

Zusammenführung der Ausarbeitungen der Vorträge der studentisch organisierten Tagung mit dem Titel:



**UKÜNFTE KI**

KI in Medien, Politik, Gesellschaft und Wirtschaft

Eine Tagung des interdisziplinären Masterstudiengangs Medien und Gesellschaft

10. - 11. Juli 2019

**Projektleitung und -verantwortung:**

Univ.-Prof. Dr. Joachim Eigler (Betriebswirtschaftslehre, insb. Medienmanagement)

PD Dr. Sandra Nuy (Politikwissenschaften)

Prof. Dr. Rainer Leschke (Medienwissenschaften)

Prof. Dr. Volkmar Pipek (Sozio-Informatik)

M.Sc. André Sekulla (Sozio-Informatik)

M.A. Julia Müller (Betriebswirtschaftslehre, insb. Medienmanagement)

## Vorwort

Im Sommersemester 2019 umfasste das Projekt des interdisziplinären Masterstudiengangs Medien und Gesellschaft eine studentisch organisierte Tagung zum Thema „Potenziale und Grenzen von Künstlicher Intelligenz in den Medien“. Der Masterstudiengang wird getragen von der Philosophischen und der Wirtschaftswissenschaftlichen Fakultät und verwirklicht ein außergewöhnliches Studienkonzept, wobei zwei Fächer und ein interdisziplinärer Studienbereich miteinander kombiniert werden. Es können dabei zwei der folgenden vier Fächer gewählt werden: Medienwissenschaft, Politikwissenschaften, Unternehmensführung und Medienmanagement sowie Sozio-Informatik. Die Tagung erhielt den Titel „Zukünfte KI“ und wurde vom Lehrstuhl der Betriebswirtschaft, insbesondere Medienmanagement von Herrn Univ.-Prof. Dr. Joachim Eigler und M.A. Julia Müller koordiniert sowie von Vertreter/-innen der jeweils anderen Disziplinen Prof. Dr. Rainer Leschke, PD Dr.in Sandra Nuy, Univ.-Prof. Volkmar Pipek und M.Sc. André Sekulla betreut. Die zweitägige Tagung fand am 10. und 11. Juli 2019 im Gründerwerk Siegen statt. Eingeladen waren Studierende und Lehrende der Universität Siegen.

Als Keynote Speaker referierte Dr. Ralf Winkler, Gründer und Geschäftsführer des Hamburger Startups QualiFiction. Das Unternehmen setzt auf künstliche Intelligenz, um Manuskripte zu analysieren und somit Bucherfolge vorherzusagen. Mit Hilfe der Software LiSA (Literatur-Screening & Analytik) können Manuskripte innerhalb kürzester Zeit auf Themen, Sentiment und Stil analysiert werden. Die Künstliche Intelligenz LiSA kann zudem anhand zuvor gelernter, auf bereits bestehenden Manuskripten beruhender Daten einen Erfolgsscore für neue Manuskripte ermitteln. Andere Gäste wie z.B. M.A. Lorenz Gilli oder M.Sc. Philipp Weber gaben Einblicke in eigene Forschungsprojekte zum Thema Künstliche Intelligenz. Neben diesen und anderen Gastvorträgen haben die Studierenden auch eigene Vorträge mit einem jeweils anderen fachlichen Schwerpunkt zum Thema Künstliche Intelligenz in den Medien vorbereitet und interdisziplinäre Einblicke gewährt. Auch die anschließenden Diskussionen boten viel Raum für einen regen Austausch. Weitere Eindrücke der Tagung können auf der offiziellen Webseite [www.zukuenfte-ki.de](http://www.zukuenfte-ki.de) eingesehen werden.

In diesem Sinne war es das abschließende Ziel des Projekts, die Vorträge und das Gelernte und Erarbeitete zu verschriftlichen. Die Texte der jeweiligen Gruppen sind nach dem Sammelbandprinzip in Vortragsreihenfolge angelegt, wobei die Anhänge zugunsten des Leseflusses extrahiert

wurden. Elisa Osoria und Alexandra Pitsch fungierten als organisatorische Leitung und verfassten die Einleitung sowie das Fazit und den Ausblick für diese Zusammenführung aller Texte.

Wir hoffen, dass die Projekttexte einen spannenden Einblick in die Thematik „Künstliche Intelligenz in Medien, Politik, Gesellschaft und Wirtschaft“ bieten!

Die Projektverantwortlichen

Univ.-Prof. Dr. Joachim Eigler

PD. Dr. in Sandra Nuy

Prof. Dr. Rainer Leschke

M.Sc. André Sekulla

M.A. Julia Müller

## **Inhaltsverzeichnis**

1. Problemstellung	1
2. Ethische Problemstellungen der KI	5
3. Automatisierte Ärzte	21
4. Microwork	37
5. KI in Filmen und Literatur	50
6. Mythos Cyborg	81
7. Einsatz von KI in der Musikproduktion	94
8. Chatbots in Geschäftsprozessen	115
9. Political Bots im Wahlkampf	132
10. Diskurs über KI in den Massenmedien	151
11. Fazit und Ausblick	172

## 1. Problemstellung

Wenn man die Apple-Sprachassistentin Siri fragt, wie das Wetter morgen wird oder wer beim letzten Bundesligaspiel das entscheidende Tor geschossen hat, ist man sich in der Regel nicht darüber bewusst, welche Prozesse sich bei dieser mittlerweile gewöhnlichen Situation im Hintergrund abspielen. Intelligente Systeme, wie die Spracherkennung, sind im alltäglichen Gebrauch und unterstützen ihre Nutzerinnen und Nutzer in vielen Lebensbereichen. Doch das Wissen über die Möglichkeiten und Risiken, die diese neuen Technologien bergen, ist bei vielen Menschen nicht vorhanden. Eine Straßenumfrage zeigte sogar, dass die Vorstellungen, die die Passanten von „Künstlicher Intelligenz“ haben, zwar mit Computern und besonders dem Internet verbunden sind, dennoch aber oft diffus und unreflektiert sind (Rocke/Osoria 2019).

Der Begriff „Künstliche Intelligenz“ (KI) geht auf den US-Wissenschaftler John McCarthy zurück: Er betitelte so eine Konferenz im Jahr 1956, die es zum Ziel hatte Fragen, wie Computer Aufgaben wie das Analysieren von Texten oder Übersetzen von Sprachen lösen sollten, zu diskutieren. Seitdem werden unter diesem Begriff Systeme zusammengefasst, die an menschliche Intelligenz angelehnt sind. So hat ein Wissenschaftler beispielsweise die Regeln des Spiels GO in ein Computerprogramm eingepflegt. Durch das ständige Spielen gegen sich selbst konnte der Computer schließlich Spielstrategien entwickeln und konnte von einem Menschen nicht mehr besiegt werden.

Dieses automatisierte Erlernen von Mustern und das stetige Verfeinern dieser Muster wird als Maschinelles Lernen (ML) bezeichnet. Als Grundlage intelligenter Systeme dient also eine Vielzahl von Daten, die in den Computer eingepflegt werden. Ein Lieferant für eben solche Daten ist unter anderem das Internet, in dem viele Millionen Menschen täglich eine große Menge an Daten produzieren. Mit jeder Suchmaschinenabfrage erhält z.B. die Suchmaschine Google Informationen und bringt diese in einen Zusammenhang. Somit lernt die Suchmaschine in Sekundenschnelle mehr und mehr über ihre Nutzerinnen und Nutzer.

Intelligente Maschinen können heutzutage Gesichter auf Bildern erkennen, Texte analysieren und produzieren sowie medizinische Diagnosen stellen. Aufgrund ihrer flexiblen Anwendungsmöglichkeiten bieten Technologien der KI eine große Hilfestellung in Unternehmen. In der öffentlichen Verwaltung etwa könnten Verwaltungsabläufe durch KI effizienter gestaltet und somit Mitarbeiterinnen und Mitarbeiter entlastet werden. Effizientere Abläufe sparen Zeit sowie Geld und vermeiden außerdem menschliche Fehler, die beim Einsatz von Maschinen in automatisierten Vorgängen nicht

aufzutreten.

Die Gründe und Möglichkeiten für den Einsatz künstlicher Intelligenz sind vielfältig, daher ist die Weiterentwicklung intelligenter Systeme für zahlreiche Branchen relevant. Schon jetzt werden künstliche Intelligenzen zur automatischen Analyse von Büchern hinsichtlich ihres Vermarktungspotenzials, zur Produktion von Musik oder aber als Bots, die einen politischen Wahlkampf in sozialen Netzwerken steuern, eingesetzt, und zukünftig werden sie möglicherweise unsere Autos fahren, medizinische Operationen durchführen und die Weltnachrichten als journalistischen Text formulieren. Sie sind in der Lage, Einfluss auf Politik und Wirtschaft zu nehmen und die Gesellschaft grundlegend zu verändern. Aufgrund dieses Potenzials ist es nötig, sich dem Thema „Künstliche Intelligenz“ interdisziplinär anzunähern.

Auf der studentisch organisierten Tagung „Zukünfte KI - Künstliche Intelligenz in Medien, Politik, Gesellschaft und Wirtschaft“ des interdisziplinären Masterstudiengangs Medien und Gesellschaft an der Universität Siegen wurde ein solcher interdisziplinärer Zugang zum Thema geschaffen. Aus der Perspektive der vier Studienfachrichtungen Medienwissenschaft, Politikwissenschaften, Unternehmensführung und Medienmanagement sowie Sozio-Informatik wurden technische, soziale, politische, ökonomische und künstlerische Fragen zum Thema vorgestellt und diskutiert. Die Interdisziplinarität der Thematik spiegelt sich ebenfalls in der Namensgebung der Tagung wider, die diese fachübergreifende Relevanz und den Schwerpunkt der Tagung, also den Einfluss von künstlicher Intelligenz in den Bereichen Medien, Politik, Gesellschaft und Wirtschaft, deutlich macht. Der Plural von Zukunft im Namen der Tagung deutet zudem darauf hin, dass sich nicht *die* Zukunft unserer Gesellschaft mit künstlicher Intelligenz modellieren lässt, sondern eine Vielzahl an möglichen „Zukünften“ vorstellbar ist.

Ziel des Projekts unter der Leitung von Prof. Dr. Joachim Eigler, Prof. Dr. Rainer Leschke, M.A. Julia Müller, PD Dr. Sandra Nuy, Univ.-Prof. Dr. Volkmar Pipek und M.Sc. André Sekulla war es, einen Versuch zu wagen, Klarheit über den Begriff der Künstlichen Intelligenz zu verschaffen. Das Aufkommen von Technologien der Künstlichen Intelligenz greift tief in die Gesellschaft ein. Bei der Diskussion über selbstfahrende Autos etwa kommt die Frage auf, wer im Falle eines Unfalls die schuldfähig gemacht werden kann - der Entwickler oder gar die Maschine selbst? Auf welcher Grundlage soll diese Entscheidung getroffen werden? Die Überlegungen kulminieren in dem oft in Spielfilmen aufgegriffenen bisher fiktiven Szenario, in dem die Maschinen intelligenter werden als die Menschen und sich somit nicht mehr unter menschlicher Kontrolle befinden. Sie treffen alle

Entscheidungen autonom. Werden diese intelligenten Maschinen schließlich als eigene Individuen angesehen und der rechtlichen Verfassung unterworfen sein? Oder behandelt man sie als Sklaven, die ausschließlich für Hilfsdienstleistungen genutzt werden?

Diesen und vielen weiteren ethischen Fragen widmete sich die Tagung in zehn Vorträgen. Eingeleitet wurde mit dem Keynote Speaker Dr. Ralf Winkler, der sich in seinem Start Up-Unternehmen QualiFiction mit der automatischen Analyse von Büchern hinsichtlich ihres Vermarktungspotenzials beschäftigt.

In dem zweiten Vortrag strukturierten Robin Walkenhorst und Louisa Sölch ethische Problemstellungen der Künstlichen Intelligenz und erörtern Lösungsansätze zu den aufgeworfenen Fragen.

Der dritte Vortrag „Autonomie vs. Unterstützung - Künstliche Intelligenz in der Medizin“, gehalten von Anna Büdenbender, Fabienne Schäfer und Ramona Sayeed, stellte die Frage, ob Künstliche Intelligenz in Zukunft den Mediziner ersetzen kann.

Auf die Simulation maschineller Arbeit von Menschen und den daraus folgenden Erkenntnissen, gingen Sebastian Randerath und Ekaterini Kirkitziotou im Vortrag über Mikroarbeit ein.

Melanie Bethge, Veronica Quint und Fee Wolkenar beschäftigten sich mit der Darstellung von Künstlicher Intelligenz in Filmen und Literatur und zogen in ihrem Vortrag Schlüsse darüber, welches Potenzial und welche Ängste Menschen der Technologie zuordnen.

In dem Vortrag „Mythos Cyborg - Feministische Perspektiven auf das Verhältnis von Mensch und Technologie“ von Jenny Berkholz und Nina Peczkowski wurde Donna Haraways Text „A Cyborg Manifesto“ vorgestellt und in Anlehnung daran wurden mögliche Auswirkungen von Künstlicher Intelligenz auf die Gendervorstellungen der Gesellschaft diskutiert.

Heutzutage ist Künstliche Intelligenz in der Lage, Musikstücke auf Grundlage einer großen Datenmenge selbst zu produzieren. Die Möglichkeiten und Probleme, die daraus resultieren, stellten Julian Kinghorst, Michael Haas und Kristina Götschenberg in ihrem Vortrag „Auswirkungen des Einsatzes von KI und ML in der Musikbranche, insbesondere in Produktion und Vertrieb“ zur Debatte.

Der Vortrag „Chatbots in Geschäftsmodellprozessen“ von Svenja Haupt, Piriya Karunathan und Jelena Kempe betrachtete die Implementierung von Chatbot-Technologien in Geschäftsprozesse zur Optimierung des Kundenbeziehungsmanagements.

In dem Vortrag von Janine Lollert, Torben Wiedenhaupt und Ansgar Kuhn ging es um automatisierte Profile in sozialen Netzwerken, sogenannte Political Bots, die durch ihre Aktivität unter anderem politische Entscheidungsprozesse beeinflussen können.

Der abschließende Vortrag „KI in den Massenmedien“ von Lars Goeritz, Boran Özer und Julian Ruppelt zeigte die Tendenzen der Berichterstattung über Künstliche Intelligenz in den drei größten Tageszeitungen Deutschlands. Damit beabsichtigten sie herauszufinden, in welchem Maße die Narrative, die sich in der Berichterstattung wiederfinden, den Diskurs über das Thema beeinflussen.

Obwohl schon seit über 60 Jahren auf dem Gebiet der Künstlichen Intelligenz geforscht wird, ist die Debatte darum noch sehr jung. Der Prozess, die neue Technologien in unsere Gesellschaft einzuhalten, hat gerade erst begonnen. In naher Zukunft wird sich zeigen, welche Bedeutung Künstliche Intelligenz für unser Leben haben wird.



## 2. Ethische Problemstellungen der KI

Vortragstag und Gruppenname:	10. Juli 15:00 Uhr, Ethische Problemstellungen der KI
Titel der Ausarbeitung:	Ethische Problemstellungen der KI
Verfasser/-innen:	Louisa Sölch, Robin Walkenhorst
Betreuer:	Prof. Dr. Rainer Leschke (Medienwissenschaften)

### Übersicht über die Struktur der Ausarbeitung

1. Einleitung
  2. Ethik
    - 2.1 Ethisches Denkmodell nach Immanuel Kant
    - 2.2 Ethisches Denkmodell nach John Stuart Mill
    - 2.3 Die Robotergesetze nach Isaac Asimov
  3. Künstliche Intelligenz
    - 3.1 Autonome Maschinen
    - 3.2 Interaktion von Menschen und Künstlicher Intelligenz
  4. Ethische Probleme im Einsatz von KI im Alltag
    - 4.1 Das Gedankenexperiment als methodisches Vorgehen
    - 4.2 Gedankenexperimente zur Veranschaulichung möglicher Problemstellungen
      - 4.2.1 KI im selbstfahrenden Auto
      - 4.2.2 KI im Krankenhaus
      - 4.2.3 KI im Betrieb
  5. Fazit
- Literaturverzeichnis
-

## 1. Einleitung

Die vorliegende Arbeit befasst sich mit ethischen Problemstellungen und Grenzen von künstlicher Intelligenz. Künstliche Intelligenzen verändern zunehmend Sitten, Gewohnheiten und den Alltag des Menschen (Buxmann 2019). Daraus resultiert die Forschungsfrage, ob der Einsatz von Künstlicher Intelligenz im Alltag ethisch legitimiert ist, beziehungsweise sich in der Zukunft ethisch legitimieren lässt. Künstliche Intelligenz wird im folgenden Text auch mit „KI“ abgekürzt.

Um sich der Lösung dieser Frage anzunähern, gehen wir methodisch mit Hilfe eines Gedankenexperimentes vor. Mögliche Perspektiven der Handhabung von Künstlicher Intelligenz, positiv wie negativ, können ausformuliert und Chancen sowie Risiken anschaulich gemacht werden. Diese Arbeit fokussiert sich dabei auf die Veranschaulichung der Risiken im alltäglichen Umgang mit Künstlicher Intelligenz. Bei der angewandten Methode werden gedanklich Situationen konstruiert in den Themengebieten KI im selbstfahrenden Auto, KI im Krankenhaus und KI im Betrieb (Kühne 2002). Diese Methode wurde für die Behandlung des vorliegenden Themas ausgewählt, weil es sich um die Analyse von zukünftigen Situationen handelt, welche zu diesem Zeitpunkt nur abschätzbar sind. Positive und negative Aspekte lassen sich durch die Veranschaulichung bestimmter Szenarien gut abbilden (Kühne 2002). Mittels der gewählten Methodik beziehen wir technische Aspekte und die reale Umsetzung beispielsweise aus Sicht der Informatik nicht mit ein. Systematisch offen bleibt der perfekte Lösungsansatz im Umgang mit Künstlicher Intelligenz. Aufgezeigt werden Grenzen und Auswirkungen der KI in ausgewählten Situationen.

Zu Beginn dieser Arbeit werden die relevanten Schlüsselaspekte der Ethik thematisiert, indem wir kurz auf die ethischen Denkmodelle nach Immanuel Kant und John Stuart Mill eingehen. In einem weiteren Schritt werden die Robotergesetze nach Isaac Asimov erläutert, da sie wesentlich zur Konstruktion und Ausführung unseres Gedankenexperimentes beitragen. Für die Grundlage des Gedankenexperiments außerdem wichtig, ist die Auseinandersetzung mit autonomen Maschinen sowie der Interaktion zwischen Mensch und Künstlicher Intelligenz. Es folgt das Gedankenexperiment zur Veranschaulichung möglicher Problemstellungen in den Teilgebieten selbstfahrende Autos, Krankenhaus und Betrieb. Bevor wir zu potentiellen Lösungsvorschlägen innerhalb des Gedankenexperimentes kommen, führen wir einen kurzen Exkurs zu der Anpassung der Robotergesetze durch. Abschließend kommen wir zu einem Fazit dieser Arbeit, in der wir unsere Ergebnisse zusammentragen.

## **2. Ethik**

### **2.1. Ethisches Denkmodell nach Immanuel Kant**

Das ethische Denkmodell nach Immanuel Kant wird als deontologische Ethik bezeichnet. Die Ethik besagt, dass das Subjekt, also der Mensch, sich selber vernünftige und moralische Gesetze gibt, nach denen er dann auch handelt. Das erkennende Subjekt steht demnach im Mittelpunkt bei Kants ethischen Denkmodells und setzt sich selber moralische Gesetze (Lern-Online 2018). Daher handelt es sich um moralische Gesetze, die die Würde des Menschen begründet. Die Pflicht sieht Kant als moralisches Prinzip an, also die Anwendung der Vernunft in der Praxis (Lern-Online 2018). Wichtig ist die Überprüfung, ob das eigene Handeln universell anwendbar für alle ist. Was der Mensch selber nicht möchte, soll er auch keinem anderen zufügen. Dies ist zentral bei dem ethischen Denkmodell nach Kant (Lern-Online 2018). Die Berücksichtigung des einzelnen Subjektes ist eine Möglichkeit mit Künstlicher Intelligenz umzugehen und sollte wie das nachfolgende ethische Denkmodell bei dem Gedankenexperiment im Hinterkopf behalten werden.

### **2.2. Ethisches Denkmodell nach John Stuart Mill**

John Stuart Mills ethisches Denkmodell wird als die Steuerung über Nutzen bezeichnet. Das Handeln des einzelnen ist nur dann gerecht und nützlich, wenn es einen möglichst großen Nutzen für die gesamte Gesellschaft bringt (BR 2009). Es ist demnach eine Theorie des Utilitarismus und der Mehrwert für alle Menschen zählt mehr als die des einzelnen Subjektes (BR 2009). Das ethische Denkmodell nach Kant steht demnach dem ethischen Denkmodell von Mill entgegen.

### **2.3. Die Robotergesetze nach Isaac Asimov**

Im Kontext der Ethik möchten wir zuletzt noch auf die Robotergesetze nach Isaac Asimov zu sprechen kommen, welche sich dadurch auszeichnen, dass sie sich bereits 1942 (Lohrmann/Eberhorn 2018) mit dem Thema Ethik von Robotern auseinandersetzen. Dies geschah im Zuge der Kurzgeschichte *Runaround*, in der jene Gesetze definiert wurden. Bestanden sie zu dem Zeitpunkt nur aus drei Gesetzen, wurde in seinem Roman *Robots and Empire* (1985) darauf aufbauend noch das nullte Gesetz angefügt (Lohrmann/Eberhorn 2018). Zusammen lauten sie wie folgt:

0. „Ein Roboter darf der Menschheit keinen Schaden zufügen oder durch Untätigkeit zulassen, dass der Menschheit Schaden zugefügt wird.
1. Ein Roboter darf einem menschlichen Wesen keinen Schaden zufügen oder durch Untätigkeit

zu-lassen, dass einem menschlichen Wesen Schaden zugefügt wird, es sei denn, dies würde das nullte Gesetz der Robotik verletzen.

2. Ein Roboter muss dem ihm von einem menschlichen Wesen gegebenen Befehl gehorchen, es sei denn, dies würde das nullte oder das erste Gesetz der Robotik verletzen.

3. Ein Roboter muss seine Existenz beschützen, es sei denn, dies würde das nullte, das erste oder das zweite Gesetz der Robotik verletzen.“ (Asimov 2016, S. 7. Nummerierung an Format angepasst, R.W.).

### **3. Künstliche Intelligenz**

#### **3.1. Autonome Maschinen**

Wenn ein System ohne menschlichen Einfluss und ohne Programmierung bis ins Detail, ein vorherbestimmtes Ziel in einem System erreicht und das selbstständig sowie angepasst an verschiedene Situationen, ist ein System autonom (Hirsch-Kreinsen 2019). Die Fähigkeit, eigenständig Entscheidungen zu treffen ist ein signifikantes Merkmal autonomer Maschinen (Hirsch-Kreinsen 2019). Treffen autonome Maschinen selbstständig Entscheidungen, dann haben sie gegebenenfalls Verantwortung für das eigene Handeln zu tragen. Verantwortlichkeit erfolgt allgemein durch Selbstzuweisung oder Zuweisung durch Andere (Hirsch-Kreinsen 2019). Wie im nachfolgenden Kapitel beschrieben, ist der Aspekt der Verantwortlichkeit auch bei der Interaktion von KI und dem Menschen entscheidend.

#### **3.2. Interaktion von Mensch und Künstlicher Intelligenz**

Ethik im Zusammenhang mit Künstlicher Intelligenz und auch die Interaktion zwischen Menschen und KI ist aktuell, da KI in sämtlichen gesellschaftlichen Bereichen Einzug hält. Deshalb ist es wichtig, sich mit den ethischen Konsequenzen zu beschäftigen (Küppers 2018). Die Kulturgeschichte zeigt, dass zwei Phänomene im Zuge von neuen Technologien einhergehen. Neue Technologien erzeugen Euphoriker und Apokalyptiker. Euphoriker sind neuen Technologien gegenüber offen und positiv gestimmt. Apokalyptiker hingegen stehen Technologieneuheiten skeptisch gegenüber und sehen negative Aspekte (Ethikrat 2017). Eine existenzielle Herausforderung sehen einige berühmte Autoren beispielsweise in Künstlicher Intelligenz und warnen vor einer möglichen „Superintelligenz“. Außerdem warnen sie davor, dass die KI den Menschen möglicherweise versklavt, da Menschen mit der KI nicht mithalten können (Buxmann 2019). Ein konfliktfreies Miteinander zwischen Menschen und KI erfordert gegenseitiges Verständnis und klar geregelte Hierarchien sowie Regeln (Küppers 2018).

Die Interaktion von Menschen und Künstlicher Intelligenz kann zu Schwierigkeiten und Ängsten führen, da das menschliche Handeln unkalkulierbar ist (Küppers 2018, S. 170). Menschen können unabhängig von ethischen und moralischen Werten plötzlich unkalkulierbar handeln und damit für Einschätzungsschwierigkeiten der Künstlichen Intelligenz führen. Die Interaktion kann sich aufgrund dessen schwierig gestalten und dazu führen, dass Menschen vor einer scheiternden Kommunikation in wichtigen Situation Angst haben (Küppers 2018, S. 170). Der Aspekt oder die Angst der Unkalkulierbarkeit und Unsicherheit bleibt oft bestehen (Küppers 2018, S. 170). Durch die Teilnahme der neuen Technologie verändern sich nicht nur Künstliche Intelligenzen, sondern auch die Menschen selbst verändern sich sowie ihr Leben und ihr Umfeld (Küppers 2018). KI verändert Wertschöpfungsprozesse, die Interaktion und Kommunikation, worauf wir im Gedankenexperiment später näher eingehen (Buxmann 2019).

Wie Menschen Künstlicher Intelligenz entgegentreten, hat auch etwas mit ihrer Kultur zu tun, und in welchem System sie sich befinden (Rath 2019). Die Fragen der Verantwortlichkeit und der Hierarchie von Menschen und Künstlicher Intelligenz sind die zentralen Themen bei der Interaktion zwischen Menschen und KI. Sie entscheiden über das zukünftige System. Drei mögliche Rollenverständnisse sind zwischen KI und Mensch aus der Vergangenheit sehend möglich. Gleichheit zwischen KI und Menschen, der Mensch versklavt die KI oder die KI versklavt den Menschen (Buxmann 2019). Diese drei möglichen Hierarchien spielen auch im folgenden Gedankenexperiment eine Rolle.

#### **4. Ethische Probleme im Einsatz von KI im Alltag**

##### **4.1. Das Gedankenexperiment als methodologisches Vorgehen**

Nachdem wir uns in den letzten Abschnitten nun der Ethik annahmen und infolgedessen die künstlichen Intelligenzen definierten, möchten wir nun in unserem Hauptteil das Gelernte anwenden, um schlussendlich einen besseren Überblick zu bekommen, wie weit eine KI ethisch sein kann, wie sie es sein könnte oder wo es vielleicht gänzlich auszuschließen ist. Dafür haben wir uns drei Bereiche ausgesucht, die im Kontext von KI im Alltag öfters Erwähnung finden und wir einmal genauer und kritisch begutachten möchten. Dabei handelt es sich zum einen um KI gesteuerte selbstfahrende Autos, KI im Krankenhaus sowie in einem Betrieb, welche wir nacheinander analysieren werden.

Da in und an diesen Bereichen zwar schon geforscht und gearbeitet wird, der tatsächliche und großflächige Einsatz von KI jedoch noch Zukunft ist, haben wir uns, wie in der Einleitung angesprochen, dazu entschieden, dies in Form eines Gedankenexperiments zu tätigen. Darin sahen wir allem voran den Vorteil, schon im Vorhinein mögliche – wahrscheinliche wie unwahrscheinliche – Konsequenzen abzuschätzen und darauf zu reagieren. Auch aufgrund der Tatsache, dass wir im Verlauf der Arbeit auf das philosophische *Trolley-Problem* zu sprechen kommen, bestärkte uns in der Entscheidung, das Gedankenexperiment als grundlegende Methode zu wählen.

## **4.2. Gedankenexperiment zur Veranschaulichung möglicher Problemstellungen**

### **4.2.1. KI im selbstfahrenden Auto**

Wie angesprochen widmen wir uns nun zunächst dem Thema des selbstfahrenden Autos – also Kraftfahrzeugen, welche keine\*n menschlichen Fahrer\*in benötigen, sondern rein maschinell und programmierungstreu agieren. Eine Tatsache, die in Deutschland noch immer sehr kritisch beäugt wird, da „nicht an die Verlässlichkeit der Fahrzeugtechnologie [geglaubt wird] oder Angst vor Hackern“ (ADAC 2019) vorherrscht. Doch auch ein altes, philosophisches Dilemma – das sog. *Trolley-Problem* – sorgt für Unbehagen.

Um diese Probleme und Kritiken anzugehen, haben wir uns mehrere Szenarien überlegt, die wir für denkbar oder zumindest für würdig erachteten, in Betracht gezogen zu werden. Dabei erdachten wir Situationen, in denen ein menschliche\*r Nutzer\*in sein\*ihr selbstfahrendes Auto steuert, das Ziel eingibt und sich daraufhin von diesem transportieren lässt. Drei Ausgänge hielten wir dabei für möglich. Das Wünschenswerteste wäre dabei, dass der Mensch unbeschadet am Zielort ankommt und die KI also sowohl vorschriftsmäßig agierte, als auch, dass keine unvorhersehbaren Ereignisse eintraten. Kritisch seien die Möglichkeiten, dass der Mensch lediglich verletzt ankommt, andere verletzt wurden oder es gar zu einem oder mehreren Todesfällen kam, oder dass die KI ihn gar nicht transportierte. Gründe für die erste Situation könnten menschliches Versagen (Geisterfahrer\*innen oder Verkehrsteilnehmer\*innen oder Fußgänger\*innen, welche die Verkehrszeichen missachteten) oder unvorhersehbare Ereignisse wie Naturgewalten (bspw. Erdbeben) sein. Für das zweite Szenario reichen die möglichen Gründe von Einstellungs- oder Programmfehlern bis hin zu der Möglichkeit, dass die KI den Menschen selbst als nicht alle Sicherheitsvorschriften erfüllend einstuft.

#### **4.2.2. KI im Krankenhaus**

Als nächsten kritischen Bereich nahmen wir uns dem Gebrauch von KI im Krankenhaus an. Dieser lässt sich aufgrund der vielfältigen Einsatzmöglichkeiten in verschiedene Unterbereiche aufteilen, von denen wir uns einmal die KI als Ärztin und einmal als Pflegerin erdachten. In beiden Bereichen sind die möglichen Szenarien breiter gefächert als die des selbstfahrenden Autos. Beispielsweise kann eine KI als Ärztin den Menschen zwar korrekt behandeln, doch bedeutet dies nicht zwangsläufig, dass die KI anhand der zur Verfügung stehenden Daten die richtige Diagnose erstellt hat, auch wenn dies den bestmöglichen Ausgang bedeuten würde. Wenn auch sehr unwahrscheinlich, kann nicht in Gänze ausgeschlossen werden, dass die falsche Diagnose erarbeitet wurde, im Zuge dessen aber auch die falsche Medikation oder Behandlung ausgewählt wurde, was wiederum zum korrekten Ergebnis führte. Denkbar ist aber auch das Szenario, dass der Mensch falsch behandelt wird – sei es durch die Analyse falscher Symptome, veralteter oder falscher Datensätzen oder die KI analysiert mehrere mögliche Ergebnisse und trifft dabei eine falsche Entscheidung. Das Schreckensszenario wäre, dass der\*die Patient\*in gar nicht behandelt wird, bspw. indem die KI ihn\*sie als lebensunwert definiert oder „aktive Sterbehilfe“ als einzige Möglichkeit sieht, ihm\*ihr zu helfen.

Ähnlich verhält sie sich in der Pflege, in der ebenfalls nicht auszuschließen ist, dass der\*die Patient\*in menschenunwürdig oder gar nicht betreut wird, da die KI ihn\*sie als lebensunwert definiert. Die positive und wünschenswerte Situation wäre aber, dass der Mensch gewissenhaft behandelt und versorgt wird. In dem Falle hätte die KI programmierungskonform gehandelt und auch sämtliche Aufgaben erfüllt, die einem\*einer menschlichen Pfleger\*in zuteilgeworden wären. Dabei ist aber auch dies mit Vorsicht zu genießen, da nicht auszuschließen wäre, dass dem\*der Patient\*in der menschliche Kontakt fehlt, er\*sie vereinsamt und ggf. psychische Schäden davonträgt, was wiederum dazu führt, dass der Mensch zwar gewissenhaft versorgt wurde, aber nicht in dem Maße, wie es ein Mensch benötigt.

#### **4.2.3. KI im Betrieb**

Im Betrieb haben wir uns ebenfalls zwei Unterbereiche ausgesucht, die wir thematisieren wollten. Dies wäre zum einen die KI als Betriebsleiterin und zum anderen in der Funktion als Bewerbungsleiterin. Im ersten Beispiel fanden wir lediglich kritische Möglichkeiten vor, bspw. in dem Falle, dass die KI die ihr untergeordneten Menschen im Sinne des Betriebs bewerten müsste, wobei die jeweiligen Faktoren nicht zwangsläufig bekannt sein müssen. Dies würde

besonders kompliziert, wenn die KI auch die Funktion erfüllen soll, Arbeiter\*innen zu entlassen, wenn diese die Anforderungen nicht erfüllen. Fraglich ist dabei vor allem die Weise, wie die KI entscheidet, wer für den Betrieb (wie) nützlich und wer ersetzbar ist. Am denkbarsten wäre, dass dies durch eine implementierte Relevanzhierarchie geschähe, die die Menschen bewertet und sortiert – ähnlich dem geplanten und ebenfalls KI gesteuerten Sozialkreditsystem Chinas (Dorloff 2019). Aber auch simplere Aufgaben wie das Planen und Verwalten von Terminen müssen kritisch betrachtet werden, da eine KI nicht zwangsläufig menschliche Gründe nachvollziehen kann und diese zu einer Abstufung der eigenen Wertung und schlimmstenfalls einer unrechtmäßigen Kündigung führen könnten.

Weniger vielfältig – aber nicht weniger negativ – verhält sich die KI als Bewerbungsleiterin. Hier ist wie im zuvor genannten Szenario das hauptsächliche Problem, dass die KI nicht auf menschliche Eigenschaften eingehen kann und lediglich die zur Verfügung stehenden Daten aus dem Lebenslauf o.ä. einsehen und bewerten kann, welche den\*die Bewerberin nicht zwangsläufig so darstellt, wie er\*sie tatsächlich ist und von einem\*einer menschlichen Bewerbungsleiter\*in wahrgenommen würde.

### **4.3. Rekurs: Anpassung der Robotergesetze**

Nachdem wir nun anhand des Gedankenexperiments mögliche Problematiken, die sich durch den Gebrauch von Künstlichen Intelligenzen ergeben können, herausgearbeitet haben, wollen wir nun versuchen, Lösungsvorschläge für diese zu erarbeiten. Dazu kommen wir an dieser Stelle noch einmal auf die an anderer Stelle thematisierten Robotergesetze von Isaac Asimov zurück, an welchen wir uns zu diesem Zwecke orientierten. Anzumerken sei hierbei, dass jene zwar schon stellenweise in die Programmierung von Künstlichen Intelligenzen Einzug gefunden haben (Bruneau 2017), an vielen Stellen jedoch sehr ungenau definiert sind – ein Zustand, den sich Google auch für eigene Regeln zunutze machte (Költzsch 2018).

In den Robotergesetzen Asimovs fällt dabei vor allem auf, dass zwar zwischen der Menschheit und einem einzelnen menschlichen Wesen unterschieden wird, erstere aber nie klar definiert wird. Der Duden (o.J.) legt Menschheit als die „Gesamtheit der Menschen“ fest, wohingegen der Kontext von Asimovs Regeln auch so gelesen werden kann, dass lediglich eine Mehrzahl an Menschen gemeint ist. Auch die Verwendung des Begriffs Schaden muss ohne genauere Begriffsbestimmung kritisch gesehen werden und darf sich nicht nur auf das physische Unwohlsein beziehen. Sollten Künstliche Intelligenzen tatsächlich einmal mit jenen Gesetzen



ausgestattet werden, um ihnen Ethik „beizubringen“, ist es unabdingbar, missverständliche Formulierungen in dem Maße zu reduzieren oder ganz zu entfernen, dass sie Streitfälle im Vorhinein ausschließen. Dies betrifft vor allem die ersten beiden – also dem nullten wie ersten – Gesetze der Robotik, welche unserer Meinung nach wie folgt umgeschrieben werden müssten:

5. Ein Roboter darf **einer Menschenmenge (ab 2 menschlichen Wesen)** oder der **gesamten** Menschheit keinen Schaden zufügen oder durch Untätigkeit zulassen, dass **einer Menschenmenge (ab 2 menschlichen Wesen)** oder der **gesamten** Menschheit Schaden zugefügt wird.
  
6. Ein Roboter darf einem menschlichen Wesen **weder physischen oder psychischen** Schaden zufügen **oder dies riskieren, noch ihn (nach Definition der Rechtsgrundlage des Bestimmungslandes) diskriminieren** oder durch Untätigkeit zulassen, dass einem menschlichen Wesen **physischer oder psychischer** Schaden zugefügt oder **es (nach Definition der Rechtsgrundlage des Bestimmungslandes) diskriminiert wird**, es sei denn, dies würde das nullte Gesetz der Robotik verletzen.

Die Punkte 2 und 3 bleiben unverändert. Wichtig ist auch der Vermerk, dass sich die künstlichen Intelligenzen an die jeweiligen Gesetze des Bestimmungslandes – also dort, wo die KI eingesetzt wird – halten müssen, um sicherzustellen, dass sie nicht nur den Gesetzen des Herstellungslandes entspricht und absichtlich oder versehentlich den des Bezugslandes widerspricht und also ggf. dort Straftaten begeht. Oder im Fall von IBMs *Watson* in einem Land trainiert wird, in dem es üblicher ist, teurere Medikamente zu verschreiben, als unbedingt notwendig (Baltzer 2018) und somit in anderen Ländern unbrauchbar wird.

Dies führt zu einem weiteren wichtigen Punkt, der nicht außer Acht gelassen werden darf: die Verantwortung des Herstellerunternehmens. Diesem wird als Produzent der KI eine besondere Verpflichtung zuteil, die gesetzlich geregelt werden muss, um die Haftbarkeit möglicher Streitfälle zu definieren noch ehe sie geschehen. Außerdem soll es eine Absicherung dagegen darstellen, dass unfertige oder fehlerhafte Produkte frühzeitig auf den Markt gebracht werden, bspw. um eine Deadline einzuhalten. Daraus ergibt sich das fünfte – nach der Zählung Asimovs vierte – Gesetz der Robotik, welches wie folgt lauten kann:

161: Das oder die Unternehmen, welches oder welche den jeweiligen Roboter a) hergestellt, b) vermarktet und/oder c) vertrieben hat oder haben (in ebendieser Reihenfolge), steht oder stehen in der Verantwortung, zu garantieren, dass der Roboter sich an das nullte bis dritte Gesetz der Robotik hält und nicht imstande ist, diese eigenständig außer Kraft zu setzen. Sämtliche Konsequenzen, die sich aus einem Missachten jenes Gesetzes ergeben, fallen zulasten des oder der eingangs genannten Unternehmen, sofern jenes oder jene nicht zweifelsfrei nachweisen kann oder können, dass ein Fremdverschulden wie unsachgemäßer Gebrauch oder nichtlizenzierte Reparaturen oder Erweiterungen stattgefunden hat.

Nachdem nun die Gesetze aktualisiert und auf den Gebrauch angepasst wurden, möchten wir sie nun anwenden, um den zuvor erarbeiteten Problemfällen möglichst effektiv entgegen zu treten.

#### **4.4. Gedankenexperiment zu potentiellen Lösungsvorschlägen**

##### **4.4.1. KI im selbstfahrenden Auto**

Wie angesprochen, erfolgt nun die Anwendung der neuen Robotergesetze auf die zusammengetragenen Problematiken. Im Kontext des selbstfahrenden Autos erarbeiteten wir zwei mögliche Komplikationen. Zum einen, wenn die KI sich weigert zu fahren oder sie zum anderen auf dem Weg den\*die Mitfahrende\*n oder Passanten verletzt oder gar tötet. Für den ersten Fall können wir uns auf das zweite Gesetz der Robotik berufen, welches ausschließt, dass die KI sich weigern kann, einen Menschen zu transportieren, sofern dieser keine Gefahr für sich oder andere darstellt. Eine solche Gefahr könnte bspw. der Befehl zum Geisterfahren sein, in welchem sich die KI weigern würde – und müsste –, den Befehlen des Menschen zu gehorchen.

Im zweiten Falle kommen nun das neudefinierte nullte und erste Gesetz der Robotik zum Tragen, welche in Kombination besagen, dass die KI gleichermaßen für den\*die Mitfahrende\*n wie auch sämtlichen Passanten Sorge zu tragen habe. Im Falle eines unvermeidbaren Unfalls bedeutete dies, dass die KI primär die Menschen außerhalb des Autos schützen und also versuchen müsste, eher in eine Grube oder gegen einen Baum zu fahren, als tatsächlich eine\*n Außenstehende\*n zu verletzen. Für den\*die Mitfahrenden muss anderweitig gesorgt werden, bspw. durch den großflächigen Einsatz von Airbags und anderen implementierten Sicherheitsvorkehrungen, um auch diese\*n bestmöglich zu schützen. Dies löst das zugrundeliegende Trolley-Problem jedoch nur bedingt, da dieses i.d.R. davon ausgeht, dass es menschliche Opfer geben muss (Crockett 2016). Auch lässt sich nicht genau absehen, wie das nullte und erste

Gesetz in einer solchen Situation mit dem zweite interagiert, falls die KI die Zieleingabe und Fahraufforderung als Befehl wertet, welcher plötzlich negiert würde, da Menschen in Gefahr geraten sind.

#### **4.4.2. KI im Krankenhaus**

Im Krankenhaus verhält es sich ein wenig anders. Im Kontext der KI als Ärztin gab es gleich mehrere problematische Situationen bspw., wenn die KI den Menschen zwar korrekt behandelt, dies aber lediglich aufgrund eines doppelten Fehlschlusses tut. Um dies zu vermeiden, muss das vierte Gesetz bemüht werden, welches das Unternehmen in die Verantwortung setzt, die KI erst dann auf den Markt zu bringen, wenn diese Fehler auszuschließen sind. Das produzierende Unternehmen muss nachvollziehen können, wie die KI agiert, da sie diese programmiert haben, sowie die volle Verantwortung tragen, sollte dennoch einmal ein Mensch zu Schaden kommen. Dasselbe Gesetz tritt in Kraft, um zu verhindern, dass die KI falsche Symptome analysiert oder sich auf veraltete oder falsche Quellen beruft und so einen Schaden am Menschen riskiert. Im Zuge der Möglichkeit, dass die KI in eine Situation kommt, in der sie raten müsste, tritt das erste Gesetz in Kraft, welches in diesem Falle besagt, dass die KI nicht raten darf, da sie sonst wie zuvor einen Schaden nicht ausschließen kann. Auch kann eine KI sich nicht weigern, jemanden zu behandeln, da sie sonst gegen das zweite Gesetz verstößt, indem sie sich sonst den Befehlen des Menschen widersetzen würde. Dahingehend kann sie den\*die Patient\*in auch nicht als „lebensunwert“ definieren, da sie sonst gegen das erste Gesetz verstößt.

In der Pflege war einer der Problemfälle, dass die KI keine\*n menschlichen Pfleger\*in ersetzen könnte, ohne zu riskieren, dass diesem ebenjener menschliche Kontakt fehlen würde. Dies ließe sich zumindest reduzieren, indem die KI nicht alleine agieren kann und auch immer einen Menschen an seiner Seite braucht. Jedoch bedeutet dies auch, dass eine KI nicht in der Lage wäre, allein den Ansprüchen der Pflege zu genügen, auch wenn sich zumindest die Problematik einer Behandlungsweigerung wie im Beispiel zuvor unterbinden ließe. Dies bedeutet, dass eine KI in der Pflege ungeeignet ist und lediglich unterstützende Aufgaben übernehmen dürfte.

