Bereich der Marketingethik ist eine unbewusste oder eine rationalitätsumgehende Einflussnahme hinreichende Bedingung für das Vorliegen von unmoralischer Manipulation des Konsumenten (vgl. Aylsworth 2020, Crisp 1987, Gatignon/Le Nagard 2015, Villarán 2015). Wertenbroch, et al. (vgl. 2020) spricht im Kontext von KI sogar von einem Autonomieverlust des Konsumenten zugunsten einer fremdbestimmten Heteronomie, in der KI unter Zuhilfenahme moderner Machine Learning Algorithmen Konsumententscheidungen diktiert. Neben Forschern befürchten auch zentrale Stakeholder die Provokation eines multiplikativen Entscheidungsfehlers mit resultierenden Konsumententscheidungen, die weniger im Interesse des Konsumenten, als im ökonomischen Kalkül KI-operierender Unternehmen liegen. Beispielsweise warnt Minevich (2020) seine Leser auf Forbes.com in der Rolle des Konsumentenadvokaten: „algorithms can manipulate our buying and selling patterns and transactions.“ Auch Regulierungsbehörden scheinen auf diese Befürchtungen zu reagieren. Seit 2017 haben mehr als 60 Länder entsprechende Richt- oder Leitlinien zur Eindämmung von „AI harms“ entwickelt (Engler 2022). Im Oktober 2022 einigte sich die EU-Kommission auf einen „Digital Services

Act“ (European Parliament 2022). Dieses Gesetz über digitale Dienste verbietet unter anderem die Nutzung von KI in Onlineplattformen „that deceives or manipulates the recipients (...) or impairs the ability of the recipients of their service to make free and informed decisions“ (European Parliament 2022). Aufbauend darauf wurde im Juni 2023 der „EU AI Act“ angekündigt, der die Nutzung von KI zur „[c]ognitive behavioral manipulation of people or specific vulnerable groups“ verbieten möchte (European Parliament 2023). Parallel dazu arbeitet auch das Weiße Haus an einer US-amerikanischen „AI Bill of Rights“ (The White House 2022).

Ein zentrales Problem ist dabei, dass Forschung um Ethik und KI im Marketing bis dato keine Antwort darauf geben kann, wann oder welche KI den Konsumenten auf unmoralische Weise manipuliert (vgl. Donath 2021, Hermann 2022, Kaplan/Haenlein 2019). Statt Antworten auf diese Grundsatzfrage zu suchen, konzentrieren sich Forscher vielmehr auf Symptombekämpfung einzelner Problemscheinungen von KI.

Zum einen befassen sich Marketingforscher mit der Milderung negativer Konsumentenreaktionen, nachdem problematisches Verhalten von KI beobachtet wird (vgl. Dietvorst/Simmons/Massey 2016, Garvey/Kim/Duhachek 2023). Beispielsweise kann das Anthropomorphisieren von KI-Agenten (z. B. Sprachassistenten oder Chatbots) zu einem vertrauensvolleren Mensch-KI-Verhältnis beitragen, sodass sogar objektiv minderwertigen KI-generierten Produktempfehlungen Vertrauen geschenkt wird (vgl. Guha, et al. 2022, Uysal/Alavi/Bezençon 2022). Dieser erste Forschungszweig scheint manipulative Kapazitäten von KI zu bestärken, anstatt einzudämmen.

Ein zweiter Forschungszweig befasst sich mit Reaktanz von Konsumenten und folgenschwerer Datenpannen, ausgelöst durch den massiven Bedarf an persönlichen Daten zur effektiven Nutzung von KI (vgl. Cukier 2021, Martin/Borah/Palmatier 2017, Puntoni, et al. 2021). In diesem Kontext wird beispielsweise die Verwendung von Metadaten als Lösungsansatz vorgeschlagen (vgl. Kopalle, et al. 2022, Wieringa, et al. 2021). Demnach sollen nach dem Training einer KI personenbezogene Daten gelöscht und lediglich relevante Modellparameter gespeichert werden (vgl. Kopalle, et al. 2022, Wieringa, et al. 2021).