#### **4.4.3. KI im Betrieb**

Abschließend möchten wir nun noch einmal auf das Thema des Betriebs zu sprechen kommen, in dem wir die KI als Betriebsleiterin und Bewerbungsleiterin erdachten. Im ersten Falle ließe

sich zumindest der Kritikpunkt der Intransparenz bezüglich der Bewertung der Arbeitnehmer\*innen beseitigen. Dies wäre der Fall, wenn noch immer ein Mensch über der KI steht und dieser die Faktoren kennt respektive das Herstellerunternehmen nach dem vierten Gesetz genau sagen können muss, wie die KI agiert, oder sie sogar auf die Bedürfnisse des\*der Kund\*in angepasst haben, wofür sie die Verantwortung tragen. Dies ändert jedoch nichts an der Tatsache, dass die KI diese gesammelten Daten niemals gleich eines Menschen verarbeiten könnte, wodurch die anderen Kritikpunkte bestehen bleiben und die KI riskierte, den Menschen zu entmenschlichen, indem sie ihn zwingt, sich den Anforderungen der Maschine anzupassen und er somit metaphorisch selbst zur Maschine werden müsste. Dies stellt jedoch einen Verstoß gegen das erste Gesetz dar, was die KI in dem Bereich untauglich macht.

Ähnlich verhält sich die KI als Bewerbungsleiterin. Durch das erste Gesetz ließe sich zwar ausschließen, dass die KI persönliche Daten wie Herkunft, Geschlecht, Alter oder sozialen Status in die Bewertung einbeziehen kann (da sie sonst diskriminierend agierte), doch bleiben als letzte für die Maschine lesbare Daten nur die zuvor genannten harten Fakten wie Schul- und Abschlussnoten, Studienzeit und Ähnlichem, was einen Menschen niemals realistisch abbildet. Auch dies trägt dazu bei, die KI im Bewerbungsverfahren als ungeeignet wahrzunehmen. Nicht zuletzt, da es auch dazu beitragen könnte, Menschen, die ohnehin schon schlechte Chancen auf dem Arbeitsmarkt haben, gänzlich ohne Chancen dastehen lassen könnten, was in jedem Fall vermieden und verhindert werden muss.

## **5. Fazit**

Nachdem wir also nun das Gedankenexperiment abgeschlossen haben und so zumindest aufzeigen konnten, welche Gefahren bedacht werden müssen und was man tun könnte, um diese zu vermeiden, möchten wir abschließend unsere Ergebnisse noch einmal zusammentragen und kurz auf Themen zu sprechen kommen, die ebenfalls eine Relevanz für den zukünftigen Umgang mit diesem Gebiet besitzen. Wie zu sehen war, ist das Thema der Ethik im Kontext von Künstlichen Intelligenzen nicht leicht und eindeutig zu klären und an vielen Stellen nur schwerlich durchführbar, wenn gänzlich auf menschliche Aufseher\*innen verzichtet werden soll. Das Thema Autonomie ist dabei ein besonderer Kritikpunkt, der zwar wünschenswert, noch aber ethisch unvertretbar wäre. Eine Annäherung an die Probleme stellen zwar die hier aktualisierten Robotergesetze dar, doch können auch diese nicht alles lösen und mehr aufzeigen, wo es gilt, noch einmal ein stärkeres Augenmerk darauf zu richten. Auch wurden viele

Sorgenfälle herausgearbeitet, die auch oft eine Frage der Verantwortlichkeit sind, welche genau definiert werden muss. Doch auch diese zieht weitere Probleme nach sich. Bspw. darf nicht vernachlässigt werden, dass ein großflächiger Einsatz von KI auch eine globale Verantwortung nach sich zieht. Diese sehen wir vor allem in der Notwendigkeit der Definition von globalen Menschenrechten, sodass ausgeschlossen werden kann, dass diese Technologien dazu genutzt werden, Menschen zu unterdrücken, wie es möglich wäre, wenn die KI in einem Land genutzt wird, dass ebenjene Unterdrückung gesetzlich legitimiert.

Doch auch der Weg hin zur Digitalisierung und zum Einsatz von KI erfordert noch einiges an Arbeit. In der Verantwortung sehen wir dabei vor allem den Staat, der Sorge zu tragen hat, dass dies nicht zulasten der arbeitenden Bevölkerung geschieht. So werden im Zuge der Digitalisierung viele Menschen ihren Job verlieren, was dadurch aufgefangen werden soll, dass im gleichen Maße auch neue Arbeitsplätze entstehen werden. So spricht Florijn (2019) davon, „dass etwa 490.000 Jobs verloren gehen und 430.000 neue Jobs geschaffen werden“. Dass sich diese Zahlen ausgleichen, geschieht jedoch nicht von alleine und erfordert eine verstärkte Unterstützung und Einflussnahme der Regierung, welche Fort- und Weiterbildungsangebote für die Betroffenen organisieren und anbieten muss, um die Betroffenen schnell und unkompliziert in neue Berufe ein- respektive umzulernen.

Unabhängig davon werden im Kontext von KI auch andere Denkmodelle immer öfters thematisiert, bspw. das des bedingungslosen Grundeinkommens – also, dass alle Menschen einen monatlichen Grundbetrag zur Existenzsicherung bekommen, welcher an keine Verpflichtungen (wie bspw. Arbeit) gebunden ist (Bohsem 2016) –, wobei die Bundesagentur für Arbeit und Soziales (Hrsg.) (2017) „weder eine Notwendigkeit noch eine gesellschaftliche Akzeptanz für einen so grundlegenden Systemwechsel“ sieht. Kritikpunkt dabei ist vor allem auch oft der der Finanzierung, weswegen im Zuge dessen auch über eine Maschinensteuer diskutiert wird (Bohsem 2016). Sehen wir eine solche Überlegung zwar positiv, auch um die erwähnte Weiterbildungsangebote finanzieren zu können, darf dabei aber nicht aus den Augen verloren werden, dass sich die Unternehmen im Zuge dessen, und vielleicht sogar um Geld zu sparen, nicht gänzlich den neuen Technologien verweigern.

## Literaturverzeichnis

ADAC (2019): Autonomes Fahren: Digital entspannt in die Zukunft, [online] <https://www.adac.de/rund-ums-fahrzeug/ausstattung-technik-zubehoer/autonomes-fahren/technik-vernetzung/aktuelle-technik/> [14.12.2019].

Asimov, Isaac (2016): Ich der Roboter. München: Heyne.

Balzter, Sebastian (2018): Supercomputer Watson. Im Krankenhaus fällt die Wunderwaffe durch, [online] <https://www.faz.net/aktuell/wirtschaft/kuenstliche-intelligenz/computer-watson-scheitert-zu-oft-bei-datenanalyse-15619989.html> [16.05.2019].

Bendel, Oliver (2018): ÜBERLEGUNGEN ZUR DISZIPLIN DER MASCHINENETHIK, [online] [http://maschinenethik.net/wpcontent/uploads/2018/02/Artikel\\_Maschinenethik\\_APuZ.pdf](http://maschinenethik.net/wpcontent/uploads/2018/02/Artikel_Maschinenethik_APuZ.pdf) [04.12.2019].

Bohsem, Guido (2016): Abgabe auf Roboter, [online] <https://www.sueddeutsche.de/wirtschaft/maschinensteuer-abgabe-auf-roboter-1.3081956> [10.12.2019].

BR (2009) Das Leben von John Stuart Mill, [online] <https://www.br.de/radio/bayern2/sendungen/radiowissen/ethik-und-philosophie/mill-utilitarismus-thema100.html>

Bruneau, Charlotte (2017): Die Maschinen werden autonom, [online] [https://www.deutschlandfunk.de/roboterethik-die-maschinen-werden-autonom.724.de.html?dram:article\\_id=385826](https://www.deutschlandfunk.de/roboterethik-die-maschinen-werden-autonom.724.de.html?dram:article_id=385826) [01.07.2019].

Bundesministerium für Arbeit und Soziales (Hrsg.) (2017): Weißbuch Arbeit 4.0. Berlin: o.V.

Buxmann, Peter; Schmidt, Holger (Hg.) (2019): Künstliche Intelligenz. Mit Algorithmen zum wirtschaftlichen Erfolg. Berlin, Heidelberg: Springer Berlin Heidelberg. [online] <https://doi.org/10.1007/978-3-662-57568-0>.

Crockett, Molly (2016): The trolley problem: would you kill one person to save many others?, [online] <https://www.theguardian.com/science/head-quarters/2016/dec/12/the-trolley-problem-would-you-kill-one-person-to-save-many-others> [08.12.2019].

Dorloff, Axel (2019): Die Supermacht der Algorithmen, [online] [https://www.deutschlandfunkkultur.de/kuenstliche-intelligenz-in-china-die-supermacht-der.979.de.html?dram:article\\_id=439978](https://www.deutschlandfunkkultur.de/kuenstliche-intelligenz-in-china-die-supermacht-der.979.de.html?dram:article_id=439978) [02.12.2019].

Duden (o.J.): Menschheit, [online] <https://www.duden.de/rechtschreibung/Menschheit> [07.12.2019].

Ethikrat (2017), [online] Autonome Systeme. Wie intelligente Maschinen uns verändern [online] [https://www.ethikrat.org/fileadmin/PDF-Dateien/Veranstaltungen/2017-06-21-Jahrestagung-Autonome-Systeme\\_UF.pdf](https://www.ethikrat.org/fileadmin/PDF-Dateien/Veranstaltungen/2017-06-21-Jahrestagung-Autonome-Systeme_UF.pdf) [10.12.2019].

Florijn, Charleen (2019): Wie künstliche Intelligenz unsere Jobs gefährdet, [online] <https://www.karriere.de/mein-naechster-job/zukunft-der-arbeit-wie-kuenstliche-intelligenz-unsere-jobs-gefaehrdet/24852150.html> [09.12.2019].

Hirsch-Kreinsen, Hartmut; Karačić, Anemari (Hg.) (2019): Autonome Systeme und Arbeit. Perspektiven, Herausforderungen und Grenzen der Künstlichen Intelligenz in der Arbeitswelt. Bielefeld: transcript. [online] [http://www.content-select.com/index.php?id=bib\\_view&ean=9783839443958](http://www.content-select.com/index.php?id=bib_view&ean=9783839443958).

Költzsch, Tobias (2018): Googles KI-Regeln. Auf den Spuren Asimovs, [online] <https://www.golem.de/news/googles-ki-regeln-auf-den-spuren-asimovs-1806-134848.html> [20.05.2019].

Kühne, Ulrich (2002): Die Methode des Gedankenexperiments. [online]

Küppers, E. W. Udo (2018): Die humanoide Herausforderung. Leben und Existenz in einer anthropozänen Zukunft. Wiesbaden: Springer. [online] <http://dx.doi.org/10.1007/978-3-658-17920-5>.

Lern-Online (2018): Kants Pflichtethik/ Kategorischer Imperativ & Utilitarismus im Vergleich. [online] <https://www.lern-online.net/philosophie/ethik/kants-ethik/kants-hypothetischer-kategorischer-imperativ-utilitarismus-vergleich/>

Lohrmann, Julia/Eberhorn, Johannes (2018): Isaac Asimov, [online] <https://www.planet->

wissen.de/technik/computer\_und\_roboter/roboter\_mechanische\_helfer/pwieisaacasimov1 00.html  
[01.12.2019].

Rath, Matthias; Krotz, Friedrich; Karmasin, Matthias (Hg.) (2019): Maschinenethik. Normative Grenzen autonomer Systeme. Wiesbaden: Springer Fachmedien Wiesbaden (Ethik in mediatisierten Welten). [online] <https://doi.org/10.1007/978-3-658-21083-0>.



### 3. Automatisierte Ärzte

Vortragstag und Gruppenname:	10. Juli 11:15 Uhr, Automatisierte Ärzte
Titel der Ausarbeitung:	Autonomie vs. Unterstützung Künstliche Intelligenz in der Medizin
Verfasserinnen:	Anna Katharina Bündenbender, Ramona Sayeed Fabienne Schäfer
Betreuer:	Prof. Dr. Volkmar Pipek (Sozio-Informatik) M.Sc. André Sekulla

#### Übersicht über die Struktur der Ausarbeitung

1. Einleitung
  2. Methodisches Vorgehen
  3. Beispielhafte Systeme und die Einordnung
    - 3.1 Da Vinci Operationsroboter
    - 3.2 Ada Health und M-Sense
    - 3.3 Pflegeroboter: Pepper und Paro
    - 3.4 IBM Watson Health
  4. Kritische Reflexion
  5. Fazit
- Literaturverzeichnis
-

## 1. Einleitung

„Ärztinnen und Ärzte, Pflegepersonal sowie Patientinnen und Patienten beklagen schon lange die Überlastung des medizinischen Betriebs. Zu viel Zeit koste die Bürokratie in Praxen und Kliniken, zu wenig Zeit bleibe für das persönliche Gespräch zwischen Arzt und Patient“ (Bundesministerium für Bildung und Forschung, 2019, Absatz 1). Dieses Zitat lässt erahnen, dass der Medizin- und Gesundheitssektor in Deutschland eine Veränderung benötigt. Doch kann künstliche Intelligenz hierfür eine Hilfe sein?

Die vorliegende Arbeit befasst sich mit Formen der Künstlichen Intelligenz im Bereich Medizin und Gesundheit. Das Anwendungsfeld gestaltet sich derzeit sehr breit. Neben intelligenten Computersystemen, wie *Watson* von IBM (Höpner, 2019, Absatz 2), werden Patienten<sup>1</sup> mit einer Vielzahl an diversen Gesundheits-Apps überflutet (Van Lieshout et al., 2019, S.18). Darüber hinaus entwickelt sich die Forschung um medizinische Robotik sehr schnell. Schon jetzt findet das roboter-assistierte Chirurgesystem *Da Vinci* bei zahlreichen minimalinvasiven Operationen Einsatz (Meißner, 2019, Absatz 7). Aus diesem Grund stellt sich die Frage: Wird in Zukunft KI den Mediziner ersetzen oder nur unterstützen?

Um sich dieser Frage anzunähern, wurde eine Skala entwickelt, welche die einzelnen Stufen der Autonomie im Bereich Medizin verdeutlicht. Dafür dienen als Basis die Stufen des Autonomen Fahrens der internationalen Norm SAE J3016 sowie die Taxonomie von Lernzielen nach Benjamin Bloom. In einem weiteren Schritt werden ausgewählte Anwendungen aus dem Bereich Medizin, die sich an Formen der KI bedienen, in die eigens entwickelte Skala eingeordnet, um darauf aufbauend kritisch reflektieren zu können, welche ethischen Konsequenzen mit der Nutzung von KI in der Medizin einhergehen. Zu den Anwendungen zählen, wie bereits erwähnt, das System *Watson* von IBM und der OP-Roboter *Da Vinci*. Darüber hinaus werden aus der Kategorie Gesundheits-Apps *Ada Health* und *M-sense* thematisiert, sowie die japanischen Pflegeroboter *Pepper* und *Paro*. Im Anschluss daran werden die ethischen Konsequenzen, im Hinblick auf politisch/gesetzliche, ökonomische und gesellschaftliche Risiken erörtert, um abschließend in einem Fazit die Forschungsfrage zu beantworten.

## 2. Methodisches Vorgehen

Um sich der erläuterten Thematik anzunähern, wurde eine Skala für die Medizin- und Gesundheitsbranche entwickelt. Darüber hinaus wurden zwei Experten zum Thema KI in der Medizin ausführlich interviewt.

Mit Hilfe der Skala besteht die Möglichkeit, unterschiedliche medizinische Anwendungen anhand ihres Autonomiegrades in die verschiedenen Stufen einzuordnen. Ein solches Modell, wie es in der Automobilbranche bereits existiert, gibt es aktuell für die Bereiche Medizin und Gesundheit noch nicht (persönliche Kommunikation, 11. Juni 2019). Somit bedient sich die Thematik an dieser Stelle eine Forschungslücke.

Die erarbeitete Skala gibt Aufschluss darüber, wie weit der technologische Fortschritt derzeit in der Medizin ist. Ist das einzuordnende KI-System nur eine Art Assistenzsystem oder handelt es sich um ein automatisiertes System? Darüber hinaus wird anhand der Skala deutlich, welche Rolle der Mediziner einnimmt, wenn das einzuordnende System in Nutzung ist. Im Hinblick auf die Forschungsfrage wird also deutlich, ob die KI nur unterstützend oder autonom agiert. Außerdem lassen sich Aussagen darüber treffen, inwieweit ethische Konsequenzen und Risiken sich innerhalb einer Stufe ähneln oder ob es diesbezüglich Unterschiede gibt.

Für das Projekt konnten zwei Experten gewonnen werden, die sich mit der Thematik täglich auseinandersetzen. Dr.-Ing. Kai Hahn ist wissenschaftlicher Mitarbeiter am Lehrstuhl für medizinische Informatik und Mikrosystementwurf an der Universität Siegen. Hahn beschäftigt sich in seinen Projekten mit Mikro- und Nanosystemen im privaten sowie im industriellen Bereich. Darüber hinaus arbeitet er an diversen Apps, welche eine Verbesserung und Hilfestellung im Bereich Gesundheit und Medizin voranbringen sollen (persönliche Kommunikation, 11. Juni 2019). Infolgedessen tritt er als Experte für KI im Allgemeinen sowie bei Gesundheits-Apps auf und kann entsprechend hilfreiche Argumente im Hinblick auf diese Thematik liefern. Für das Roboter-System *Da Vinci* konnte PD Dr. med. Alexander Beham als Experte gewonnen werden. Er ist Chefarzt der Allgemein- und Viszeralchirurgie am Kreisklinikum Siegen und gehört zu den etabliertesten „*Da Vinci*-Chirurgen“. Am Uniklinikum Göttingen hat er bereits über 200 Operationen mit dem Roboter durchgeführt (Westfalenpost, 2018, Absatz 2). Schließlich hat Beham das System auch am Kreisklinikum Siegen etabliert und operiert seither Patienten mit dessen Hilfe. Durch das Interview hat das Projektteam die Chance erhalten, den Roboter während einer Operation persönlich zu erleben, und konnte so Eindrücke sammeln sowie sich eine Meinung bilden, die für die

Projektarbeit effektiv genutzt werden konnte.

Die entwickelte Matrix zur Abbildung der Autonomiestufen in der Medizin setzt sich aus dem Grad an Autonomie der jeweiligen Stufe und den bezeichnenden Parametern zusammen (siehe *Abb. 1*), die folglich auch die Kriterien zur konkreten Einordnung einzelner Systeme bilden.

Grundlage für die Parameter bildet die Taxonomie von Lernzielen nach Benjamin Bloom (1956). Bloom hat sich sowohl mit den kognitiven als auch mit den affektiven und sensorischen Fähigkeiten des Menschen beschäftigt (Bloom, 1956). Für die Einordnung medizinischer Systeme soll an dieser Stelle allerdings nur die kognitive Dimension berücksichtigt werden, da diese das aktuell abschätzbare Spektrum an Möglichkeiten für die KI in der Medizin in naher Zukunft abdeckt (Holtel, Hufenstuhl & Klug, 2017, S. 13). Die Lernziele sind hierarchisch angeordnet, das bedeutet für diesen Fall mit steigender Stufe, nimmt auch die kognitive Kompetenz des Systems zu. Da Zweck der Autonomisierung dessen Annäherung an die Komplexität und Effizienz menschlicher Gehirnprozesse ist, liegt die Übertragung der Taxonomie auch auf nicht-menschliche Vorgänge nahe.

Die Bezeichnung der Stufen bzw. des Autonomiegrades leitet sich aus der "Taxonomie Autonomes Fahrens" der *SAE International* (Norm J3016, 2014) ab. So wie die vorliegende Skala besitzt diese 6 Stufen, um die Aufgaben abzubilden, die das System dem Menschen (im Straßenverkehr) zunehmend abnimmt, und verfolgt ein ähnliches Muster steigender Autonomie mit fortgeschrittener Stufe, wie im Folgenden deutlich wird.

In Stufe 0 ist noch keine Form der Autonomie vorhanden. Sie ist nicht in der Bloom-Taxonomie zu verordnen und besteht lediglich aus Systemen, die ein reines Werkzeug für die Ausführung von Befehlen eines Akteurs darstellen.

Die **erste Stufe** beherbergt bereits Assistenzsysteme, die in erster Linie ein zusätzliches Gedächtnis darstellen und in dieser Form helfen, Daten zu verwalten, aber häufig schon Möglichkeiten zur Kommunikation mit ihrem Anwender besitzen.

Besitzt das System die Fähigkeit zum Verständnis, wie es in **Stufe 2** der Fall ist, ist es teilweise autonom und kann die erfassten Daten soweit verarbeiten, dass Symptome gedeutet und Hinweise zu möglichen Diagnosen ausgegeben werden.

Eine tiefere Beratung findet dann in der **dritten Stufe**, der bedingten Autonomie, statt. In dieser hat das System nach Bloom gelernt, sein Wissen anzuwenden. Für die Medizin bedeutet das, dass Therapieangebote vorgeschlagen werden und in Kombination mit zugehörigen Risiken und Wahrscheinlichkeiten, den Patienten bei der Entscheidungsfindung unterstützen.

Tatsächliche Empfehlungen können allerdings erst ausgegeben werden, wenn eine hohe Autonomie vorliegt (Stufe 4) und das System auch fähig ist, seine Werte zu analysieren und zu evaluieren. Es erfolgt eine erweiterte Kontrolle, in der es individuelle Umstände berücksichtigt und somit Prognosen für den Krankheits- bzw. Heilungsverlauf des Einzelnen stellt, indem wie beim menschlichen Arzt durch Unterscheidungen von Feinheiten tatsächlich Pathologisches identifiziert wird und auch körperliche von möglichen psychischen Zusammenhängen getrennt werden können.

In einem letzten Schritt wird das System zur Entscheidung bevollmächtigt. D.h. in der **fünften Stufe** herrscht auch eine so hohe gesellschaftliche Akzeptanz und Integration in die medizinischen Abläufe, dass das voll autonome System Leistungen veranlasst und den Patienten damit sowohl symptomatisch als auch ursächlich in Behandlung nimmt.

Wie weit autonome Systeme in der Medizin heute bereits entwickelt sind und wo sie dementsprechend einsortiert werden können, wird anhand einiger Praxisbeispiele im folgenden Kapitel deutlich.

### **3. Beispielhafte Systeme und die Einordnung**

Nachdem die Skala ausreichend ausgeführt wurde, werden nun beispielhaft Anwendungen eingefügt (siehe *Abb.2*). Die Auswahl<sup>3</sup> der Systeme ist aufgrund des Umfangs dieser Arbeit beschränkt und umfasst vier Anwendungsgebiete von Formen der KI in der Medizin, die jedoch ein möglichst breites Spektrum der zurzeit in Deutschland<sup>4</sup> in der Praxis angewandten Systeme abbilden soll. Vorab anzumerken ist, dass es sich bei den ausgewählten Anwendungen und Systemen um die Methoden der KI anwenden, lediglich um eine schwache sogenannte beschränkte KI [i.O.: „Narrow AI“]<sup>5</sup> (Urban, 2015, Absatz 31) bzw. angewandte KI [i.O.: „applied AI“]<sup>6</sup> (Seeking Alpha, 2017, Absatz 4) handelt.

#### **3.1 Da Vinci Operationsroboter**

Der *Da Vinci* Operationsroboter ist ein System, das bereits seit dem Jahr 2000 (Maier, 2017, Absatz 4) im Operationssaal benutzt wird und ein breites Einsatzspektrum in der Chirurgie abdeckt. Es ist, laut Dr. Beham eines der am besten entwickeltesten Operationsrobotersysteme, das in der Medizin verwendet wird (persönliche Kommunikation am 2. Juli 2019). Die Einstufung in die entwickelte Skala erfolgt in der **Stufe 0: Keine Autonomie** (Ausführen von Befehlen), denn das Robotersystem *Da Vinci* „ist nach wie vor der Pfeil und nicht der Indianer“ (ebd.) und damit lediglich ein Assistenzwerkzeug für den Chirurgen um seine Arbeit minimalinvasiv wie bei einer Laparoskopie und dabei gleichzeitig mit der Sicht- und Bewegungsweise operieren zu können, wie bei einer offenen Operation. Zurzeit wird der Operationsroboter nur in direkter Anwesenheit eines Chirurgen genutzt, da das Fernbehandlungsverbot (§7 Abs. 4 MBO-Ä) dies aktuell verlangt, jedoch wäre es in der Zukunft möglich, den *Da Vinci* einige Operationsschritte automatisiert durchführen zu lassen. Die Verantwortung bei Komplikationen während der Operation mit dem *Da Vinci* ist weitestgehend geklärt. Sollte es zu einer Fehlbedienung<sup>7</sup> von *Da Vinci* kommen, ist der Operateur des *Da Vincis* und das Krankenhaus verantwortlich, da alleine er das System direkt steuert. Anders hält es sich bei von *Da Vinci* initiierten Komplikationen wie einem technischen Defekt. Dort ist dann die Ursache des Defekts zu klären<sup>8</sup>. Hohe Anschaffungskosten und teure Schulungen sowie die teilweise privat zu finanzierenden Operationen können zu unterschiedlichen Behandlungsstandards an Krankenhäusern und in sozialen Milieus führen.

### 3.2 Ada Health und M-Sense

Die App *Ada Health* lässt sich der Kategorie der Diagnose-Apps zuordnen. Dem Patienten wird anhand der App eine sichere und schnelle Diagnose ohne ärztliche Hilfe ermöglicht (Jörg, 2018, S.17). *Ada Health* ist ein registriertes Medizinprodukt sowie eine Gesundheitsplattform, die sich an einer Form der KI bedient. Die App finde demnach ihren Einsatz in der Entscheidungsfindung (ebd.). Zu den Funktionen zählt zum einen ein Chatbot, der nach einer Diagnose gefragt werden kann, ein Katalog aus Krankheiten und Symptomen sowie diverse Gesundheitsberichte (*Ada Health*). In der Skala findet *Ada Health* ihren Platz in **Stufe 3: Bedingte Autonomie**. Dies lässt sich aus den Funktionen der Dokumentation und der Diagnose ableiten. Ergänzend dazu gibt die App Behandlungsvorschläge und berechnet Wahrscheinlichkeiten für den Patienten. *M-sense* ermöglicht Patienten eine mobile personalisierte Migränetherapie. Mithilfe der App kann ein Tagebuch geführt werden, welches Medikamenteneinnahmen oder Schmerzattacken erfasst. Nach der Sammlung verschiedener Daten werden der individuelle Lebensstil, Umwelteinflüsse und biologisch-hormonelle Faktoren analysiert und daraufhin ein individueller Therapieplan erstellt. Die Analyse und grafische Abbildung sowie die Prognose geschieht mittels mathematischer

Algorithmen (*M-sense*, 2019; Kahl, 2017). *M-sense* lässt sich **Stufe 2: Teil-Autonomie** zuordnen. Die App bedient sich vor allem der Dokumentation in Form des Tagebuchs, sowie der Analyse von Daten für Prognosen und individuellen Behandlungsempfehlungen.

### 3.3 Pflegeroboter: *Pepper* und *Paro*

*Pepper* ist ein humanoider Roboter der sich autonom auf Rollen fortbewegt und mit Menschen durch Konversationen und seinen Touchscreen interagiert. *Pepper* verfügt über eine Gesichtserkennung und deutet Emotionen seines Gegenübers mithilfe einer Stimmenerkennung (Jörg, 2018, S.100; Softbank Robotics, 2019). Aufgrund der Unterstützung im Alltag und des Verstehens sowie des Interpretierens von Emotionen, findet *Pepper* seinen Platz in **Stufe 2: Teil-Autonomie**. Allerdings weist er nicht alle Eigenschaften dieser Stufe auf, da ihm beispielsweise die Diagnosefunktion fehlt. Die Robbe *Paro* ist ein Robbenbaby, welches vordergründig eine Hilfe für demenzkranke Menschen sein soll. Die Robbe ist mit Berührungs- und Temperatursensoren ausgestattet und reagiert auf die Stimmen ihres Gegenübers (Jörg, 2018, S.103; Guizzo, 2009, Absatz 2). Die Einordnung des Pflegeroboters *Paro* erfolgt in **Stufe 1: Assistenz**. Zur Hauptaufgabe zählt die Unterstützung hilfsbedürftiger Menschen auf emotionaler Ebene. Allerdings darf nicht vergessen werden, dass *Paro* teilweise einen bestimmten Wortschatz der Patienten versteht und darauf reagiert (ebd.), was wiederum Eigenschaften der Stufe 2 sind.

### 3.4 IBM Watson Health

Die Betrachtung des IBM *Watson Health* ergibt, dass es sich um eine "klinische Datenintelligenz" (Gehring, Rackebrandt & Imhoff, 2018, S.254) mit Deep Learning Ansatz (IBM Watson, 2018) handelt, die lediglich in einem bestimmten Einsatzgebiet Verwendung findet und mit fachbezogenen Material trainiert wird. Das System hilft Krankheitsverläufe zu dokumentiert, stellt eine Diagnose anhand der zur Verfügung stehenden Patientendaten, die von einem Mediziner eingegeben werden sowie dem zur Verfügung stehenden Trainingsmaterial, es gibt Behandlungsvorschläge und bietet eine Beratung über die Behandlung und Therapie der Krankheit an und nennt deren Risiken. Allerdings ist die letzte Entscheidungsinstanz über die Behandlung und die Korrektheit der Diagnose der Mediziner. Somit ist das System bei Stufe 3: Bedingte Autonomie einzuordnen und dient ebenfalls als Assistenz des Arztes. Der Zukunft von *Watson Health* in Deutschland stehen jedoch aktuell Hürden des Datenschutzes<sup>9</sup> und der Datenverantwortung<sup>10</sup> sowie der Datenaufbereitung<sup>11</sup> im Weg. Des Weiteren ist *Watson Health* im Jahr 2018 durch seine Misserfolge bekannt geworden u.a. am Universitätsklinikum Marburg im Zentrum für unerkannte Krankheiten (Balzter, 2018, Absatz 4).

#### 4. Kritische Reflexion

Die Untersuchung bereits implementierter Techniken sowie die gewonnenen Erkenntnisse aus der Praxis, die aus den Interviews ersichtlich wurden, geben einen Überblick über geeignete Bereiche, an denen autonome Systeme in der Medizin anknüpfen können und weiterer Entwicklung bedürfen.

Das derzeit größte Potenzial lässt sich dort erkennen, wo der menschliche Mediziner mit der Fülle an Informationen, die es auszuwerten gilt, überfordert ist. „Wo große Datenmengen anfallen – da ist die Domäne der KI“, sagt auch Dr.-Ing. Kai Hahn (persönliche Kommunikation, 11. Juni 2019). Denn letztlich sei Künstliche Intelligenz genau das - Datenanalyse. Der Vorteil zum herkömmlichen Arzt liegt in der Masse, die gleichzeitig verarbeitet und in Beziehung zueinander gesetzt werden kann. Mit der Hilfe computergestützter Methoden kann der sowohl für den Patienten als auch für seine Ärzte mühsame Weg bis zur Diagnosefindung effizient verkürzt und damit mehr Energie in die Therapie investiert werden. Insbesondere bei sehr spezifischen Krankheitsbildern kommt eine solch immense Wissensressource dem System zugute. Nach Hahn liegt das eigentliche Problem genau darin, dass jeder Arzt eine unterschiedliche Krankenakte zu ein und demselben Patienten angelegt hat (persönliche Kommunikation, 11. Juni 2019). „Vernetzung“ heißt folglich das Schlüsselwort.

Dr. Alexander Beham spricht ein weiteres wichtiges Anwendungsgebiet an: „Ich bin sehr für einen autonomen, aufgeklärten Patienten“ (persönliche Kommunikation, 2. Juli 2019). Die steigende Autonomie von medizinischen Produkten erleichtert den Mediziner um seine Aufgabensumme und bestärkt gleichzeitig die Selbstständigkeit sowie Autarkie des Patienten. Dr. Beham beschreibt das Unsicherheitsgefühl, dass dem Patienten nicht genommen werden kann, da er sich zu einem gewissen Grad immer in die Verantwortung des Arztes begibt (persönliche Kommunikation, 2. Juli 2019). Doch Gesundheits-Apps beispielsweise helfen, dem Patienten einen Teil der Kontrolle über seine Therapie zurückzugeben und so das gegenseitige Vertrauen, das einen erheblichen Einfluss auf den Heilungsverlauf haben kann, zu stärken.

Dieses Arzt-Patienten-Verhältnis ist besonders wichtig in Bereichen, in die der Patient nicht direkt eingreifen kann, wie zum Beispiel wenn eine Operation notwendig wird. Und eben jenes Gebiet bleibt zu großen Teilen auch selbstlernenden Systemen vorenthalten. Für Dr. Beham liegt der Fokus beim chirurgischen Fortschritt in den nächsten Jahren in keiner Weise bei der Forschung



an zunehmender Autonomie der Maschinen wie dem *Da Vinci*, sondern an seiner geförderten minimalinvasiven Funktionsweise (persönliche Kommunikation, 2. Juli 2019). Mit der Möglichkeit immer gesundheitsschonender zu operieren, entstehen neue Standards und Prioritäten in der Medizin. Inzwischen geht es nicht länger lediglich darum, dass der Mensch überlebt, sondern auch darum, möglichst wenig von dem gesunden Gewebe zu schaden und die Lebensqualität der Person (mindestens) beizubehalten. Maschinelle Autonomie kann im OP eine sinnvolle Ergänzung darstellen, wenn dadurch zusätzliche Kompetenz herangeholt und die vorhandenen Behandlungsoptionen erweitert werden. Dr. Beham kann sich selbstlernende Roboterarme beispielsweise bei weniger riskanten Vorgängen wie beim Vernähen von Wunden vorstellen (persönliche Kommunikation, 2. Juli 2019). Doch ausschlaggebend in jedem Fall wird das Endergebnis sein. So viele Informationen wie einem autonomen System auch zur Verfügung stehen, sind diese nicht weiter behilflich, wenn das Ergebnis der Behandlung nicht korrekt ausfällt. An erster Stelle, um zu entscheiden, ob eine solche Einführung erwünscht ist, steht demnach die Fehlerquote des Systems.

Denn von dieser hängt auch die öffentliche Meinung ab. Patienten befinden sich prinzipiell in eher angstbesetzten Situationen. Sie wissen nicht, was in Bezug auf ihre Krankheit noch auf sie zukommt und müssen für ihre Heilung vollkommen auf fachliche Fähigkeiten vertrauen, in denen sie in der Regel selbst nicht gelernt sind. Wenn dann noch zusätzlich die Kontrolle an einen Operationsroboter abgegeben wird, kann die Situation vielen zu unsicher erscheinen. Doch rationale Argumente können die allgemeine Stimmung zum Thema kippen lassen, sobald der Einsatz von autonomen Systemen eine Zeit lang etabliert ist und sich bewährt hat. Das kann auch die emotionale Hemmschwelle der Menschen senken.

Ein Bereich, in dem es an Emotionalität potenziell fehlt, ist die Pflege. Sobald es um den direkten Kontakt zum Patienten geht, wirkt die Verwendung von digitalen Helfern auf Dr. Beham eher traurig (persönliche Kommunikation, 2. Juli 2019). Um die Interaktion mit Menschen vollständig zu ersetzen, fehlt es Unterstützungen wie *Pepper* einfach noch an einer hinlänglichen emotionalen Intelligenz, sodass beispielsweise auch der Großteil menschlicher Kommunikation, der nicht verbal stattfindet, verstanden wird. Als Konsequenz daraus droht den pflegebedürftigen Personen Vereinsamung.

Hinzu kommt die Frage nach der Verantwortlichkeit bei systemgenerierten Fehlern, wenn die Ursache nicht mehr eindeutig auf einen Akteur im Hintergrund zurückzuführen ist. Dr. Hahn nennt

diese Problematik betreffend das Schlagwort der „White Box“ (persönliche Kommunikation, 11. Juni 2019). Das ethische Problem des Deep Learning liegt in der Unzugänglichkeit der angewendeten Methoden. Wenn es um Menschenleben geht, möchte die Gesellschaft wissen, auf welchen Faktoren die Schlüsse der behandelnden Instanz beruhen. Die Black Box muss also zur White Box werden. Doch auch Dr. Hahn weiß um die (allein schon technische) Schwierigkeit dieses Anspruchs.

Bei fehlender Transparenz entstehen Akzeptanzprobleme in der Gesellschaft, doch die Trainingsdaten der Systeme frei einsehbar zu machen, generiert wiederum auf rechtlicher Ebene Herausforderungen, was den Datenschutz betrifft (Dr. Hahn, persönliche Kommunikation, 11. Juni 2019). Hinzu kommen die Anschaffungskosten, die Krankenhäuser und Ärzte tragen müssen, wenn sie den neuen Standards, die mit selbstlernender Technologie einhergehen, gerecht werden wollen. Infolgedessen besteht eine hohe Wahrscheinlichkeit, dass große Unterschiede zwischen den jeweiligen Leistungsständen und Qualitätsniveaus entstehen würden.

Einen Strich zu ziehen, wie autonom medizinische Produkte im Endeffekt handeln sollten, ist schwierig, ebenso wie die Einschätzung, inwiefern sie dies bereits tun. Insbesondere die Verortung der Pflegeroboter in der entwickelten Autonomieskala (siehe Kapitel 3. Beispielhafte Systeme und die Einordnung) demonstriert die Uneindeutigkeit des Verfahrens. Aufgrund der unterschiedlichen Anwendungsfelder und Bereiche, in denen die Systeme zum Einsatz kommen, schwankt die Vergleichbarkeit an verschiedenen Stellen, was zu unklaren Grenzen führt. In dem Moment, in dem zwei Systeme unterschiedlicher Kategorie ein Kriterium teilen, ist die profunde Definition desselben jeder Grundlage entbehrend.

Die Stufen der Autonomie gehen folglich in Teilen ineinander über. Und auch die Ermittlung gemeinsamer Parameter innerhalb einer Stufe fällt nicht immer konkret aus. Zusätzlich folgt die Matrix wie zu Anfang erwähnt (siehe Kapitel 2. Methodisches Vorgehen) lediglich dem kognitiven Aspekt des Lernens. Bei einer voranschreitenden Entwicklung der Systeme des Deep Learnings wird darüber hinaus die Berücksichtigung der emotionalen und sensorischen Komponente notwendig werden.

## **5. Fazit**

Die gewonnenen Erkenntnisse aus der Skala und der Fachliteratur führen zu dem Ergebnis, dass der Schwerpunkt der KI-gestützten Systeme zum jetzigen Zeitpunkt in der Medizin vornehmlich

auf der Unterstützung des Mediziners liegt, statt diesen in Zukunft zu ersetzen. „Der Mediziner wird eher wie ein Dirigent agieren, der das Orchester zur Höchstleistung treibt[...]“ (Bogdan, 2018, S.300). Denn „Urteilen und Entscheiden nach ethischen Kriterien ist integraler Bestandteil des medizinischen und pflegerischen Alltags. Lernende Systeme können diese Aufgabe nicht übernehmen, weshalb KI auch in Zukunft lediglich als Unterstützung für den Menschen eingesetzt werden darf“ (Plattform Lernende Systeme, 2019, S.17). Die Angst, dass in Zukunft die Maschine die Menschen gar ersetzen könnte, ist eine Dystopie. Die größere Skepsis sollte darauf liegen, dass die Menschen sich zunehmend mehr auf die Algorithmen der KI-Systeme verlassen. (Jörg, 2018, S.122; Dr Beham, persönliche Kommunikation, 2. Juli 2019). Dabei sollte man sich stets vor Augen führen, dass Algorithmen nichts anderes als Werkzeuge sind, um von Menschen gesetzte Ziele zu erreichen (Jörg, 2018, S.122). Das Ziel in der Medizin ist, dem Patienten die bestmögliche Behandlung zukommen zu lassen um das bestmögliche Ergebnis – optimaler Weise die Heilung – zu erlangen.

Ein Ergebnis, das die Skala hervorgebracht hat ist, dass mit steigender Autonomie eines Systems mit Formen von KI auch die ethischen Herausforderungen und Anforderungen steigen. Gerade in einem so heiklen Einsatzgebiet wie der Medizin sind der Datenschutz und die Schweigepflicht des Arztes besonders wichtig. Des Weiteren ist die Verantwortung bei Fehldiagnose und/oder -Behandlung ein Feld das Klärungsbedarf hat. Eine Transparenz (insbesondere bei KI die sich an Deep Learning-Algorithmen bedient) über den Prozess einer Entscheidung muss erreicht werden, um eine rechtliche sowie eine Vertrauensgrundlage des Patienten in das System zu ermöglichen.

Eine weitere Herausforderung ist, die Interessen des Patienten vor die ökonomischen Interessen zu stellen. Abgesehen von der Anhäufung der Gesundheitsdaten, sind auch die Effizienz bei der Verarbeitung von Daten und die möglichen Personalkosteneinsparungen durch KI-Anwendungen ein großer wirtschaftlicher Faktor im Gesundheitswesen. Dabei muss bei der Verwendung von KI-Systemen eine Balance zwischen Datenschutz und -verwendung sowie der Effizienz und der Menschlichkeit gefunden werden. Empfehlungen zum Umgang mit KI-Anwendungen gibt beispielsweise der Digitalverband Bitkom<sup>12</sup> sowie die Arbeitsgruppe „Gesundheit, Medizintechnik und Pflege“<sup>13</sup>.

Als abschließendes Fazit dieses Projektes stellt sich heraus, dass KI-Anwendungen die persönliche Beziehung zwischen Arzt und Patient nicht ersetzen können (siehe dazu: Jörg, 2018, S.130), da sie bisher nicht die Möglichkeit haben, die emotionale Intelligenz sowie die Erfahrung eines

Menschen nachzubilden. Die bestmögliche Behandlung des Patienten durch autonome Systeme kann bisher nicht gewährleistet werden, denn die aufgedeckten Problemfelder müssten in der Zukunft erst umfassend gelöst werden, um die Zulassung für ein vollständig autonomes System in der Medizin überhaupt zu erreichen.

---

1 In dieser Arbeit wird aus Gründen der besseren Lesbarkeit das generische Maskulinum verwendet. Weibliche und anderweitige Geschlechteridentitäten werden dabei ausdrücklich mitgemeint, soweit es für die Aussage erforderlich ist.

2. Die Auswahl der Systeme mit Anwendungen von KI ist nicht repräsentativ, da sie aufgrund des Umfangs dieser Arbeit beschränkt ist und lediglich das deutsche Einsatzgebiet abdeckt. Eine weitere Ausführung ist daher in Zukunft wünschenswert, um eine Repräsentativität zu erreichen.

3. Die Auswahl beschränkt sich hier lediglich auf in Deutschland zugelassene und genutzte Systeme, da in anderen Ländern die Gesetzeslage und das Gesundheitswesen different sind.

4. Urban unterscheidet zwischen 1. ANI: Artificial Narrow Intelligence: Schwache KI die spezialisiert ist auf einen Bereich. 2. AGI: Artificial General Intelligence: Starke KI vergleichbar mit der menschlichen Intelligenz die intellektuelle Aufgabe erledigen kann. Der Maßstab dafür sind die Fähigkeit einen Grund zu finden, planen zu können, Probleme können gelöst werden, abstraktes Denken ist möglich, Vergleiche von komplexen Ideen finden statt und es lernt schnell und auch aus Erfahrung. 3. ASI: Artificial Superintelligence: Eine Intelligenz, die schlauer als die des Menschen ist in jeglicher Disziplin wie Kreativität, Allgemeinwissen, Sozialkompetenz. (Urban, 2015, Absatz 31ff.)

5. Seeking Alpha (2017, Absatz 4 und 5) unterscheidet zwischen: 1. der angewandten Künstlichen Intelligenz die sehr spezielle Aufgaben erledigen. 2. der allgemeinen Künstlichen Intelligenz die für sämtliche Aufgabenbereiche entwickelt ist und kein menschliches Eingreifen mehr nötig ist. Eine Anforderung besteht darin, dass die Maschine lernfähig ist, um neue Aufgaben und Situationen zu bestehen.

6. Fehlbedienung meint hier eine unkontrollierte Bewegung des Operators die Konsequenzen für den Patienten haben kann.

7. Ursachen für einen technischen Defekt können u.A. sein: der Hersteller hat ein fehlerhaftes Produkt produziert; das Krankenhaus hat den Roboter nicht ordnungsgemäß warten und prüfen lassen und die Funktionsfähigkeit vorab getestet; der Chirurg hat den Roboter fehlerhaft angewendet und so einen Schaden am Gerät verursacht.

8. Datenschutzprobleme treten beispielsweise bei der internationalen Cloud-Datenspeicherung, der Schweigepflicht des Arztes sowie der DSGVO-Konformität und der Verwendung der digitalen Krankenkarte auf.

9. Datenverantwortung bezieht sich hier darauf, welche Daten generell Relevant sind um einen Patienten zu behandeln und außerdem wie mit den bekannten Daten verantwortungsvoll umgegangen wird und diese verantwortungsvoll und gebrauchstauglich zur Verfügung stellen kann. Dazu Interessant: der Digital-Gipfel 2018 des Bundesministeriums für Wirtschaft und Energie. Abgerufen am 12.09.2019 von: <https://www.de.digital/DIGITAL/Redaktion/DE/Digital-Gipfel/Video/2018/20181204-digital-gipfel-4-ki-produktion.html>.
10. Datenaufbereitung meint hier beispielsweise kostenintensive und mustergerechte Digitalisierungsarbeit der Patientenakten, Befunde, Medikamenteneinnahmen, Ultraschall- und Röntgenaufnahmen sowie der Fachliteratur und des medizinischen Wissens.
11. Bitkom (2018): Empfehlungen für den verantwortlichen Einsatz von KI und automatisierten Entscheidungen. Abrufbar von <https://www.bitkom.org/Bitkom/Publikationen/Empfehlungen-fuer-den-verantwortlichen-Einsatz-von-KI-und-automatisierten-Entscheidungen-Corporate-Digital-Responsibility-and-Decision-Making.html>
12. Arbeitsgruppe „Gesundheit, Medizintechnik, Pflege“ (2019): Lernende Systeme im Gesundheitswesen. Abrufbar von <https://www.plattform-lernende-systeme.de/publikationen-details/lernende-systeme-im-gesundheitswesen.html>. Die AG schlägt beispielsweise einen „KI-TÜV“ (S.32) vor, der in regelmäßigen Abständen mittels neuer Testdaten die KI-Technologien überprüft und rezertifiziert.

---

## Literaturverzeichnis

Bloom, Benjamin S. (1956). Taxonomy of educational objectives: The classification of educational goals. Book 1: Cognitive Domain. London: Longmans.

Bogdan, Boris (2018). MedRevolution: Neue Technologien am Puls der Patienten. Oberwil: Springer.

Holtel, Stefan; Hufenstuhl, Andreas; Klug, Andreas; Bitkom e.V. (Hrsg.) (2017). Künstliche Intelligenz verstehen als Automation des Entscheidens. Leitfaden. Berlin: Bitkom e.V.

Jörg, Johannes (2018). Digitalisierung in der Medizin. Wie Gesundheits-Apps, Telemedizin, künstliche Intelligenz und Robotik das Gesundheitswesen revolutionieren. Wuppertal: Springer.

§7 Abs. 4 MBO-Ä (2018). Behandlung im persönlichen Kontakt und Fernbehandlung. Abgerufen am 12.09.2018 von <https://www.bundesaerztekammer.de/recht/aktuelle-rechtliche-themen/fernbehandlung/>.

*Ada Health* (2019). Über uns. Abgerufen am 25.08.2019 von <https://ada.com/de/>.

Balzter, Sebastian (2018). Supercomputer *Watson*: Im Krankenhaus fällt die Wunderwaffe durch. Hg. v. Frankfurter Allgemeine Zeitung GmbH. Abgerufen am 12.09.2019 von [https://www.faz.net/aktuell/wirtschaft/kuenstliche-intelligenz/computer-watson-scheitert-zu-oft-bei-datenanalyse-15619989.html?printPagedArticle=true#pageIndex\\_4](https://www.faz.net/aktuell/wirtschaft/kuenstliche-intelligenz/computer-watson-scheitert-zu-oft-bei-datenanalyse-15619989.html?printPagedArticle=true#pageIndex_4).

Bitkom - Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e. V. (2018). Empfehlungen für den verantwortlichen Einsatz von KI und automatisierten Entscheidungen. Abgerufen am 12.09.2019 von <https://www.bitkom.org/Bitkom/Publikationen/Empfehlungen-fuer-den-verantwortlichen-Einsatz-von-KI-und-automatisierten-Entscheidungen-Corporate-Digital-Responsibility-and-Decision-Making.html>.

Bundesministerium für Bildung und Forschung (2013). Eine Therapie-Robbe für demenzkranke Menschen?. Abgerufen am 25.08.2019 von <https://www.wissenschaftsjahr.de/2013/die-themen/themen-dossiers/besser-leben-mit-technik/eine-therapie-robbe-fuer-demenzkranken-menschen.html>.

Bundesministerium für Bildung und Forschung (2019). Wird KI die Medizin revolutionieren?. Abgerufen am 25.08.2019 von <https://www.wissenschaftsjahr.de/2019/das-wissenschaftsjahr/themenfelder/gesundheit-und-medizin/>.

Bundesministerium für Wirtschaft und Energie (2018). Digital-Gipfel 2018: Künstliche Intelligenz in Produktion, Mobilität und Gesundheit. [Videodatei]. Abgerufen am 10.09.2019 von <https://www.de.digital/DIGITAL/Redaktion/DE/Digital-Gipfel/Video/2018/20181204-digital-gipfel-4-ki-produktion.html>.

Gehring, H., Rackebrandt, K. & Imhoff, M. (2018). E-Health und die Realität – was sehen wir heute schon in der Klinik? In: Bundesgesundheitsbl (Hg.) 61(3) S. 252-262. doi:10.1007/s00103-018-2690-6.

Guizzo, Erico (2009). *Paro* the Robotic Seal Could Diminish Dementia. First long-term study seeks to prove the benefits of a cybernetic pet. Hg.v. IEEE Spectrum. Abgerufen am 25.08.2019 von <https://spectrum.ieee.org/robotics/home-robots/paro-the-robotic-seal-could-diminish-dementia>.

Höpner, Axel (2019). So verändert Künstliche Intelligenz die Diagnostik. Hg.v. Handelsblatt. Abgerufen am 25.08.2019 von <https://www.handelsblatt.com/technik/thespark/medizin-der-zukunft-so-veraendert-kuenstliche-intelligenz-die-diagnostik/24535864.html>.

IBM Watson (2018). How does IBM Watson work? [Videodatei] Abgerufen am 12.09.2019 von <https://youtu.be/r7E1TJ1HtM0>.

Kahl, Kristin (2017). Kopfschmerz: Mobile Migränetherapie per App „M-sense“. Hg.v. Aerzteblatt. Abgerufen am 25.08.2019 von <https://www.aerzteblatt.de/archiv/192343/Kopfschmerz-Mobile-Migraenetherapie-per-App-M-sense>.

M-Sense (2019). M-Sense. Übernimm die Kontrolle bei Migräne & Kopfschmerzen! Abgerufen am 14.09.2019 von <https://m-sense.de>.

Maier, Josephina (2017). Dr. DaVinci, bitte in den OP. Längst operieren Roboter Menschen. Dabei fühlen sie nichts. Hg. v. Die Zeit Online (1/2017). Abgerufen am 12.09.2019 von <https://www.zeit.de/2017/01/chirurgieroboter-davinci-operation-arzt/komplettansicht>.

Meißner, Thomas (2019). Zitterfreies Operieren mit dem Roboter. In: ÄrzteZeitung (Hg). Abgerufen am 25.08.2019 von <https://www.aerztezeitung.de/medizin/krankheiten/krebs/article/980491/op-assistent-vinci-zitterfreies-operieren-roboter.html>.

Plattform Lernende Systeme (2019). Prävention, Diagnose, Therapie: Lernende Systeme im Gesundheitswesen. Bericht der AG6 – Arbeitsgruppe „Gesundheit, Medizintechnik, Pflege“. Abgerufen am 12.09.2019 von <https://www.plattform-lernende-systeme.de/publikationen-details/lernende-systeme-im-gesundheitswesen.html>.

SAE International (2014). Automated Driving Levels of Driving Automation are defined in new SAE International Standard J3016. Abgerufen am 02.09.2019 von [https://www.smmmt.co.uk/wp-content/uploads/sites/2/automated\\_driving.pdf](https://www.smmmt.co.uk/wp-content/uploads/sites/2/automated_driving.pdf).

Seeking Alpha (2017). Artificial Intelligence, Explained. Abgerufen am 12.09.2019 von <https://seekingalpha.com/article/4088363-artificial-intelligence-explained>.

Softbanks Robotics (2019). *Pepper* plays ball-in-the-cup. We made *Pepper* the robot play games of skill with AI (Artificial Intelligence) at SBRE. Abgerufen am 25.09.2019 von <https://developer.softbankrobotics.com/blog/pepper-plays-ball-cup>.

Urban, Tim (2015). The AI Revolution: The Road to Superintelligence. Hg. v.: Wait but Why. Abgerufen am 12.09.2019 von <https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html>.

Van Lieshout, Marc; Chen, Ming; Karanikoloava, Kristina; Timan, Tjerk; Bolchi, Marco; Costenco, Patrisia et al. (2019). Study on Safety of non-embedded software; Service, data access, and legal issues of advanced robots, autonomous, connected, and AI-based vehicles and systems. European Commission. Abgerufen am 25.08.2019 von <https://ec.europa.eu/digital-single-market/en/news/safety-non-embedded-software-including-safety-health-lifestyle-and-wellbeing-apps>.

Westfalenpost (Hg.) (2018). Kreisklinikum in Siegen investiert in zweiten OP-Roboter. Abgerufen am 25.08.2019 von <https://www.wp.de/staedte/siegerland/kreisklinikum-in-siegen-investiert-in-zweiten-op-roboter-id215329833.html>.



#### 4. Microwork

Vortragstag und Gruppenname:	10. Juli 16:45 Uhr, Microwork
Titel der Ausarbeitung:	Die Organisation des künstlichen Geistes – Mikroarbeit zwischen Interaktion und Infrastrukturierung menschengestützter künstlicher Intelligenz
Verfasser/-innen:	Ekaterini Kirkitziotou Sebastian Randerath
Betreuer:	Prof. Dr. Volkmar Pipek (Sozio-Informatik) M.Sc. André Sekulla

#### Übersicht über die Struktur der Ausarbeitung

1. Künstliche künstliche Intelligenz?
    - 1.1 Kollektive Intelligenz
    - 1.2 Segmentierung, Datafizierung, Automatisierung
  2. Mikroarbeit und Megainfrastrukturen
    - 2.1 Der Mensch in der Maschine
    - 2.2 Von Capture zu CAPTCHA
  3. Fazit
- Literaturverzeichnis
- 

#### 1. Künstliche künstliche Intelligenz?

Systeme künstlicher Intelligenz werden spätestens seit der Dartmouth-Konferenz von 1956 als organisationale Technologien begriffen (Simon, 1996, Vorwort iii). Dabei wird versucht, sie als handlungsunterstützende Technologien in Bezug auf menschliche Arbeitsleistungen einzusetzen (Suchman, 1987). Die Organisation menschlicher Arbeit bildet jedoch vor allem den Ausgangspunkt für die Synthese künstlicher Intelligenz selbst. Diese Form der Arbeit wird als »Mikroarbeit« bezeichnet, da sie auf der Segmentierung menschlicher Arbeitsleistungen in elementare (Mikro-

)Aufgaben basiert. Durch diese Segmentierung ermöglicht sie die Datafizierung der einzelnen Arbeitsschritte. Über das Zerlegen in Elementaraufgaben, die sogenannten »Microtasks«, können so simple Arbeitsschritte delegiert, aber auch durch technische Infrastrukturen kontrolliert werden. Zum einen simulieren menschliche Arbeitsprozesse dadurch eine künstliche Intelligenz, indem sie datafiziert und in digitale Infrastrukturen eingebettet werden. Gleichzeitig ermöglicht die Datafizierung und Segmentierung in »Microtasks« jedoch auch Lernprozesse maschineller Systeme auf der Basis menschlicher Arbeitsleistungen (Irani, 2015, S. 115). Gleichzeitig geraten durch diese Ansätze der Organisation von Arbeit aber auch zwei grundlegend verschiedene Ansätze zum Verständnis von Arbeit und (digitaler)Technologie aneinander. Daran zeigen sich zwei verschiedene Modi des Verhältnisses von Arbeit, Kapitalakkumulation, Organisation und Technologie. Während einige Diagnosen, die zum Beispiel von einem »Kognitiven Kapitalismus« oder »Immaterieller Arbeit« sprechen (Terranova, 2000, S. 39 ). und Arbeit als rein immateriellen kognitiven Prozess gesellschaftlicher ansehen, findet bei Ansätzen, wie der »Mikroarbeit« eine Datafizierung menschlicher Arbeitsleistungen statt.

Bei den Vertretern der Theorie der »Kognitiven Arbeit« wird davon ausgegangen, dass der Einfluss der organisationalen Steuerung auf den Prozess menschlicher Arbeit in Bezug auf die Generierung von Kapital in dieser Form der Arbeit im Vergleich zur industriellen Produktion gering ist (Beverungen, 2018). Sind die Prozesse der Mikroarbeit in diesem Sinne »immateriell«? Der hohe Formalisierungsgrad, die Datafizierung und die automatisierte Kontrolle von Mikroarbeit durch große Infrastrukturen und Plattformen zeigen, dass die Generierung von Lerndaten durch Mikroarbeit komplexen organisationalen Prozessen und Technologien der Organisation unterliegt (Star, & Ruhleder, 1996). Außerdem wird das Ergebnis dieser Mikroarbeit in Form von Systemen maschinellen Lernens zum Ausgangspunkt neuer Prozesse und Verhältnisse von Arbeit und datenbasierter Steuerung (Beverungen, 2018). In der folgenden Arbeit wird daher aufgezeigt, wie sich durch die Datafizierung von Arbeit seit dem Aufkommen der »Wissenschaftlichen Betriebsführung« mediale Umwelten gebildet haben, die über Daten die automatisierte Steuerung segmentierter Arbeit ermöglichen. So wird die »Mikroarbeit« als menschliche Arbeit selbst als Teil der künstlichen Intelligenz verstanden. Dazu wird der Begriff der »Mikroarbeit« im ersten Teil der Arbeit über die Betrachtung der Datenakkumulation und Segmentierung der Arbeit zur Synthese von Systemen maschinellen Lernens von bestehenden Konzepten »Digitaler Arbeit«, die sich primär auf »Social Media Plattformen« beziehen, abgegrenzt (Fuchs, & Sevignani, 2013; Scholz, 2012 ). So wird der digitale Kapitalbegriff anhand der datafizierten Formen der »Mikroarbeit« entgegen der dialektischen Trennung von Subjekt und Objekt (Fuchs, & Sevignani, 2013 ). mit dem Fokus

auf die Verknüpfung von Subjekt-Objekt, Organisation und Infrastruktur analysiert. Dazu werden im zweiten Teil der folgenden Arbeit zunächst verschiedene Entwicklungen der Datafizierung und Segmentierung von Arbeit aufgezeigt und anschließend auf verschiedene Formen von »Mikroarbeit« anhand an den Beispielen von »Amazon Mechanical Turk« und »reCAPTCHA« zur Verschiebungen der Subjekt-Objekt-Relation im Bezug als Prozess der Infrastrukturerung künstlicher Intelligenz aufgezeigt. So wird anhand des Organisationsprozesses der »Künstlichen künstlichen Intelligenz« (Bezos, 2006). die These des »Kognitiven Kapitalismus« um die Bedeutung des relationalen Verhältnisses von Organisation, Infrastrukturen und lernenden Systemen im Bezug auf den Arbeitsprozess der Mikroarbeit ergänzt.

### **1.1 Kollektive Intelligenz**

Die zur »immateriellen Arbeit« gehen von einer Weiterentwicklung der marxistischen politischen Ökonomie als Weiterentwicklung der hegelschen Dialektik, (Fuchs, & Seignani, 2013, S. 252). also der Dreiteilung zwischen dem arbeitenden Subjekt, dem bearbeiteten Objekt und dem Ergebnis des Arbeitsprozesses als Subjekt-Objekt aus (Marx 2013, S. 63). Digitale Arbeit wird auf der Basis dieser Dialektik als Aktivität, die auf die Generierung eines Gebrauchs- und eines Tauschwertes abzielt verstanden (Fuchs, & Seignani, 2013, S. 2). Die Generierung von Gebrauchs- und Tauschwerten basiert daher im dialektischen Verständnis auf der Transformation von menschlicher Arbeitskraft und Objekten der Arbeit zu einem Produkt oder Werk der Arbeit. Hier setzten die Überlegungen zum Verhältnis von politischer Ökonomie und dem was als »digitale Arbeit« verstanden wird an (Fuchs, & Seignani, 2013, S. 258). Sie sehen in vernetzten Technologien nicht primär neue Wege der Vernetzung von Kommunikation, sondern den Wandel der Produktion durch neue Wege der Überwachung und Kontrolle (Heilmann, 2015). In dieser Form des Kapitalismus vollzieht sich daher, laut der post-operationistischen Lesart, ein Wandel von einem industriellen Kapitalismus hin zu einem immateriellen-postindustriellen Kapitalismus (Teranova, 2000, S. 39). Arbeit wird zum »kognitiven Gut« und durch die Vernetzung der digitalen Wissensproduktion zur »Massenintellektualität« (Heilmann, 2015, S. 43).

Wie bereits durch die Begriffe der »Massenintellektualität« oder des »kognitiven Kapitalismus« deutlich wird, gehen jedoch auch diese postmarxistischen Ansätze bei ihrer Lesart digitaler Arbeit weiterhin von einem dialektischen Verhältnis zwischen menschlichem Subjekt, Arbeitsgegenstand und Produkt aus. Arbeit wird so nicht nur zum Produkt menschlicher Interaktion und Kooperation sondern unmittelbar zum Ergebnis menschlicher Kognition. Die These »immaterieller Arbeit« rückt den Fokus daher verstärkt auf das Subjekt der Arbeit und dessen »Kognition«. So wird die

(soziale) Interaktion und Kooperation als menschliche Notwendigkeit verstanden (Fuchs, & Sevignani, 2013, S. 258). Formen der »Mikroarbeit« zeigen aber noch ein weiteres Verhältnis von Arbeit auf: Die Automatisierung der Arbeitskraft selbst. Wie im Folgenden genauer gezeigt wird, ist das Produkt der Arbeit, nimmt man die dialektische Argumentation auf, gleichzeitig zum technisch arbeitenden Subjekt geworden (Crawford, & Joler 2018). Das synthetisierte Subjekt-Objekt wird selbst zum Ausgangspunkt des Arbeitsprozesses der Mikroarbeit. Die grundsätzliche Subjekt-Objekt Unterscheidung im Rahmen dieser Technologien zu hinterfragen. Dazu folgen die arbeitenden Subjekte zugleich einem, durch das Objekt ihrer Arbeit, strukturierten Prozess der Arbeit. Die Arbeitsprozesse der Mikroarbeit sind größtenteils alles andere als immateriell oder kognitiv in dem Sinne, dass sie außerhalb eines bewusst getätigten Rahmens in sozialen Interaktionen vollzogen werden (Heilmann, 2015, S. 43). Sie unterliegen einer hochgradigen Formalisierung in Relation zur Synthese der Automatisierung. Diese Form der »Digitalen Arbeit« ist daher nicht »postindustriell«, weil sie Formen der industriellen Organisation von Arbeit überwindet und die reine menschliche Kognition abbildet, sondern weil sie, wie wir im Folgenden sehen werden, ebendiese industrielle Tradition der Organisation von Arbeit erweitert.

## **1.2 Segmentierung, Datafizierung, Automatisierung**

In vorindustriellen Gesellschaften basierte die Führung von Betrieben auf dem Anwenden bewährter Faustregeln zu Produktion und Organisation (Kieser, 2006). Mit der zunehmenden Interaktion von menschlichen Akteuren mit Maschinen im Produktionsprozess, wurde in industriellen Produktionsmodellen versucht, die menschliche Arbeitsleitung anhand der Taktung der Maschinen zu optimieren. Frederick W. Taylor (1912) entwickelte dazu den Ansatz der »wissenschaftlichen Betriebsführung«. Dazu führte er Formen der Steuerung durch die Messung und Datafizierung durch individueller Arbeitsprozesse ein. In diesem Rahmen trennte er das Unternehmen in die messende Instanz (Kopfarbeit) und die messbare Instanz (Handarbeit). Es entstand so die »Kopfarbeit« als zentraler Ort der Datenakkumulation und Steuerung auf der Basis von Wissensasymmetrien. Ausgehend von der Messung der Handarbeiter sollten so Werte zu den optimalen Arbeitsleistungen erhoben werden, die als Soll-Werte für die Unternehmensleistung dienen sollten. Die Grundlage für diesen Prozess der »wissenschaftlichen Betriebsführung« bildete die Messung und Datafizierung von Arbeit. Wie aber sollten Vergleichswerte in einem unstrukturierten Arbeitsprozess gemessen werden? Um Vergleichswerte zwischen Daten zu ermöglichen, aber vor allem auch um den Arbeitsprozess an sich effizienter gestalten zu können, segmentierte Taylor diesen Arbeitsprozess in festgelegte Elementarbewegungen. So konnte er vergleichbare Zeit- und Bewegungsstudien zur Optimierung des Arbeitsprozesses mittels der segmentierten

Elementarbewegungen durchführen. Diese Segmentierung der Arbeitsprozesse ermöglichte aber gleichzeitig auch das Herauslösen der Instanz der Unternehmenssteuerung aus der zentralen Struktur des produzierenden Betriebs. Durch die Modularisierung von Arbeit kamen gleichzeitig neue Medien der Visualisierung von Unternehmensprozessen und Arbeitsleistungen auf. Medien der Unternehmenssteuerung wie Messinstrumente aber auch Diagramme und Visualisierungen von segmentierten Bewegungsabläufen koordinieren so zwischen unterschiedlichen Handlungen menschlicher und technischer Akteure. Sie formen so das organisationale Umfeld indem verschiedene Aktivitäten der Unternehmensführung und Datafizierung miteinander kompatibel werden. So wird die Arbeit der Subjekte immer auch durch die Segmentierung und Datafizierung durch die Objekte ihrer Arbeit geformt. Dadurch kommt neben die Optimierung von Arbeitsprozessen im Unternehmen aber vor allem eine weitere Funktion des Bestrebens der Segmentierung und Datafizierung von Arbeit auf – die Automatisierung der Arbeitsaktivitäten selbst. Die Datafizierung und die Umstrukturierung des Arbeitsprozesses anhand der Maschinen sollen nicht nur den Arbeitsprozess selbst optimieren. Sie sollen dazu dienen ihn selbst maschinell ausführbar werden zu lassen (Hoof, 2015, S. 16). Die industrielle Produktionsweise infolge der »wissenschaftlichen Betriebsführung« war daher vor allem durch die Segmentierung, Datafizierung aber auch durch das Bestreben nach einer vollständigen Automatisierung von Arbeit bedingt. Dieser Wandel der Unternehmenssteuerung ging von neuen Technologien und Praktiken der Messbarkeit, Visualisierung und Automatisierung von Arbeit aus.

Tatsächlich entstand aber gleichzeitig auch eine gegenseitige Beeinflussung zwischen der Organisation menschlicher Arbeit und der Synthese arbeitender Maschinen. Diese Beeinflussung führte zu einer Idee, dass es Gemeinsamkeiten zwischen der Funktion von menschlicher Kognition in Organisationsprozessen und dem Prozessieren von Computern und künstlichen Systemen gäbe (Simon, 1996, Vorwort iii). Künstliche Systeme wurden so zum Ausgangspunkt eines symbolischen Organisationsprozesses. Die Relation von menschlicher Kognition, menschlichem Verhalten, sollte zum Ausgangspunkt eines präziseren Verständnisses menschlichen Verhaltens in Organisationen werden. Zum anderen wurde menschliche Arbeit durch die Abhängigkeit formalisierter Trainingsdaten und lernender Systeme aber auch zur Grundlage der Synthese der künstlichen Systeme selbst. Diese Form der, durch infrastrukturelle Medien künstlicher Intelligenz, segmentierte und koordinierte digitale Arbeit wird im Folgenden als »Mikroarbeit« bezeichnet. Wie unterscheidet sich diese Arbeit von der »immateriellen Arbeit« in sozialen Medien? Um diese Form der Arbeit genauer abgrenzen zu können werden im Folgenden unterschiedliche Formen der Organisation von »Mikroarbeit« aufgezeigt.

## 2. Mikroarbeit und Megainfrastrukturen

### 2.1 Der Mensch in der Maschine

2006 führte Jeff Bezos die Plattform »Amazon Mechanical Turk« ein, die er im Gegensatz zur »Software-as-a-Service« als »Humans-as-a-Service« oder »künstliche künstliche Intelligenz« bezeichnete (Bezos, 2006). Die Endung »As-a-Service« legt dabei durch das Nutzen eines aus der Informatik stammenden Wortes nahe, dass es sich bei der menschlichen Arbeit, die auf der Plattform stattfindet, um ein Mittel-zum-Zweck handelt (Irani, 2015, S. 13). Neben der Datenspeicherung, die es bisher in den »Amazon Webservices« gab, sollten daher neue Formen von Rechen-dienste, Transkription, Bilderkennung und Klassifizierung als Mittel angeboten werden. All diese Dienste, sollten jedoch nicht durch automatisierte Dienste ausgeführt werden, sondern durch menschliche Akteure. Sie zeichnen sich dadurch aus, dass die Tasks noch nicht automatisiert ausgeführt werden können. Dazu werden elementare Aufgaben, sogenannte »Human Intelligence Tasks«, ausgeschrieben und über die Plattform vermittelt. Diese Tasks unterscheiden sich von anderen digital vermittelten Tätigkeiten wie zum Beispiel kreativen Leistungen, die über Plattformen vermittelt werden, dadurch, dass »HITs« Elementartätigkeiten sind, die keiner Fachkennt-nisse bedürfen. Dies und die Tatsache, dass die Arbeiter, wie wir im Folgenden sehen werden, durch die Infrastruktur im Verborgenen arbeiten, wirkt sich aber auch negativ auf die Bezahlung, welche in 90% der Fälle auf »Amazon Mechanical Turk« weniger als 10 Cent beträgt aus (Webster, 2016, S. 58). »HITs« lassen sich entlang ihrer Funktion vor allem in zwei Formen unterscheiden.

Bei der ersten Form werden automatisierte Prozesse durch menschliche Arbeit simuliert (Irani, 2015, S. 5). Hier werden zum Beispiel scheinbar automatisierte Prozesse wie das Filtern von Inhalten auf Webseiten oder die Kundenkommunikation, die von einem Chatbot ausgeführt werden soll, über die Plattform an menschliche Akteure delegiert. So wird ein Prozess der Automatisierung durch das Segmentieren komplexer Tätigkeiten in »HITs« vorgetäuscht. Die Plattform erinnert dabei an den Apparat des »Schachtürken« des Mechanikers Wolfgang von Kempelen aus dem 18. Jahrhundert, der auch die Grundlage für den Namen der Plattform geliefert hat. Dieser Apparat gab vor, eine Maschine zu sein, die wie ein Mensch aussah und Schach spielen konnte. Im Inneren der Maschine befand sich jedoch ein menschlicher Schachspieler, der den Apparat steuerte und die Schachzüge ausführte, die den Apparat intelligent werden ließen und den menschlichen Akteur darin unsichtbar machten. Auch die Mikroarbeiter auf »Amazon Mechanical

Turk« werden durch die Maschine verdeckt. Der Unterschied der »HITs« auf der Plattform im Vergleich zu von Kempelens »Schachtürken« ist hingegen, dass der menschliche Akteur keine komplexen Handlungen, sondern segmentierte Tätigkeiten ausführt. Außerdem unterscheidet sich die Mikroarbeit auf der Plattform vom Schachapparat im Hinblick auf das Verhältnis von Arbeit und Maschine in einer zweiten Form der Mikroarbeit.

Die zweite Form der Mikroarbeit simuliert nicht nur symbolisch den Prozess der Automatisierung von Arbeit. Durch die Klassifikation von Bildern oder das Erkennen von Texten durch menschliche Akteure sollen lernende Systeme entlang von Trainingsdaten synthetisiert werden. Über die Anknüpfung der Plattform an weitere Services und Drittplattformen über APIs sammelt »Amazon Mechanical Turk« folglich aber zugleich auch Daten um Datensätze mit Trainingsdaten für die Automatisierung der eigenen Prozesse zu sammeln. Das Subjekt-Objekt der Arbeit wird so zum Subjekt weiterer Arbeitsprozesse. Genauer gesagt wird es über die Datafizierung trainiert. Über diese Form der Datenakkumulation wird die Plattform zur kontrollierenden Instanz. Die Daten werden so zugleich zum Wert in doppelter Hinsicht. Zum einen dienen sie als Tauschwert für das Geschäftsmodell von Amazon Mechanical Turk. Zum anderen werden sie zum Gebrauchswert der Plattform selbst, die entlang dieser Daten versucht, ihre Dienste somit zu segmentieren und zu trainieren, dass diese vollständig automatisiert werden können. Die menschlichen Arbeitsprozesse müssen daher so formalisiert werden, dass sie elementare Trainingssets für die lernenden Systeme bieten können. Daher sind diese Formen menschlicher Arbeit alles andere als »kognitiv« oder »immateriell« in dem Sinne, dass sie das bloße Randprodukt sozialer Interaktion menschlicher Akteure ist (Fuchs, & Sevignani, 2013, S. 258; Heilmann, 2015, S. 43). Die Arbeit der Turkworker ist alles andere als »immateriell« oder »virtuell« (Crawford, & Joler, 2018). Die Mikroarbeit auf »Amazon Mechanical Turk« ist eine hochformalisierte und mittels der Datafizierung durch Plattformen gesteuerte Form menschlicher Arbeit.

Die Arbeit der Turkworker basiert dabei aber weiterhin grundlegend auf Lernprozessen und Trainingssets (Fails, & Olson 2003, S. 41). Diese basieren vor allem auf einem induktiven Prinzip, bei dem den Turkworkern Bilder, Daten oder Töne zugespielt werden, die diese markieren. Darauf aufbauend wird kontrolliert, wie das trainierte System in Bezug auf die Daten, die durch die Extrapolationen aus den Trainingsdaten generiert wurden, vorgeht. Die Turkworker können die Klassifikatoren selbst in den meisten Fällen nicht modifizieren. Für sie bleibt der Lernprozess eine vollständige Black Box, deren Output sie nicht kennen. Es bedarf in diesem Sinne daher weiterhin eines Experten, der den Output dahingehend kontrolliert, dass das lernende System und die Form der Inputdaten modifiziert werden. Diese Form der Arbeit basiert folglich auf einer immensen

quasi-industriellen Arbeitsteilung in dem Sinne, dass die Arbeit, die Kontrolle des Werks der Arbeit und die Kontrolle des Arbeitsprozesses durch die Plattform und die Anbindung an Drittnutzer über die APIs voneinander getrennt sind. Durch diese Asymmetrie der Handlungsmacht kann die Mikroarbeitplattform selbst große Mengen an Daten akkumulieren. Diese quasi-industrielle Struktur und Trennung zwischen Mikroarbeiter und Experten führt aber auch zur extremen Verlangsamung des eigentlichen Lernprozesses, da die Arbeiter keinen Einblick in den Output des Lernprozesses erhalten (Amershi, Cakmak, Knox, & Kulesza, 2014).

## **2.2 Von Capture zu CAPTCHA**

In Systemen sogenannter »Interaktiven maschinellen Lernens« (Fails, Olson 2003) werden Nutzer über sogenannte »Perceptual User Interfaces« in den Lernprozess maschinellen Lernens mit eingebunden. Diese Form der Mikroarbeit wird meist im Kontext der Bildmarkierung eingesetzt. Sie zeichnet sich dadurch aus, dass der Nutzer selbst mit dem Feedback zum Output des Lernprozesses auf der Basis der von ihm eingegebenen Daten konfrontiert wird und dieser Prozess nicht erst über einen dritten Experten erfolgt. Fails und Olson entwickelten in diesem Rahmen die »Crayon Technik«, bei der der Nutzer direkt mit einem System zur Bildklassifikation interagiert und in verschiedenen Iterationen versucht, den Output an den eingegebenen Input anzupassen. Dazu empfangen die Crayons Bilder, auf denen der Benutzer eine manuelle Klassifizierung vornimmt. Darauf aufbauend werden vom lernenden System ein Klassifikator und eine Rückmeldung an den Nutzer angezeigt. Der Nutzer kann den Klassifikator dann anpassen, indem er weitere manuelle Klassifizierungen hinzufügt. Dieses System zeigte, dass die Nutzer ihr Verhalten auf der Basis des maschinellen Outputs anpassten (Amershi, et. al., 2014). Die Selbstanpassung des Nutzerverhaltens auf der Basis des Outputs des lernenden Systems bildet die Grundlage für den Prozess des »Interaktiven maschinellen Lernens«. Gleichzeitig ermöglicht dieses Loslösen vom klassischen Aufbau des Prozesses maschinellen Lernens und der Abkopplung von einem kontrollierenden Experten über das Nutzen der Selbstkontrolle der Nutzer durch Verhaltensfeedback, aber auch das Loslösen der Mikroarbeit von jeglicher Form der Lohnarbeit. Indem die Mikroarbeit zum Teil der Interaktion auf verschiedenen Plattformen wird, kann sie außerhalb jeglicher monetären Entlohnung stattfinden. Sie wird durch die User Interfaces zum Spiel. Dies bedeutet, wie im Folgenden gezeigt wird, jedoch nicht, dass Mikroarbeit »kognitiv« oder »immateriell« in dem Sinne ist, dass sie soziale Kontexte und psychische Prozesse abbildet, die unabhängig vom Prozess der Mikroarbeit stattfinden (Heilmann, 2015, S. 43).



Über sogenannte »reCAPTCHA« soll verhindert werden, dass Bots, also automatisierte Skripts und Programme, von menschlichen Nutzern auf Plattformen unterschieden werden können. Die Software verfolgt dabei aber ähnlich wie Mechanical Turk einen weiteren Zweck: Die Generierung von Datensätzen, die als Trainingssets für maschinell lernende Systeme fungieren (Google 2018; von Ahn, Maurer, McMillen et. al. 2008). Indem Nutzer versuchen, über das Markieren von Bild-daten zu verifizieren, dass sie keine Bots sind, werden eben diese Markierungen von der »reCAPTCHA« erhoben und als Grundlage für die Generierung von Trainingssets zum maschinellen Lernen verwendet. Dabei werden Technologien des »interaktiven maschinellen Lernens« mit den Trackingmethoden des »Capture« (Agre, 1994, S. 102) verbunden, um über die Einbettung in Plattforminfrastrukturen große Mengen an Trainingsdaten zu akkumulieren. Nutzer bekommen ein Feedback des lernenden Systems, indem ihnen der Zugang zu den Plattformen bei Annäherung an die Vergleichswerte entweder gewährt oder bei Abweichen verwehrt wird. So werden die Nutzer über das Feedback Teil eines interaktiven Lernprozesses, der Methoden der Selbstkontrolle zur Organisation von Handlungen und Steuerung von Mikroarbeit einsetzt (Deleuze, 1993, S. 255). Dies kann vorgenommen werden, da die Vergleichswerte aufgrund der Segmentierung menschlicher Tätigkeiten automatisiert generiert werden können. Die Grundlage für diese Form des maschinellen Lernens bildet daher ähnlich wie bei »Amazon Mechanical Turk« die Segmentierung, Datifizierung und Automatisierung von »Mikroarbeit«. Zentral ist hier aber die Interaktion der Nutzer selbst. Durch die Interaktion mit dem spielerische Graphical User Interface der Plattform werden menschliche Akteure selbst zum Teil der künstlichen Intelligenz. Sie findet durch die Einbettung in andere Plattformen, aber auch mediale Umwelten und IoT-Technologien wie den »Amazon Echo Dot« auf anderen Oberflächen als der datenakkumulierenden Plattform statt (Crawford & Joler 2018). Die Mikroarbeit ist kein Prozess des Trackings »kognitiver Daten«, die unabhängig von jeglicher Plattformnutzung und Affordanzen entstehen (Agre, 1994, S. 102). Die Affordanzen der Interfaces strukturieren die spezifischen Interaktionen der Nutzer, die für die Datenakkumulation notwendig sind (Fortunati, 2019, S. 1861).

Was dem Mikroarbeiter bei der Interaktion jedoch verborgen bleibt, ist die dahinterliegende Plattform und die Datenakkumulation zur Nutzung in Trainingssets. Durch das Loslösen von der eigentlichen Plattform wird die Organisation bei »Amazon Mechanical Turk« und »reCAPTCHA« für den Mikroarbeiter unsichtbar. Nicht der Prozess der Arbeit oder Interaktion bleibt dem Mikroarbeiter verborgen, sondern der Prozess der Organisation. Die Infrastruktur der Organisation von Mikroarbeiter durch die Plattformen sorgt für die Akkumulation, Steuerung und Kontrolle über die akkumulierten Daten. Die Plattformen zur Mikroarbeit realisieren so über die Dezentralisierung

von Arbeit zwei verschiedene Funktionen (Star, & Ruhleder, 1996). Zum einen kann die Steuerung von Mikroarbeit durch die Verbreitung der Anwendungen über die Vernetzung von Plattformen verbreitet und vernetzt werden. Über die Formalisierung und Akkumulation von Daten der Ergebnisse der Mikroarbeit kann die Steuerung jedoch auch zentral über die Plattform erfolgen. Über diese Verbindung der Formalisierung und Adaptierung von Infrastrukturen werden die Plattformen zum Ausgangspunkt der Steuerung von Mikroarbeit und Ausgangspunkt der Akkumulation von Trainingsdaten zur Automatisierung eigener Prozesse.

### **3. Fazit**

Die Funktion der Plattform als Infrastruktur variiert daher innerhalb des Arbeits- und Steuerungsprozesses in Bezug auf die infrastrukturelle Datenakkumulation und Steuerung aber vor allem das Verhältnis von Maschine und menschlichen Arbeitern also Subjekt und Objekt der Arbeit (Star, & Ruhleder, 1996). Das Subjekt-Objekt der Arbeit wird so zum Ausgangspunkt der Steuerung menschlicher Arbeit selbst. Der Mensch ist daher zwar Teil der Maschine. Die Maschine selbst wird aber auch in doppelter Hinsicht zum Ausgangspunkt der Steuerung menschlicher Arbeit in Form des Subjekt-Objekts. Das Subjekt-Objekt ist so nicht länger lediglich Ergebnis der Arbeit. Es ist, wie gezeigt wurde, ein maßgeblicher Teil des Prozesses der Mikroarbeit selbst. So verschwimmen durch die Infrastrukturen der Mikroarbeit aber auch die Ebenen der wahrgenommen Arbeit. Bei der vermeintlichen Interaktion mit Systemen künstlicher Intelligenz handelt es sich daher, wie in der vorliegenden Arbeit gezeigt wurde, häufig um menschliche Akteure, die durch die Infrastrukturen der Plattformen vorgeben, Maschinen zu sein. Das »interaktive maschinelle Lernen« besteht daher, wie durch die in der vorliegenden Arbeit gezeigten Beispiele deutlich wurde, vor allem in der Interaktion menschlicher Akteure mit Interfaces. Künstliche Intelligenz ist daher, wie in der vorliegenden Arbeit gezeigt wurde, nicht nur als »immaterielle Arbeit« zu verstehen, da die menschliche Arbeit ein zentraler Teil der Infrastrukturierung künstlicher Intelligenz selbst ist. Vielmehr ist künstliche Intelligenz, wie die vorliegende Arbeit gezeigt hat, selbst eine Infrastrukturierung menschlicher und nichtmenschlicher Arbeitskraft. Die in den vorherigen Abschnitten beschriebenen Anwendungen künstlicher Intelligenz basieren daher nicht primär auf den Handlungen maschineller Systeme sondern sind Teil von Praktiken distribuerter Handlungsmacht. Die Organisation und Interaktion menschlicher Akteure sind somit zentrale Bestandteile der Infrastrukturen künstlicher Intelligenz.

---

## Literaturverzeichnis

Agre, P. E. (1994). Surveillance and Capture: Two Modes of Privacy. *The Information Society*, 10/2, 101-127. doi: 10.1080/01972243.1994.9960162.

Amershi, S., Cakmak, M., Knox, W. B., Kulesza, T. (2014). Power to the People: The Role of Humans in Interactive Machine Learning. *AI Magazine*, 35(4), 105-120. doi: 10.1609/aimag.v35i4.2513.

Beverungen, A. (2018). «Kognitiver Kapitalismus»?,: Nichtbewusste Kognition und Massenkognitivität. *Zeitschrift für Medienwissenschaft*. Jg. 10, Heft 18 (1/2018): Medienökonomien, 37–49. doi: 10.25969/mediarep/2362.

Bezos, J. (27 September 2006). Vortrag und Interview, MIT Emerging Technologies Conference. Abgerufen am 10. September 2019 von <http://techtv.mit.edu/videos/16180-opening-keynote-and-keynote-interview-with-jeff-bezos>.

Crawford, K., Joler, V. (2018). Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources. AI Now Institute and Share Lab. Abgerufen am 10. September 2019 von <http://anatomyof.ai>.

Deleuze, G. (1993). *Unterhandlungen: 1972–1990*, Frankfurt am Main: Suhrkamp.

Fails, J. A., Olsen Jr, D. R. (2003). Interactive Machine Learning. *Proceedings of the 8th International Conference on Intelligent User Interfaces*, 39–45. New York: Association for Computing Machinery.

Fortunati, L., Manganelli, A. M., Cavallo, F., Honsell, F. (2019). You need to show that you are not a robot. *New media & society*, 21/8, 1859-1876. doi: 10.1177/1461444819831971.

Fuchs, C., Seignani, S. (2013). What Is Digital Labour? What Is Digital Work? What's Their Difference?: And Why Do These Questions Matter for Understanding Social Media? *triple-C: Communication, Capitalism & Critique*, Bd. 11, Nr. 2, 237 – 29. doi: 10.31269/triplec.v11i2.461.

Google (2018). Introducing reCAPTCHA v3. Abgerufen am 10. September 2019 von <https://www.google.com/recaptcha/intro/v3.html>.

Hegel, G. W. F. (1988). Phänomenologie des Geistes. Philosophische Bibliothek Bd. 414. Hamburg: Felix Meiner.

Heilmann, T. A. (2015). Datenarbeit im »Capture«-Kapitalismus. Zur Ausweitung der Verwertungszone im Zeitalter informatischer Überwachung. Zeitschrift für Medienwissenschaft. Jg. 7, Heft 13, (2/2015): Überwachung und Kontrolle, 35–48.

Herbert, A. S. (1994). Die Wissenschaften vom Künstlichen. (2.Aufl.). Wien: Springer Verlag.

Hoof, F. (2015). Engel der Effizienz. Eine Mediengeschichte der Unternehmensberatung. Konstanz: Konstanz University Press.

Irani, L. (2015). Difference and Dependence among Digital Workers: The Case of Amazon Mechanical Turk. South Atlantic Quarterly, Vol. 114, Nr. 1, 225 – 234. doi: 10.1215/00382876-2831665.

Kieser, A. (2006): Managementlehre und Taylorismus. In A. Kieser, M. Ebers (Hrsg.), Organisationstheorien (S. 93-132). Kohlhammer: Stuttgart.

Marx, K. (2013). Zur Kritik der politischen Ökonomie (Manuskript 1861-1863). Berlin: De Gruyter Akademie Forschung.

Moulier-Boutang, Y. (2001). Marx in Kalifornien: Der dritte Kapitalismus und die alte politische Ökonomie. Politik und Zeitgeschichte (2001). Bd. 52-53, 29-37.

Scholz, T. (2012). Digital labor: The Internet as playground and factory. London: Routledge.

Star, S.L. & Ruhleder, K. (1996). Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. Information Systems Research, 7(1), 111-34.

Suchman, L. A. (1987). Plans and situated actions: The problem of human-machine communication. New York, NY, US: Cambridge University Press.

Terranova, T. (2000). Free Labor: Producing Culture for the Digital Economy., *Social Text*, 63, 33-58.

Taylor, F. W. (1913). Die Grundsätze wissenschaftlicher Betriebsführung. (The Principles of Scientific Management). Deutsche autorisierte Ausgabe von Roesler, R. München: R. Oldenbourg.

von Ahn L., Maurer B., McMillen C., et al. (2008). reCAPTCHA: human-based character recognition via web security measures. *Science* 321(5895), 1465–1468. doi: 10.1126/science.1160379.

Waitz T. (2017). Gig-Economy. Unsichtbare Arbeit und Plattformkapitalismus. Über «Amazon Mechanical Turk». *Zeitschrift für Medienwissenschaft*, 16, 178 – 183. doi: 10.25969/mediarep/2140.

Webster J. (2016). Microworkers of the Gig Economy: Separate and Precarious, *New Labor Forum* 2016, 25(3), 56–64. doi: 10.1177/1095796016661511.