Ein dritter Forschungszweig um das moralische Spannungsfeld von Marketing-KI befasst sich mit technologischen Ansätzen und Fairnessmetriken zur Prävention algorithmischer Diskriminierung (vgl. Friedman/Nissenbaum 1996, Lambrecht/Tucker 2019). Beispielsweise soll die Aufteilung von Werbebudgets nach Geschlecht der Werbeempfänger verhindern, dass intelligente Targeting-Algorithmen aus Kostengründen vermehrt Männer mit Karriereanzeigen des naturwissenschaftlichen Bereichs bespielen (vgl. Lambrecht/Tucker

2019). Erneut scheint ein Workaround entwickelt zu werden, anstatt die moralische Legitimität der Nutzung von KI für selektive Werbekampagnen im Bildungsbereich zu adressieren.

Ein vierter Forschungszweig befasst sich mit erklärender KI, die intransparente KI analysieren und deren Entscheidungspfade ausweisen soll (vgl. Rai 2020, Ribeiro/Singh/Guestrin 2016). Obwohl zweifelsohne ein erhebendes Unterfangen, erschreckt die Tatsache, dass die komplexen Überlegungen eines KI-Agenten mittlerweile als „black box“ gelten und mit menschlicher Intelligenz allein nicht mehr nachvollzogen werden können (Ribeiro/Singh/Guestrin 2016, S. 97). Insgesamt bleibt die Frage offen, wann oder welche KI den Konsumenten unmoralisch manipuliert.

Infolge dieser Wissenslücke fordern Forscher dazu auf das Verhältnis von KI und Moral im Marketing genauer auszuleuchten (vgl. Donath 2021, Kaplan/Haenlein 2019, Martin 2019). Beispielsweise ruft Donath (2021) in ihrem Kommentar im *Journal of Marketing* unter Verweis auf die ausgeprägte Fähigkeit von KI zur gezielten Beeinflussung von Entscheidungsprozessen auf: „Now is a propitious and critical time for those who care about the ethics of AI in marketing to think deeply about how such persuasive abilities, should they come to fruition, ought to be used“ (S. 162).

Leider ist eine Antwort aufgrund eines vorgelagerten Problems – des generell ungeklärten Verhältnisses von Marketing und Manipulation – alles andere als trivial. Für die Marketingdisziplin ist deshalb die Zeit gekommen, sich „der Frage nach den Grenzen zwischen effektiver Marktbearbeitung und Manipulation“ anzunehmen (Klarmann 2020, S. 627), um sich ein Urteil über den (un-)moralischen Einsatz von KI erlauben zu können. Aufgrund zahlreicher kontroverser Perspektiven mangelt es der Marketingdomäne sogar an einer einheitlichen Definition des Manipulationsbegriffs (vgl. Gatignon/Le Nagard 2015, Sunstein 2016). Berechtigterweise stellt Wertenbroch (2016) in seinem Editorial deshalb die fundamentale Frage: „Manipulation and Marketing: The Elephant in the Room?“ (S. 209). Als Antwort hierauf befasst sich die vorliegende Dissertation mit den beiden übergeordneten Zielen herauszufinden, (1) wann Marketing-KI überhaupt manipulativ ist und (2) wie Marketing-KI moralischer gestaltet werden kann.

Ein wichtiger Aspekt dieser Dissertation ist die Tatsache, dass es gemäß der Natur ethisch-philosophischer Fragestellungen nicht die eine, allgemeingültige Antwort geben kann (vgl. Frey/Schmalzried 2013, S. 60 ff.). Manipulationswahrnehmungen entsprechen zunächst wortgetreu dem, was sie sind – subjektive Wahrnehmungen eines einzelnen Betrachters. Universalistische Urteile darüber,