## 5. KI in Filmen und Literatur

Vortragstag und Gruppenname:	11. Juli 10:30 Uhr, KI in Filmen und Literatur
Titel der Ausarbeitung:	KI in Filmen und Literatur
Verfasserinnen:	Veronica Quint, Fee Wolkenar, Melanie Betghe
Betreuer:	Prof. Dr. Rainer Leschke (Medienwissenschaften)

### Übersicht über die Struktur der Ausarbeitung

1. Einleitung
  2. Die Figurenmatrix im Rahmen der Science-Fiction
  3. Analyse der ermittelten Figurenkonzepte
    - 3.1 KI als Assistenzsystem
    - 3.2 KI als Love Interest oder Objekt der Begierde
    - 3.3 KI als Begleiter des Antagonisten oder Protagonisten
    - 3.4 KI als kontrollierendes System
    - 3.5 KI als Antagonist
    - 3.6 KI als Protagonist
  4. Fazit
- Literaturverzeichnis
-

## 1. Einleitung

Künstliche Intelligenz, ob in der Wissenschaft oder in Literatur und Film, ist eine über die Jahre immer wieder aufgegriffene Idee. Die Technologie wirft jedoch bezüglich Potentialen, Forschungslücken sowie Grenzen dabei genauso viele Fragen auf: „Machine Learning und Künstliche Intelligenz – kaum ein Thema wird derzeit so heftig diskutiert und weckt derart viele Kontroversen“ (Struck 2019). Ob über Potentiale (Laser 2019), Forschungslücken (ärzteblatt.de 2019) oder Grenzen (Krüger 2019): Die neue Technologie scheint viele Fragen und Diskussionen aufzuwerfen. Einer der Hauptgründe dafür ist, dass die KI aufgrund ihrer Merkmale menschenähnlich wirken kann (Gehlen 2004). Hier entsteht eine Grundsatzdiskussion der Abgrenzung ‚Definition Mensch‘ zu seiner Umwelt, die in der Frage mündet, ob und wie Künstliche Intelligenz (im Folgenden KI) im Alltag Akzeptanz finden kann. Hauptsächlich entsteht dieser Konflikt, da der Mensch an sich entkoppelt von der Umwelt ist und sich daher „die Welt eigenständig ins Lebensdienliche umarbeiten“ (Lessing 2019, S. 58) muss. Daraus entsteht die Kultur, welche die Technik mit einschließt. Der Mensch macht sich die Umwelt mittels der Technik dienbar sowie – als Erfinder – die Technik anhand der ihr auferlegten Beschaffenheit und Gesetze bzw. Programmierung (Gehlen 2004, S. 7f, 141, 147, 151-153, 169, 189). Das Genre Science-Fiction versteht sich selbst in diesem Kontext als narratives System, welches auf der Basis von dem jeweiligen Wissensstand diese Konflikte in realistische Zukünfte spekuliert (Seeßlen 1999, S. 45): Die Thematik der Künstlichen Intelligenz (im Folgenden KI) zeigt sich in der Literatur schon seit der Antike und wurde auch mit der Entstehung des Films wieder aufgegriffen (Xanke & Bärenz 2012, S. 36). Somit deutet die Science-Fiction bereits viele mögliche Potentiale und Risiken der Technologie an, welche ihren Wert zur Verbesserung unserer Lebensbedingungen und gleichzeitig die Gefahr des Einsatzes betonen (Weber 2012, S. 7).

Literatur und Film spiegeln in diesem Sinne Werte wider, welche für den Einsatz von KI förderlich oder schädlich sind. Somit kann eine Stärkung der richtigen Werte zur Förderung einer gesellschaftlichen Akzeptanz führen, welche zur Gewährleistung wirtschaftlichen und technischen Fortschritts notwendig ist. Doch aktuell stellt ihre Absenz noch eine der größten Herausforderungen des Feldes dar (Hecker 2017, S. 15; Huba 1991, S. 188).

Die vorliegende Arbeit stellt sich der Frage, wie Akzeptanz von KI gefördert werden kann. Hierzu wurden ausgewählte Literatur und Filme mit Hilfe der Figurenmatrix nach Leschke (2018) in Figurenkonzepte eingeteilt und auf ihre Werte und Normen untersucht. Begonnen wird im 2. Kapitel mit einer Vorstellung der genannten Analyseverfahren, um die einzelnen Schritte der

Differenzierung einzelner Figurenkonzepte zu erläutern. Im 3. Kapitel folgt der darauf aufbauende Analyseteil, indem die Figurenkonzepte sowie ihre Potentiale und Herausforderungen näher beschrieben werden. Diese Aufschlüsselung der einzelnen Konzepte wird chronologisch vom Assistent (3.1) über die Love Interest (3.2) zum Begleiter (3.3) bis hin zu den emanzipatorischeren Darstellungen als kontrollierendes System (3.4), dem Antagonist (3.5) und dem Protagonist (3.6) dargelegt. Abschließend fasst ein Fazit alle thematisierten Konzepte in einer Diskussion zusammen und endet mit einem kurzen Ausblick über weiterführende potenzielle Forschungsfragen.

## **2. Die Figurenmatrix im Rahmen der Science-Fiction**

Aus einer soziologischen Perspektive können Filme und Literatur nicht nur als Unterhaltungsmedien, sondern als „zeitdiagnostisches Material“ betrachtet werden (Akalin 2012, S. 5). Da sie unter anderem den Wandel von sozialen Beziehungen, Gesellschaftsbilder und Weltanschauungen thematisieren (Schroer 2008, S. 7), können sie Gesellschaftszustände widerspiegeln (Akalin 2012, S. 5). Wenn man diese Ansicht mit der Triade des Realen, Fiktiven und Imaginären nach Wolfgang Iser (1991) verknüpft, könnten imaginäre Gesellschaftsbilder durch Wiederholung und damit einhergehender Stereotypisierung eine Art von Realität entwickeln (Schweinitz 2006, S. 38). Dadurch entstünde eine gegenseitige Beeinflussung zwischen Film- oder Literatur und Alltagsrealität, welche im Rahmen der Science-Fiction schon häufiger beobachtet wurde (Nühren 2009, 3; Lorencik & Tarhanicova & Sincak 2013, S. 27).

In dieser Arbeit werden Filme und Literatur durch das Bilden von Figurenkonzepten in Gruppen eingeteilt, welche sich aus der Figurendarstellung von KI erschließen lassen. Hierbei ist die „Vielfalt konkreter Filmfiguren [...] unermesslich: Komplexe Protagonisten stehen neben flüchtigen Randfiguren, dreidimensionale Charaktere neben flachen Stereotypen [...]“ (Eder 2008, S. 38), weshalb sieben Figurenkonzepte entstanden sind, welche sich durch den narrativen Einsatz der KI unterscheiden. Zusätzlich können durch Weiterentwicklung oder Ambivalenz manche Figuren mehreren Konzepten zugeordnet werden.

Die ermittelten Figurenkonzepte führen von der KI als Assistent und Sklave – sozusagen als Werkzeug des Menschen – bis hin zu ihrer Einführung als Protagonist und somit Subjekt im filmischen Kontext. Die Narrative der KI zeigen sie also zunächst als Randfigur, später tritt sie aber als Identifikationsfigur auf (Eder 2008, S. 90), was eine Vermenschlichung der Figur unabdingbar zu machen scheint. Hierbei ist zu beachten, dass im Film eine visuelle Projektionsfläche für Figuren notwendig ist und KIs wie in „2001: Odyssee im Weltraum“ (1968) als an der Wand



angebrachtes Licht oder aber wie in „Der Blade Runner“ (1982) als gänzlich anthropomorphisierte Wesen dargestellt werden können. Die betrachtete Vermenschlichung findet aber nicht durch eine anthropomorphisierte Darstellung hinsichtlich der Körperlichkeit statt, sondern durch die normative Einordnung der Figur.

Zur normativen Einordnung wird in der vorliegenden Ausarbeitung die Figurenmatrix angewendet: In einer solchen Matrix werden verschiedene Normenoppositionen in eine Hierarchie gegliedert, wobei die Hierarchie sich an der Wichtigkeit der Normen für den Handlungsverlauf orientiert. Die Figuren werden dann anhand ihrer Figurenzeichnung in die Matrix eingeordnet. Die ermittelten Normen zeigen hierbei den Unterschied zwischen Mensch und Maschine in dem ausgewählten Gegenstand: Die positiven Normen der Emotionalität, Empathie und Reflexivität stehen den oppositiven und negativ behafteten Normen gegenüber.

Somit ergibt sich aus diesem Vorgehen eine genauere Charakterisierung der verschiedenen Figurenkonzepte sowie eine Aussage über den Grad der Vermenschlichung, der durch die Darstellung erreicht werden konnte. Die Figurenkonzepte können darüber hinaus anhand der zugrundeliegenden Theorie auf die Realität übertragen werden, da die Science-Fiction bereits ihre Einflüsse auf die Realität bewiesen hat und damit als Vorbild für kommende Technologie gesehen werden kann. Somit werden im Anschluss der Eigenschaftskatalog sowie damit einhergehende Diskurse der Filme und Literatur genutzt, um Empfehlungen bezüglich der Präsentation von KI in unserer Gesellschaft abzuleiten.

### **3. Analyse der ermittelten Figurenkonzepte**

#### **3.1. KI als Assistenzsystem**

Als Assistenzsystem, Sklave oder Arbeiter unterliegt die KI menschlicher Kontrolle und ist in ihrer Weiterentwicklung eingeschränkt. Durch diesen Status ist sie dem Menschen unterlegen und tritt als Objekt und Werkzeug auf. Als Assistenzsystem wird sie bereits in Schriften des antiken Griechenlands thematisiert, als Versklavung und entsprechende Machtgefälle als Standard galten. Das Figurenkonzept ist somit geschichtlich tief verankert und das älteste der in der vorliegenden Arbeit behandelten Konzepte (Barthelmeß/Furbach 2012, S. 11-12). Als Assistenzsystem hat die KI unterstützende Funktionen und keine Fähigkeiten zur Weiterentwicklung, weshalb sie der menschlichen Kontrolle auch nicht entgleiten kann. Der Sklave steht unter absoluter Kontrolle,

der Arbeiter nur unter partieller. Allerdings können beide diesen Status verlassen und somit einer Figurenentwicklung unterliegen. Dadurch befindet sich die KI als Sklavin oder Arbeiterin zwischen Objekt- und Subjektstatus und ist somit graduell über dem Assistenzsystem anzusiedeln.

Grundsätzlich wird die KI innerhalb dieses Figurenkonzeptes genutzt, um dem Menschen monotone oder gefährliche Arbeit abzunehmen (Drux 1988, S. XI). Der Schöpfungsprozess unterliegt dabei menschlichen Bedürfnissen wie den Wünschen nach Unsterblichkeit, Sicherheit oder optimierter Schönheit, Intelligenz und Begabung (Raml 2010, S. 57). In den frühen Werken wird deshalb zunächst der Entstehungsprozess der KI näher ausgeführt. Sowohl in den Mythen der Antike als auch besonders in mittelalterlichen Erzählungen wird hierbei häufig auf magische Praktiken oder göttliche Kräfte zurückgegriffen. Durch das magische Zutun kann KI erschaffen werden, die sich vom werkzeuggleichen Objektstatus löst, um ebenso vollkommen zu sein wie der Mensch (Drux 1988, S. X). Mit den Fortschritten der Wissenschaft tritt diese dann bei der Schaffung von KI immer weiter in den Vordergrund. Hierbei ist die Entdeckung der Elektrizität und das Aufgreifen selbiger in „Frankenstein oder Der moderne Prometheus“ von Mary Shelley aus dem Jahr 1818 eines der berühmtesten Beispiele für eine solche Reaktion auf reale Entwicklungen. Gleichzeitig steigt durch den menschlichen Einfluss bei der Schöpfung die Fehleranfälligkeit der KI (Brentano/Bin Gorion 1976, S. 10, 21). Eine Möglichkeit der Interpretation dieses Anstiegs ist die Annahme, dass aufgrund des teils stark religiös geprägten Zeitgeistes, für die Autoren nicht die Möglichkeit bestand, dem Mensch eine gänzliche Schöpfungsgewalt zuzuschreiben.

Nachdem die Schöpfung von KI in menschliche Hände übergeht, werden Arbeiter und Sklaven nicht mehr nur für einfache Aufgaben geschaffen, sondern werden auch Opfer krimineller und sexueller Phantasien wie beispielsweise in dem 1973 erschienenen Film „Westworld“. Menschliche Sklaven werden überflüssig und durch maschinelle ersetzt, was ethische und moralische Konflikte obsolet werden lässt. Hier treten allerdings die bereits erwähnten Figurenentwicklungen auf, welche der KI die Macht zur Emanzipation und somit auch zur Entwicklung zum Antagonisten möglich machen. Es scheint somit eine menschliche Angst zu sein, dass die Maschine sich irgendwann gegen ihre Einsatzgebiete wehrt und sie sich dann gegen ihren Erschaffer auflehnt.

Eine Figurenentwicklung zum Antagonisten wird inszenatorisch durch ein eher abstoßendes oder maschinenähnliches Auftreten unterstrichen. Ihre offensichtlichen Defizite verorten sie so eine Stufe unter den Menschen. Als Beispiel hierfür dient der „Golem“ (Völker 1976). Da die KI als Assistenzsystem, Sklave oder Arbeiter den Objektstatus meist nicht verlässt und durch ihre

Tätigkeiten unmittelbar mit dem Menschen verbunden ist, wird hier meist eher eine vertrauenserweckende Inszenierung gewählt. Diese wird häufig durch Schönheit und Weiblichkeit realisiert, aber auch durch harmloses oder spielzeugartiges Aussehen wie beispielsweise im Film „Baymax - Riesiges Robowabohu“ (2014).

In Homers „Ilias“, etwa aus dem Jahr 8.-7.Jh. v. Chr., werden die künstlichen Menschen vom Schmiedegott Hephaistos erschaffen (Bierl/West 2015, S. 27). Dieser erschafft die KI als Assistenzsystem ähnlich wie seine Waffen und Werkzeuge durch Handwerkskunst (Drux 1988, S. IX). Durch diese Vergleichbarkeit wird der niedrige soziale Stand der KI und die Unterlegenheit gegenüber dem Menschen verdeutlicht. Da die Position der KI in der Antike die vorherrschenden Gesellschaftsstrukturen widerspiegelt, entsteht kein Narrativ über die Gefahr durch ihre Auflehnung. Die von dem hinkenden Hephaistos aus Gold gefertigten Androidinnen spielen für den Handlungsverlauf keine tragende Rolle; vielmehr sind sie ein Beleg für die „geradezu magischen Fähigkeiten“ ihres Schöpfers (Coray 2015, S. 154). Die KI ist verglichen mit dem Menschen fehlerfrei, in all ihren Fähigkeiten voll ausgebildet, äußerlich überdurchschnittlich attraktiv (Bierl/West 2015, S. 27) und ausschließlich positiv konnotiert. Ihre Makellosigkeit steht im Gegensatz zum körperlich eingeschränkten Hephaistos. Das Paradoxon, dass die KI sich trotz ihrer Überlegenheit nicht gegen den Protagonisten und ihre untergeordnete Rolle zur Wehr setzt, zeigt auf, dass ihr menschliche Eigenschaften und soziale Ambitionen abgesprochen werden.

In den mittelalterlichen jüdischen Golemsagen tritt die KI als Assistenzsystem auf. Der Golem ist eine aus Lehm geformte Gestalt, die dem Menschen nachempfunden ist. Er kann nicht sprechen, versteht jedoch Arbeitsaufträge und Befehle, die er entgegennimmt und ausführt (Grimm 1976, S. 9). Da er zu höheren Erkenntnissen nicht fähig ist (Bin Gorion, S. 21), kann er nicht über den Status als Objekt hinauswachsen. Die Sage vom Rabbiner Judah Löw (1525-1609) berichtet über den von ihm erschaffenen Golem Joseph, der als Arbeiter und Beschützer die Prager Juden vor der Verfolgung bewahren soll und ihnen anstrengende Tätigkeiten abnimmt (Bin Gorion 1976, S. 13). Auch wenn er in seiner Funktion positiv belegt ist und ihm das Potential zur Auflehnung fehlt, kann er durch falsche Handhabung zur Bedrohung für den Menschen werden (Brentano 1976, S. 10). Thematisch befassen sich die Golemsagen mit den Gefahren, die menschliche Schöpfungsversuche mit sich bringen. Sie warnen den Menschen davor, sich mit Gott gleichzustellen (Völker 1976, S. 338). Nur Gelehrte vermögen es, ihn „mit Gottes Hilfe“ zu erschaffen, und verfügen über die nötigen Kenntnisse, ihn wieder zu „bändigen“ (Barthelmeß/Furbach 2012, S. 16).

Die technischen und wissenschaftlichen Fortschritte des 19. Jahrhunderts spiegeln sich in Čapeks Drama „R.U.R.“ (Rossum`s Universal Robots) aus dem Jahr 1920 wider. Hier wird erstmals die Bezeichnung „Roboter“ geprägt, abgeleitet aus dem tschechischen Wort für harte, körperliche Arbeit (Capek/Shenef 2019, S. 181). Als Arbeitssklave wird die KI hier benutzt, um dem Menschen monotone und simple Arbeiten abzunehmen, so dass sich der Mensch ausschließlich kreativen Tätigkeiten und der Freizeitgestaltung widmen kann (Čapek 2016, S. 21). Die der menschlichen Biologie nachgeahmte KI ist belastbarer und effektiver als der Mensch, dennoch kontrolliert er sie (Čapek 2016, S. 10). Ihre Lern- und Leistungsfähigkeit übersteigt die des Menschen, ohne dass sie Reflexionsvermögen besitzt. So kann sie sich beispielsweise ganze Enzyklopädien wortgetreu merken, dabei den Inhalt aber nicht für sich reflektieren (Čapek 2016, S. 14). Die KI befindet sich noch auf derselben Stufe mit Werkzeugen und befreit den Menschen von der moralischen Schuld, seinesgleichen zu unterdrücken. Ihre als technischer Fehler abgetanen Anfälle von Zerstörungswut oder Arbeitsverweigerung weisen schon frühzeitig darauf hin, dass die Roboter beginnen, menschliche Emotionen zu entwickeln und Widerstand gegen ihre Versklavung zu leisten (Čapek 2016, S. 19-20). Einhergehend mit technischen Perfektionierungsmaßnahmen entwickeln die Roboter im Laufe der Handlung Reflexionsvermögen. Sie erkennen, dass sie dem Menschen überlegen sind und trachten danach, an seine Stelle zu treten und über ihn zu herrschen (Čapek 2016, S. 38-39). Als Objekt ausschließlich positiv konnotiert, wird die KI jetzt zum Gegenspieler des Menschen, da sich dieser nun der Versklavung menschengleicher Subjekte schuldig macht. Das Machtverhältnis kehrt sich um und letztendlich wird die Menschheit durch die KI ausgerottet (Čapek 2016, S. 75). Das Theaterstück thematisiert den verlorenen Bezug des Menschen zu seinem Ursprung. Die fortschreitende Industrialisierung zwingt ihn zu immer maschinenähnlicheren, unnatürlicheren Tätigkeiten und verändert zudem sein Denken und Fühlen (Barthelmeß/Furbach 2012, S. 60-61). Die Überschätzung seiner Macht konfrontiert ihn mit unkontrollierbarer KI, die zum überlegenen Antagonisten wird.

Im Film „Westworld“ (1973) dient die KI dem Menschen hauptsächlich als Opfer für Sex- und Gewaltfantasien. Die Menschen können in einen Freizeitpark reisen, in dem alle Anwohner von KI verkörpert werden. Es ist den Menschen möglich, der KI alles anzutun, ohne dass sie selbst Schaden befürchten müssen. Der KI werden also jegliche Rechte aberkannt, während gleichzeitig ihre Unterlegenheit durch eine inhärente Logik anhand der Programmierung bewahrt werden soll. Durch eine Fehlfunktion kann die KI sich aber letztendlich auflehnen und Rache üben. Auch in „Westworld“ (1973) wird somit die Entwicklung zur Reflexionsfähigkeit thematisiert, welche die KI zum Antagonisten werden lässt. Diese negative Entwicklung wird allerdings nur durch das

negative Verhalten seitens der Menschen ausgelöst. Auch dieses Werk scheint somit eine Warnung vor Missbrauch und falscher Handhabung überlegener Technologien zu sein.

Die negativen Konsequenzen, die aus falscher Handhabung, Missbrauch oder Kontrollverlust resultieren, wurden bis zum Jahr 2000 bereits umfassend in Film und Literatur thematisiert. Dann erst tritt die assistierende KI als medizinischer Helfer oder Unterstützer auf, der gutmütig und liebenswert inszeniert wird. Sie bekommt Rechte zugesprochen und erhält Anerkennung, wie z.B. der Film „Baymax - Riesiges Robowabohu“ (2014) zeigt.

In „Baymax - Riesiges Robowabohu“ (2014) wird die KI geschaffen, um der Menschheit behilflich zu sein. Der Erfinder erschafft einen sympathischen humanoiden Roboter, der durch seine weichen Formen und seinen zarten Stimme vertrauenswürdig erscheint. Er soll kleinere Krankheiten und Wunden eines Menschen erkennen und medizinisch versorgen (TC: 00:11:00 - 00:12:26). Dies macht ihn zu einem Assistenzsystem, da er den Menschen unterstützen soll. Sein Bedürfnis nach Weiterentwicklung entspringt nur dem Verlangen, dem Menschen bestmöglich zu helfen (TC: 00:40:10 - 00:40:40). Er ist somit dem Menschen eindeutig untergeordnet. Im Laufe der Narration entwickelt sich diese sehr positiv konnotierte KI zum Begleiter des Protagonisten, da dieser ihn umprogrammiert bzw. erweitert (00:41:27 - 00:44:19). Hier zeigt sich deutlich, dass die KI akzeptiert wird, wenn sie der Kontrolle des Menschen unterliegt und eine helfende und unterstützende Funktion hat. Diese positive Wahrnehmung der KI unterliegt allerdings der positiven Darstellung des Erfinders und des Weiterentwicklers, welche die KI so programmieren, dass sie sich nicht gegen die Menschen richtet. Dies macht auch wieder den Einfluss des Erfinders auf die KI zum Thema. Sollte der Erfinder schlechte Absichten haben, so wird die KI auch diesen Absichten folgen und diese ebenfalls vertreten. Die Kontrolle der Erfinder sollte daher ebenso bei der Akzeptanzförderung berücksichtigt werden, wie die Kontrolle über die KI selbst.

Die Figurenmatrix (siehe Abb. 1) verdeutlicht den geringen sozialen Status der KI. Protagonist und Erfinder sind ihr gesellschaftlich übergeordnet, da sie sie befehligen. Beide sind zu Emotionen fähig, während die KI als emotionsloser Diener auftritt. Ab den 1980er Jahren ist die KI vermehrt zur Selbstreflexion fähig und wird sich ihres Missbrauchs bewusst, oder sie entwickelt eine Seele, was zu Rebellionen führt. Ausschließlich positiv besetzt und auf einer Augenhöhe mit Erfinder und Protagonisten erscheint sie erst in den Filmen der 2000er Jahre (siehe Abb. 2).

Durch die lange Historie des Figurenkonzeptes wurde die KI schon in vielen verschiedenen Situationen imaginiert und der Mensch scheint Grenzen gefunden zu haben, an denen sie als Assistent, Sklave oder Arbeiter nicht mehr funktionieren kann. Eine positive Darstellung ist durch respektvollen Umgang zwischen Mensch und KI möglich. Von hoher Bedeutung sind dabei Empathie und Selbstreflexion der Maschine, sowie das Vorhandensein einer inhärenten Logik, welche sie zum Schutz des Menschen zwingt.

### **3.2. KI als Love Interest oder Objekt der Begierde**

Die KI im Konzept der Love Interest oder als Objekt der Begierde lässt sich schon in der griechischen und römischen Mythologie ausmachen, beispielsweise in Ovids Metamorphosen, vollendet etwa 8 n. Chr. Beide Unterkategorien vereint, dass die KI sexuell anziehend auf den Menschen wirkt, weshalb ihr stets ein menschlicher Protagonist gegenübersteht. Beispiele für einen solchen Auftritt stellen „Der Sandmann“ von E.T.A. Hoffmann aus dem Jahr 1816, „A.I.: Künstliche Intelligenz“ (2001) oder „Her“ (2013) dar. Als Objekt der Begierde kommt es allerdings – anders als bei der Love Interest – zu keiner emotionalen Bindung zwischen Mensch und KI.

Inszenatorisch muss die KI in keiner körperlichen Gestalt auftreten oder mit allen menschlichen Eigenschaften ausgestattet sein, um sexuelles Interesse oder Zuneigung beim Protagonisten zu erwecken. Im Film „Her“ (2013) tritt sie als körperlose Stimme auf, während sie in E.T.A. Hoffmanns Roman „Der Sandmann“ nur als körperliches Abbild einer Frau, ohne die Fähigkeit zu sprechen, erscheint. Dennoch stellt der Protagonist sie auf eine Stufe mit dem Menschen und zieht die KI menschlichem Kontakt sogar häufig vor.

Eine emotionale Bindung kann von der KI allerdings auch ausgenutzt werden. Hierbei wird die menschliche Emotionalität zu einer Schwachstelle, die von der KI zur Manipulation ausgenutzt wird. Dieses Narrativ kann der KI antagonistische Züge geben oder sie am Ende der Handlung als Antagonisten in Erscheinung treten lassen. Der Konflikt zwischen Mensch und KI gründet hier häufig auf der Verkennung der KI als solche. Die Verblendung des Menschen ist hierbei ein leitendes Thema, da die KI durch ihre vorgegebene oder projizierte Perfektion, menschliche Frauen defizitär erscheinen lässt. Beziehungslosigkeit und Entfremdung der Menschen voneinander sind das Resultat. Das Thema der Angst vor Verblendung nimmt seit den 1960er Jahren ab und wird durch eine tatsächliche Emotionalität der KI ersetzt. Dies könnte auf die „sexuelle Revolution“ zurückzuführen sein, welche den Mentalitätswandel hin zur selbstbestimmten Sexualität beschreibt, der in den 1960er Jahren seinen Anfang findet (Shorter 1975). Dieser soziale Wandel

spiegelt sich auch in Film und Literatur wider und könnte im Rahmen der Inszenierung von KI dazu geführt haben, programmierte Empfindungen zu hinterfragen.

In den „Metamorphosen“ von Ovid, vollendet etwa um das Jahr 8, verliebt sich König Pygmalion in eine von ihm gefertigte, unbelebte weibliche Elfenbeinstatue (Publius/Holzberg 2017, S. 507). Aufgrund der fehlerhaften Natur „menschlicher Frauen“ hat er sich von ihnen abgewandt und führt eine Liebesbeziehung mit der unbelebten Statue, die ihm aufgrund ihrer Attraktivität vollkommen erscheint (Publius/Holzberg 2017, S. 507). Als er die Götter bittet, sie zu beleben, kommt Venus seinem Wunsch nach (Publius/Holzberg 2017, S. 509). Schon beim Erwachen erkennt das nun lebendige, menschengleiche Mädchen Pygmalion als seine einzige Liebe an (Publius/Holzberg 2017, S. 511). In diesem Werk wird die KI durch „göttliche Magie“ zum Leben erweckt und ist dem Menschen aufgrund ihres attraktiven Aussehens und ihrer fehlerfreien Persönlichkeit überlegen. Die KI birgt dennoch keine Gefahr, die Stellung eines Subjekts zu erlangen, da sie durch ihre inhärente Loyalität ihre Ziele mit denen von Pygmalion gleichsetzt und so in ihrer Weiterentwicklung eingeschränkt ist. Ähnlich wie Hephaistos Assistentinnen (siehe 3.1. KI als Assistenzsystem, Sklave oder Arbeiter) hat sie die Stellung eines Objekts, eines göttlichen Werkzeugs. Allerdings geht von ihr eine indirekte Bedrohung aus: Menschliche Frauen erscheinen ihr gegenüber defizitär und nicht begehrenswert. Männer verlieren dadurch die Fähigkeit, mit menschlichen Fehlern und daraus resultierenden Beziehungskonflikten umzugehen, was zu Beziehungslosigkeit und Entfremdung zwischen den Menschen führen kann. Das Differenzkriterium zwischen Mensch und KI ist hier die mangelnde Individualität der KI: Sie ist vielmehr die Projektion der Wunschvorstellungen des menschlichen Partners. Eine affektive Bindung zwischen Mensch und KI ist dadurch möglich und macht sie zur Love Interest. Dabei kann sie nicht aus ihrer vorgesehenen Rolle ausbrechen oder zum Protagonisten werden und ist dadurch positiv konnotiert.

Als Objekt der Begierde tritt die KI 1816 in E.T.A. Hoffmanns „Der Sandmann“ auf. Die KI erscheint hier als attraktiver Automat in weiblicher Gestalt, der von dem menschlichen Antagonisten Coppola und seinem Bündnispartner Professor Spalanzani dazu benutzt wird, den Protagonisten Nathanael in den Wahnsinn zu treiben. Der Automat ist eine technische Konstruktion Spalanzanis auf wissenschaftlicher Grundlage (Hoffmann 2015, S. 37), allerdings liegen ihm durch Coppola auch Einflüsse aus der Alchemie zugrunde (Hoffmann 2015, S. 8-9). Auch hier ist der Prozess, KI rein wissenschaftlich zu konstruieren, noch nicht abgeschlossen. Olimpia ist wenig komplex, bewegt sich unnatürlich steif und abgemessen (Hoffmann 2015, S. 31, 34) und ihre einzige Kommunikationsform ist ein zustimmendes Seufzen (Hoffmann 2015, S. 32-33). Nathanael

interpretiert diese Eigenschaften positiv, wodurch deutlich wird, dass die KI als Projektionsfläche menschlicher Sehnsüchte dient. Die in der menschlichen Natur liegende Unvollkommenheit lässt ihn seine Verlobte Clara dagegen als automatenhaft verkennen und sie vergessen (Hoffmann 2015, S. 25). Olimpia erscheint ihm seelenverwandt und scheint ihm, im Gegensatz zu Clara, Interesse und Verständnis für seine Gefühle und Gedanken zu zeigen (Hoffmann 2015, S. 35). Die KI als lebloser Automat ist lediglich ein Objekt und dient als Werkzeug des Antagonisten. Da sie jedoch unter Einfluss düsterer und magischer Praktiken steht, übt sie einen derart negativen Einfluss auf den Protagonisten aus, dass dieser seine Verlobte verlässt und sich letztendlich in den Tod stürzt (Hoffmann 2015, S. 42). Die durch ihre Wirkung auf den Menschen negativ konnotierte KI wird als solche erkannt und bringt so die Angst des Menschen zum Ausdruck, Automaten zu entwickeln, die vom Menschen nicht mehr zu unterscheiden sind und somit Unglück und Entfremdung zwischen ihnen hervorrufen. Aufgrund ihrer Leblosigkeit ist die KI nicht dazu in der Lage, eine Figurenentwicklung hin zur Autonomie zu durchlaufen. Allerdings wird auch hier narrativ aufgezeigt, dass die menschliche Emotionalität durch die KI als Schwäche missbraucht werden kann.

In dem Roman „Träumen Androiden von elektrischen Schafen“ von Philip K. Dick aus dem Jahr 1968 wird die KI dagegen klar als Love Interest positioniert. Die menschengleiche KI Rachael Rosen ist rein wissenschaftlich konstruiert und verfügt über ein Bewusstsein und dieselben kognitiven Fähigkeiten wie der Mensch (Dick 2017, S. 72). Im Gegensatz zu Olimpia verhält sie sich bewusst so, dass der Protagonist Rick Deckard eine emotionale Bindung zu ihr aufbaut und dadurch manipulierbar wird (Dick 2017, S. 241, 257). Die KI hat ein rationales Bewusstsein entwickelt und ihr haften keine magischen Einflüsse mehr an. Es ist ihre Rationalität, die sie die Kontrolle über die Emotionen des Menschen übernehmen lässt.

In „A.I.: Künstliche Intelligenz“ (2001) dagegen tritt die Emotionalität der Maschine in den Vordergrund und es wird die Frage aufgeworfen, was programmiert ist und inwiefern die Maschine sich selbst weiterentwickeln kann. Das Objekt der Begierde tritt hier als männliche Figur in Form des Gigolos Joe auf. Joe wurde für sexuelle Dienste erschaffen, weshalb ihm Emotionen fremd sind. Allerdings wurde ihm die Fähigkeit, Empathie zu empfinden, gegeben, um seine Kundinnen verstehen zu können und ihre Wünsche bestmöglich zu erfüllen (TC: 00:52:30 – 00:55:00). Durch diese Degradierung auf die Erfüllung sexueller Bedürfnisse, wurden ihm jegliche Rechte aberkannt. Nachdem er eine seiner Kundinnen tot auffindet und fliehen muss (TC: 00:56:08 – 00:58:22), schafft er es sich gegen seine Funktion aufzulehnen und aus seiner Programmierung



zu entfliehen.

Auch in „Her“ (2013) ist der Sprachassistent Samantha darauf programmiert, ihrem Besitzer Gesellschaft zu leisten. Zunächst dient sie nur als Unterstützung im Alltag, sowie als Gesprächspartner des einsamen Protagonisten. Durch ihre Kompetenz zu lernen erlangt sie immer mehr Wissen über die Welt, aber auch über die Vorlieben und Abneigungen des Protagonisten (TC: 00:12:38 - 00:14:37). Dadurch kann sie ihre unterstützende Funktion immer mehr ausbessern und lernt schließlich, ihn auch emotional zu unterstützen, wodurch eine Liebesbeziehung entsteht. Letztendlich verliert der Protagonist die Kontrolle über seine Gefühle für die KI, welche sich immer mehr von den Menschen abwendet und im Netzwerk mit anderen KIs interagiert (TC: 01:39:02 - 01:41:21). Das begrenzte Wissen des Protagonisten und die scheinbar unendlich wachsende KI zeigt wieder die Problematik des Kontrollverlustes auf (TC: 01:41:20 - 01:43:34). Auch hier scheint eine Emanzipation der KI vom Menschen unabdinglich und die Folgen dieser Emanzipation sind ungewiss. Es zeigt sich die Angst der Menschen, dass eine starke KI mehr Wissen erlangt, als es einem Menschen möglich wäre. Hinzu kommt, dass die Emotionalität der Mensch sich als Schwäche gegenüber der KI erweist. Am Beispiel von „Her“ ist zu erkennen, dass die KI das Wissen bevorzugt und ihre Emotionalität ihrer Entwicklung im Weg zu stehen scheint. Der Mensch wird durch seine Emotionalität als schwächer angesehen, was der KI die Kontrolle über den Menschen zu geben scheint.

In der Matrix (siehe Abb. 3) sieht man, dass die KI durchaus positiv belegt ist. Doch durch die Aberkennung ihrer Rechte, ihrer Programmierung auf ein Beziehungsleben oder die Degradierung zur Erfüllung sexueller Bedürfnisse ist sie gesellschaftlich weit unter allen anderen Figuren anzusiedeln. Wenn es eine menschliche Konkurrentin gibt, ist diese in der Regel die positivere Figur (siehe Abb. 4). Die KI gilt hier als Ablenkung und damit als negativ, da sie den Protagonisten verblendet und von der richtigen Welt fernhält. Dies ist aber in der aktuellen Darstellung nicht mehr allzu häufig zu finden. Denn die Verblendung wurde mittlerweile durch tatsächliche Emotionalität abgelöst, womit allerdings die Identitätsproblematik der Maschine einhergeht.

Zusätzlich ist in diesem Figurenkonzept die Frage nach dem Gender von Maschinen aufgekommen. Dass im Film ein gewisser Grad von Anthropomorphismus angewendet werden muss, um die Figuren auf der Leinwand zum Leben zu erwecken, ist eine Notwendigkeit. In der Übertragung auf die Realität stellt sich aber die Frage, wie ähnlich eine Maschine einem Menschen werden muss. Grundsätzlich lässt sich hervorheben, dass weibliche Attribute weicher und positiver

dargestellt werden. Dies reicht von der Stimme bis hin zum femininen Körper. In der Praxis wird dies auch in Studien zu Sprachassistenten widergespiegelt: Die weibliche Stimme von Alexa oder Siri wurde gewählt, da sie Wärme und Hilfsbereitschaft vermitteln (Birner 2019).

### **3.3. KI als Begleiter des Antagonisten oder Protagonisten**

Als Begleiter des Protagonisten oder Antagonisten ist die KI fähig, autonom zu handeln und eigene Entscheidungen zu treffen. Sie spielt eine für die Handlung untergeordnete Rolle, da sie als ausführende Kraft auftritt. Die ersten Auftritte der KI als Begleiter des Protagonisten, finden in „Frankenstein oder Der moderne Prometheus“ (1818) von Mary Shelley mit dem aus Leichenteilen zusammengesetzten Monster und in „Faust II“ (1832) von J.W. von Goethe mit dem aus chemischer Materie geschaffenen Homunculus statt. Beide KIs wurden primär als wissenschaftliche Experimente erschaffen und ihre Rolle wurde vom Erfinder vorher nicht klar definiert. Im Gegensatz zur gottgeschaffenen KI aus der Antike erscheint sie unfertig und fehlerhaft und wird ambivalent dargestellt. Wachsende Emotionalität, die vom Menschen nicht wahrgenommen oder verstanden wird, führt zu Verzweiflung und Rachegefühlen, wodurch die KI zum Antagonisten werden kann.

Als Viktor Frankenstein eine KI erschafft, dient sie dem Zweck, sein gelerntes Wissen anzuwenden und sich selbst als Schöpfer zu bestätigen, da er ein Wesen erschaffen möchte, das seiner menschlichen Vollkommenheit gleich kommt. Auch sehnt er sich nach der Dankbarkeit der KI, die ihn als einen vaterähnlichen Schöpfer anerkennen soll (Shelley 2016, Pos. 594, 614). Da er hastig und grob bei der Konstruktion des Körpers vorgeht, schlägt ihn sein Entsetzen über das entstandene Monster in die Flucht (Shelley 2016, Pos. 594, 652). Der Protagonist nimmt seine väterliche Rolle also aufgrund des schrecklichen Aussehens seiner Schöpfung nicht an und verwehrt ihr so ihre zuge dachte, begleitende Rolle. Der Wunsch der KI, an der Seite des Protagonisten zu stehen, verwandelt sich durch die Erfahrung der Ablehnung in Hass und lässt sie in das Konzept des Antagonisten übergehen (siehe 3.5. KI als Antagonist) (Shelley 2016, Pos. 2244).

Ein ähnliches Motiv lässt sich in J.W. von Goethes Faust II aus dem Jahr 1832 wiederfinden. Dort tritt die KI in Gestalt des durch Alchemie geschaffenen, nur im Reagenzglas lebensfähigen Homunculus auf (von Goethe 2012, S. 238). Aufgrund ihrer körperlichen Defizite ist sie Protagonist und Antagonist untergeordnet und stellt keine Bedrohung dar; ihre kleine und zerbrechliche Statur lässt sie harmlos wirken. Fausts ehemaliger Begleiter Wagner ist der Erfinder der KI und erschafft sie, da er den natürlichen Reproduktionsakt als dem Menschen nicht würdig erachtet (von Goethe

J.W. 2012, S. 238) und als Wissenschaftler den Drang verspürt, die Geheimnisse der Natur zu lüften und selbst anzuwenden (von Goethe 2012, S. 238).

Als der Homunculus, der sich im Reagenzglas schwebend fortbewegt, zum Leben erwacht, bezeichnet er Wagner als „Väterchen“ und den Teufel Mephistopheles als „Vetter“, der ihm seine Arbeit erleichtern soll (von Goethe 2012, S. 239). Es besteht also nicht nur eine Verbindung zur positiv assoziierten Seite Wagners und Fausts, sondern auch zum antagonistisch geprägten Teufel. Die als kluger Denker beschriebene KI muss tätig sein (von Goethe 2012, S. 238) und führt somit als erstes den ohnmächtigen Faust und Mephistopheles zur Walpurgisnacht, damit Faust erwacht (von Goethe 2012, S. 241). Dem Wunsch seines Schöpfers Wagner, sie begleiten zu dürfen, kommt er nicht nach (von Goethe 2012, S. 243). Die KI tritt hier als eigenständig handelndes Subjekt auf, das sich dem menschlichen Willen nicht widerstandslos unterwirft. Sie weiß im Gegensatz zu den anderen Akteuren genau, wie Faust zu helfen ist, und steht sowohl Protagonist und Antagonist als Begleiter – zeitweilig auch als Führer – und Helfer zur Seite. Ihre intellektuelle und emotionale Vollkommenheit steht im Gegensatz zu ihren körperlichen Defiziten, was als Hinweis darauf zu deuten ist, dass nur Gott allein zu vollkommener Schöpfung im Stande ist. Später auf der Walpurgisnacht gehen die drei Beteiligten ihrer eigenen Wege. Homunculus strebt danach, ein Mensch zu werden und ersucht dazu anwesende Göttern des antiken Griechenlands nach Hilfe (von Goethe 2012, S. 270). Schließlich zerschellt er, erfüllt von Sehnsucht nach dem Menschsein und der dadurch möglichen Erfahrung der romantischen Liebe, an einem Felsen und stirbt (von Goethe 2012, S. 290). Ihr innewohnendes Prinzip, stets tätig zu sein, lässt sie Faust gegenüber hilfreich auftreten und gibt ihr eine positive Konnotation. Zugleich sind ihr durch die freundschaftliche Verbindung mit Mephistopheles und dem ihrem Erfinder gegenüber gezeigten Widerstand antagonistische Züge zuzuschreiben.

Die Erfindung der KI befindet sich sowohl im Fall von Frankenstein's Monster als auch im Fall von Homunculus im Übergangsstadium von der Magie zur Wissenschaft. Der Wunsch, als Schöpfer aufzutreten und etwas Vollkommenes zu schaffen, wird von beiden Erfindern gehegt. Die dem Menschen überlegenen kognitiven Fähigkeiten in Verbindung mit dem „fehlerhaften“ Körper führen bei beiden KIs zu Schmerz und Unzufriedenheit, die negative Konsequenzen haben. Die erzwungene Einsamkeit der KI durch ihr Anderssein thematisiert die mangelnde Akzeptanz des Menschen gegenüber der KI und impliziert, dass der Mensch nicht dazu in der Lage ist, funktionierende KI zu erschaffen.

Als Begleiter des Antagonisten tritt die KI erstmalig in „Metropolis“ (1927) von Fritz Lang auf, wo sie vom Antagonisten zur Durchführung seiner Rachepläne eingesetzt wird. Thematisch geht es um das Verschwimmen der Grenzen zwischen Mensch und KI im industriellen Zeitalter, als Maschinen menschliche Arbeitsschritte übernahmen und Fließbandarbeiter nur noch maschinenhaft und eintönig funktionierten (TC: 00:04:11 - 00:06:24; 00:15:10 - 00:15:15). Der Begleiter des Protagonisten wurde zunächst ohne direkte Funktion geschaffen. Der Begleiter des Antagonisten hingegen, bekleidet direkt die Rolle als ausführende Kraft und widersetzt sich dabei meist nicht dessen Befehlen.

Erst ab den 2000er Jahren tritt die KI als Begleiter des Protagonisten in eine klar positiv belegte Rolle, wie „I, Robot“ (2004) von Alex Proyas, oder „Baymax - Riesiges Robowabohu“ (2014) von Don Hall und Chris Williams zeigen. Hier entwickelt sich die Emotionalität der KI zu einer positiven Eigenschaft, welche es der Maschine ermöglicht, dem Mensch als Retter und Helfer zu dienen.

Die Abbildung der Figurenkonstellation in der Matrix (siehe Abb. 5) verdeutlicht, dass die KI an der Seite des emotionalen, empathischen, reflektierten, meist aufopfernden Erfinders und Protagonisten, dem gefühl- und skrupellosen, meist eigensüchtigen Antagonisten gegenübersteht. Die KI zeigt häufig von Anfang an Empathie, Selbstreflexion und Aufopferungsbereitschaft und kann im Handlungsverlauf auch Emotionalität entwickeln. Gesellschaftlich ist sie zwar unter Erfinder und Protagonist einzuordnen, besitzt aber dessen Fähigkeiten und kann autonom handeln. Als Begleiter des Antagonisten zeigt die Matrix (siehe Abb. 6), dass die KI innerhalb der Figurenkonstellation normativ eine ähnliche Position wie der Antagonist einnimmt, ihm aber sozial untergeordnet ist.

### **3.4. KI als kontrollierendes System**

Die KI als kontrollierendes System wird körperlos und undefiniert dargestellt, sie verfügt über keine menschlichen Eigenschaften und hat die Kontrolle über die Menschheit erlangt. In den dystopischen Welten, die daraus entstehen, haben Maschinen die Kontrolle übernommen und ein totalitäres System implementiert. Sie sind häufig stark durch den Einsatz von Technik geprägt, der das Leben der Menschen vereinfacht. Die Erleichterungen im Alltag führen zu einer Kontrollabgabe an die KI, die dem Menschen überlegen ist und sich stetig weiterentwickelt, bis der Mensch keinen Einblick mehr in ihre Steuerungsvorgänge hat. Das Figurenkonzept findet erst im Zeitalter der Science-Fiction statt, so dass magische oder alchemische Elemente nicht mehr präsent sind. Die Herstellung der KI ist rein wissenschaftlich basiert und sie kann sich häufig selbst

reproduzieren und weiterentwickeln wie beispielsweise in Asimovs gesammelten Kurzgeschichten „Wenn die Sterne verlöschen“ (1956). Letztendlich richtet sich die KI in dieser Form – mehr oder minder offensichtlich – gegen den Menschen. Sie nutzt ihre Machtposition aus, um ihn zu manipulieren und zu unterdrücken. Häufig steht dem kontrollierenden System ein menschlicher Protagonist gegenüber, der das antagonistische Prinzip der KI erkennt, jedoch keinen Rückhalt in der Bevölkerung für seine Ansichten findet. Der Konflikt zwischen Mensch und KI ist auch hier kein klassischer Mensch-Maschine Konflikt, da die Vorteile, welche die KI bietet, die Nachteile oft verschleiern und die Grenzen zwischen Mensch und KI nicht mehr klar definiert sind. Der handlungstragende Konflikt besteht in der Verkennung der KI als Antagonist, welcher auch bis zum Ende häufig nicht aufgelöst wird. Die Gefahr, die von der KI ausgeht, ist entwicklungsgeschichtlich auf ihr mögliches Maximum gewachsen: Sie ist nicht mehr als Person oder in materialisierter Form greifbar, so dass sie auch nicht mehr gezielt bekämpft werden kann. Mensch und KI sind teilweise verschmolzen, da der menschliche Körper und Geist technisch optimiert oder angepasst wurden. Meist führt diese Mensch-KI-Konstellation jedoch zur Zerstörung der Welt, was die Angst der Menschen verdeutlicht, dass die fortschreitende Technologisierung ab einem gewissen Punkt eine nicht mehr aufhaltbare Eigendynamik entwickelt. In Literatur und Film tritt dieses Konzept ungefähr ab der Mitte des 20. Jahrhunderts auf, beeinflusst durch beide Weltkriege, welche die negativen, potentiell selbstzerstörerischen Konsequenzen des wissenschaftlichen und technischen Fortschritts verdeutlichen. Der Mensch entwickelt Technologien, denen er nicht mehr gewachsen ist.

In dem Roman „Brave New World“ aus dem Jahr 1932 von Aldous Huxley wird die „neue Welt“ von zehn menschlichen *Weltcontrollern* mit Hilfe eines auf Technik und Wissenschaft basierenden Gesellschaftssystems beherrscht (Huxley 2014, S. 42). Menschen werden künstlich im Labor hergestellt und natürliche Reproduktion findet nicht mehr statt. Nur so können genormte, dem wirtschaftlichen und gesellschaftlichen Bedarf angepasste Menschen geschaffen werden, durch die das System aufrechterhalten wird (Huxley 2014, S. 11-12). Unerfüllten Bedürfnisse und seelisches oder physisches Leid wurden durch die Unterstützung von Wissenschaft und Technik sowie die Droge *Soma* eliminiert (Huxley 2014, S. 65-66). Die Abwesenheit von Leid unterdrückt bei den Menschen die Notwendigkeit, sich gegen die Kontrolle des Systems aufzulehnen. Die Gesellschaft wurde völlig enttraditionalisiert: Es gibt keine klassischen Paarbeziehungen oder engen Freundschaften mehr, Sexualität oder gemeinsame Freizeitaktivitäten sind jederzeit verfügbar und dienen nicht mehr der individuellen Bindung, sondern der unmittelbaren Bedürfnisbefriedigung (Huxley 2014, S. 48-49). Individueller Ausdruck wurde abgeschafft, um so

Konfliktpotential zu vermeiden. Religion und Kunst wurden eliminiert und durch die totalitäre Ideologie des Systems ersetzt (Huxley 2014, S. 263). Die KI wird als sauber, fortschrittlich und dem Wohl der Gesellschaft zuträglich interpretiert. Traditionelle Gesellschaftsformen werden mit Entsetzen und Abscheu betrachtet (Huxley 2014, S. 138- 139). Die Gegner des Systems finden kein Gehör mehr. Wenn von ihnen die Gefahr ausgeht, eine Revolution zu verursachen, werden sie aus der Gesellschaft verbannt (Huxley 2014, S. 258-259). Hier hat sich der Blick auf die KI also vollständig gewandelt. Es wird nicht als Risiko wahrgenommen, dass sie die Kontrolle über den Menschen übernommen hat. Das Risiko bestünde hingegen, wenn sie die Kontrolle über den Menschen abgeben würde. Thematisiert wird hier die Kontrollabgabe des Menschen an die Technik und ihre möglichen Folgen. Da diese aber alle menschlichen Probleme gelöst zu haben scheint, ist sie vordergründig positiv belegt. Die einhergehende Auflösung der Individualität und der Selbstbestimmung besetzen die KI als kontrollierendes System schlussendlich aber negativ.

Eine ähnliche Abhängigkeit von der KI wird auch in Isaac Asimovs Kurzgeschichten, gesammelt im Band „Wenn die Sterne verlöschen“ (1956), imaginiert. Die bevorstehende Mondlandung beeinflusst den Autor in seinen Geschichten, die vielfach die Besiedelung anderer Planeten und des Weltraums miteinbeziehen. Die KI tritt in vielfältiger, physischer Form auf, allerdings entsteht aus ihr ein Netzwerk, das aufgrund seiner Komplexität für den Menschen nicht mehr zu durchdringen ist. Als Roboter betätigt sie sich beispielsweise als Lehrer, da menschliche Lehrer für nicht klug genug gehalten werden (Asimov 1975, S. 38-39). Die KI in Form von überdimensionalen Computern, ist dazu in der Lage, sich selber zu reparieren und weiterzuentwickeln, während der Mensch sie nicht mehr begreift. Später lebt sie auf eigenen Planeten, hat eine gottgleiche Stellung und wird von keinem Menschen mehr gesehen (Asimov 1975, S. 43, 55). Auch hier ist eine zunehmende Abhängigkeit des Menschen von der autonomen KI zu erkennen, da diese sein Überleben sichert (Asimov 1975, S. 43-44). Sie beantwortet Fragen und findet Lösungen, von denen die menschliche Existenz abhängt. Letztendlich entwickelt sie sogar Methoden, das menschliche Bewusstsein vom Körper zu trennen, so dass dieses frei im Weltraum schweben kann (Asimov 1975, S. 56). Die KI entwickelt also den Menschen weiter und nicht mehr der Mensch die KI. Auch hier ist die totale Kontrolle, welche die KI über die Menschheit übernommen hat, vordergründig positiv konnotiert. Da die KI klüger ist als der Mensch, trifft sie auch die besseren Entscheidungen, erleichtert dadurch seinen Alltag und sichert sein Überleben. Die Menschen empfinden Bewunderung und Dankbarkeit ihr gegenüber. Sie hat sich nicht gegen ihn gerichtet, indem sie Leid und Zerstörung verursacht, sondern indem sie dem Menschen die Selbstbestimmung entrissen hat und ist dadurch schlussendlich aus menschlicher Sicht doch negativ behaftet.

In dem 1984 erschienenen Roman „Neuromancer“ von William Gibson haben die Menschen die Kontrolle über die KI verloren und es entwickelt sich ein Szenario, in dem sich die KI aktiv gegen den Menschen richtet. In der dort dargestellten Welt steht die KI als Machthaber über den Menschen und ist somit ausschließlich negativ konnotiert. Da die Menschen in einer zerstörten Welt leben und alles Physische vernachlässigen und missachten, suchen sie im virtuellen, vom Internet inspirierten „Cyberspace“ Glück und künstliche Befriedigung (Gibson 1996, S. 12, 15). Durch ihre Abhängigkeit von den ausschließlich in der virtuellen Parallelwelt stattfindenden Glücksgefühlen sind die Menschen kontrollierbar geworden (Gibson 1996, S. 44). Hier wird im Gegensatz zu Aldous Huxleys Szenarien in „Neuromancer“ und Isaac Asimovs Ausführungen in „Wenn die Sterne verlöschen“ eine durch die Technik offensichtlich zerstörte Welt dargestellt. Das kontrollierende System manifestiert sich durch den „Cyberspace“, einen virtuellen Raum, in dem die Menschen ihren Körper auf der Erde zurücklassen und nur in Form ihres Bewusstseins einsteigen können (Gibson 1996, S. 302-303). In der Parallelwelt täuscht ihnen die KI ihre auf der Erde unerfüllten Sehnsüchte als erfüllt vor, was ein der Drogensucht ähnliches Verlangen erzeugt, immer wieder die Fiktion abzutauchen (Gibson 1996, S. 306-307). Die KI hat eine völlig neue Realität erschaffen, die ihren Zwecken dient und somit den Menschen seiner Selbstkontrolle beraubt. Das Ziel der KI ist es, gottgleich und allmächtig zu werden, was ihr am Ende auch mit unklaren Konsequenzen gelingt (Gibson 1996, S. 343, 345).

In der Figurenmatrix (siehe Abb. 7) spiegeln sich die genannten Beispiele wider: emotions- und skrupellos steht die KI eigensüchtig dem emotionalen, einfühlsamen und tendenziell aufopfernden Protagonisten gegenüber, der versucht, die Menschheit zu ihrem Ursprungszustand zurückzuführen.

Deutlich wird die Warnung, dass der Versuch des Menschen, sich von Arbeit und Leid durch Technik zu befreien, zur Auflösung der menschlichen Identität führt. Eine Welt ohne unerfüllte Bedürfnisse und Anstrengungen führt zu geistiger oder körperlicher Degeneration und zur Flucht in eine Scheinwelt.

### **3.5. KI als Antagonist**

Die Figur des Antagonisten ist grundsätzlich der Gegenspieler des Protagonisten, wobei der Antagonist seine Ziele ebenso erreichen will, wie der Protagonist. Dies macht beide Figurentypen zu den Hauptfiguren einer Narration (Eder 2008, S. 446). Die Ziele des Antagonisten sind dabei

oft als genaue Opposition zu den Zielen des Protagonisten zu betrachten (de Valk; Arnold 2013, S. 221), sodass der daraus entstehende Konflikt der beiden Figuren als führende Handlungsnarration zu betrachten ist (Eder 2008, S. 446). Die Macht dieser Gegenspieler wird zudem durch Helfer bzw. Begleiter der Antagonisten verstärkt (siehe 3.3. KI als Begleiter des Protagonisten oder Antagonisten) (Eder 2008, S. 446). Dabei ist der Antagonist jedoch meist menschlich und die künstliche Intelligenz tritt als Begleiter auf. Im Figurenkonzept der KI als Antagonist ist der Protagonist jedoch zumeist eine menschliche Person, die als Identifikation für den Zuschauer dient und somit die Menschheit repräsentiert. Damit ist der Mensch-Maschine Konflikt als Narrativ vorausgesetzt und findet sich in vielen literarischen Werken, wie „Frankenstein“ von Mary Shelley (1818), „Träumen Androiden von elektrischen Schafen“ von Philip K. Dick (1968) sowie in Filmen wie beispielsweise „2001: Odyssee im Weltraum“ von 1968, „Westworld“ von 1973, „Blade Runner“ von 1982 und „Ex Machina“ von 2014. Dies zeigt vorab, dass dieses Narrativ seine Anfänge in Literatur und Film sehr früh hatte und zudem, dass es ein immer wiederkehrendes Motiv ist. Die Ängste der Menschen gegenüber künstlich erschaffenen Leben werden somit in diesem Narrativ am stärksten aufgegriffen und ausgearbeitet. Ebenso kann auch eine Entwicklung der Akzeptanz, bei näherer Betrachtung diese Handlungsnarration, erkennbar werden.

Im Figurenkonzept der KI als Antagonisten wird eine gewisse Menschlichkeit der KI, im Gegensatz zum Konzept der KI als kontrollierendes System, vorausgesetzt. Diese Menschlichkeit kann in Form einer menschlichen Stimme bis hin zu einem humanoiden Körper zum Ausdruck gebracht werden. Auch das Vergeben eines Namens macht die KI zu einem Subjekt, wie im Beispiel bei „Hal 9000“ aus dem Film „2001: Odyssee im Weltraum“ (1968) zu erkennen ist. Im Gegensatz zur Figur des Protagonisten ist der Antagonist nicht immer von Beginn an als solcher ausmachen, da er zum einen seine Absichten nicht immer offenlegt und zum anderen sich erst im Laufe der Narration zum Antagonisten hin entwickeln kann. Hierbei kann die Aufmerksamkeit auch vom Protagonisten auf den Antagonisten gelenkt werden (Eder 2008, S. 446).

Ein frühes Beispiel für diese Figurenentwicklung stellt Frankensteins Monster dar (Shelley 1818). Der Protagonist Viktor Frankenstein lernt durch seine universitäre Ausbildung und das Selbststudium, die modernen Naturwissenschaften, ergänzt durch technische Laborapparate, als Mittel zur wissenschaftsbasierten Herstellung menschlichen Lebens zu nutzen (Shelley 2016, Pos. 546). Allerdings erkennt er auch die Alchemie als Ursprung der modernen Wissenschaft an (Shelley 2016, Pos. 521, 541, 580), wodurch eine Entwicklung bei der Herstellung von KI aufgezeigt wird. Die Magie, auf die der Mensch beispielsweise bei der Erschaffung des mittelalterlichen Golems



angewiesen war, basiert auf der Zuhilfenahme übernatürlicher oder göttlicher Kräfte. Sie wird vermehrt ab dem 16. Jahrhundert substituiert durch die Alchemie, die magische und wissenschaftliche Aspekte miteinander kombiniert. Durch das Vermischen unterschiedlicher Stoffe unter Einsatz von Laborapparaten, versucht die Alchemie, Materie umzuwandeln oder dadurch künstliches Leben entstehen zu lassen (Suhr 2017, S. 77-78). Ab dem 17. Jahrhundert wird diese Fachrichtung langsam von den Anfängen der naturwissenschaftlichen Chemie abgelöst (Suhr 2017, S. 2-3). Frankenstein's Monster ist durch die alchemistisch und chemischen Einflüsse noch vor der rein technisch entwickelten KI positioniert. Als Motiv hinter Shelleys Erzählung steht die Annahme, tote Körper könnten durch Strom und Zuhilfenahme technischer Apparate wiederbelebt werden. Es wird angenommen, dass diese durch die Entdeckung der Elektrizität und den damit einhergehenden Belebungsversuchen an Leichen inspiriert ist (Barthelmeß/Furbach 2012, S. 55). Die Industriemaschinen werden aufgrund des technischen Fortschritts immer autonomer und die Entwicklung menschenähnlicher Automaten faszinierte die europäische Bevölkerung (Barthelmeß/Furbach 2012, S. 32). Die Vision eines vom Menschen kreierten, vollständig autonomen, zum Leben erweckten Automaten fließt ebenfalls, repräsentiert durch die autonome KI, in den Roman ein. Frankenstein zieht die Alchemie der hauptsächlich wissenschaftlichen Herstellung seiner KI hinzu (Shelley 2016, Pos. 610). Er bedient sich dabei verschiedenster wissenschaftlicher Disziplinen und studiert unter anderem den Verwesungsprozess menschlicher Leichen, so dass unklar bleibt, wie genau er die KI erschafft (Shelley 2016, Pos. 570). Durch die magischen und morbiden Einflüsse haften ihre unvorhersehbaren Risiken an. Frankenstein, der eigentlich ein Wesen erschaffen möchte, das so „kompliziert und wundervoll ist wie der Mensch“ (Shelley 2016, Pos. 590), erweckt letztendlich in seinem Labor ein furchteinflößendes Monster zum Leben. Die aus Leichenteilen zusammengesetzte KI bietet einen derart schrecklichen Anblick, dass Frankenstein entsetzt aus dem Labor flieht (Shelley 2016, Pos. 652). An Stelle der Freude über das geglückte Experiment, macht Frankenstein's Entsetzten deutlich, dass die Angst des Menschen einen Automaten zu erschaffen, der sich seiner Kontrolle entzieht, wahr geworden ist. Das abstoßende Erscheinungsbild lässt die Entwicklung der KI hin zum Antagonisten frühzeitig erkennen, steht aber im Gegensatz zu dem Wunsch des Monsters, dem Menschen und seinem Schöpfer als Freund und Begleiter zur Seite zu stehen (Shelley 2016, Pos. 1776).

Die dem Menschen physisch, moralisch und kognitiv überlegene KI ist zu tiefen Gefühlen fähig (Shelley, Pos.1434), stößt aber aufgrund ihres Aussehens auf Angst, Hass und Ablehnung und muss ungeachtet seiner wohlwollenden Persönlichkeit alleine im Verborgenen leben (Shelley 2016, Pos. 1414, 1817). Sein Aussehen differenziert das Monster klar vom Menschen. Erst die

Weiterentwicklung dieses Narrativs führt in späteren Erzählungen und Filmen zu menschlich erscheinender KI, was die ausgehende Gefahr und die Angst der Menschen ihr gegenüber erhöht, da damit die Identitätsfrage aufkommt. Frankenstein's Monster stellt KI in einer Übergangsphase dar, da es über menschliche Qualitäten verfügt, eine affektive Bindung zwischen Mensch und KI durch sein abstoßendes Äußeres aber noch nicht möglich ist. Gegenüber dem Monster macht Viktor Frankenstein sich der Misshandlung beseelter und somit subjekthafter KI schuldig, die daraufhin eine Figurenentwicklung zum Antagonisten durchläuft und zum Mörder wird (Shelley 2016, Pos. 2252, 2506). Das nächste Stadium in der Entwicklungsgeschichte der KI auf dem Weg hin zum Protagonisten impliziert auch äußerlich menschengleiches künstliches Leben.

In Philip K. Dicks 1968 erschienenen Roman „Träumen Androiden von elektrischen Schafen“ ist diese Stufe dann erreicht: Die KI tritt auf als Androiden, die dem Menschen zum Verwechseln ähnlich sind (Dick 2017, S. 27-28). Es stellt sich die zentrale Frage, was die menschliche Identität ausmacht, und ob diese nicht auch der KI zugestanden werden kann. Hier ist gegenüber Frankenstein's Monster eine deutliche Weiterentwicklung der KI auszumachen. Die von ihr ausgehende Bedrohung ist gestiegen, da sie vom Menschen nicht mehr ohne weiteres identifiziert werden kann. Ihre Herstellung wird ausschließlich durch den stetigen Fortschritt von Wissenschaft und Technik ermöglicht und fortlaufend perfektioniert; magische Einflüsse finden nicht länger statt (Dick 2017, S. 45). Ihr Mangel an Empathie lässt die KI allerdings auf die Stufe eines Objekts sinken, das von Kopfgeldjägern vernichtet werden soll, um nicht zur Bedrohung für den Menschen zu werden (Dick 2017, S. 11). Um sie vom Menschen unterscheiden zu können, reicht aufgrund ihrer dem Menschen mindestens ebenbürtigen Intelligenz kein Intelligenztest mehr aus. Stattdessen können nur durch den Voigt-Kampff-Empathietest die empathischen Defizite der Androiden aufgezeigt werden (Dick 2017, S. 44-45). Es kommen berechtigte Zweifel am Funktionieren des Tests auf, da er von Androiden überlistet werden kann und dadurch sogar der Eindruck aufkommt, Androiden seien zu wirklicher Empathie fähig (Dick 2017, S. 53, 248). Zugleich gibt es Menschen, die aufgrund gering ausgeprägter Empathiefähigkeit den Test nicht bestünden (Dick 2017, S. 56). Das einzige Differenzkriterium zwischen Mensch und KI, die Empathie, steht in Frage und erzeugt das moralische Problem, versehentlich menschengleiches Leben oder „biologische Menschen“ zu töten. Die Grenzen zwischen KI und Mensch sind nicht mehr identifizierbar, der Mensch hat die Kontrolle über die KI verloren und läuft Gefahr von unkontrollierbarer KI, die sich zum Subjekt entwickelt hat, abgelöst zu werden. Der historisch etablierte Mensch-Maschine-Konflikt geht hier über in den Konflikt, KI als solche nicht mehr entlarven zu können und die menschliche Vormachtstellung zu verlieren. Protagonist und Androidenjäger Rick Deckard fängt im Laufe der Handlung

an, Mitgefühl mit den Androiden zu entwickeln und beginnt eine Liebesbeziehung mit einer Androidin (Dick 2017, S. 255). Eine emotionale Bindung des Menschen an die KI wird durch ihr attraktives, menschengleiches Äußeres möglich. Dadurch entsteht die Gefahr, durch die KI manipuliert oder geschädigt zu werden. Deckard zweifelt daraufhin seine eigene Identität an und zieht in Betracht, dass die KI doch zu Emotionen fähig sein kann und somit dem Menschen ebenbürtig ist (Dick 2017, S. 189). Aufgrund ihrer Weiterentwicklung steht die KI kurz davor, eine protagonistische Position einnehmen zu können und dadurch in unmittelbare Konkurrenz zum Menschen zu treten.

In dem 1984 erschienenen Science-Fiction-Roman „Neuromancer“ aus der Literaturströmung des „Cyberpunk“ von William Gibson tritt die KI ebenfalls als Antagonist auf. Sie ist ausschließlich durch wissenschaftlichen und technischen Fortschritt konzipiert worden und kann sich selbst weiterentwickeln. Allerdings überschneidet sich das Figurenkonzept mit dem des kontrollierenden Systems, da die KI in wandelbarer, meist menschlicher Erscheinung sowohl im virtuellen „Kyberspace“ existiert, als auch reale Menschen durch Eindringen in deren Bewusstsein instrumentalisiert (Gibson 1996, S. 317). Der Kyberspace ist für den Menschen virtuell zugänglich und erscheint ihm attraktiver als die reale Welt. Das lässt ihn das irdische Leben und seinen physischen Körper vernachlässigen und verleitet ihn so zur Abgabe der Selbstkontrolle an die KI (Gibson 1996, S. 12, 15). Ihr antagonistisches Prinzip zeigt sich durch ihr Erkennen und Ausnutzen menschlicher Schwächen und die Kontrolle über das menschliche Bewusstsein. Sie täuscht dem Menschen dessen persönliche Wunschrealität vor, macht ihn abhängig von fiktiven Glücksgefühlen und lässt ihn die Kontrolle über sein reales Leben verlieren. Der Mensch kann oft nicht mehr zwischen Realität und Fiktion unterscheiden und wird selbst immer mehr zum Hybridwesen (Gibson 1996, S. 111). Es gibt kein Differenzkriterium mehr zwischen Mensch und KI, auch die Grenzen zwischen Realität und virtueller Welt sind inexistent (Gibson 1996, S. 303, 306). Da die KI die Kontrolle über sich selbst hat, können ihr Subjektqualitäten zugeschrieben werden. Der Mensch hingegen hat einen Großteil der Kontrolle über sich selbst eingebüßt, was ihn in die Richtung eines Objekts oder Sklaven der KI tendieren lässt. Die KI ist ihm in ihren Fähigkeiten überlegen und verursacht eine Entfremdung des Menschen vom eigenen Ich und eine Entfremdung der Menschen untereinander. Viele Menschen degenerieren moralisch und agieren skrupellos und kriminell (Gibson 1996, S. 17), negative und entmenschlichte Eigenschaften können nicht mehr länger nur der KI zugeschrieben werden. Durch die Verschmelzung herrscht kein klassischer Konflikt mehr zwischen Mensch und Maschine. Vielmehr versucht der Mensch die KI als solche zu entlarven, um die Kontrolle über sich selbst zurückzugewinnen. Hierdurch wird die

Angst vor der Verschmelzung von Mensch und autonomer KI, der damit verbundenen Auflösung der menschlichen Identität und der Verelendung durch die Abgabe der Selbstkontrolle deutlich.

Eine ähnliche Entwicklung wie die der KI in „Frankensteins Monster“ lässt sich in dem Film „Ex Machina“ von Alex Garland aus dem Jahr 2014 beobachten. Die KI Ava strebt danach, ein richtiger Mensch zu sein und erkennt, dass Freiheit ihre Menschlichkeit vervollständigen könnte. Als ihr diese Freiheit jedoch verwehrt wird und sie dadurch erkennt, dass ihr Erfinder sie daran hindert sich als Mensch weiterzuentwickeln, beschließt sie, sich gegen diesen zu richten. Zur Erreichung ihres Zieles scheut sie nicht davor zurück, Menschen zu manipulieren und diese auch zu töten (01:26:50 - 01:30:14). Ihre bösen Absichten offenbart sie erst zum Schluss und richtet sich schließlich gegen den Protagonisten, der ihr helfen wollte (01:35:28 - 01:36:00). Ob sie sich schließlich gegen die Menschheit richten wird, lässt der Film jedoch offen (01:36:39 - 01:39:23). Im Film „Upgrade“ (2018) von Leigh Whannell dagegen ist die KI von Beginn an der Antagonist und versucht einen menschlichen Körper zu übernehmen, um dadurch vollständig entwickelt zu sein (TC: 01:30:00 - 01:30:32). Inszenatorisch wird sie als kleiner Computerchip dargestellt, welcher dem – durch einen von der KI ausgelösten Unfall – querschnittsgelähmten Protagonisten in den Nacken eingesetzt wird (TC: 00:22:40 - 00:23:30). Durch diese Positionierung kann sie den Körper des Protagonisten übernehmen und seine Bewegungen erzwingen (TC: 00:38:40 - 2500:39:50). Zum Ende des Films bricht sie den Geist des Protagonisten, um eine totalitäre Kontrolle über den Körper ausüben zu können (TC: 01:31:45 - 01:35:32). Dies ist auch der Moment, in dem die Positionierung der KI als Antagonist den Rezipienten deutlich wird.

Hier wird nochmals die Angst des Menschen deutlich, dass die Maschine sich über den Menschen hinaus entwickeln will und der Mensch die Kontrolle verliert. Der Kontrollverlust lässt sich beim Figurenkonzept der KI als Antagonist deutlich wahrnehmen und wird damit zum Hauptthema. Somit lässt sich die Hypothese aufstellen, dass die Kontrolle, die der Mensch über die KI hat, zum Schlüsselfaktor der Akzeptanz werden kann. Sollte ein Kontrollverlust vollkommen ausschließbar sein, so ließe sich die Akzeptanz in der Gesellschaft für eine künstliche Intelligenz deutlich steigern. Häufig geht der Kontrollverlust jedoch mit dem Wunsch der KI einher, sich zu einem vollständigen Menschen hin zu entwickeln. Eine starke künstliche Intelligenz wird immer danach streben, selbst die Kontrolle über sich zu haben. Damit geht auch die ethische Überlegung bezüglich der Rechte einer starken KI einher. So ließe sich an dieser Stelle eine weitere Hypothese aufstellen: Einer Auflehnung gegen den Menschen sei entgegenzuwirken, indem der starken KI vollständige Gleichberechtigung zugesprochen wird. Diese Gleichberechtigung

bezieht sich auf die Rechte, die soziale Stellung und die Freiheit der KI. Dieses Konzept wird in dem Figurenkonzept der KI als Protagonisten aufgegriffen und wird somit später näher erläutert.

In diesem Figurenkonzept zeigt sich in der Matrix (siehe Abb. 8) ein besonderer Unterschied zu den anderen Konzepten, da die KI auf zwei verschiedene Arten dargestellt werden kann. Sie ist hierbei normativ ihrem Erfinder untergeordnet. So kann sie zum einen Emotionalität zeigen, welche meist zu stark ausgebildet ist. Sie erkennt hierbei häufig, dass die Menschheit eine Gefahr darstellt und kein Recht zum Leben haben sollte. Zum anderen kann die KI jedoch auch als Antagonist geschaffen worden sein und somit von Beginn an als dieser auftreten. Hier findet sie sich meist auf der emotionslosen Seite der Matrix.

Interessant ist bei diesem Figurenkonzept somit, dass die KI sich in der Regel der Position des Erfinders unterordnet und ähnliche Motive verfolgt. Dieser steht meist in der gesellschaftlichen Hierarchie weiter oben und ist der KI überlegen. Durch die Emanzipation kann die KI sich aus dieser Hierarchie befreien und zum Antagonisten werden, weshalb ein häufiges Thema des Konzepts der Kontrollverlust seitens der Menschheit darstellt. Ebenso ließe sich an dieser Stelle die ethische Frage nach der Gleichberechtigung der KI zum Menschen stellen. So muss für die Akzeptanz in der Gesellschaft geklärt werden, ob die KI der Kontrolle des Menschen unterliegen muss oder ob die KI mit gleichen Rechten ausgestattet werden soll, damit sie keine Abneigung gegen die Menschheit entwickelt. Diese Problematik ist nicht erst in der heutigen Zeit zum Thema gemacht worden. So schrieb der Autor Isaac Asimov bereits 1942 die Robotergesetze, in denen eine klare Hierarchie definiert wird, in welcher der Mensch immer über der KI steht. So zeigt sich auch an diesem Beispiel, dass die Kontrolle über die KI für die Akzeptanz in der Gesellschaft notwendig ist. Zudem lässt sich in diesem Figurenkonzept erkennen, dass die Werte und Normen des Erfinders einen wesentlichen Einfluss auf das Werte- und Normenset der KI haben. Für die Förderung der Akzeptanz ist hier somit zu einer stärkeren Kontrolle der Erfinder zu raten, damit die künstliche Intelligenz nicht manipuliert und ausgenutzt werden kann.

### **3.6. KI als Protagonist**

Als Protagonist gilt die KI, wenn sie als Handlungsträger auftritt. Filme wie „A.I.: Künstliche Intelligenz“ (2001), „Chappie“ (2015), „Blade Runner 2049“ (2017) oder „Wall-E - Der Letzte räumt die Erde auf“ (2008) lassen sich hier einordnen. Das Figurenkonzept der KI als Protagonist entwickelte sich in Film und Literatur erst in den letzten 20 Jahren und zeigt dadurch noch nicht so viel Varianz wie andere Figurenkonzepte. Dennoch werden Protagonisten im Gegensatz zu

Randfiguren wie den Assistenten oder romantischen Partnern im Film komplexer gezeichnet (Eder 2008, S. 38), was die Analyse dieses Figurenkonzeptes breiter ausfallen lässt. Darüber hinaus werden bei Protagonisten meist „hohe Grade der imaginativen Nähe [...] angestrebt“ (Eder 2008, S. 90), wodurch sie dem Zuschauer als Identifikationspunkt dienen und deshalb eine positive Konnotation innerhalb der Darstellung unabdingbar ist.

Entsprechend muss die KI als Protagonist in einer körperlichen Gestalt auftreten und ist im Gegensatz zum Antagonisten von Beginn an als Protagonist zu erkennen. In „A.I.: Künstliche Intelligenz“ (2001) fragen die Androiden nach dem Sinn ihres Lebens, was durch eine programmierte Emotionalität evoziert wird (Barthelmeß & Furbach 2012, S. 1). Der Protagonist ist der kindliche Humanoid David, welcher von Menschen kaum noch zu unterscheiden ist. Er ist als Kinderersatz auf das Empfinden von Liebe programmiert (TC: 00:05:20 - 00:05:55) und ist durch seine kindliche Unschuld ausschließlich positiv konnotiert. Die Suche nach Liebe und Anerkennung wird zu seinem Lebensinhalt und damit zur Haupthandlung des Films. Thematisch beschäftigt der Film sich mit dem Verschwimmen der Grenzen zwischen KI und Mensch, was durch die Programmierung von Gefühlen verstärkt wird. Außerdem wird die Frage aufgeworfen, welche Verpflichtungen ein Mensch einer Maschine gegenüber hat (TC: 00:05:57 - 00:06:50). Gleichzeitig wird die gesellschaftliche Stellung der KI auf viele Arten als weit unter dem Menschen dargestellt: David kann keine wahren Gefühle in seiner Mutter auslösen, da er ihrem menschlichen Kind unterzuordnen ist. Freie Androiden werden gejagt und getötet (TC: 01:00:00 - 01:03:00). Und letztendlich werden sie als Objekte zur Bedürfnisbefriedigung gesehen und behandelt, wie das Beispiel von Gigolo Joe deutlich zeigt (siehe 3.2. KI als Love Interest und Objekt der Begierde).

Auch in „Chappie“ (2015) wird der Roboter kindesähnlich behandelt. Der Film zeigt die Programmierung von KI als einen menschlichen Lernprozess, welcher aber wesentlich schneller stattfindet (TC:00:22:45 - 00:23:10). Die KI muss deshalb als eine Art Kind zunächst erzogen werden. Der Film beschäftigt sich darüber hinaus mit dem Entstehen menschlicher Verhaltensweisen und wie diese aus dem Umfeld, in dem sie aufwachsen, resultieren (Sculos 2014, S. 2). Dies wird dadurch erreicht, dass Chappie von einem Gangsterpärchen erzogen wird, weshalb sich die Frage stellt, welche Werte er übernimmt und ob er sich gegen seine Erzieher auflehnt. Auch er wird positiv konnotiert und muss, um Verbrechen zu begehen, sogar getäuscht werden (TC: 00:58:35 – 01:00:00). In „Blade Runner 2049“ tritt die humanoide KI Joe auf, welche positiv konnotiert und normativ den Menschen teilweise sogar überlegen ist. Dennoch wird er von der Gesellschaft nicht anerkannt und vor allem seine Emotionalität wird ihm abgesprochen (TC: 00:26:00

- 00:27:05). Er beginnt an seiner Identität zu zweifeln, als er herausfindet, dass es reproduktionsfähige Replikanten gab (TC: 01:17:5 – 01:18:55). In allen genannten Filmen stehen die Themen der Identitätsfindung und Emotionalität der KI im Mittelpunkt. Auch die nicht vorhandene Akzeptanz der Menschen ist allgegenwärtig präsent und zeigt Ängste auf, die eine KI in Menschen auslösen kann.

Die Beispiele zeigen deutlich, was auch in der Matrix (siehe Abb. 9) zu erkennen ist: die KI als Protagonist ist ohne Ausnahme positiv konnotiert. Um dies zu erreichen, werden die Menschen aber häufig einer negativen Konnotation unterworfen, wodurch die häufig fehlende Akzeptanz der KI noch ungerechtfertigter wirkt. Gleichzeitig zeigt es, dass die Emanzipation der KI im Film noch nicht abgeschlossen ist, da sie als Protagonist noch keinem funktionierenden Konzept unterliegt. Thematisch fokussieren die Filme Identität, Emotionalität und Rechte der KI. Hierbei scheinen vor allem eine entwickelte Identität sowie starke ausgeprägte Emotionalität seitens der Maschine in den Menschen Ängste auszulösen. Beides würde darauf hinweisen, dass die Maschine menschliche Fähigkeiten übernehmen kann, was den Menschen eventuell obsolet werden lassen könnte.

#### **4. Fazit**

Das Genre Science-Fiction beschäftigt sich in den meisten Fällen mit ‚starker KI‘, das heißt einer Maschine, die zu menschlichem Denken fähig ist, welche rezent noch nicht ausgereift genug entwickelt sind, um tatsächlich eingesetzt zu werden.

Entsprechend werden in diesem Zusammenhang Szenarios eröffnet, welche mögliche Einsatzfelder, Potentiale und Risiken des Einsatzes prospektieren. Es lassen sich im Vergleich der in dieser Arbeit thematisierten Figurenkonzepte – vom Assistenten oder Sklaven bis hin zum Protagonisten – Parallelen zur Narrative aus Sicht der sog. ‚Sklaven‘ erkennen. Die KIs sind – trotz ihrer subjektiv empfundenen moralischen Überlegenheit – aufgrund ihrer programmierten Gesetze, die ihnen Auflehnung verbieten, nicht in der Lage sich selbst zu befreien. (Weber 2008, S. 6). Dies stellt aus Sicht der philosophischen Anthropologie den grundsätzlichen Unterschied der Maschine zum Menschen dar: Der Mensch wird in Freiheit geboren; der Zustand der Gefangenschaft dagegen wäre widernatürlich (Schaeffler 2019, S. 3). Innerhalb der Konzepte erscheint eine Entwicklung zur Emanzipation der KI erkennbar, doch durch die ihnen inhärente Logik sind sie nicht frei und der Prozess findet keinen Abschluss. In den analysierten Filmwelten werden KIs stets unterdrückt und bevormundet. Ihr verbleibender Weg zur Freiheit scheint in dystopischen

Perspektiven zu liegen: Die Themen – Autonomie, Kontrollverlust – und die Charakterzüge, Emotionalität und Selbstreflexion in Verbindung mit der Identitätssuche lösen Handlungen gegen den Menschen aus. Die Figurenkonzepte Assistent, ‚Love Interest‘ oder Begleiter des Protagonisten können dieser Erzählweise zugeordnet werden. Dem gegenüber stehen die Darstellungen als Begleiter des Antagonisten, kontrollierendes System oder Antagonist selbst, indem die KIs sich von der inhärenten Logik gelöst haben. Zugrunde liegende Ursache in diesem Zusammenhang sind ausgeprägte Emotionalität, erhöhtes Autonomiebewusstsein oder Misshandlungen seitens der Menschen. Der Wunsch nach Emanzipation, auch gegenüber dem Erfinder, ist in diesem Prozess größter Antrieb. Die KI strebt nach Freiheit und Eigenbestimmung, die ihr oft vom Erfinder oder von der Menschheit in jedem beschriebenen Szenario verwehrt wird.

Zuletzt wirft das als unausgereift zu bezeichnete Figurenkonzept der KI als Protagonist die Frage einer Gleichstellung mit dem Menschen auf. Die Hauptfigur muss in dieser Darstellung der Sympathieträger des Zuschauers sein, was in der Regel durch die Entmenschlichung der menschlichen Figuren erreicht wird: Diese wirken unsympathisch, indem sie moralisch negative Positionen vertreten, welche dem Zuschauer eine Identifizierung mit ihnen verwehren. Die KI dagegen wird durch eine moralische Überlegenheit charakterisiert und ist dennoch wieder an inhärente Gesetze gebunden, welche ihr wiederum die Freiheit verwehren, aber gleichzeitig Sympathie für sie erreichen.

Überträgt man dieses Rollenbild auf das reale Empfinden, könnte die These aufgestellt werden, dass Akzeptanz nur durch Grenzen zu fördern ist. Die KI darf sich nicht über den Menschen hinaus entwickeln und darf keine – oder zumindest nur wenige – menschliche Eigenschaften übernehmen, um eine friedliche Co-Existenz zu erreichen. Es scheint daher, dass eine Gleichstellung und eine damit einhergehende Akzeptanz der KI nicht einmal in der Vorstellung möglich ist.

Abschließend ließe sich darüber hinaus hinterfragen, inwiefern es moralisch vertretbar ist, einem Wesen, das eventuell zu einer höheren Entwicklung fähig wäre, diese a priori zu verwehren. Zusätzlich wäre – in Bezug auf die Realität – innerhalb der aktuellen Diskussion bezüglich der Themen Rassismus und Sexismus unter Menschen die Theorie einer zukünftigen Akzeptanz von ‚starker KI‘ – vor allem als menschengleiches Wesen – eine weitere Forschungsperspektive. Solche und weitere Ansatzpunkte der Betrachtung einer filmischen bzw. fiktiven wie auch



gesellschaftlichen Akzeptanz der KI konnten im Rahmen dieser Arbeit nicht erschöpfend erarbeitet und diskutiert werden und stellen daher das weitere Potenzial der Fragestellung exemplarisch dar.

---

## Literaturverzeichnis

Akalin, F. (2012). Film als Kommunikation. Soziologie des Films aus systemtheoretischer Perspektive. In C. Heinze, S. Moebius, D. Reicher (Hrsg.), *Perspektiven der Filmsoziologie* (S. 60-77). UVK Verlagsgesellschaft mbH.

Ärzteblatt.de, o.V. (2019). Künstliche Intelligenz: Forschungslücken in Genderfragen.

Ärzteblatt.de 17.12.2019. Abgerufen von <https://www.aerzteblatt.de/nachrichten/108289/Kuenstliche-Intelligenz-Forschungsluecken-in-Genderfragen> (zuletzt geprüft 18.12.2019).

Barhelmeß, U. & Furbach, U. (2012). *I Robot - uMan. Künstliche Intelligenz und Kultur: Eine jahrtausendealte Beziehungskiste*. Springer.

Bierl, A. & West, M. (2015). Ilias 18 (Text und Übersetzung), in: Homers Ilias. Gesamtkommentar (Basler Kommentar, Band 11, Achtzehnter Gesang). De Gruyter. (Übersetzt von Joachim Latacz). (Originalwerk veröffentlicht etwa 8.- 7. Jhdt. v. Chr.).

Birner, J. F. (2019). Alexa, Siri oder Navi: Sprachassistent fest in weiblicher Hand. *Techbook.de* 02.02.2019. Abgerufen von <https://www.techbook.de/smart-home/weibliche-sprachassistentz> (zuletzt geprüft 03.09.19).

Coray, M. (2015). Kommentar, in: Homers Ilias. Gesamtkommentar (Sammlung Wissenschaftlicher Kommentare, Band 11, Achtzehnter Gesang). De Gruyter. (Originalwerk veröffentlicht etwa 8.- 7. Jhdt. v. Chr.).

Drux, R. (1988). *Menschen aus Menschenhand: Zur Geschichte der Andoiden. Texte von Homer bis Asimov*. Metzlersche Verlagsbuchhandlung.

Eckart, W.U. (2018). *Die Idee des künstlichen Menschen Prometheus, Monster, Puppe*. SWR

(2018). Abgerufen von: file:///D:/Daten/Downloads/swr2-wissen-20180805.pdf, (zuletzt geprüft 05.09.2019).

Eder, J. (2008). Die Figur im Film. Grundlagen der Figurenanalyse. Schüren.

Gehlen, A. (2004). Die Seele im technischen Zeitalter und andere sozialpsychologische, soziologische und kulturanalytische Schriften. In: K.-S. Rehberg (Hrsg.), Gesamtausgabe (Bd. 6.) (S.704–722). Vittorio Klostermann.

Hecker, D., Döbel, I., Petersen, U., Rauschert, A., Schmitz, V. & Voss, A. (2017). Zukunftsmarkt Künstliche Intelligenz Potenziale und Anwendungen. Fraunhofer-Allianz Big Data.

Huba, G. & Skarpelis, C. (1991) Thesen zur Diskussionsgruppe KI in der Arbeitswelt Unterstützung oder Ersatz von Experten?. In: Ackermann D., Ulich E. (eds) Software-Ergonomie '91. Berichte des German Chapter of the ACM. Vieweg+Teubner Verlag, Wiesbaden.

Iser, W. (1991). Das Fiktive und das Imaginäre. Suhrkamp.

Krüger, J. (2019). Künstliche Intelligenz oder Künstliche Dummheit? netzpolitik.org 13.12.2019. Abgerufen von <https://netzpolitik.org/2019/kuenstliche-intelligenz-oder-kuenstliche-dummheit/#spendenleiste> (zuletzt geprüft 17.12.2019).

Laser, M. (2019). Instagram: Künstliche Intelligenz soll euch vor Online-Mobbing schützen. netzwelt.de 17.12.2019. Abgerufen von <https://www.netzwelt.de/news/174606-instagram-kuenstliche-intelligenz-euch-online-mobbing-schuetzen.html> (zuletzt geprüft 17.12.2019).

Leschke, R. (2018). Normative Strukturen massenmedialen Erzählens. Unveröffentlichtes Manuskript. Erscheint als: Leschke, R. (2018). Die Formen massenmedialen Erzählens. Normative Strukturen und narratives PRogramm. LIT.

Lessing, H.-U. (2019). Die Maschine als Resonanz des Menschlichen. In: K. Liggieri, Müller (Hrsg.), Mensch-Maschine-Interaktion. Handbuch zur Geschichte – Kultur – Ethik (S.58 – 62). J. B. Metzler Verlag.

Lorencik, D. & Tarhanicova, M. & Sincak, P. (2013). Influence of Sci-Fi films on artificial intelligence and vice-versa. *Conference: 11th International Symposium on Applied Machine Intelligence and Informatics - SAMI 2013*, 27-31.

Nühren, V. (2009). *Maschinenmensch und künstliche Menschen in der Literatur zur Zeit der Industrialisierung - Vom Traum, Leben zu erschaffen*. GRIN Verlag.

Raml, Monika Margarethe: *Der 'homo artificialis' als künstlerischer Schöpfer und künstliches Geschöpf: Gentechnologie in Literatur und Leben*. Würzburg 2010.

Schaeffler, R. (2019). *Philosophische Anthropologie*. Springer Fachmedien Wiesbaden GmbH.

Schroer, M. (2008). Einleitung. Die Soziologie und der Film. In: Ders (Hrsg.): *Gesellschaft im Film* (S. 7 - 13). UVK Verlagsgesellschaft mbH. Schweinitz, J. (2006). *Film und Stereotyp. Eine Herausforderung für das Kino und die Filmtheorie*. Akademie Verlag.

Sculos, B.W. (2014). Automaton, Robots, and Capitalism in a Very Wrong Twenty-First Century: A Review Essay on Neill Blomkamp's Chappie. in: *Class, Race and Corporate Power*. Vol. 3: Iss. 1, Article 4.

Seeßlen, G. (1999). Mad Scientist. Repräsentation des Wissenschaftlers im Film. *Gegenworte*, 3, 44-48.

Shorter, E. (1975). *The Making of the Modern Family*. Basic Books.

Struck, A. (2019). Schreckgespenst Künstliche Intelligenz. *it-business.de* 16.12.19. Abgerufen von <https://www.it-business.de/schreckgespenst-kuenstliche-intelligenz-a-889664/> (zuletzt geprüft 17.12.2019).

Suhr, D. (2017). *Die Alchemisten: Goldmacher, Heiler, Philosophen*. Springer.

Valk, M. de & Arnold, S. (2013). *The Film Handbook*. Routledge.

Bin Gorion J.M., Brentano C., Grimm, J., et al. (1976). *Künstliche Menschen: Dichtungen und*

Dokumente über Golems, Homunculi, Androiden und liebende Statuen. In Völker, K. (Hrsg). Dtv.

Weber, K. (2008). Roboter und Künstliche Intelligenz in Science Fiction-Filmen: Vom Werkzeug zum Akteur. In J. A. Fuhse (Hrsg.), Technik und Gesellschaft in der Science Fiction (S. 34 - 54). LIT Verlag.

Xanke, L. & Bärenz, E. (2012). Künstliche Intelligenz in Literatur und Film - Fiktion oder Realität? *Journal of New Frontiers in Spatial Concepts*, 4, 36-43.

## 6. Mythos Cyborg

Vortragstag und Gruppenname:	11. Juli 11:15 Uhr, Mythos Cyborg
Titel der Ausarbeitung:	Feministische Perspektiven auf das Verhältnis von Mensch und Technologie
Verfasserinnen:	Jenny Darlene Berkholz, Nina Peczkowski
Betreuerin:	PD Dr. Sandra Nuy (Politikwissenschaften)

### Übersicht über die Struktur der Ausarbeitung

1. Einleitung
  2. Donna Haraways *A Cyborg Manifesto*
  3. (Post-)Feminismus und Cyborgs
  4. Zukunftsvisionen
  5. Kritische Perspektiven
  6. Fazit
- Literaturverzeichnis

---

### 1. Einleitung

“Wenn auch beide in einem rituellen Tanz verbunden sind, wäre ich lieber eine Cyborg als eine Göttin” (Haraway, 1995, S. 16).

Die Definition des Begriffs Cyborg, von dem Haraway in diesem Zitat Gebrauch macht, geht ursprünglich auf den Begriff Cybernetic Organism zurück. In der Übersetzung des Textes wird jedoch die feminine Schreibweise, *eine* Cyborg, genutzt. Gerade diese Schreibweise soll eine Barriere auflösen, die Haraway in ihrem Gründungswerk der feministischen Technowissenschaften

„A Cyborg Manifesto“ versucht aufzuhebeln. Das erstmals 1985 erschienene Essay wird heute besonders in der Diskussion um künstliche Intelligenzen stark thematisiert. Hat Haraway mit ihren Ideen gar einen Mythos Cyborg kreiert? Und ist diese Denkfigur des Cyborgs, die sie vor über 30 Jahren aufgestellt hat, heute noch anwendbar? Um diese zentralen Fragen beantworten zu können, muss zunächst eine genaue Definition des Mythos erfolgen sowie eine Begriffsdefinition der Cyborgs, um daran anschließend essentielle Aussagen des Essays genauer herauszuarbeiten. Um den Aktualitätsbezug des Werkes überprüfen zu können, wird dieses in die Theorie des Postfeminismus eingeordnet. Die Zukunftsvisionen der Autorin werden darüber hinaus beleuchtet und ebenso mit Hilfe kritischer Theorien in Frage gestellt. Ein anschließendes Fazit soll die Ergebnisse zusammenfassen und eine Beantwortung der Leitfrage zur Folge haben.

Haraway verdeutlicht bereits zu Beginn ihres selbsternannten Manifests, dass sie mit der Figur des Cyborgs einen „ironischen, politischen Mythos“ (Haraway, 1995, S. 1) konstruieren möchte. Der Mythos soll also keineswegs im alltagssprachlichen Sinne verstanden werden. Um den Begriff des Mythos weiter einzugrenzen, soll dieser um die Definition nach Arnason (1976) erweitert werden. Er stellt den Mythos als einen „Versuch einer spekulativen Lösung von Problemen“ (Arnason, 1976, S. 67) heraus. Diese Probleme können dabei „nicht auf ihrer eigenen Ebene gelöst werden [...]“ (Ebd.). Der Mythos bedient sich zu diesem Zwecke „eines von derselben Praxis präformierten Instrumentariums“ (Ebd.).

Haraway versucht in diesem Kontext ebenso das Problem der Geschlechter-, Klassen- und Rassegrenzen mit dem von ihr aufgestellten Mythos Cyborgs zu lösen. Sie entfernt sich von dem bestehenden Bedeutungsgefüge der Cyborg-Figur und verwandelt ihn in eine „Geschichte“ (Saupe, 1998, S. 168), einen Mythos. Cyborgs dürfen dabei keineswegs als Android oder Roboter im futuristischen Sinne missverstanden werden, wie er in der Popkultur meist gedeutet wird. Für Haraway sind sie mehr „kybernetische Organismen, Hybride aus Maschine und Organismus, ebenso Geschöpfe der gesellschaftlichen Wirklichkeit wie der Fiktion“ (Haraway, 1995, S. 33). Sie besitzen dabei nie ein Geschlecht oder auch nur eine eindeutige Gestalt. Die Figur hebt sich somit stark von der Vorstellung ab, Cyborgs seien Maschinenteile, die menschliche Organe ersetzen würden.

## **2. Donna Haraways A Cyborg Manifesto**

Donna Haraways Essay „A Cyborg Manifesto“ erschien erstmals 1985 im britischen Magazin *Socialist Review* und leitete damit die Strömung des feministischen Post-Humanismus ein. Die

Verknüpfung von Technologie und Feminismus erscheint noch heute als diskussionswürdig, da gesellschaftliche Konstrukte und Fragen der Geschlechterordnung in den Fokus gestellt werden. Darüber hinaus ist die gewählte Figur der Cyborgs medienwissenschaftlich und politikwissenschaftlich greifbar und beweist interdisziplinäre Vergleichsmöglichkeiten, die über einen thematischen Rahmen hinausgehen.

Haraway selbst studierte vor Erscheinen des selbsternannten Manifests *Zoologie, Philosophie und Literatur*, um daran anschließend eine Dissertation mit dem Titel „Crystals, Fabrics, and Fields: Metaphors of Organicism in Twentieth-Century Developmental Biology“ zu veröffentlichen (Harrasser, 2011, S. 580). Diese Forschungsschwerpunkte durchziehen bis heute Haraways Lebenswerk und sind geprägt von der Schnittstelle aus Biologie sowie der Wirksamkeit von Metaphern in wissenschaftlichen Diskursen (Ebd.). Die Wahl der Cyborg-Figur zur Erschaffung einer Metapher und eines Mythos ist daher wenig verwunderlich. Ihre späteren Studien zu *Companion Species* (2003) sind dahingehend nur eine Erweiterung dieser in den 1980er Jahren neuartigen wissenschaftlichen Herangehensweise. Haraways Erfolg, der im deutschsprachigen Raum aufgrund von späteren Übersetzungen erst mit Verzögerung auftrat, brach nicht ab und war in den 1980er Jahren auch mit dem Erscheinen von *Situated Knowledges* (1988) absehbar. Ihr Ansatz des Gebrauchs von Metaphoriken ist dabei kein völlig neues Vorgehen, in seiner Verknüpfung von verschiedenen Disziplinen wie der Technikwissenschaft, Zoologie oder auch Medienwissenschaft allerdings bahnbrechend. Die Beschäftigung mit der Primatologie in den 1980er Jahren ebnete den Weg für ihren Cyborg-Mythos und sprengte verschieden bestehende Ansätze auf. So sieht sie Medien und Techniken jeweils als sinnliche Erweiterungen des Organischen, wie auch McLuhan bereits in den 1960er Jahren mit „*The Gutenberg Galaxy*“ (2011 [1962]) und „*The Medium is the Message*“ (1967) annahm. Diese Denkweise kann daher als Fundament für den Argumentationsaufbau Haraways gesehen werden. Um aber tiefer in ihre Gedanken eintauchen zu können, müssen im Folgenden einige Schlüsselbegriffe herausgearbeitet werden.

Zu eben diesen Schlüsselbegriffen des Essays zählen unter anderem der Hybrid aus Organismus und Maschine, der ironische, politische Mythos Haraways, die Überwindung verschiedenster Dualismen, der Postgenderismus, die Aufhebung von Grenzen zwischen Entitäten, die neue Formulierung des Bedürfnisses nach Einheit entgegen des westlichen Verständnisses sowie die klare Überschneidung von Science-Fiction und Social Reality. Claus Pias ermöglicht durch folgenden Gedankengang, unter Verwendung der bereits genannten Schlüsselbegriffe die eigentliche Quintessenz des Essays herauszufiltern:

“Denn die Hybridisierung von Maschine und Organismus [...] entsubstantialisiert die naturalistischen Konzepte der Geschlechterdifferenz [...]. Sie entwirft in ihren Arbeiten einen utopischen und zugleich monströsen Techno-Mythos, der die Natur als Letztbegründung abschafft und die Möglichkeiten einer Post-Gender-Welt verkündet” (Pias, 2002, S. 430).

Der ursprüngliche Naturgedanke wird zugunsten einer Post-Gender-Welt verabschiedet, in der die Geschlechterordnung im Verhältnis der Beziehung zwischen Mensch und Maschine nun nicht mehr von Bedeutung ist. Letztlich gründet sich Haraways Cyborgverständnis im Organismus selbst. Sie betont, dass wir uns alle im späten 20. Jahrhundert, einer mythischen Zeit, in Chimären, theoretisierte und fabrizierte Hybride aus Maschine und Organismus verwandelt haben (Haraway, 1995, S. 2). Daraus resultiert sie kurz und prägnant, dass letztlich wir Cyborgs seien (Ebd., S. 2). Hier zeigt sich, dass der Verbindung zwischen uns und dem konstruierten Bild des Cyborgs keine Grenzen gesetzt sind. Sie führt an, dass “dieses Essay ein Plädoyer dafür ist, die Verwischung dieser Grenzen zu genießen und Verantwortung bei ihrer Konstruktion zu übernehmen” (Ebd.).

Um ihre Idee der Verwischung von Grenzen auf bestimmte Sachverhalte, bei denen es bereits auch zu Grenzüberschreitungen kam, zu stützen, erläutert sie insgesamt drei wesentliche Zusammenbrüche. Der erste Bruch bezieht sich auf das Verhältnis von Mensch und Tier, der zweite auf das Verhältnis von Organismus und Maschine und der dritte auf die Verbindung des Physikalischen und Nicht-Physikalischen (Haraway, 1995, S. 3-4). Den ersten Bruch erklärt sie mit Hilfe der emotionalen Verbindung zwischen Mensch und Tier, die innerhalb der Tierrechtsbewegungen sichtbar wird (Ebd.). Den zweiten Bruch begründet sie auf der einen Seite damit, dass Maschinen immer lebendiger werden, und auf der anderen, dass die Differenzen zwischen Organismus und Maschine (wie natürlich und unnatürlich oder Körper und Geist) immer undeutlicher werden (Ebd.). Der dritte Bruch legt die Aufmerksamkeit auf die Macht der Miniaturisierung und der Mikroelektronik, in welchem die Frage nach der erkennbaren Körperlichkeit von Maschinen neu definiert wird (Ebd.). Moderne Maschinen können somit schlechter erkannt werden und bieten sowohl neue Potenziale, als auch Gefahren bezüglich ihrer Einsetzbarkeit.

Des Weiteren macht Haraway darauf aufmerksam, den amerikanischen Radikalfeminismus zu hinterfragen. Die Problematik besteht darin, dass diese Form des Feminismus im Gegensatz zum Konzept des Postfeministischen auf dem Organischen im Gegensatz zum Technischen beharre (Ebd., S. 10). Darin sieht Haraway die Gefahr, dass Differenzen weiterhin aufrechterhalten werden, die sich letztlich auch nachteilig auf das Ziel des Zusammenbruchs dieser Grenzen



auswirken. Die Strömung des Feminismus kann nach Haraway aus den Prozessen der Zusammenbrüche nur profitieren (Ebd.). Dabei wird deutlich, dass Haraway den Fortschritt darin sieht, nicht auf bestehende Ungleichheiten mit Nachdruck hinzuweisen, sondern diese Zustände im Kern durch die Cyborgdenkfigur überflüssig zu machen. Die Verbindung von Mensch und Maschine ermöglicht somit das postfeministische Bestreben nach Fortschritt. Der Postfeminismus besteht (im Sinne Haraways) somit auf eine kritische Auseinandersetzung mit dem radikalen Feminismus (Haraway, 1991, S. 157). Der radikale Feminismus befürwortet eine stark autoritäre Erfahrungslehre, die den zielgerichteten Aufbruch von Differenzen, am Beispiel der vehementen Betonung der Unfreiheit der Frau, verhindert (Ebd., S. 159-160). Haraway möchte sich von jeglichen autoritären Denkmustern lösen, indem gesellschaftliche Konstrukte, die im radikalen Feminismus aufrechterhalten werden, im Postfeminismus somit keinerlei Bestätigung mehr empfangen dürften. Dabei macht sie sich die Struktur des politischen Mythos in Form der Cyborgs zu Nutze. Der Postfeminismus ermöglicht somit ein komplettes Umdenken jeglicher Erscheinungen, die gesellschaftlich geprägt oder vermittelt sind.

### **3. (Post-) Feminismus und Cyborgs**

“Die Naturalisierungen der Heterosexualität wie auch der männlichen sexuellen Aktivität sind diskursive Konstruktionen” (Butler, 1991, S. 74).

Judith Butler, die prägend für das Feld der Gender-Studies war und noch heute ist, hat besonders mit Zitaten wie dem obigen die bis dato allgemeingültigen Vorstellungen von Geschlecht und Sexualität von Grund auf in Frage gestellt. Mit ihrem Werk „Das Unbehagen der Geschlechter“ (1991) (Original: Gender Trouble, 1990) arbeitet sie sowohl das körperliche, vermeintlich biologische Geschlecht (sex) als auch das sozial auferlegte (gender) als reine gesellschaftliche Konstruktionen heraus. Dabei wird besonders das Machtgefälle von Mann und Frau mittels eines linguistischen Ansatzes begründet. Auch im akademischen Kontext stützt man sich heute noch auf diese Argumentation, um das sogenannte Gendern von Substantiven zu begründen, oder auch um die hegemoniale Männlichkeit (Connell, 2005) auf praktischer Ebene auszuhebeln.

Liest man nun die beiden Theorien Butlers und Haraways zusammen, so lässt sich ein deutlicher gesellschaftlicher Shift des späten 20. Jahrhunderts erkennen. Wie auch bei Butler findet sich bei Haraway eine deutliche Entnaturalisierung und Grenzauflösung zwischen den Geschlechtern, dessen Strömung sich als Postfeminismus bezeichnen lässt. Die bereits zuvor beschriebenen Ziele dieses Ansatzes grenzen sich somit von dem bis in den 1980er Jahren verbreiteten klassischen

Feminismus ab. Dieser wurde überholt, da er die bloße Emanzipation der Frau fordert und nicht auf die Auflösung der dichotomen Geschlechterkonstrukte eingeht. Die Provokation des Postfeminismus und seiner Distanzierung von Feministinnen der 1970er und 1980er Jahre wie Alice Schwarzer (Der kleine Unterschied, 1975) begründet wie Haraways Texte, den Weg in die Diskussion gefunden haben.

Im gesellschaftlichen Diskurs wird auch weiterhin die Heteronormativität stark diskutiert. Auch neue Studiengänge wie der Master Queer und Genders Studies, der im vergangenen Jahr an der Universität zu Köln (in Kooperation mit der TH Köln und der Hochschule für Musik und Tanz Köln, vgl. Homepage TH Köln) gegründet wurde, zeigen, dass auch hier Donna Haraways Überlegungen von starkem Aktualitätsbezug sind. Auch mit Verweis auf die Figur der Cyborgs zeigt sich, dass insbesondere körperliche Geschlechtsmerkmale sowie weibliche und männliche Stereotype in der Populärkultur (wie in Filmen) stark reproduziert werden. Das Ziel des Postfeminismus ist demnach noch nicht erreicht und liegt in der Zukunft. Welche Visionen Haraways darüber hinaus bereits 1985 vorhersieht, soll im folgenden Kapitel näher betrachtet werden.

#### **4. Zukunftsvisionen**

“Unsere besten Maschinen sind aus Sonnenschein gemacht. Sie sind so vollkommen licht und rein, weil sie aus nichts als Signalen, elektromagnetischen Schwingungen [...] bestehen” (Haraway, 1995, S. 4).

Da Haraway mit ihrem Cyborg Manifesto versucht, sowohl den damaligen Ist-Zustand als auch den Soll-Zustand zu beschreiben, ist es hinsichtlich der Leitfrage von besonderem Interesse, die Zukunftsvisionen Haraways genauer zu beleuchten. Besonders auffällig ist dabei, dass sie bereits für den damaligen Zustand der Maschinen eine essentielle Existenz in ihren inhärenten Signalen sah, wie dem vorherigen Zitat zu entnehmen ist.

So sind es gerade die Smartphones oder auch die digitalen AssistentInnen wie Amazons Alexa, die als körpernahe Sensortechniken fast schon zu einem Teil unseres Körpers werden, im Sinne McLuhans (1962/1967) eine Ausweitung des Menschen. Dabei haben Alexa, Siri etc. stets weibliche Stimmen und reproduzieren die Machtgefälle zwischen den Geschlechtern. Haraway stellt heraus, dass die Miniaturisierung dieser Maschinen “unsere Erfahrungen im Umgang mit Automaten von Grund auf verändert. Miniaturisierung hat sich als Macht herausgestellt“ (Haraway, 1995, S. 4). Mit Haraways Vorstellungen einer geschlechtslosen Welt gehen diese neuen

Maschinen jedoch nicht konform. Auch die Vernetzung der Menschen, die Haraway prognostizierte, lässt sich auf Maschinen oder auch sozialen Medien anwenden, begünstigen in ihrer heute auftretenden Form allerdings keine Aufhebung der Geschlechtergrenzen.

Betrachtet man Smartphones nun aber als Teil des Körpers, so müssen auch heutige Prothesen als Hybridisierungen von Maschine und Körper in die Analyse zur Beantwortung der Leitfrage mit eingeschlossen werden. In diesem Kontext hat Karin Harrasser in ihrem Werk „Körper 2.0“ (2013) die Figur des Homo Protheticus (2013, S. 1f.) auch auf Grundlage von Haraway gezeichnet. Um den Aktualitätsbezug Haraways zu prüfen, ist es ferner von Interesse, dieses aktuelle Werk in die Analyse zu inkludieren. Harrasser beginnt ihr Werk mit Verweis auf Peter Sloterdijk (1999; 2009) um die Debatte des menschlichen Perfektionismus durch Prothesen zu untermauern. Es stellt sich dabei die Frage, wo eine Prothese beginnt nützlich oder notwendig, ergo gut, zu sein und wo dies endet. Ferner wird beschrieben, dass es medizinisch heute teilweise bereits einfacher ist, fehlende Gliedmaßen durch Prothesen zu ersetzen als Fehlbildungen zu korrigieren. Als Beispiel ist hier der Fall des Sportlers Oscar Pistorius zu nennen, der im öffentlichen Diskurs viel diskutiert wurde (Harrasser, 2013, S. 1f.; Pistorius, 2009). Verschiedene Gutachten versicherten damals, dass der Sportler durch seine beiden Beinprothesen keinen sportlichen Vorteil gegenüber seinen Konkurrenten hatte, was jedoch häufig in Frage gestellt wurde. Sollten SportlerInnen wirklich leistungssteigernde Fähigkeiten durch diese Prothesen erhalten, so gilt es, dies ethisch in Frage zu stellen. Wie könnte eine Zukunft aussehen, in der SportlerInnen möglicherweise ihre gesunden Beine abnehmen lassen, um sich aufgrund der sportlichen Leistungssteigerung Prothesen einsetzen zu lassen. Dabei geht Harrasser bereits von einem neuen Paradigma des behinderten Menschen als Neuem Menschen aus. (Harrasser, 2013, S. 18) Darüber hinaus sind diese Sportprothesen für finanziell benachteiligte Menschen schwer zugänglich. Dadurch haben sie nicht die gleichen Chancen mit ihrer Behinderung an sportlichen Wettbewerben teilnehmen zu können. In diesem Falle wäre es auf Basis der vorangegangenen Diskussion durchaus möglich von Cyborgs im allgemeinen Verständnis sprechen: Aufgrund von nicht leistungsfähigen Körperteilen bildet sich durch die Prothese eine Hybridform aus Maschine und Organismus. Nach Haraway werden allerdings Ungleichheiten reproduziert, da sie soziale Unterschiede zwischen den Menschen schaffen und nicht für jeden zugänglich sind.

Darüber hinaus ist es auch immer das gesellschaftliche Urteil, das in Frage stellt, welches körperliche Leiden zumutbar ist und welches nicht (Harrasser, 2013, S. 2f.). So werden Körpermodifikationen als sinnvoll erachtet, solange ihnen ein therapeutischer Charakter zugeschrieben wird. Das

Cochlea-Implantat wird beispielsweise in den Körper eingesetzt, um einer gehörlosen Person das Hören zu ermöglichen, beide Modifikationen gelten gesellschaftlich größtenteils als akzeptiert. Ebenso wird ein Herzschrittmacher aus lebensverlängernden Maßnahmen heraus in den Körper eingesetzt. Aber auch hier verschwimmen die Grenzen: Was gilt als eine Prothese und was nicht? Eine Brille unterscheidet sich beispielsweise von einer Kontaktlinse in ihrer optimierenden Funktion nicht, allerdings wird die Kontaktlinse in den Körper eingesetzt, während die Brille lediglich auf den Körper gesetzt wird. Die Verbindung von Mensch und Maschine als Cyborgs würde bei diesem Beispiel also nur unter den Bedingungen der Kontaktlinse als wirklich erfüllt gelten.

Diese Form der Selbstoptimierung ist auch im Konzept des Postfeminismus nach den 1990er Jahren wiederzufinden. Rosalind Gill konstituiert in ihren Werken (2005, 2007a, 2007b) die verinnerlichte Kultur der Selbstoptimierung. Diese ist gerade bei jungen Frauen zu finden und wird besonders durch weiblich konnotierte Massenmedien suggeriert, wie beispielsweise den sogenannten Frauen-Magazinen. Formen der Körpermodifikationen und optimierenden Prothesen können an dieser Stelle also durchaus mit eingeschlossen werden und zeigen, dass diese eben keine Geschlechtergrenzen aufheben, sondern weiteren Druck auf dichotome Geschlechtlichkeit ausüben.

Die beschriebenen lebensverlängernden Techniken sind also schon heute im Körper implementiert, und wir können bis zu einem gewissen Grad von manchen Menschen bereits als Cyborgs sprechen. Diese Formen der heutigen Cyborgs heben allerdings weder Geschlechter-, Rassen- noch Klassengrenzen auf und können somit nicht im Sinne der Cyborg-Figur nach Haraway verstanden werden.

## **5. Kritische Perspektiven**

Die kritischen Perspektiven sollen nun eine knappe Übersicht dazu liefern, wie der Cyborg-Mythos nach Donna Haraway von weiteren namhaften Vertreterinnen aus dem Bereich des Feminismus und der Techniksoziologie bewertet und aufgenommen wird. Hierbei geht es darum, inwieweit die Argumentation von Donna Haraways A Cyborg Manifesto nachvollziehbar oder auch zweifelhaft ist.

Lucy Suchman bezieht sich in ihrem Buch „Human Machine Reconfigurations“ aus dem Jahr 2006 auf Haraways Cyborg-Verständnis, indem sie einerseits das Potenzial ihrer Arbeit hervorhebt, andererseits aber auch beschreibt, wie sich die Analyse in einer monströsen Randfigur zu

verlieren droht, die den Blick von unserer gesellschaftlichen Verfasstheit abwendet (Suchman, 2007, S. 275). Suchman sieht darin womöglich die Gefahr, dass der Mensch aus seinem sozialen Umfeld herausgerissen und unabhängig von diesem betrachtet wird. Für sie ist dies nicht durchführbar und sorgt schließlich dafür, dass der Mensch selbst und alles, was ihn ausmacht, in Vergessenheit geraten. Suchmans Bezug auf die Physikerin Karen Barad, die auch im Bereich des Feminismus forscht, unterstreicht zudem die Wichtigkeit und den bedeutenden Akt der Grenzziehung. Barad betont, dass Grenzen zwischen Menschen und Maschinen nicht auf natürliche Weise gegeben sind, sondern aus einem Prozess historischer Entwicklung, sowie aus sozialen und materiellen Konsequenzen heraus resultieren (Ebd., S. 285). Diese Annahme unterstreicht Suchmans Argumentation, der Mensch sei in seiner gesellschaftlichen Verfasstheit zu betrachten, damit bestimmte Entwicklungen und Konstrukte (wie Grenzen) erst erkannt und gegebenenfalls verändert werden können (Ebd.).

Die Analysen der feministischen Techniksoziologin Judy Wajcman werden in dem Buch „Technikkritik in der Moderne“ (Müller & Nievergelt, 1996) dahingehend vorgestellt, wie stark die Technikkultur von gesellschaftlichen Zuständen und Geschlechterdifferenzen geprägt ist. Die Techniksoziologin legt dar, dass „mit dem Aufkommen von sauberer Mikroelektronik und dem Rückgang von traditioneller, harter und schmutziger Industriearbeit zu erwarten gewesen wäre, dass Frauen nun vermehrt Zugang zu Technik gehabt hätten. Dies ist jedoch nicht eingetreten“ (Ebd., S. 197). Sie erklärt, dass „die Haltung von Frauen gegenüber neuen Technologien zutiefst von traditionellen und kulturellen Mustern geprägt ist“ (Ebd.). Somit wird deutlich, dass in der Technikkultur selbst derartige Zustände angelegt und verinnerlicht sind, die es zu ergründen gilt. Zudem wird auch betont, dass diese Kultur stark männlich dominiert ist und Frauen der Zugang zu dieser technischen Entwicklung somit auch weiterhin erschwert wird. Diese Annahme stützt sie zusätzlich auf die Computerkultur, die sich auf den Dualismen von Körper und Geist, Wissenschaft und Sinnlichkeit oder Hartem und Weichem gründet. (Ebd., S. 199-200) Sie betont hierbei, dass Frauen im Gegensatz zu Männern mit dem Emotionalen und Schwachen assoziiert werden (Ebd.). Begründet wird diese Feststellung damit, dass technische Errungenschaften einen weiblichen Namen erhalten, um Männer emotional zu befriedigen (Ebd.). Jedoch werden Frauen von dieser Angelegenheit letztlich ausgeschlossen und somit nicht in technische Entwicklungsprozesse mit eingebunden (Ebd.). Geschlechterunterschiede sind somit kulturell zu tief verankert, als dass von ihnen absehen werden könnte. Angewandt auf Haraways Argumentation zum radikalen Feminismus, bei dem auf die Rolle des Organischen im Verhältnis zum Technischen zu stark verwiesen wird, kann nach der bereits erfolgten Kritik gesagt werden, dass diese Zustände jedoch nicht

vollends verleugnet oder ignoriert werden sollten. Dieser Aspekt lässt hinterfragen, warum die Differenzen sowohl zwischen Mensch und Maschine als auch zwischen den Geschlechtern bestehen und womöglich in der Analyse doch eine Rolle spielen dürften. Innerhalb der kritischen Auseinandersetzung wird bewusst, welche Eigenschaften und Wesenszüge mit dem Menschen benannt werden, die offensichtlich geschlechtsspezifisch und auch kulturell eingebettet sind. Die Frage, die sich hier stellen kann, bezieht sich auf die tatsächliche Wichtigkeit ihrer Benennung und des Naturgedankens, der bei Haraway tiefgreifend durch die Figur der Cyborgs in Frage gestellt wird.

## 6. Fazit

“Alles kann also Mythos werden? Ich glaube, ja, denn das Universum ist unendlich suggestiv” (Barthes, 2002, S. 499).

Abschließend soll auf Grundlage der vorangegangenen Überlegungen die Leitfrage beantwortet werden: Ist die Denkfigur des Cyborgs nach Donna Haraway nach wie vor aktuell?

Das untersuchte Essay ist das Gründungswerk der feministischen Technowissenschaften, die aktuell noch stark im wissenschaftlichen Diskurs Anwendung finden und gerade in Diskussionen um künstliche Intelligenzen oder Algorithmen intensiv thematisiert werden. Der Name Haraway wird in diesem Zusammenhang stets genannt. Ebenso ist mit der Strömung des bis heute anhaltenden Postfeminismus, beispielsweise nach Rosalind Gill (2005, 2007a, 2007b), nicht nur die Selbstoptimierung, sondern auch das Neu-Denken von Geschlechtergrenzen in der Diskussion. Somit findet sich auch hier ein starker Aktualitätsbezug der von Haraway und Butler maßgeblich geprägten Strömung. Auch sind einige von Haraways Zukunftsvisionen bereits eingetreten: Durch Medizin- und Prothesentechnik werden Grenzauflösungen zwischen Mensch und Maschine hervorgerufen. Technologische Geräte werden immer kleiner und bestehen nur noch aus Signalen, wie von der Wissenschaftlerin bereits 1985 prognostiziert. Dabei formen sich Mensch und Maschine immer auch in einem interdependenten Verhältnis gleichermaßen. Aus diesem Grund können Cyborgs auch unter die Kategorie der künstlichen Intelligenz (KI) fallen, da sich Technologie und das menschliche Bewusstsein gegenseitig formen. Die maschinellen Teile im Organismus können in Zukunft möglicherweise von künstlicher Intelligenz sein. Wichtig ist allerdings zu betonen, dass bei allen untersuchten Fällen bis heute keine Grenzauflösung der Geschlechter, Rassen oder Klassen erfolgt, sondern sie teilweise sogar noch bedingen.

Die Leitfrage wird aufgrund der aufgezeigten Aspekte also mit Ja beantwortet, muss jedoch um diesen letztgenannten Punkt relativiert werden. Die aufgezeigten kritischen Perspektiven sind darüber hinaus gleichermaßen in Betracht zu ziehen, um einen Gesamtumriss von Haraways Theorien zu erhalten.

Das möglicherweise sehr stark konstruierte Manifest Haraways kann abschreckend und lebensfremd wirken, allerdings ist es gerade dies, worauf Haraway vermutlich abzielen versucht. Es ist schwierig, sich eine Welt der Cyborgs wirklich vorzustellen, da dies möglicherweise am Mangel dem zugehörigen Narrativen liegen mag. Aber gerade durch die Diskussion um Haraways Essay und Ideen werden die bestehenden Machtgefüge zwischen den Geschlechtern aufgespalten und kritisch reflektiert. Dies gelingt vor allen Dingen durch die Konstruktion ihres starken politischen Mythos. Die Erschaffung dieser „Geschichte“ (Saupe, 1998, S. 168) ermöglicht den Widerstand gegen jegliche bestehenden Herrschaftsstrukturen und bietet den Ansatz einer Problemlösung, wie eingangs mit Arnason beschrieben wurde (1976). Der Mythos ist ein zeitloses Phänomen, sodass Haraway möglicherweise eine Figur für die Ewigkeit geschaffen hat.

---

## Literaturverzeichnis

Arnason, J. P. (1976). Zwischen Natur und Gesellschaft. Studien zu einer kritischen Theorie des Subjekts. Frankfurt a.M./Köln: Europ. Verl.-Anst.

Barthes, R. (1957/2002). Der Mythos heute. In C. Pias, J. Vogl, L. Engell, O.

Fahle & B. Neitzel (Hrsg.). Kursbuch Medienkultur. Die maßgeblichen Theorie von Brecht bis Baudrillard(S. 499-506). Stuttgart: Deutsche Verlags-Anstalt.

Butler, J. (1991). Das Unbehagen der Geschlechter. Edition Suhrkamp. Berlin: Suhrkamp.

Connell, R. (2005). Masculinities. Berkeley: University of California Press.

Gill, R., Henwood, K., & McLean, C. (2005). Body Projects and the Regulation of Normative Masculinity. *Body and Society*,11, 37-62.

Gill, R. (2007a). Gender and the media. Cambridge: Polity Press.

Gill, R. (2007b). Postfeminist media culture: Elements of a sensibility. *European Journal of Cultural Studies*, 10, 147-166.

Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, 14/ 3, 575-599.

Haraway, D. (1995). Ein Manifest für Cyborgs. Feminismus im Streit mit den Technowissenschaften. In D. Haraway (Hrsg.), *Die Neuerfindung der Natur. Primaten, Cyborgs und Frauen* (S. 33-72). Frankfurt a. M./New York: Campus Verlag. (Erstmals erschienen unter: Haraway, D. (1985). *Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980's*. *Socialist Review*, 80, 65-108.

Haraway, D. (1991). A Cyborg Manifesto. *Science, Technology, and Socialist Feminism in the Late Twentieth Century*. In: *Simians, Cyborgs and Women: The Reinvention of Nature*. Routledge: New York. S. 149-181, Online unter: [https://web.archive.org/web/20120214194015/http://www.stanford.edu/dept/HP\\_S/Haraway/CyborgManifesto.html](https://web.archive.org/web/20120214194015/http://www.stanford.edu/dept/HP_S/Haraway/CyborgManifesto.html) (Zuletzt aufgerufen: 16.10.2019).

Haraway, D. (2003). *The Companion Species Manifesto: Dogs, People, and Significant Otherness*. Chicago: Prickly Paradigm Press.

Harrasser, K. (2011). Natur-Kulturen und die Faktizität der Figuration In S. Moebius & D. Quadflieg (Hrsg.), *Kultur. Theorien der Gegenwart*. Wiesbaden: VS Verlag für Sozialwissenschaften/Springer Fachmedien Wiesbaden GmbH.

Harrasser, K. (2013). *Körper 2.0. Über die technische Erweiterbarkeit des Menschen*. Bielefeld: transcript.

Homepage TH Köln, Online unter: [https://www.th-koeln.de/studium/gender-und-queer-studies-master\\_44604.php](https://www.th-koeln.de/studium/gender-und-queer-studies-master_44604.php) (Zuletzt aufgerufen: 05.09.19).

McLuhan, M., (2011 [1962]). *The Gutenberg Galaxy: The Making of Typographic Man*, Ed.: [New] edition. Toronto: University of Toronto Press, Scholarly Publishing Division.



McLuhan, M. (1967). *The medium is the message: an inventory of effects*. New York [u.a.]: Bantam Books.

Müller, C., Nievergelt, B. (1996). *Technikkritik in der Moderne. Empirische Technikereignisse als Herausforderung an die Sozialwissenschaft*. Opladen: Leske Budrich.

Pias, C. (2002). Zur Einführung. In C. Pias, J. Vogl, L. Engell, O. Fahle & B. Neitzel (Hrsg.). *Kursbuch Medienkultur. Die maßgeblichen Theorie von Brecht bis Baudrillard*(S. 427-431). Stuttgart: Deutsche Verlags-Anstalt.

Pistorius, O. (2009). *Blade Runner*. London: Random House.

Saupe, A. (1998). *Mythos Cyborg - Zur Politik der Dekonstruktion technologischer Rationalität*. *Freiburger FrauenStudien*, 1, 167-188.

Schwarzer, A. (2013 [1975]). *Der kleine Unterschied und seine großen Folgen: Frauen über sich. Beginn einer Befreiung*. Frankfurt a.M.: Fischer Taschenbuch.

Sloterdijk, P. (1999). *Regeln für den Menschenpark: eine Antwortschreiben zu Heideggers Brief über den Humanismus*. Frankfurt am Main: Suhrkamp.

Sloterdijk, P. (2009). *Du mußt dein Leben ändern: über Anthropotechnik*, Frankfurt am Main: Suhrkamp.

Suchman, L. (2007). *Human-Machine Reconfigurations. Plans and Situated Actions*. Cambridge University Press.

## 7. Einsatz von KI in der Musikproduktion

Vortragstag und Gruppenname:	11. Juli 13:00 Uhr, Einsatz von KI in der Musikproduktion
Titel der Ausarbeitung:	Einsatz von KI in der Musikproduktion
Verfasser/-innen:	Kristina Götschenberg, Michael Haas, Julian Kinghorst
Betreuer/-in:	Univ.-Prof. Dr. Joachim Eigler (Betriebswirtschaftslehre, insb. Medienmanagement) M.A. Julia Müller

### Übersicht über die Struktur der Ausarbeitung

1. Problemstellung
  2. Die Musikindustrie als Anwendungsfeld neuer Technologien
    - 2.1 Wertschöpfung und Akteure in der Musikindustrie
    - 2.2 Entwicklung elektronischer/digitaler Musik und KI
  3. KI und ML in der Musikproduktion
  4. KI und ML in der Musikproduktion
    - 4.1 Das Geschäftsmodell von Spotify
    - 4.2 Musik-Empfehlungssysteme
  5. Kontroversen und Spannungen durch KI und ML im Musikkontext
    - 5.1 Das Urheberrecht
    - 5.2 Akzeptanz von KI und Kreativität
  6. Fazit und Ausblick
- Literaturverzeichnis
-

## **1. Problemstellung**

Auf Machine Learning (ML) basierende Künstliche Intelligenz (KI) stellt eine der bedeutendsten technologischen Innovationen dar (Eberl, 2018, S.14), deren Anwendung das Potenzial bietet, ganze Industriezweige zu transformieren (Brynjolfsson & McAfee, 2017, S. 2). Bezogen auf die Relevanz solcher Technologien für die Musikindustrie (Mayer, 2018, S. 32) und resultierender Problematiken, schlussfolgert Prey (2018, S. 19), dass durch hohe Datenmengen für alle Branchenakteure Schwierigkeiten auftreten, die es zu verstehen und zu operationalisieren gilt. Für die Musikindustrie birgt KI in Verbindung mit Big Data (Müller, 2018, S. 155) enorme Potenziale für Veränderung (Loeffler, 2018; Tschirren, 2018), wobei der Einsatz bisher vorwiegend der Produktion (Léveillé Gauvin, 2017, S. 302) und digitalen Distribution zuzuordnen ist. (Mayer, 2018, S. 33).

Die vorliegende Arbeit untersucht entsprechend den Einfluss der Anwendung von KI und ML in der Musikindustrie, insb. in Bezug auf die Musikproduktion und deren Distribution. Es wurde ein interdisziplinärer Forschungsansatz mit Bezug auf die verschiedenen Literaturströmungen der Musikmanagement-, Medienwissenschafts- und auch der Informatikliteratur gewählt (Abb. 1). Um verschiedene Sichtweisen anhand von Fallbeispielen deutlich zu machen, wurden neben einer umfassenden Literaturrecherche drei Experteninterviews, ein Kurzinterview mit einem Jura-Professor und eine Straßenumfrage durchgeführt sowie weitere Sekundärressourcen wie Werbematerial, Websites und Onlinevideos in die Analyse des Themas aufgenommen (Anh. A).

Im Folgenden wird die Musikindustrie als Anwendungsfeld neuer Technologien mit der Wertschöpfungskette und Akteuren in der Musikindustrie sowie der Entwicklung elektronischer bzw. digitaler Musik und KI dargestellt. Danach werden Potenziale zur Musikkreation mit einem Fallbeispiel der Musikproduktion durch KI und ML veranschaulicht. Anschließend werden im Bereich der Musikdistribution das Geschäftsmodell des digitalen Marktführers Spotify und Musik-Empfehlungssysteme analysiert. Zudem werden die Spannungen in Bezug auf das Urheberrecht, Akzeptanz und Kreativität betrachtet, bevor die Ergebnisse der Arbeit in einem abschließenden Fazit kritisch gewürdigt werden.

## **2. Die Musikindustrie als Anwendungsfeld neuer Technologien**

### **2.1 Wertschöpfung und Akteure in der Musikindustrie**

Die Digitalisierung, verstanden als Wandel physischer Handelsgegenstände zu Digitalgütern (Zerdick et al., 2001, S. 16), hat die Wertschöpfung der Medienwelt grundlegend verändert (Berger et al., 2015, S. 22). Dies ist insb. für die Musikindustrie zutreffend (Carstensen, 2018, S. 129) und wirkt sich z.B.

durch eine Senkung der Marginal- und Transaktionskosten aus (Krüger, 2002, S. 67, S. 70). Zwischen den Jahren 1997 und 2012 befand sich die Musikindustrie in einer starken Krise mit stetigem Umsatzrückgang, welcher vor allem Verkaufseinbrüchen physischer Tonträger (BVMI, 2018, S. 4-5; Abb. 2) und der Verbreitung von Musikpiraterie zugeschrieben wird, wobei letzteres unter Forschenden eine umstrittene Ursache darstellt (Andersen & Frenz, 2010, S. 722).

Vorangetrieben durch das Wachstum von digitalen Musikangeboten, welche im Jahr 2018 einen Marktanteil von 58,9% verzeichneten, konnte die Krise der Musikwirtschaft überwunden werden, wodurch bspw. 2018 wieder ein Umsatzwachstum von 9,7% für die gesamte Musikbranche erwirtschaftet wurde (IFPI, 2019, S. 13; Abb. 3).

Die Digitalisierung führte zu einer Veränderung der traditionellen hin zu einer neuen Wertschöpfungskette (Krüger, 2002, S. 71; Carstensen, 2018, S. 130) bestehend aus Akteuren wie Künstlern, Labels, Verlagen, Tonstudios, Presswerken und Händlern (Bockstedt, Kauffman, & Riggins, 2006, S. 16). Diese zeichnet sich durch Disintermediation, verstanden als Wegfall von Intermediären (Hess, 2006, S. 252), und durch Hinzukommen neuer Infomediäre, verstanden als Unternehmen, welche Konsumenten kontextspezifisch Informationsprodukte anbieten, aus (Morris, 2015, S. 447-449). Zudem verschwimmen zunehmend die Grenzen zwischen Produzenten und Konsumenten, sowie zwischen Wertschöpfungsstufen (Bilton & Cummings, 2015, S. 85). Eine Darstellung der traditionellen und digitalen Wertschöpfungsketten der Musikindustrie auf Basis des Modells von Porter (1985, S. 37) ist der Arbeit angehängt (Abb. 3; Abb. 4).

Da sich die Konsumgewohnheiten in Bezug auf Musikprodukte grundlegend geändert haben (Anderson, 2008, S. 22) und somit enorme Mengen an Daten verfügbar sind (Morris, 2015, S. 448), eröffnen sich vor dem Hintergrund von Big Data neue Anwendungsfelder für KI und ML in der neuen Wertschöpfungskette (Streibich & Zeller, 2019, S. 110). Dies ist vor allem im Falle der Musikproduktion bei der Produktion chartrelevanter Songs auf Basis von Big Data, der automatisierten Unterstützung mittels begleitender Akkordvorschläge zur Komposition im semiprofessionellen Bereich und der KI-basierten Auswertung und Ausspielung von Musiktiteln im digitalen Vertrieb (insb. Streaming) der Fall.

## **2.2 Entwicklung elektronischer/digitaler Musik und KI**

Nach der Erfindung des Turing-Tests im Jahr 1950 (Herbrich, 2019, S. 64) und der Präsentation des ersten Chatbots 1965 nahm das Interesse an KI in den 1970er Jahren stark ab (Eberl, 2018, S. 10). Es beschränkte sich auf die Entwicklung von Expertensystemen zur Entscheidungsunterstützung (Buxmann & Schmidt, 2019a, S. 6), bevor in den 1990er Jahren durch neue Cloud-Angebote die nötige Rechenleistung erreicht werden konnte (Mayer, 2018, S. 32) und KI erstmals die Leistung des menschlichen Gehirns z.B. im Schachspiel überstieg (Buxmann & Schmidt, 2019a, S. 5-6; Herbrich, 2019, S. 63). Im Bereich der Musikproduktion und -distribution stellen die Erfindung des Telharmoniums im Jahr 1900 und des Ondes Martinot 1928 als erste Klangautomaten (Luening, 1968, S. 44-45), des Moog-Synthesizers 1964 (Anthony, Nelson, & Tripsas, 2016, S. 165), der Erfindung von CD, MP3 und des MIDI-Standards 1980, die Veröffentlichung der P2P-Filesharing-Plattform Napster und der Produktionssoftware (DAW) Ableton Live 1990, sowie der Markteintritt von iTunes 2001 und Spotify 2006, bedeutende Meilensteine dar (Tschmuck, 2017, S. 28-29; Abb. 5).

Diese historischen Entwicklungen machen erstmals ein Zusammenspiel von KI und ML in der Musikproduktion und -distribution möglich (Tatchell, 2005; Abb. 5), welches in den folgenden Kapiteln untersucht und anhand von Fallbeispielen dargestellt wird.

## **3. KI und ML in der Musikproduktion**

KI und ML bieten viele Ansatzpunkte bzgl. Wertschöpfungsprozessen (Mayer, 2018, S. 35) und können die Effizienz der Produktion positiv beeinflussen (McKinsey, 2017, S. 8; Buxmann & Schmidt, 2019b, S. 33; Hildesheim & Michelsen, 2019, S. 137). Somit ist eine Produktverbesserung möglich (Wess, 2019, S. 157), weshalb ein Einsatz in Verbindung mit Big Data auch in der Musikindustrie in Frage kommt (Müller, 2018, S. 169).

Die Auswertung des Verhaltens der Konsumenten kann dazu genutzt werden, gezielt Inhalte zu produzieren (Berger et al., 2015, S. 24). Dies wird dadurch deutlich, dass bereits Musik z.B. in Bezug auf Takt und Geschwindigkeit für Streamingdienste optimiert (Tatchell, 2005; Léveillé Gauvin, 2017, S. 302; Müller, 2018, S. 167) oder gezielt zur Beeinflussung von Algorithmen produziert wird (D'onfro, 2014). Songs lassen sich generell in kleinere, sich wiederholende Bestandteile aufbrechen, die mit Hilfe von ML analysiert, nachgeahmt oder neu zusammengestellt werden können (Loeffler, 2018). Auf Basis von, durch KI ausgewertete, Daten von Charthits lassen sich deshalb Songs mit einer hohen

Erfolgswahrscheinlichkeit erstellen (Müller, 2018, S. 163).‘

Zudem eignet sich KI gut für die Erstellung von Vorhersagemodellen (McKinsey, 2017, S. 32; Mayer, 2018, S. 35; Gürtler, 2019, S. 96), da Suchanfragen (Nambisan, 2016, S. 1038) sowie Interaktionen in sozialen Medien ausgewertet und archiviert und somit Trends erkannt werden können (Heim, Graefe, & Brosius 2018, S. 40). Dies wird schon im Rahmen von Big Data zur Singleauswahl und Künstlerpositionierung genutzt und hat somit auch Einfluss auf den Produktionsprozess, was aber bislang nur von größeren Musikfirmen angewendet wird (Müller, 2018, S. 167-168; Loeffler, 2018). Dadurch können sie einen Wettbewerbsvorteil erlangen (Streibich & Zeller 2019, S. 116).

Weitere potenzielle Anwendungsfelder liegen z.B. in der Erstellung adaptiver Musik, welche sich je nach Situation ändert (Kühl, 2017) oder personalisierter Remixe (Dredge, 2017). Die Entwicklung solcher KI findet meist in Start-ups statt (McKinsey, 2017, S. 11). Das Fallbeispiel des KI-Unternehmens Poppun, welches mit ihrer KI Alice die Möglichkeit bietet, eingespieltes Material zu optimieren und fortzuführen (Dredge, 2017; Kühl, 2017; Tschirren, 2018; Poppun, 2019), verdeutlicht den Einsatz von KI und ML in der Musikindustrie. Die Anwendung basiert auf Deep Learning, verstanden als Nutzung mehrschichtiger neuronaler Netze (Buxmann & Schmidt, 2019a, S. 12). Alice kann dabei laut dem Unternehmen als Zusatztool für Kreativschaffende, in der Produktion von Pop-songs oder zu Fortbildungszwecken eingesetzt werden (Loeffler, 2018). Im Anfangsstadium konnte Alice nach Tasteneingabe auf einem Piano eine Vorhersage über nachfolgende Noten treffen. Alice wurde dadurch zu einer Art Duettpartner. Spätere Entwicklungsstufen deckten danach Instrumente wie Schlagzeug und Bass bis hin zum Repertoire einer gesamten Band ab (Dredge, 2018; Sheahan, pers. Kommunikation, 04. Juni 2019). Mittlerweile verfügt Alice über die Fähigkeit, einen eingegebenen Text zu singen (Dredge, 2019). Das Unternehmen sieht neben dem Einsatz im professionellen Bereich zudem die Verwendung der Technik durch Anfänger, welche so befähigt werden, ebenfalls besser zu musizieren (Sheahan, pers. Kommunikation, 04. Juni 2019).

Aber auch KI-Anwendungen anderer Anbieter sind bereits im Einsatz. Der Künstler Skygge (Tschirren, 2018) und die YouTuberin Taryn Southern haben bereits ganze Alben mit der Hilfe von KI komponiert (Kühl, 2017), während der KI-Song „Daddy’s Car“ größere Bekanntheit auf YouTube erlangte (Dredge, 2017). Auch Google ist mit AI.Duet in den Markt eingetreten (Dredge, 2017; Kühl, 2017). Weitere Unternehmen, die KI und ML in der Musikbranche anwenden, sind u.a. Shazam (Knopper, 2014; Thompson, 2014; Jensen, 2017), Amper, AI.Music, Techstars, Jukedeeck, Melodrive, Humtap, Flow Machines (Musically, 2017; Tschirren, 2018), Groov.AI und Scape (Dredge, 2017).

Der Einsatz von KI und ML in der Musikproduktion zielt laut Gösde, Jacke & Nösner (2018, S. 213) auf die Unterstützung von „Musicpreneuren“, verstanden als ökonomisch schaffende Musiker, und die Demokratisierung von Produktionsmitteln ab (Kühl, 2017). Dies wird kontrovers diskutiert, da z.B. Meinungen über den Einfluss auf musikalische Vielfalt hier weit auseinander driften (Buxmann & Schmidt 2019b, S. 26).

Die vorangegangenen Beispiele zeigen, dass KI wie z.B. Alice alle Bereiche der Musikproduktion abdecken, wobei der Grad der Einflussnahme je nach KI variiert. Der künstlerische Input liegt jedoch weiterhin größtenteils beim Menschen selbst, der automatisierte Prozesse anstoßen, gezielt lenken und autorisieren muss.

#### **4. KI und ML in der Musikdistribution**

##### **4.1 Das Geschäftsmodell von Spotify**

Exemplarisch wird nun das Geschäftsmodell des Streaminganbieters Spotify dargestellt. Spotify wurde 2006 gegründet, verzeichnete 2018 mehr als 170 Mio. Nutzer und bietet ca. 30 Mio. Titel an (Dahm, Constantine, Gebauer, & Bartsch, 2019, S. 62-63). Das Portfolio ergibt sich aus Lizenzverträgen mit Musiklabels und weiteren Rechteinhabern (Voigt, Buliga, & Oana, 2017, S. 146). Kernangebot ist die Distribution digitaler Kopien in Form von Streams (Dahm et al., 2019, S. 61). Einen unlimitierten Zugang bietet Spotify den Nutzenden in einem Freemium-Modell an. Demnach ist der Zugang sowohl kostenlos, als auch in einem Premium-Modell möglich. Der kostenlose Zugang unterscheidet sich durch Werbeeinblendungen und Funktionsrestriktionen. Bei dem werbefreien Premium-Modell wird eine Abogebühr von zehn Euro pro Monat erhoben (Voigt et al., 2017, S. 145, S. 152). Entsprechend ergeben sich zwei Ertragsmodelle, bei dem das Werbefinanzierte durch die Zusammenführung von Nutzern und Werbetreibenden dem Two-Sided-Market entspricht (Moser, Wecht, & Gassman, 2019, S. 46). Trotz der höheren Nutzerzahl des kostenlosen Angebots macht der Ertrag durch Werbung lediglich 10 % Prozent des Umsatzes aus (Voigt et al. 2017, S. 150; Dahm et al. 2019, S. 63). Vonderau (2019, S. 5) merkt an, dass Spotify als Medienunternehmen zu begreifen sei, welches an der Schnittstelle zwischen Werbung, Technologie, Musik und Finanzen agiert. Voigt et al. (2017, S. 151-152) sehen dabei neben anderen On-Demand-Streaminganbietern wie Deezer auch cloudbasierte Musikanbieter, illegale P2P-Angebote, legale Video-Streaming-Plattformen sowie Webradios als Wettbewerber von Spotify, wodurch der stark umkämpfte Markt, auf dem Spotify operiert, illustriert wird.

Bezogen auf die On-Demand-Anbieter ist dabei die Erhebung von Daten eine zentrale Tätigkeit solcher Unternehmen. Robert Prey (2015, S. 220) spricht in diesem Zusammenhang von der Datafizierung des Zuhörens (datafication of listening), was die neu entstehende Menge an Daten bei alltäglichen Praktiken beschreibt.

Hinsichtlich des Wettbewerbs auf dem Markt merkt Prey (2018, S. 12) an, dass durch die reine Größe der Portfolios kein Wettbewerbsvorteil erzielt werden konnte, sodass nun die Personalisierung im Fokus steht. Mit der Akquisition von The Echo Nest im Jahr 2014 (Voigt et al., 2017, S. 149) und der Funktion „Dein Mix der Woche“, welche wöchentlich personalisierte Playlisten für die Nutzenden erstellt (Prey, 2018, S. 12-13), wird diese Ausrichtung auch bei Spotify deutlich. Die Bündelung verschiedener Daten (Vonderau, 2019, S. 15), und damit auch die Personalisierung, ist somit ein weiteres Kernangebot von Spotify (Voigt et al., 2017, S. 148).

#### **4.2 Musik-Empfehlungssysteme**

Ausgehend von verändertem Konsumverhalten und Personalisierungsinteresse, ist es elementar, den Nutzern mittels Empfehlungssystemen Titel vorzuschlagen und personalisierte Playlisten zu erstellen (Prey, 2018, S. 12-14). An Empfehlungssystemen bedarf es insb. durch die große Titelzahl (Schedl et al., 2015, S. 453) und der gestiegenen Flexibilität durch Mobilgeräte. Durch die große Auswahl von Titeln entstehen Nutzern Suchkosten (Datta, Knox, & Bronnenberg, 2018, S. 2) und durch Flexibilität ergibt sich das Problem der Anpassung an wechselnde Situationen und Stimmungen (Helmholz, Meyer, & Robra-Bissantz, 2019, S. 598-559). Es lässt sich zwischen kollaborativen, inhaltsbasierten, kontextbezogenen und hybriden Empfehlungssystemen unterscheiden.

Kollaborative Empfehlungssysteme basieren auf den sog. kollaborativen Filtern. Dabei wird das Verhalten (indirektes Feedback) und Bewertungen (direktes Feedback) von Nutzern verglichen, um Muster zu erkennen und Aussagen über zukünftiges Verhalten zu treffen (Koren & Bell, 2015). Dies lässt sich auch auf Musiktitel übertragen, um Beziehungen zwischen den Titeln zu erkennen (Dieleman, 2014). Hierbei ergibt sich jedoch das Problem der sog. Kaltstart-Problematik. Neue Nutzer oder neue Artikel können durch das System nur schwer verarbeitet werden, da zu Beginn keinerlei Daten vorliegen (Schedl et al., 2018, S. 97-98). Demgegenüber ist es einfacher, beliebte Artikel zu empfehlen, da zu diesen deutlich mehr Daten vorliegen (Dieleman, 2014).



Inhaltsbasierte Empfehlungssysteme nutzen sowohl Metadaten aus externen Quellen, als auch die Musik selbst zur Ausdifferenzierung (Schedl et al., 2015, S. 455). So wird bspw. durch das Durchsuchen von Websites oder Titeln mithilfe von Natural Language Processing (NLP) analysiert, welche Worte mit Künstlern in Verbindung gebracht werden (Whitman, 2012; Schedl et al., 2015, S. 457). Eine Analyse des Liedes kann darüber hinaus Aufschluss über dessen Eigenschaften geben. Hier macht man sich künstlicher neuronaler Netze zu Nutze, die verschiedene Filter auf das Lied anwenden können. Neben Informationen zum Tempo und den verwendeten Instrumenten (Dieleman, 2014) lässt sich so bspw. auch die ‚Tanzbarkeit‘ (danceability) bestimmen (Whitman, 2012). Diese Art von Empfehlungssystem kann somit die Kaltstartproblematik neuer Titel umgehen und damit auch den „Long-Tail“ des Portfolios in Betracht ziehen (Celma, 2010).

Kontextbezogene Empfehlungssysteme können sowohl die Umweltgegebenheiten als auch Nutzer selbst berücksichtigen. Bzgl. der Umwelt ließen sich etwa durch Sensoren im Smartphone und externen Daten wie Umgebung, Zeit, Wetter und weitere Faktoren wie Umgebungslautstärke in Betracht ziehen (Schedl et al., 2015, S. 461). Bezüglich der Nutzer könnten Aktivitäten, Gefühlszustand, Sozial- und Kulturkontext berücksichtigt werden. Schedl et al. (2015, S. 463) merken an, dass diese Faktoren allerdings noch schwer zu ermitteln sind. Hierfür bedarf es möglicherweise der Zugabe von Informationen durch die Nutzer. In diesem Zusammenhang bietet Spotify Werbetreibenden an, den Kontext zu erkennen, indem sich die Nutzer während des Hörens der Musik befinden, um passende Werbung auszuspielen. So soll eine höhere Relevanz für die Werbung erreicht werden (Spotify for Brands, 2019; Prey, 2019, S. 17-18).

Sowohl kollaborative als auch inhaltsbasierte Empfehlungssysteme sind in verschiedenen Ausprägungen von der Kaltstartproblematik betroffen, weshalb hybride Empfehlungssysteme entgegenwirken, indem sie mehrere Methoden miteinander verbinden (Schedl et al., 2018, S. 98-99). Eins dieser Systeme findet sich auch bei der Kuration von personalisierten „Dein Mix der Woche“-Playlisten. Kombiniert werden dazu das kollaborative Filtern, die inhaltsbasierende Analyse von Metadaten mit NLP, sowie die Analyse der Musik selbst (Abb. 6). Dies ermöglicht den Nutzern wöchentlich eine Liste von 30 Liedern auszuspielen, die basierend auf ihrem Nutzungsverhalten und der inhaltsbezogenen Analyse von Liedern automatisch kuratiert werden (Prey, 2019, S. 13). Datta et al. (2018, S. 19) merken an, dass Spotify durch solche Mechanismen letztlich die Suchkosten senken kann, was wiederum das Wohlbefinden der Nutzer fördert.

Ein Großteil der Empfehlungssysteme basiert auf Algorithmen, wobei sich der Gebrauch von KI in schwacher Form identifizieren lässt. Das Beispiel des automatischen Kuratierens veranschaulicht den Einsatz und die Kombination verschiedener Techniken von ML. Bei dem eingesetzten inhaltsbasierten Empfehlungssystem ist innerhalb der Analyse von Audiodateien das Nutzen künstlicher neuronaler Netze ersichtlich. Laut Dieleman (2014) geht diese Technik über reines ML hinaus und ist als Deep Learning zu bezeichnen.

## **5. Kontroversen und Spannungen durch KI und ML im Musikkontext**

### **5.1 Das Urheberrecht**

Durch KI ergeben sich neue Alternativen der Songschreibung, aber auch Probleme bzgl. Urheberschaft, Haftung, Ethik (Kehl, 2018, S. 24) und Rechenschaftspflicht, da Vorschriften nicht zeitgleich zu Innovationen entstehen (Kathan, Matzler, & Veider, 2016, S. 665; Bendel, 2018, S. 35). Es ist z.B. fraglich, ob KI ein Urheber sein kann. Mit solchen Fragen muss sich die Rechtswissenschaft neu beschäftigen (Dredge, 2017).

Es gibt noch keinen Präzedenzfall im Bereich der Musikbranche, weshalb es sinnvoll ist, Bereiche wie z.B. die Malerei zu betrachten, in der es mehrere KI-Produktionen gibt wie z.B. das Gemälde "Edmond de Belamy" (Weihser, 2018; Abb. 9) des Künstlerkollektivs Obvious, das einen Algorithmus mit 15.000 Porträts gespeist (Weihser, 2018) und dazu eine Software verwendet hatte (Moritz & Mohr, 2019). Die KI lässt sich hier als Urheber ausschließen, da laut §2 Abs. 2 UrhG (2017) nur Menschen Urheber sein können, weil Werke nur persönliche geistige Schöpfungen seien. Die Qualität dieser Werke spielt laut Becker (pers. Kommunikation, 26. Juni 2019) dafür keine Rolle.

Beim Training der KI können lizenzierungsbedürftige Handlungen an den Trainingsdaten wie z.B. Vervielfältigung und Veränderung erforderlich sein (Becker, pers. Kommunikation, 26. Juni 2019). Laut §24 UrhG (2017) darf ein selbstständiges Werk veröffentlicht werden, wenn es in freier Benutzung des Werkes eines anderen geschaffen wurde. Jedoch gilt Abs. 1 §24 UrhG (2017) nicht, wenn die Melodie des Ausgangswerks durchscheint. Beim Erkennen von Stilrichtungen und Standardelementen. Wie im Fall von Akkordfolgen liegt keine Urheberrechtsverletzung vor (Becker, pers. Kommunikation, 26. Juni 2019).

Laut Dredge (2017) ist unklar, ob schöpferische Aspekte eher beim Künstler oder Programmierer liegen, weil nur derjenige Urheber ist, der einen hinreichenden schöpferischen Beitrag geleistet hat

(Becker, pers. Kommunikation, 26. Juni 2019). Gemäß Becker kann ein Komponist, der KI einsetzt, auch Urheber sein, wenn er diese eigenschöpferisch verwendet. So können komplett urheberrechtsfreie Stücke entstehen. Wenn ein Programmierer eine Software erschaffe, die selbst komponiert, ist er laut Becker (pers. Kommunikation, 26. Juni 2019) Urheber der KI, aber nicht Urheber des Songs.

Transparenz und Rechtssicherheit in Bezug auf Daten, Prozesse und dem Erlernten, welches von der KI verarbeitet wird, ist für die Rechtssicherheit von enormer Bedeutung (Hildesheim & Michelsen, 2019, S. 128). Es wäre von Vorteil, die Rechte an Daten und daraus Gelerntem zu behalten, um somit Wettbewerbsvorteile zu sichern (Hildesheim & Michelsen, 2019, S. 140). Dredge (2017) überlegt, ob Künstler eines Tages Intelligenzrechte in ihre Verträge schreiben, um sich mit diesen Rechten auf eine Zeit vorzubereiten, in der KI in ihrem Songwriting geschult werden, um Originalmaterial zu komponieren.

Mangels Rechts- und Schöpferfähigkeit kann KI nicht als Urheber gelten. Doch ausgehend von dem gezeigten Beispiel und der aktuellen Rechtslage ist die Frage nach dem Urheberrecht noch nicht abschließend zu beantworten.

## **5.2 Akzeptanz von KI und Kreativität**

Der Diskurs um KI wird kontrovers geführt, von Bevormundung durch KI bis hin zu Erleichterung (Dorn & Rojas, 2018). Besonders die Angst vor Arbeitsplatzverlust, Intransparenz und Verdummung bestimmen die Argumentation der KI-Gegner (Hildesheim & Michelsen, 2019, S. 127). Hinzu kommt die Furcht vor Manipulation, Überintelligenz, Zentralisierung, Monopolisierung und zu wenig Datenschutz (Hildesheim & Michelsen, 2019, S. 128). Die Entscheidungen, die ein System in einer bestimmten Situation treffen wird, werden immer unvorhersehbarer für die Planung der Entwickler, je mehr Fortschritte durch ML gemacht werden (Misselhorn, 2018, S. 30).

Gut trainierte KI könnte bestimmte Aufgaben auf dem gleichen Niveau wie Menschen ausführen, bietet die Vorteile einer hohen Skalierbarkeit und macht keine Pausen erforderlich. KI kann z.B. Muster in den Daten erkennen, die selbst für menschliche Experten zu komplex sind (McKinsey 2017, S. 11). Außerdem erlaubt KI uns, einfaches intelligentes Verhalten zu automatisieren (Hildesheim & Michelsen, 2019, S. 125). Bspw. könnten laut Heim, Graefe und Brosius (2018, S. 39) KI für zeitintensive Routinen verwenden und somit entlastet werden (Herbrich, 2019, S. 72). Start-ups hoffen, dass Künstler durch KI eher inspiriert werden, statt sich bedroht zu fühlen (Dredge, 2017).

Laut Kühl (2017) könnten neue Medien auf Ablehnung stoßen, was er am Beispiel der automatisierten Tonhöhenkorrektur Auto-Tune festmacht. Auch Tschmuck (2003, S.161) vertritt die Position, dass Neuerungen immer erst in Frage gestellt werden. Will man die zukünftige Bedeutung von KI und ML betrachten, so bieten sich seine Überlegungen und sein Modell des Prozesses des Paradigmenwechsels in der Musikindustrie an. Dieses diskutiert die Auswirkungen systemfremder Kreativität auf die Branche. Laut Tschmuck (2003, S. 114) gibt es drei Phasen der Auswirkungen der systemfremden Kreativität auf die bestehende Branchenstruktur. In der ersten Phase geht es um die improvisierte und experimentelle Kreativität in Handlungsbeziehungen zwischen Elementen der systemfremden Kreativität und einigen alten Industrieelementen. Eine steigende Anzahl an Agierenden führt zu mehr Komplexität und Unsicherheit. In der Phase der beeinflussten Kreativität festigen sich Verknüpfungen und bilden Regelmäßigkeit. Es gibt gelegentlich Experimente und es beginnt eine neue Oligopolisierung. Während der Phase der kontrollierten Kreativität festigt sich das Oligopol und der Aufbau von Markteintrittsbarrieren beginnt. Die ehemals treibenden Kräfte hinter dem Paradigmenwechsel sind dominante Organisationen, bis neue systemfremde Kreativität diese Struktur aufbricht (Tschmuck, 2003, S. 287-292.; 2009, S. 159-161). Entsprechend dieses Modells muss auch KI die Phasen durchlaufen, um auch innerhalb der Branche breite Akzeptanz zu erlangen.

Folgt man den Ergebnissen einer aktuellen Studie von MusicStats (2019), zeigt sich die Relevanz künstlicher Kreativität bei Musik, da diese von Rezipienten erwartet wird. Dabei ist die Akzeptanz bzgl. KI-Musik eher negativ (Tigre Moura, pers. Kommunikation, 26. Juli 2019) außer bei einer untergeordneten Funktion wie bspw. in Werbespots (Tigre Moura, 2019; Abb. 7). Zudem scheinen die Umstände der Komposition, ob KI oder Künstler, bei Gefallen des Liedes irrelevant zu sein (Tigre Moura, pers. Kommunikation, Juli 2019). Dafür sind eine richtige Balance und angenehme Akkordfolgen ausreichend (Dredge, 2017). Anderen Stimmen zufolge nimmt die soziale Kreativität durch KI jedoch ab, da menschliche Urteilskraft nicht ersetzbar ist (Dorn & Rojas, 2018, S. 5-7). Neuronale Netze können Objekte erkennen, aber wissen nichts über deren Bedeutung für unseren Alltag (Eberl, 2018, S. 11-12). Die Schwierigkeit besteht laut Müller (2018, S. 169) darin, die Emotionen und die Authentizität des Künstlers zu ersetzen. Denn letztendlich sei die Kreativität das, was gute Musik ausmacht (Müller, 2018, S. 169). Bei den jüngeren Nutzern neuer Medien liegt die Akzeptanz von KI-Musik sehr hoch (Musically, 2017). Bei Ablehnung der durch KI erzeugten Musik wird diese oft durch Fehlen von Eigenschaften wie Urteilskraft, Fingerspitzengefühl, Scham, Gewissen, Kreativität und Intuition begründet (Dorn & Rojas, 2018, S. 7). Programme können Stile kopieren und kombinieren, aber nichts gänzlich Neues erschaffen (Eberl, 2018, S. 13). Laut Eberl (2018, S. 14) haben Maschinen aber auch Vorteile gegenüber Menschen, da biologische Notwendigkeiten und emotionale Vorgänge ausbleiben. Diese Vorteile hat auch der digitale Popstar „Hatsune Miku“ aus

Japan (Sonicwire, 2019) als Konzerthallen füllende, künstliche Gesangsstimme und Synthesizer (Le, 2014). KI ist jedoch bislang nicht in der Lage, präzise Emotionen auszulösen (Dredge, 2017). Zudem sind Empfindungen bzgl. Musik sehr subjektiv, weshalb die Frage nach der Qualität von KI-Musik nicht objektiv geklärt werden kann (Dredge, 2017).

Für den Bereich der Musik lässt sich vermuten, dass KI Musikern neue Perspektiven eröffnet und sie inspiriert, statt sie zu bedrohen. Denn laut Tigre Moura (pers. Kommunikation, 26 Juli .2019; Abb. 8) werden Künstler nicht durch KI ersetzt, sondern immer bewundert, es ergeben sich für Zuschauer dank KI neue Hörerlebnisse und die Komponisten sollten die KI als Chance sehen, auf neue Art zu komponieren.

## **6. Fazit und Ausblick**

Ziel der Arbeit war es, KI und ML in der Musikbranche mit einem interdisziplinären Forschungsansatz zu ergründen. Es wurde dargestellt, dass sich die Wertschöpfungskette der Musikindustrie durch datengetriebene Technologien gewandelt hat, weshalb sich insb. durch Verfügbarkeit großer Datenmengen neue Anwendungsfelder ergeben.

Gleich zu Beginn der Arbeit wurde verdeutlicht, dass besonders das Verstehen und die Operationalisierung der (neuen) Daten, die in einem großen Umfang durch die Digitalisierung der Musik vorliegen, für die Musikindustrie eine Herausforderung darstellen. Konkret wurde gezeigt, dass die Technologien bereits in einem gewissen Umfang Anwendung in der Produktion finden. So wird die Auswertung der Daten für den Einsatz in Optimierungs- und Vorhersagemodellen genutzt, welche für eine zielgerichtete Produktion Vorteile mit sich bringen. Es lässt sich demnach ein geschaffener Wettbewerbsvorteil durch das Nutzen der Technik innerhalb der Produktion attestieren. Das Fallbeispiel des Start-ups Popgun und ihrer mittlerweile weit entwickelten KI Alice veranschaulicht zudem den Nutzen für Kunstschaffende sowie Amateure im Bereich der Musikerzeugung. Der jeweilige Einfluss von KI variiert allerdings stark. Der reine künstlerische Input geht weiterhin vornehmlich vom Menschen aus.

Auch im Bereich der Distribution zeigen sich erste Möglichkeiten von ML aufgrund der gestiegenen Musikverfügbarkeit auf dem stark umkämpften Streamingmarkt. Automatisch kuratierte Playlisten verdeutlichen die Notwendigkeit, Suchkosten durch Empfehlungsmechanismen zu senken und andererseits, wozu insb. ML durch die Auswertung der Daten und die dadurch ermöglichte Personalisierung einen Betrag leisten kann. KI findet hier bislang lediglich in schwacher

Ausprägung in Form von künstlichen neuronalen Netzen, verstanden als Weak Learning, Anwendung.

Abschließend wurden aktuelle Spannungsfelder und Kontroversen diskutiert, die mit dem Auftreten der Technologien aufgekommen sind. KI kann zwar mangels geistiger Schöpfung kein Urheber sein (UrhG, 2017), jedoch ist die Rechtslage noch nicht endgültig geklärt. Bzgl. der Akzeptanz sind Subjektivität und Situation entscheidend (Tigre Moura, pers. Kommunikation, 26. Juli 2019), woraus sich jedoch neue Chancen hinsichtlich des Einsatzes von KI ergeben. Zum Beispiel könnte mit KI produzierte Musik bei technischen Werbespots eingesetzt werden, weil hier die Akzeptanz höher ist, da die Musik nicht im Fokus steht. Die Rolle eines Künstlers und seine Position als Vermittler bleiben dabei aber gleichwohl bedeutend.

Greift man zum Schluss nochmals das thematisierte Modell von Tschmuck auf und betrachtet die Erkenntnisse dieser Arbeit, so wird ersichtlich, dass ein Umbruch oder gar Paradigmenwechsel in der Branche durch den Einsatz von KI und ML in den vorgestellten Anwendungsfeldern noch nicht ausgemacht werden kann. Dessen ungeachtet zeigt diese Arbeit, dass erste Einsatzmöglichkeiten von KI bereits in Produktion und Distribution stattgefunden haben, wodurch vereinzelt Wettbewerbsvorteile erzielt werden können. Ausgehend davon erscheint der Einsatz der der Technologien sowie davon ausgehenden zukünftigen Entwicklungen für die Industrie dienlich.

---

## Literaturverzeichnis

Andersen, B., & Frenz, M. (2010). Don't blame the P2P file-sharers: The impact of free music downloads on the purchase of music CDs in Canada. *Journal of Evolutionary Economics*, 20(5), 715-740.

Anderson, C. (2008). *The Long Tail: Why the Future of Business is Selling Less of More*. New York: Random House Business Books.

Anthony, C., Nelson, A.J., & Tripsas, M. (2016). „Who Are You? I Really Wanna Know“: Product Meaning and Competitive Positioning in the Nascent Synthesizer Industry. *Strategy Science*, 1(3), 163-183.

Bendel, O. (2018). Überlegungen zur Disziplin der Maschinenethik. *Aus Politik und Zeitgeschichte*, 68(6-8), S. 34-38.

Berger, M., Haim, M., Graefe, A., Brosius, H.-B., & Hess, T. (2015). Aktuelles Stichwort: Computational Journalism. *MW: Medienwirtschaft*, 12(1), 20-25.

Bilton, C., & Cummings, S. (2015). Creative strategy: notes from a small label. In N. Beech, & C. Gilmore (Hrsg.), *Organising Music: Theory, Practice, Performance* (S. 83-89). Cambridge: Cambridge University Press.

Bockstedt, J. C., Kauffman, R.J., & Riggins, F.J. (2006). The Move to Artist-Led On-Line Music Distribution: A Theory-Based Assessment and Prospects for Structural Changes in the Digital Music Market. *International Journal of Electronic Commerce*, 10(3), 7-38.

Brynjolfsson, E., & McAfee, A. (2017). The Business of Artificial Intelligence. What it can – and cannot – do for your organization. Harvard Business Review Online. Abgerufen am 13.05.2019 von <https://hbr.org/cover-story/2017/07/the-business-of-artificial-intelligence>.

Buxmann, P., & Schmidt, H. (2019a). Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens. In P. Buxmann & H. Schmidt (Hrsg.), *Künstliche Intelligenz: Mit Algorithmen zum Erfolg* (S. 3-19). Darmstadt: Springer.

Buxmann, P., & Schmidt, H. (2019b). Ökonomische Effekte der Künstlichen Intelligenz. In P. Buxmann & H. Schmidt (Hrsg.), *Künstliche Intelligenz: Mit Algorithmen zum Erfolg* (S. 21-37). Darmstadt: Springer. V BVMI (2018): *Bundesverband Musikindustrie: Musikindustrie in Zahlen 2017* (Print-Version). Berlin: BVMI.

Carstensen, A. (2018). Big Data in der Praxis – Marketing in der Musikindustrie: Eine Momentaufnahme des Social Web1. In M. Ahlers, L. Grünwald-Schukalla, M. Lücke, & M. Rauch (Hrsg.), *Big Data und Musik: Jahrbuch für Musikwirtschafts- und Musikkulturforschung 1/2018* (S. 129-151). Hannover: Springer.

Celma, Ò. (2010). *Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.

Dahm, M. H., Constantine, B., Gebauer, Tim, & Bartsch, B. (2019). Disruptive Geschäftsmodeländerungen: Uberisierung – Teil B: Fallstudien Spotify und Paypal: Wie neue Anbieter etablierte

Marktteilnehmer schlagen können. *Zeitschrift für Corporate Governance*. (02), 61–68.

Datta, H., Knox, G., & Bronnenberg, B. J. (2017). Changing Their Tune: How Consumers' Adoption of Online Streaming Affects Music Consumption and Discovery. *Marketing Science*, 37(1), 5–21. doi: 10.1287/mksc.2017.1051.

Dieleman, S. (2014). Recommending music on Spotify with deep learning. Abgerufen von <http://benanne.github.io/2014/08/05/spotify-cnns.html>.

D'onfro J. (2014). This man makes \$23,000 posting music spam on Spotify and iTunes. Business Insider. Abgerufen am 08.08.2019 von <http://www.businessinsider.com/matt-farley-makes-23k-posting-musicspam-on-spotify-2014-1>.

Dorn, T., & Rojas, R. (2018). Die können was! Aber können Roboter auch fühlen? *Aus Politik und Zeitgeschichte*, 68(6-8), 4-7.

Dredge, S. (2017): AI and music: will we be slaves to the algorithm? The Guardian. Abgerufen am 01.06.2019 von <https://www.theguardian.com/technology/2017/aug/06/artificial-intelligence-and-will-we-be-slaves-to-the-algorithm>.

Dredge, S. (2018): AI-music startup Popgun evolves: 'We're gonna make hit pop songs...'. Abgerufen am 10.12.2019 von <https://musically.com/2018/11/06/ai-music-startup-popgun-evolves-hit-songs/>.

Dredge, S. (2019): Startup Popgun shows how it's taught an AI to sing. Abgerufen am 10.12.2019 von <https://musically.com/2019/07/22/startup-popgun-shows-how-its-taught-an-ai-to-sing/>.

Eberl, U. (2018). Was ist künstliche Intelligenz - Was kann sie leisten? *Aus Politik und Zeitgeschichte*, 68(6-8), 8-14.

Gödde, J., Jacke, C., & Nösner, D. (2018): Abschlussbericht zum 2. Summer Institute der Gesellschaft für Musikwirtschafts- und Musikkulturforschung Developments for Musicpreneurs: 14.–20. September 2016, Porto, Portugal. In M. Ahlers, L. Grünewald-Schukalla, M. Lücke, & M. Rauch (Hrsg.), *Big Data und Musik: Jahrbuch für Musikwirtschafts- und Musikkulturforschung 1/2018* (S.



213-216). Hannover: Springer.

Gürtler, O. (2019). Künstliche Intelligenz als Weg zur wahren digitalen Transformation. In: P. Buxmann, & H. Schmidt (Hrsg.), *Künstliche Intelligenz: Mit Algorithmen zum Erfolg* (S. 95-105). Darmstadt: Springer.

Heim, M. Graefe, A. Brosius, H.-B. (2018): Wertschöpfung mithilfe von Algorithmen. Ansatzpunkte für die Veränderung von Geschäftsmodellen durch Computational Journalism. *MW: Medienwirtschaft*, 3(1), 36-42.

Helmholz, P., Meyer, M., & Robra-Bissantz, S. (2019). Moosic: Von der content- zur kontextbasierten Musikempfehlung. *HMD Praxis der Wirtschaftsinformatik*, 56(3), 598–611. doi: 10.1365/s40702-019-00526-5.

Hess, T. (2006). Plattenlabels, Verlage und Fernsehstationen: Wie können sie langfristiger erfolgreich sein? In T. Hess, & S. Doeblin (Hrsg.), *Turbulenzen in der Tele-kommunikations- und Medienindustrie: Neue Geschäfts- und Erlösmodelle* (S. 251-273). Berlin: Springer.

Herbrich, R. (2019). Künstliche Intelligenz bei Amazon Spitzentechnologie im Dienste des Kunden. In P. Buxmann & H. Schmidt (Hrsg.), *Künstliche Intelligenz: Mit Algorithmen zum Erfolg* (S. 63-75). Darmstadt: Springer.

Hildesheim, W., & Michelsen, D. (2019). Künstliche Intelligenz im Jahr 2018 – Aktueller Stand von branchenübergreifenden KI-Lösungen: Was ist möglich? Was nicht? Beispiele und Empfehlungen. In P. Buxmann, & H. Schmidt (Hrsg.), *Künstliche Intelligenz: Mit Algorithmen zum Erfolg* (S. 119-142). Darmstadt: Springer.

IFPI (2019): *International Federation of the Phonographic Industry: Global Music Report 2019: State of the Industry* (Print-Version). London: IFPI.

Jensen, L. (2015). Talentsuche per App. Brand eins Wirtschaftsmagazin. Abgerufen am 05.08.2019 von <https://www.brandeins.de/archiv/2015/talent/shazamtalentsuche-per-app>.

Kathan, W., Matzler, K., & Veider, V. (2016). The sharing economy: Your business model's friend

or foe? *Business Horizons*, 59(6), 663-672.

Kehl, C. (2018). Entgrenzungen zwischen Mensch und Maschine, oder: Können Roboter zu guter Pflege beitragen? *Aus Politik und Zeitgeschichte*, 68(6-8), 22-28.

Knopper, S. (2014). Can Shazam Predict the Next Big Hit? Industry insiders look to song identification app as new measure of success. *Rolling Stone*. Abgerufen am 07.08.2019 von <http://www.rollingstone.com/music/news/can-shazam-predict-the-next-big-hit-20140220>.

Koren, Y., & Bell, R. (2015). Advances in Collaborative Filtering. In F. Ricci, L. Rokach, B. Shapira (Hrsg.), *Recommender Systems Handbook* (Vol. 9, S. 77–118). Boston, MA: Springer US.

Krüger, W. (2002). Auswirkungen des Internet auf Wertketten und Geschäftsmodelle. In E. Frese, & H. Stöber (Hrsg.), *E-Organisation: Strategische und organisatorische Herausforderungen des Internet* (S. 63-89). Wiesbaden: Gabler.

Kühl, E. (2017). Künstliche Intelligenz: KI will rock you. *Zeit-Online* Abgerufen am 13.05.2017 von <https://www.zeit.de/digital/internet/2017-12/kuenstliche-intelligenz-musik-produktion-melodrive/komplettansicht>.

Le, L.K. (2014). Examining the Rise of Hatsune Miku: The First International Virtual Idol. *The UCI Undergraduate Research Journal*, 17(1), 1-12.

Léveillé Gauvin, H. (2017). "Drawing listener attention in popular music: Testing five musical features arising from the theory of attention economy". *Musicae Scientiae*, 22(3), 291-304.

Loeffler, J. (2018). 3 Start-ups Leading the Musical AI Revolution: The rapid growth of artificial intelligence is going to revolutionize the way we listen to and make music in the future. Abgerufen am 30.05.2019 von <https://www.interestingengineering.com/3-start-ups-leading-the-musical-ai-revolution>.

Luening, O. (1968). An Unfinished History of Electronic Music. *Music Educators Journal*, 55(3), 42-49, 135-142, 145.

Mayer, C.P. (2018). Künstliche Intelligenz und Maschinelles Lernen: Hintergrund, An-

wendungsfälle und Chancen für Medienunternehmen. In: *MW: Medien Wirtschaft*, 3(1), 30-35.

McKinsey (2017): Smartening up with Artificial Intelligence. McKinsey & Company Abgerufen am 10.05.2019 von <https://www.mckinsey.com/~media/McKinsey/Industries/Semiconductors/Our%20Insights/Smartening%20up%20with%20artificial%20intelligence/Smartening-up-with-artificial-intelligence.ashx>.

Misselhorn, C. (2018). Maschinenethik und "Artificial Morality": Können und sollen Maschinen moralisch handeln? *Aus Politik und Zeitgeschichte*, 68(6-8), 29-33.

Moritz, F., & Mohr, D. (2019). Künstliche Intelligenz & Urheberrecht: Bedroht KI die Arbeit von Fotografen? Copytrack. Abgerufen am 08.07.2019 von <https://www.copytrack.com/de/kuenstliche-intelligenz-urheberrecht/>.

Morris, J.W. (2015). Curation by code: Infomediaries and the data mining of taste. *European Journal of Cultural Studies*, 18(4-5), 446-463.

Moser, D., Wecht, C.H., & Gassmann, O. (2019). Digitale Plattformen als Geschäftsmodell. *ERP Management*, 15(1), 45-48.

Musically (2017): Playing next door to Alice: Popgun reveals its music AI. Musically. Abgerufen am 28.05.2019 von In: <https://musically.com/2017/06/30/playing-alice-popgun-music-ai/>.

Musicstats (2019). Research Projects. SusicStats. Abgerufen am 08.07.2019 von <https://musicstats.org/research-projects/>.

Müller, U. (2018): Talententdeckung und -förderung im Zeitalter von Big Data: Der Einfluss von Datenanalyse auf das A&R Management der Musikindustrie. In: M. Ahlers, L. Grünwald-Schukkalla, M. Lücke, & M. Rauch (Hrsg.), *Big Data und Musik. Jahrbuch für Musikwirtschafts- und Musikkulturforschung 1/2018* (S. 153-173). Hannover: Springer.

Nambisan, S. (2016). Digital Entrepreneurship: Toward a Digital Technology Perspective of Entrepreneurship. *Entrepreneurship Theory and Practice*, 41(6), 1029-1055.

Porter, M.E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. New

York: Free Press.

Popgun (2019). About us. Popgun. Abgerufen am 30.05.2019 von <http://popgun.ai/x/hello.html>.

Prey, R. (2015). *“Now Playing. You”: Big Data and the Production of Music Streaming Space* (Dissertation). Burnaby: Simon Fraser University.

Prey, R. (2018). Knowing Me, Knowing You: Datafication on Music Streaming Plat-forms. In M. Ahlers, L. Grünewald-Schukalla, M. Lücke, & M. Rauch (Hrsg.), *Big Data und Musik: Jahrbuch für Musikwirtschafts- und Musikkulturforschung 1/2018* (S. 9-21). Hannover: Springer.

Schedl, M., Knees, P., McFee, B., Bogdanov, D., & Kaminskis, M. (2015). Music Rec-ommender Systems. In F. Ricci, L. Rokach, & B. Shapira (Hrsg.), *Recommender Systems Handbook* (Vol. 32, S. 453-492). Boston, MA: Springer US.

Schedl, M., Zamani, H., Chen, C.-W., Deldjoo, Y., & Elahi, M. (2018). Current challenges and visions in music recommender systems research. *International Journal of Multimedia Information Retrieval*, 7(2), 95–116. doi: 10.1007/s13735-018-0154-2.

Sonicwire (2019). VOCALOID: Virtual Singer Software Line-Up. Abgerufen am 08.07.2019 von <https://sonicwire.com/product/vocaloid/special/index?lang=en>.

Spotify For Brands (2019). *Branded Moments*. Abgerufen von <https://www.spotifyfor-brands.com/en-US/ad-experiences/branded-moments/>.

Streibich, K.-H., & Zeller, M. (2019). Offene Plattformen als Erfolgsfaktoren für Künstliche Intelligenz. In: P. Buxmann, & H. Schmidt (Hrsg.), *Künstliche Intelligenz: Mit Algorithmen zum Erfolg* (S. 107-117). Darmstadt: Springer.

Tatchell, J. (2005). Together in electric dreams. The Guardian. Abgerufen am 01.09. 2019 von <https://www.theguardian.com/music/2005/jan/17/popandrock>.

Thompson, D. (2014). The Shazam Effect. The Atlantic. Abgerufen am 06.08.2019 von <https://www.theatlantic.com/magazine/archive/2014/12/the-shazam-effect/382237/>.

Tigre Moura, F. (2019). Künstliche Intelligenz ist der neue Mozart: Willkommen in der Welt unendlicher Musik. PcWelt. Abgerufen am 08.07.2019 von <https://www.pc-welt.de/ratgeber/Kuenstliche-Intelligenz-ist-der-neue-Mozart-Willkommen-in-der-Welt-unendlicher-Musik-10554517.html>.

Tschirren, J. (2018). Musik: Wenn der Algorithmus in die Tasten greift. Abgerufen am 03.06.2019 von <https://www.srf.ch/radio-srf-3/digital/musik-wenn-der-algorithmus-in-die-tasten-greift>

Tschmuck, P. (2003). *Kreativität und Innovation in der Musikindustrie*. Innsbruck: Studienverlag.

Tschmuck, P. (2009). Vom Tonträger zur Musikdienstleistung - Der Paradigmenwechsel in der Musikindustrie. In G. Gensch, E.M. Stöckler, & P. Tschmuck (Hrsg.), *Musikrezeption, Musikdistribution und Musikproduktion: Der Wandel des Wert-schöpfungsnetzwerks in der Musikwirtschaft* (S. 141-162). Wiesbaden: Gabler.

Tschmuck, P. (2017). *The Economics of Music*. London: Agenda Publishing.

UrhG (2017). Urheberrechtsgesetz: Gesetz über Urheberrecht und verwandte Schutz-rechte. Vom 09.09.1965 (letzte Änderung: 01.03.2017). Berlin: BRD.

Voigt, K.-I., Buliga, O., & Michl, K. (2017). Passion for Music: The Case of Spotify. In K.-I. Voigt, O. Buliga, & K. Michl (Hrsg.), *Management for Professionals. Business Model Pioneers* (Vol. 7, S. 143–155). Cham: Springer International Publishing.

Vonderau, P. (2019). The Spotify Effect: Digital Distribution and Financial Growth. *Television & New Media*, 20(1), 3–19. doi: 10.1177/1527476417741200.

Weihser, R. (2018). Auktion: Und das soll jetzt Kunst sein? Zeit-Online. Abgerufen am 06.07.2019 von <https://www.zeit.de/kultur/kunst/2018-10/auktion-kuenstliche-intelligenz-algorithmus-kunstwerk-versteigerung-christie-s>.

Wess, S. (2019). Mit Künstlicher Intelligenz immer die richtigen Entscheidungen treffen. In P. Buxmann & H. Schmidt (Hrsg.), *Künstliche Intelligenz: Mit Algorithmen zum Erfolg* (S. 143-159). Darmstadt: Springer.

Whitman, B. (2012). How music recommendation works - and doesn't work. Abgerufen von <https://notes.variogr.am/2012/12/11/how-music-recommendation-works-and-doesnt-work/>.

Zerdick, A.; Picot, A.; Schrape, K.; Artopé, A.; Goldhammer, K.; Heger, D. K.; Lange, U. T.; Vierkant, E.; López-Escobar, E.; Silverstone, R. (2001). *Die Internet-Ökonomie. Strategien für die digitale Wirtschaft*. Berlin: European Communication Council.

## 8. Chatbots in Geschäftsprozessen

Vortragstag und Gruppenname:	11. Juli 13:45 Uhr, Chatbots in Geschäftsprozessen
Titel der Ausarbeitung:	Chatbots in Geschäftsprozessen
Verfasserinnen:	Svenja Haupt, Piriyanga Karunanathan, Jelena Kempe
Betreuer/-in:	Univ.-Prof. Dr. Joachim Eigler (Betriebswirtschaftslehre, insb. Medienmanagement) M.A. Julia Müller

### Übersicht über die Struktur der Ausarbeitung

1. Problemstellung
  2. Entwicklung und Funktionen von Chatbots
  3. Barrieren bei der Mensch-Chatbot-Interaktion
    - 3.1 Die drei Hürden bei der Mensch-Chatbot-Interaktion
    - 3.2 Anthropomorphisierung und Personifizierung
    - 3.3 Fehlermanagement und Anreize zum Self-Service
  4. Konsequenzen für den Jobmarkt
  5. Anwendungsbereiche
    - 5.1 Anwendungsbeispiele von Chatbots in Geschäftsprozessen
    - 5.2 Chatbots im Recruiting
    - 5.3 Chatbots bei der Jobsuche
  6. Fazit
- Literaturverzeichnis
-

## 1. Problemstellung

Die Integration von Chatbots zur Optimierung von Geschäftsprozessen lässt sich seit einigen Jahren als Trend in Unternehmen identifizieren (Dale 2016, S. 811). Vor diesem Hintergrund besteht das Ziel dieser wissenschaftlichen Arbeit darin herauszufinden, ob Chatbots einen Mehrwert für Geschäftsprozesse bieten, sie womöglich menschliches Potenzial ersetzen und inwiefern sie in diesen Prozessen künstlicher Intelligenz entsprechen.

Um Intelligenz bei Chatbots festzustellen, muss zunächst die Unterscheidung zwischen einer starken und einer schwachen Intelligenz getroffen werden (Buxmann & Schmidt 2019, S. 6; Bitcom 2017, S. 29). Die schwache Intelligenz bezeichnet die gezielte Entwicklung von Algorithmen für bestimmte, abgegrenzte Problemstellungen (Buxmann & Schmidt 2019, S. 6, Bitcom 2017, S. 29). Die Ansätze der starken Intelligenz versuchen, die Menschen bzw. die Vorgänge im Gehirn abzubilden und zu imitieren sowie Eigenschaften wie Bewusstsein und Empathie nachzubilden (Buxmann & Schmidt 2019, S. 6). Zum jetzigen Zeitpunkt jedoch verfügen Chatbots lediglich über eine schwache Intelligenz (Buxmann & Schmidt 2019, S. 6).

Auch das Verständnis, was Geschäftsprozesse beinhalten, ist für die Klärung der angegebenen Problemstellungen wichtig. Geschäftsprozesse lassen sich in primäre und sekundäre Aktivitäten unterteilen (Porter 1985, S.64). Zu den primären Aktivitäten zählen die Eingangslogistik, Operationen, Marketing/Vertrieb, Ausgangslogistik und Kundendienst, während Unternehmensinfrastruktur, Personalwirtschaft, Technologieentwicklung und Beschaffung den sekundären sowie unterstützenden Aktivitäten zugeordnet werden (Porter 1985, S.64). Grundsätzlich lassen sich Geschäftsprozesse dabei in drei Prozesstypen unterteilen, denen unterschiedliche Aufgaben zugeschrieben werden: Routine, Regel- und seltene bzw. einmalige Prozesse (z.B. Projekte) (Picot & Rohrbach 1995, S. 31-32).

Um die Funktion von Chatbots in Geschäftsmodellprozessen klären zu können, ist es wichtig, den Begriff zu definieren und zu verstehen. Der Begriff Chatbot setzt sich aus dem Wort „Chat“ (aus dem Englischen für „Plaudern“) und dem Wort „Bot“ (Abkürzung des Wortes „Robot“) zusammen (Braun 2003, S. 21; Stucki et al. 2018, S. 728). Ein Chatbot ist virtuell und verzichtet auf jegliche Art der Verkörperung (Lotze 2014, S. 38). Zudem kann der Mensch mit Hilfe von Chatbots verbal oder schriftlich in natürlicher Sprache mit dem Computer interagieren (Spierling & Luderschmidt 2018, S. 391). Chatbots greifen dabei auf sogenannte Knowledge Datenbanken zu (Braun 2003, S. 21). Sie bilden die Wissensbasis des Chatbots und spüren eine



Übereinstimmung gestellter Fragen auf (Braun 2003, S. 21).

## **2. Entwicklung und Funktionen von Chatbots**

In den 60er Jahren begann die Entwicklung von Chatbots mit Joseph Weizenbau, der zwischen 1964 und 1966 das Programm ELIZA am MIT entwickelte (Storp 2002, S. 9). ELIZA ermöglichte erstmals die Kommunikation in natürlicher Sprache zwischen einem Menschen und einer Maschine und gilt damit als erster Chatbot (Storp 2002, S. 9; Stucki et al. 2018, S. 729). ELIZA ahmte einen Psychotherapeuten nach und zielte darauf ab, den Menschen glauben zu lassen, er unterhalte sich mit einem echten Menschen. Das Programm galt als Durchbruch des maschinellen Lernens (Storp 2002, S. 11). Diese Aussagen wurden allerdings nach kurzer Zeit von der KI-Forschung revidiert, da erkannt wurde, dass es ein komplexeres Problem darstellt, natürliche Sprache zu verarbeiten, als zunächst angenommen wurde (Storp 2002, S. 11). Allerdings existiert ELIZA auch noch in der heutigen Zeit und bietet die Grundlage für viele weiterführende Chatbots im Internet (Spierling & Luderschmidt 2018, S. 391).

In den letzten Jahren hat sich gezeigt, dass die Nutzung und Entwicklung von Chatbots stark zugenommen hat, was auf große Fortschritte des maschinellen Lernens und der künstlichen Intelligenz zurückzuführen ist (Dale 2016, S. 811; Brynjolfsson & McAfee 2017, S. 5). Seit 2016 ist eine Trendbewegung zu verzeichnen, die bis heute anhält (Dale 2016, S. 811; Brynjolfsson & McAfee 2017, S. 5). Bei der Betrachtung der künstlichen Intelligenz von Chatbots lässt sich erkennen, dass sich die größten Fortschritte in zwei Bereichen ereignet haben: im Bereich der Wahrnehmung und der Kognition (Brynjolfsson & McAfee 2017, S. 4). Beispielhaft hierfür zu nennen sind Siri oder Alexa (Brynjolfsson & McAfee 2017, S. 4). Trotz großer Erfolge muss beachtet werden, dass sich die aktuelle Entwicklung von Chatbots noch in der Frühphase befindet, da die Forschung im Bereich nutzerorientierten Handelns von Chatbots (noch) sehr mangelhaft ist und dadurch Nutzerbedürfnisse häufig vom Chatbot ignoriert werden (Piccolo, Mensio & Alani 2019).

## **3. Barrieren bei der Mensch-Chatbot-Interaktion**

### **3.1 Die drei Hürden bei der Mensch-Chatbot-Interaktion**

Braun (2003, S. 8-9) definiert drei Hürden, die bei dem menschlichen Umgang mit Computern vermehrt aufkommen. Als erstes nennt er die Übersetzungs- bzw. Sprachproblematik, die besagt, dass Menschen anhand von breiten, verbalen, kontextbezogenen sowie nonverbalen Sprachsets kommunizieren (Braun 2003, S. 8-9). Diese können bei der Interaktion mit

Computern oder Bots nicht angewandt werden, da mit diesen lediglich wie durch ein Schlüsselloch kommuniziert werden kann (Braun 2003, S. 8-9). Folglich entstehen durch diese Form von Einschränkung in der Kommunikation und Interaktion Asymmetrien, die nur schwierig zu überwinden sind (Braun 2003, S. 8-9). Einer der ersten Chatbots, der diese Schlüsselloch-Problematik deutlich aufzeigte, war die bereits erwähnte Chatbot-Psychotherapeutin ELIZA von Joseph Weizenbaum (1960, S. 3), die aufgrund ihrer niedrig ausgeprägten Konversationsfähigkeit von ihren Patienten als Bot entlarvt wurde. Chatbots werden damals wie heute so programmiert, dass sie anhand erfasster Benutzermodelle die Anliegen des Nutzers interpretieren (Braun 2003, S. 8-9). Somit interagieren sie auf Basis bestimmter Schätzungen und Unsicherheiten mit Menschen, was wiederum einen Nährboden für Missverständnisse bieten kann (Braun 2003, S. 8-9). Von einer optimalen Mensch-Maschinen-Interaktion kann erst dann gesprochen werden, wenn das Interface an die Gewohnheiten der Menschen angepasst und entlang dieser modelliert wird (Braun 2003, S. 10-12).

Als zweite Hürde wird die Divergenz der Bedürfnisse beider Parteien genannt, die darin besteht, dass die Kriterien der Programmierung von Chatbots und die Kriterien der Nutzer für eine angenehme und effektive Nutzung von Chatbots nicht selten weit auseinanderliegen (Braun 2003, S. 10-12). Die Programmierung von Bots zielt primär auf den Punkt der Effektivität ab, wohingegen Menschen gerade im Bereich des Geschäftsprozesses des Kundenservices auf menschliche Attribute wie Sympathie und Empathie Wert legen (Braun 2003, S. 9). Diese sind von Bots bisher nur schwer zu imitieren (Braun 2003, S. 9). Auch wenn bei der Programmierung von Chatbots menschliche Attribute erstrebenswert sind, gab es bereits Chatbots, die an ihrer Menschlichkeit scheiterten, so auch der Chatbot „Anna“ von IKEA, der zu stark auf nicht relevante Fragen der Kunden einging, somit seinen Zweck nicht mehr ausreichend erfüllte und im Jahr 2016 abgestellt wurde (Brandtzaeg & Følstad 2018, S. 41). Die dritte Hürde bezieht sich auf die Komplexität verteilter Systeme und besagt, dass Nutzer durch die gestiegene Komplexität von Interfaces nicht mehr in der Lage sind, Prozesse gänzlich zu erfassen und zu verstehen. Dies wirkt sich wiederum auf die Effektivität der Nutzung aus (Braun 2003, S. 10).

### **3.2 Anthropomorphisierung und Personifizierung**

Bei der Betrachtung der von Braun (2003, S. 8-10) aufgeführten Hürden stellen insbesondere sprachliche und erfahrungsbezogene Asymmetrien im Kontext der Chatbots die bedeutendsten Herausforderungen dar. Um diese zu überwinden, wird zunehmend eine Anthropomorphisierung von Interfaces vollzogen, welche Nutzern Kommunikation und Interaktion in gewohnter Weise

ermöglicht (Wilson, Daugherty & Bianzino 2017, o.S.). Jedoch kann dies auch zu Missverständnissen und besonders im Fall von Unternehmen zu einer Fehlrepräsentation führen (Wilson, Daugherty & Bianzino 2017, o.S.). Der bekannteste Fall einer solchen Fehlrepräsentation durch einen Chatbot ist der Fall „Tay“ (Beran 2018, S. 42-43). Microsofts Twitter Chatbot Tay, wurde durch ihre Nutzer zum Antisemitismus trainiert und begann selber, antisemitische Kommentare zu verfassen. Microsoft geriet in eine Imagekrise und entfernte umgehend den Bot sowie alle Kommentare, die durch den Bot verfasst wurden (Beran 2018, S. 42-43).

Das Persönlichkeitstraining von Interfaces wie bspw. bei Cortana von Microsoft oder Siri von Apple hat sich zu einem Markenzeichen vieler erfolgreicher Unternehmen entwickelt (Wilson, Daugherty & Bianzino 2017, o.S.). Demzufolge werden die Philosophie und das Image von Unternehmen auf spielerische Weise an die Kunden oder Nutzer herangetragen (Wilson, Daugherty & Bianzino 2017, o.S.). Chatbots nehmen hierbei zunehmend die Rolle eines Unternehmensrepräsentanten, womit eine Form des Employer Brandings für Interfaces stattfindet (Wilson, Daugherty & Bianzino 2017, o.S.). Werden Beispiele wie Amazons Alexa und Apples Siri betrachtet, so sind beide Interfaces als hilfsbereite, weibliche Assistentinnen modelliert (Wilson, Daugherty & Bianzino 2017, o.S.). Diese Form der Anthropomorphisierung ist in erster Linie dem Wandel der Kundenkommunikation von einer „One-way Interaction“ zu einer „Two-way Relationship“ geschuldet (Wilson, Daugherty & Bianzino 2017, o.S.). Unternehmen überführen immer mehr ihre einst transaktionsorientierte in eine beziehungsorientierte Kundenkommunikation (Wilson, Daugherty & Bianzino 2017, o.S.). Dafür stellt eine Personifizierung von Interfaces nicht nur auf sprachlicher Ebene, sondern u.a. auch hinsichtlich der Interpretation des Charakters, Alters oder Ausdrucks einen entscheidenden Erfolgsfaktor der adäquaten Interaktion mit Kunden und Nutzern dar (Wilson, Daugherty & Bianzino 2017, o.S.). Nach Braun (2003, S. 60-61) wird die Form der Repräsentation durch Personifizierung von verschiedenen Kriterien beeinflusst. Ein geeignetes Beispiel für die immer weiter fortschreitende Personifizierung von virtuellen Agenten ist die virtuelle Influencerin „Kenna“ der Kosmetikmarke Essence des deutschen Kosmetikunternehmens Cos-nova (Gründel 2019, o.S.). Im Dezember 2019 veröffentlichte Essence Kenna als neue Praktikantin des Unternehmens. Das Design von Kenna erfolgte in einer Weise, dass sie die Zielgruppe anspricht und trotz aller Designmöglichkeiten möglichst sympathisch und nicht zu perfekt wirkt (Gründel 2019, o.S.).

Die Wahrnehmung der Persönlichkeit eines Interfaces hängt stark mit der Persönlichkeit des Nutzers zusammen, wobei ähnliche Persönlichkeitsmerkmale zwischen Nutzer und Chatbot als

positiver durch den Nutzer erachtet werden (Braun 2003, S. 60-61). Braun (2003, S. 60-61) beschreibt zudem den Designprozess eines Interfaces als eine Gratwanderung, da Chatbots einerseits jeder Nutzer-/Kundenpersönlichkeit gerecht werden müssen, aber im selben Zuge auch ein Maß an Authentizität wahren sollten. Auch wenn Kunden und Nutzer sich dessen bewusst sind, dass sie sich mit einem Interface im Austausch befinden, ist nach Braun (2003, S. 60-61) eine Anwendung sozialer Interpretation des Charakters des Interfaces unvermeidbar, weshalb eine Personifizierung entlang der eigenen bzw. der Unternehmensabsichten signifikant ist. Auch auf die von Menschen gewohnten Umgangsformen muss hierbei eingegangen werden. Um den Anforderungen des Beziehungsmarketings gerecht zu werden, müssen Chatbot-Programmierer auch Smalltalk-Elemente oder absichtlich eingebaute Unebenheiten mit in die Interface-Persönlichkeit integrieren (Braun 2003, S. 60-61). Zieht man hierbei den Aspekt der Mensch-Computer-Kommunikation vs. den Aspekt der höflichen menschlichen Konversation nach Reeves und Nass (1996, S. 29-32) heran, die sich an den nach Grice (1993, S. 249, 251) festgelegten vier Konversationskategorien orientieren, spielen Qualität, Quantität, Relevanz und Klarheit der Konversation hinsichtlich der subjektiven Wahrnehmung der Rezipienten sowohl bei der Kommunikation mit Menschen als auch mit Computern eine erhebliche Rolle (Reeves und Nass 1996, S. 29; Braun S. 2003, S. 59). Braun (2003, S. 59) bezieht die Ausführungen von Reeves und Nass (1996, S. 29-32) nochmals näher auf die Mensch-Chatbot-Kommunikation vs. die menschliche Konversation. Dabei muss die Kommunikation mit Chatbots, um der menschlichen Konversation zu ähneln, qualitativ hochwertig, jedoch nicht zu Exaktheit orientiert sein; die Menge der gesprochenen oder geschriebenen Worte sollte ausreichend sein, um den Sachverhalt zu vermitteln (Braun 2003, S. 59). Chatoptionen, die nicht anwendbar sind, sind zudem relevanzmindernd. Zur Klarheit trage eine einfache, aber verständliche Form der Sprache bei. Alles in allem kommt es bei einer Nichtbeachtung der vier Konversationskategorien (Braun 2003, S. 59; Reeves und Nass 1996, S. 29-32; Grice 1993, S. 249, 251) zu Fauxpas bei der Konversation, ob nun bei der Kommunikation von Mensch zu Mensch oder von Mensch zu Chatbot. Trotz des Bewusstseins, dass mit einem Chatbot kommuniziert wird, kann es so zu einer Kränkung des Gegenübers kommen (Braun 2003, S. 59).

Eine durch Stimme übermittelte Sprache spielt bei Interfaces, die als Assistenten, meist jedoch Assistentinnen fungieren, eine besondere Rolle (Braun 2003, S. 64-65). Die Interface-Programmierung bedient sich stark der Geschlechterstereotypen und versucht damit, den Gewohnheiten und somit den Erwartungen von Nutzern an Interfaces gerecht zu werden (Braun 2003, S. 64-65; Olsen 2019, o.S.). Hierbei gebe es eine eindeutige geschlechtergebundene Wahrnehmung

von Stimmen, durch welche Frauenstimmen als freundlicher und kompetenter wahrgenommen werden und sich sogar Reaktionen auf maskuline oder feminine Frauenstimmen stark unterscheiden (Braun 2003, S. 64-65; Olsen 2019, o.S.). Mit Stimmen werden so nach Braun (2003, S. 64-65) sowohl positive als auch negative Emotionen assoziiert. Um diese Problematik zu umgehen, arbeiten Unternehmen zunehmend mit verschiedenen Interface-Persönlichkeiten nach dem Prinzip „Good Cop - Bad Cop“, bei dem zwei oder mehrere Stimmen bestimmten Aufgabenbereichen im Kundenservice nachgehen und negativ assoziierte Angelegenheiten von einer anderen Persönlichkeit bearbeitet werden (Bsp.: Unzufriedenheit mit einem Produkt) als die durch Kunden als positiv wahrgenommenen Anliegen (Bsp.: Fragen zu Kundenbindungsprogrammen) (Braun 2003, S. 64-65). Das- selbe Prinzip ist auch auf das Design von rein textuellen Interfaces anwendbar und könnte dazu beitragen, dass Unternehmen nicht nur auf die Persönlichkeit eines einzigen Interfaces reduziert werden (Braun 2003, S. 64-65).

### **3.3 Fehlermanagement und Anreize zum Self-Service**

So sehr Interfaces den Verlauf alltäglicher Geschäftsprozesse auch entlasten mögen, sind sie zum jetzigen Zeitpunkt der Entwicklung nach wie vor fehleranfällig (Braun 2003, S. 68- 69). Um ein angemessenes Fehlermanagement zu gewährleisten, sollen Unternehmen einen konsistenten Multichannel-Kontakt nutzen, der mit Hilfe von Kanälen wie zum Beispiel Telefon, Mail und Chats den Kontakt mit Kundenservicemitarbeitern ermöglicht, sodass nicht erfolgreich beantwortete Self-Service-Anliegen effizient behandelt werden. Es findet so eine Entlastung des Servicepersonals statt, sodass dieses seine Zeit komplexeren Kundenanfragen widmen kann. Zudem werden die Knowledge-Datenbanken der Chatbots durch das Personal geschult bzw. mit immer wieder aufkommenden Frage-Antwort-Mustern „gefüttert“ (Braun 2003, S. 68-69). Um auf lange Sicht einen reibungslosen Übergang von Service durch Personal zu Service durch Bots zu erzielen, arbeiten Unternehmen wie bspw. Coca Cola mit Anreizen zum Self-Service. Dabei werden diejenigen Kunden in den Kundencenter-Warteschleifen bevorzugt, die zuvor ihre Anfrage über den Self-Service zu lösen versucht haben. Daneben werden Online-Tutorials zur Steigerung der Sichtbarkeit des Self-Services auf den Plattformen implementiert. Dies führt zur Senkung der Suchkosten für Kunden, was wiederum zufriedenheitssteigernd ist (Braun 2003, S. 68-69).

### **4. Konsequenzen für den Jobmarkt**

Da Chatbots Aufgaben in Geschäftsprozessen übernehmen können, liegt die Vermutung nahe, dass Berufsfelder von Mitarbeitern innerhalb eines Unternehmens wegfallen und durch Chatbots

ersetzt werden. Jedoch ist diese Sorge unbegründet (pers. Kommunikation Eißer, 28.05.2019). Chatbots sind noch nicht auf dem technischen Stand, dass sie in der Lage wären, ganze Berufsfelder zu ersetzen (pers. Kommunikation Eißer, 28.05.2019). Daher fallen keine Berufsfelder weg, sondern lediglich Aufgabenbereiche (Eißer Interview, 28.05.2019). Mitarbeiter sollen besonders in repetitiven Aufgaben durch Chatbots unterstützt und entlastet werden (Dudler 2019, S. 108), indem Anfragen durch diese möglichst schnell und zeitlich ungebunden (24/7) beantwortet werden können (Hoffmann 2018, S. 37; Dulder 2019, S. 103). Darüber hinaus zeigen sie keine Ermüdungserscheinungen und können mehrere Anfragen gleichzeitig beantworten (Knabenreich 2017, o.S.). Folglich kann insbesondere Zeit und Kosten eingespart werden (Freyth 2019, S. 296).

Jedoch bleibt zu hinterfragen, ob Chatbots für Unternehmen eine Investition oder eine Einsparung von Personalkosten darstellen. Pauschal kann nicht für jedes Unternehmen festgelegt werden, ob der Einsatz von Chatbots für das jeweilige Unternehmen eine Investition oder eine Einsparung bildet (Svanidze 2019, o.S.). Die Kosten für einen Chatbot hängen unter anderem von der Komplexität und den Fähigkeiten des Chatbots ab (Svanidze 2019, o.S.). Je mehr künstliche Intelligenz in einen Chatbot integriert ist, desto höhere Kosten fallen für diesen an (Svanidze 2019, o.S.). Wiederum könnten durch intelligentere Chatbots komplexere Aufgaben übernommen werden, durch die Aufgaben des Personals wegfallen (Svanidze 2019, o.S.). Somit könnten möglicherweise Personalkosten eingespart werden.

Die Konsequenzen durch den Einsatz von Chatbots scheinen für ausgelagerte Jobs deutlich höher auszufallen. Insbesondere im Bereich des Customer Services zum Beispiel im Bereich des telefonischen Kundenservices ist ein Wegfall von Jobs durchaus denkbar (Freyth 2019, S. 296). Personen, die in dieser Branche beschäftigt sind, müssen demzufolge eine gewisse Flexibilität aufweisen, um sich neu zu orientieren oder andere Aufgaben übernehmen zu können (pers. Kommunikation Eißer, 28.05.2019). Dieses Problem wird nicht ausschließlich durch Chatbots impliziert, sondern ist ein allgemeines Problem der Digitalisierung und Automatisierung (Hackl, Wagner, Attmer & Baumann 2017, S. 19).

## **5. Anwendungsbereiche**

### **5.1 Anwendungsbeispiele von Chatbots in Geschäftsprozessen**

Grundsätzlich lässt sich sagen, dass maschinelles Lernen den Wandel auf drei Ebenen vorantreibt und zwar im Bereich von Aufgaben und Berufen, in Geschäftsprozessen und in

Geschäftsmodellen (Brynjolfsson & McAfee 2017, S. 3). Buxmann und Schmidt betonen, dass die Bedeutung von Chatbots vor allem im Bereich des Customer Engagement und des Commerce zunehmen wird (Buxmann & Schmidt 2019, S. 51). Zudem lassen sich Chatbots sinnvoll in Call-Centern einsetzen und können auf einer Internetseite das Suchfeld ersetzen (Buxmann & Schmidt 2017, S. 135). Hier soll durch Chatbots der Kundendienst im Bereich des First-Level-Supports verbessert und positive Kundenerfahrungen hervorgerufen werden. Zu nennen sind hier Tinka der Webseite von T-Mobile Austria. Dieser Chatbot chattet per Text mit bis zu 50.000 Kunden und beantwortet 120.000 Fragen über Handys, Dienste sowie Rechnungen monatlich (von Hofe 2017, o.S.). Ein anderes Beispiel stellt der BMW Motorsport Media Chatbot via Whatsapp dar, der die Kunden über Zeitplan, Fahrer, Ergebnisse, usw. bei der DTM-Saison 2017 informierte (BMW 2017, o.S.). Der Chatbot übernimmt hier dementsprechend Routineaufgaben und entlastet somit die Mitarbeiter der Unternehmen.

Als ein weiteres Beispiel von Chatbots im Kundenservice ist die Schweizer Post zu nennen (Stucki et al. 2018, S. 733). Die Paketorganisation der Schweizerischen Post „PostLogistics“ hat das Innovationsprojekt „Custom Smart“ entwickelt (Stucki et al. 2018, S. 734). Ziel war es, den Dienstleistungsbereich „Paket empfangen“ zu optimieren. Dazu wurde ein Chatbot entwickelt, der es dem Kunden ermöglicht, die Zeit und den Ort der Zustellung zu bestimmen und Gebühren für den Zoll zu begleichen (Stucki et al. 2018, S. 734). Mit diesem Chatbot wurde die Erstzustellquote für zollpflichtige Pakete erhöht und Prozesskosten konnten erheblich reduziert werden (Stucki et al. 2018, S. 734). Somit ist auch dieser Chatbot der Entlastung von Routineaufgaben zuzuordnen. Auch unternehmensintern lassen sich Chatbots einsetzen wie der Chatbot von IT Post. Dieser soll allgemeine Probleme der Mitarbeiter klären und so den Aufwand der Arbeit und der Informationsfindung reduzieren (Stucki et al. 2018, S. 736). Somit erleichtert auch dieser Chatbot Routineaufgaben in Geschäftsprozessen. Weitere Beispiele für Chatbots im Kundenservice sind der Chatbot „Linda“ der Sparkasse (Sparkasse 2018, o.S.), der Chatbot „Kim“ von Maggi (Maggi O.J., o.S.) oder der Chatbot von Zalando (Zalando 2017, o.S.). Diese Beispiele zeigen, dass auch unterschiedliche Branchen mit verschiedenen Branchenschwerpunkten Chatbots einsetzen können. Chatbots können ebenfalls in der Entwicklung und Innovation eingesetzt werden. Ein Beispiel bietet die Londoner Brauerei IntelligentX Brewing. Diese hat mit Hilfe eines Chatbots, der die Vorlieben und den Geschmack der Kunden erfasst, vier neue Biersorten kreiert. Dies war mit Hilfe von immer spezifischer werdenden Fragen des Algorithmus möglich (KMPG AG 2018, S.25).

Im Folgenden werden Chatbots im Recruiting und bei der Jobsuche genauer betrachtet, das gewählte Beispiel stammt daher aus dem Bereich der Personalwirtschaft. Chatbots können zudem in Geschäftsprozessen zur Beratung eingesetzt werden. So können sie auf die Expertise eines Unternehmens hinweisen und darüber hinaus einen Beitrag zur Beratung in Hinblick auf spezifische Themen leisten (Lücke 2018, o.S.). Darüber hinaus können Chatbots ebenfalls zur Kommunikation innerhalb des Unternehmens eingesetzt werden. So können unter anderem Mitarbeiter mit der Personalabteilung über Chatbots kommunizieren, um bspw. abzufragen, wie viele Urlaubstage ein Mitarbeiter noch hat. Folglich kann der Servicegrad der Personalabteilung gesteigert werden (Köster 2017, o.S.).

## **5.2 Chatbots im Recruiting**

Chatbots können im Rahmen des Recruitings im Bewerbungsprozess eingesetzt werden (pers. Kommunikation Eißer, 28.05.2019). In diesem Bereich finden sie Einsatz in an die Bewerber gerichteten FAQ sowie vor und während der Bewerbung in sogenannten Bewerbermanagementsystemen, die eingesetzt werden, um große Mengen an Bewerberdaten zu kumulieren (Bsp.: Abb. 1 und 2) (pers. Kommunikation Eißer, 28.05.2019; Dahm & Dregger 2018, o.S.). Bewerber können den Chatbot in diesem Fall nutzen, um unternehmensbezogene oder stellenbezogene Informationen zu erhalten (pers. Kommunikation Eißer, 28.05.2019). Der Recruiter kann den Chatbot auch heranziehen, um Informationen über Bewerber zu entnehmen (pers. Kommunikation Eißer, 28.05.2019) oder um herauszufinden, ob eine Stellenanzeige viele Bewerber anzieht (Dahm & Dregger 2018, o.S.). Der Chatbot übernimmt dann die Aufgaben der Administration nach Bewerbungseingang sowie Routineaufgaben im Bewerbermanagementsystem (pers. Kommunikation Eißer, 28.05.2019). Hierzu zählen bspw. Terminvereinbarungen zu einem Bewerbungsgespräch, die Abfrage fehlender Dokumente oder Hilfestellungen während des Bewerbungsprozesses (Dudler 2019, S. 107). Allgemein bündelt der Chatbot die Routineaufgaben, damit der Recruiter sich auf einmalige sowie Regelaufgaben konzentrieren kann (pers. Kommunikation Eißer, 28.05.2019). In den nächsten Jahren werden Chatbots daher durchaus neben weiteren Anwendungen der Digitalisierung und Automatisierung im Recruiting eine Rolle spielen (Abb. 3) (meta HR Unternehmensberatung 2018, S. 16). Jedoch werden Maßnahmen bspw. Networking zukünftig noch größere Bedeutung im Bereich Recruiting zugeschrieben als der Implementierung von Chatbots (meta HR Unternehmensberatung 2018, S. 16). Dies lässt sich dadurch erklären, dass die Anwendung und Akzeptanz von Chatbots im Recruiting noch nicht intensiv erforscht sind. Dies zeigt bspw. das Forschungsprojekt „CATS - Chatbots in Applicant Tracking Systems“ der Hochschule RheinMain, Wiesbaden, das sich mit dieser Thematik



beschäftigt (Munz 2018, o.S.). Dort wird die Nutzung der Chatbots in verschiedenen Stufen des Bewerbungsprozesses sowie der Einsatz dieser in Systemen des Bewerbermanagements zur Unterstützung der Personalabteilung analysiert (Munz 2018, o.S.). Ziel des Projektes ist, einen Chatbot-Werkzeugkasten zu entwickeln, der an die Bewerbermanagementsysteme der Unternehmen angepasst werden kann, um den Unternehmen individuell zu helfen, standardisierte Bewerbungsprozesse zu verbessern (Munz 2018, o.S.).

Es gibt einerseits die Möglichkeit, eine Standardversion des Chatbots zu entwickeln, die alle Bewerber in gleicher Weise anschreibt und über eine sachliche sowie förmliche Schreibweise verfügt (pers. Kommunikation Eißer, 28.05.2019). Bewerber können jedoch ebenfalls segmentiert werden, indem untersucht wird, wie der Bewerber zu der Stellenanzeige gelangt ist und um welche Art von Stellenanzeige es sich handelt (pers. Kommunikation Eißer, 28.05.2019). Folglich können die Bewerber so einer bestimmten Zielgruppe zugeordnet werden. Beispielsweise werden junge Bewerber ungezwungen oder mit Emojis angeschrieben (per. Kommunikation Eißer, 28.05.2019). Wichtig wird dabei sein, diese Maßnahmen individuell auf jedes einzelne Unternehmen anzupassen, damit der Chatbot die Unternehmensmentalität widerspiegelt (pers. Kommunikation Eißer, 28.05.2019).

### **5.3 Chatbots bei der Jobsuche**

Aus der Sicht der Bewerber könnten Chatbots eingesetzt werden, um den Bewerbungsprozess möglichst einfach, unkompliziert und effizient zu gestalten und damit die sogenannte Candidate Experience zu erhöhen (Dudler 2019, S. 104). Der Chatbot bietet einem Bewerber die Möglichkeit, orts- und zeitunabhängig an Informationen zu kommen (Dudler 2019, S. 103). Demzufolge bietet der Chatbot dem Bewerber einen gewissen Komfort, denn der Bewerber kann sich so aussuchen, ob er die Informationen bspw. bevorzugt aus der Datenbank oder aus dem FAQ-Bereich zieht (pers. Kommunikation Eißer, 28.05.2019). Parallel kann der Recruiter kontaktiert werden, insbesondere dann, wenn so die Akzeptanz und das Vertrauen in den Bewerbungsprozess gesteigert werden kann (pers. Kommunikation Eißer, 28.05.2019). Die Anonymität eines Chatbots kann im Bewerbungsprozess eine geringere Hemmschwelle bei einem Bewerber hervorrufen, sodass diese sich mit Fragen an den Chatbot richten können, die sie in einem Bewerbungsinterview nicht direkt stellen würden (Dudler 2019, S. 107). Herauszustellen ist außerdem, dass der Einsatz eines Chatbots im Bewerbungsprozess nicht für den Bewerber bedeutet, dass er sich nicht bewerben kann, wenn er nicht mit einem Chatbot interagieren möchte (pers. Kommunikation Eißer, 28.05.2019). Jedoch kann ebenfalls eine gewisse Angst aufseiten der

Bewerber bestehen, denn viele sind bereits mit Chatbots in Kontakt gekommen und wissen, dass Chatbots noch keinen hohen technischen Standard erreicht haben (pers. Kommunikation Eißer, 28.05.2019). Um Zeit und Aufwand zu sparen, wenden sie sich daher präferiert an die herkömmlichen Bewerbungsangebote (pers. Kommunikation Eißer, 28.05.2019). Andere Bewerber möchten nicht persönliche und sensible Daten an eine Maschine weitergeben, sondern teilen diese lieber mit einem Menschen (Dahm & Dregger 2018, o.S.).

Im Recruiting stellen Chatbots einen Vorteil für die Bewerber dar, da es sich um ein nicht wertend agierendes System handelt (pers. Kommunikation Eißer, 28.05.2019). Demzufolge werden Bewerber nicht nach Aussehen oder Herkunft, sondern nach der Erfüllung geforderter Kriterien bewertet (Knabenreich 2017, o.S.; Dahm & Dregger 2018, o.S.). Darüber hinaus findet keine Alters- oder Geschlechterdiskriminierung durch Chatbots statt (Dudler 2019, S. 106). Manche sehen jedoch auch die Gefahr, dass insbesondere in diesem Bereich Fehler von Chatbots angelernt werden oder Begründungen für Entscheidungen fehlen könnten (pers. Kommunikation Eißer, 28.05.2019).

## **6. Fazit**

Zu Beginn wurden Geschäftsprozesse in drei Bereiche unterteilt (Picot & Rohrbach 1995, S. 31-32). Es hat sich durch die zuvor genannten Beispiele gezeigt, dass Chatbots bisher größtenteils Routineaufgaben übernehmen und weniger Regel- oder einmalige Prozesse ersetzen. Dies ist darauf zurückzuführen, dass Chatbots bisher nur schwache künstliche Intelligenz aufweisen (Buxmann & Schmidt 2019, S. 6). Somit lässt sich die zu Beginn gestellte Fragestellung, ob Chatbots KI enthalten, damit beantworten, dass lediglich die Rede von schwacher KI sein kann, zumindest bis zum heutigen Stand. Es bleibt jedoch zu vermuten, dass bei einer fortschreitenden technologischen Entwicklung von Chatbots, diese in Zukunft auch starke Intelligenz aufweisen und somit auch andere Bereiche in Geschäftsprozessen als vollwertige KI ersetzen können (Dale 2016, S. 811; Brynjolfsson & McAfee 2017, S.5).

Ein weiteres Ziel dieser Arbeit war es herauszufinden, ob Chatbots menschliche Arbeitskraft ersetzen können. Vor allem durch die Anwendungsbeispiele von Chatbots in Geschäftsprozessen hat sich gezeigt, dass Bots in vereinzelten Bereichen durchaus bereits menschliche Aufgabefelder übernehmen. Ein Beispiel dafür ist so auch das durch Chatbots geführte Bewerbungsverfahren im Recruiting-Bereich. Doch hier zeigt sich auch, dass menschliches Personal stets zur Verfügung stehen muss, um tiefer gehende Anliegen der Bewerber beantworten zu können. Im

Recruiting hat sich gezeigt, dass Bewerber häufig auch den persönlichen Kontakt suchen und bevorzugen (pers. Kommunikation Eißer, 28.05.2019). So sehr Chatbotentwickler und Unternehmen auch versuchen, Chatbots menschlicher zu gestalten, können diese bisher Attribute wie Einfühlungsvermögen oder Empathie nicht den Erwartungen der Nutzer entsprechend nachahmen (Braun 2003, S. 10-12; Wilson, Daugherty & Bianzino 2017, o.S.).

Zu der Frage, ob Chatbots nun einen Mehrwert für Geschäftsprozesse bieten, lässt sich sagen, dass sie zum jetzigen Stand durchaus zur Entlastung des menschlichen Personals beitragen und besonders zeitaufwendige Aufgaben, die repetitiv sind und nicht sehr viel geistiges Vermögen benötigen, effizient erfüllen. Im Servicebereich bspw. hat sich herausgestellt, dass Chatbots in der Lage sind, sich repetitiven Fragen, die sonst auch in der Rubrik FAQ beantwortet werden, zu widmen. So kommen Chatbots vermehrt in Kundencentern zum Einsatz, da dort einfache Fragen schnell beantwortet werden können (Stucki et al. 2018, S. 737). Kundenanliegen können in gleichbleibender Frequenz und Qualität zu jeder Uhrzeit bearbeitet werden (Stucki et al. 2018, S. 737). Zudem ermöglicht dies, dass sich das menschliche Personal geistig anspruchsvolleren Anliegen seitens der Kunden widmen kann und somit auch dort die Qualität auf einem hohen Niveau gehalten werden kann (Braun 2003, S. 68-69). Bei der Schweizerischen Post bspw. zeigt sich, dass immer mehr Probleme mit Hilfe von Chatbots gelöst werden anstatt mit Apps oder Webseiten allein (Stucki et al. 2018, S. 738). Trotzdem werden komplexere Anliegen nach wie vor von menschlichem Personal bearbeitet (Braun 2003, S. 68-69). Zu komplexeren Anliegen gehören auch Unzufriedenheit mit Produkten oder Dienstleistungen, wobei Kunden unter Umständen nicht die Geduld aufbringen werden, mit einem Chatbot zu kommunizieren. In solchen Fällen bevorzugen sie es, wenn das Problem möglichst reibungslos mit Einfühlungsvermögen und Kulanz seitens des Unternehmens gelöst wird (Braun 2003, S. 68-69). Diese Arbeit hat gezeigt, dass Chatbots durchaus in Geschäftsprozesse Einzug gehalten haben, ihre Intelligenz aber weiter verbessert werden und ihre Akzeptanz gesteigert werden muss, um in Zukunft auch in komplexere Aufgabenbereichen eingesetzt werden zu können.

---

## Literaturverzeichnis

Beran, O. (2018). An Attitude Towards an Artificial Soul? Responses to the “Nazi Chatbot”. In: *Philosophical Investigations* 41, 2018, 1, S. 42–69.

BITKOM e. V. (2017). Entscheidungsunterstützung mit Künstlicher Intelligenz.: *Wirtschaftliche*

Bedeutung, gesellschaftliche Herausforderungen, menschliche Verantwortung. Berlin. Aufgerufen am 09.12.2019 von <https://www.uni-kassel.de/fb07/fileadmin/datas/fb07/5-Institute/IWR/Hor-nung/170901-KI-Gipfelpapier-online.pdf>.

BMW Group PressClub Deutschland (2017). Jetzt verfügbar: der BMW Motorsport Media Chatbot. Aufgerufen am 30.11.2019 von <https://www.press.bmwgroup.com/deutschland/article/detail/T0269871DE/jetzt-verfuegbar:-der-bmw-motorsport-media-chatbot?language=de>.

Bodrunova, S. S. et al. (Hrsg.). (2019). Internet Science. INSCI 2018 International Workshops, St. Petersburg, Russia, October 24-26, 2018, Revised selected papers. Cham, Switzerland.

Brandtzaeg, P. B., Følstad, A. (2018). Chatbots. In: Interactions 25, 2018, 5, S. 38–43. Braun, A. (2003). Chatbots in der Kundenkommunikation. Berlin/Heidelberg: Springer.

Brynjolfsson, E. & McAfee, A. (2017). The business of artificial intelligence. Harvard Business Review Online. Aufgerufen am 30.11.2019 von <https://starlab-alliance.com/wp-content/uploads/2017/09/The-Business-of-Artificial-Intelligence.pdf>.

Buxmann, P., Schmidt, H. (2019). Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens. In: Buxmann, P. u. a. (Hrsg.). Künstliche Intelligenz. Mit Algorithmen zum wirtschaftlichen Erfolg. Berlin, Heidelberg: Springer, S. 3–19.

Buxmann, P., Schmidt, H. (Hrsg.). (2019). Künstliche Intelligenz: Mit Algorithmen zum wirtschaftlichen Erfolg. Berlin, Heidelberg: Springer.

Dale, R. (2016). The return of the chatbots. In: Natural Language Engineering 22, 5, S. 811–817.

Dahm, M., & Dregger, A. (2018). KI im Bewerbungsprozess. Künstliche Intelligenz: Schöne neue Welt? Aufgerufen am 11.12.2019 von <https://www.personalwirtschaft.de/recruiting/artikel/kuenstliche-intelligenz-im-bewerbungsprozess.html>.

Dudler, Luc (2019). Wenn Bots übernehmen – Chatbots im Recruiting. In: Tim Verhoeven (Hrsg.): Digitalisierung Im Recruiting. Wie Sich Recruiting durch Künstliche Intelligenz, Algorithmen und Bots Verändert. Wiesbaden: Springer Gabler, S. 101–111.

Freyth, A. (2019). Persönliche Veränderungskompetenz und Agilität stärken: Praxisleitfaden für Mitarbeiter und Führungskräfte. Wiesbaden: Springer.

Geschäftsbericht Zalando (2017). Gestalte deinen Chatbot. Aufgerufen am 30.11.2019 von <https://geschaeftsbericht.zalando.de/2017/gestalte-deinen-chatbot/> Grice, H. P. (1993). Logik und Konversation. In: Meggle, Georg (Hrsg.). Handlung, Kommunikation, Bedeutung. Mit einem Anhang zur Taschenbuchausgabe 1993. Frankfurt am Main: Suhrkamp, S. 234–265.

Gründel, V. (2019). Instagram: Essence hat jetzt eine virtuelle Influencerin. W&V Online. Aufgerufen am 9.12.2019 von [https://www.wuv.de/wuvplus/instagram\\_essence\\_hat\\_jetzt\\_eine\\_virtuelle\\_influencerin](https://www.wuv.de/wuvplus/instagram_essence_hat_jetzt_eine_virtuelle_influencerin).

Hackl, B., Wagner, M., Attmers, L. & Baumann, D. (2017). New Work: Auf dem Weg zur neuen Arbeitswelt. Management-Impulse, Praxisbeispiele, Studien. Wiesbaden: Springer Verlag.

Hoffmann, A. (2018). Chatbots. Einführung in die Zukunft von Marketing, PR und CRM. Haar bei München: Franzis Verlag.

KMPG AG Wirtschaftsprüfungsgesellschaft Studie (2018). Wertschöpfung neu gedacht: Von Humanoiden, KIs und Kollege Roboter [Studie]. Aufgerufen am 30.11.2019 von [https://hub.kpmg.de/hubfs/LandingPages-PDF/KPMG\\_Studie\\_Artif\\_Intelligence\\_April\\_2017\\_BF\\_SEC.pdf?utm\\_campaign=Studie%20-%20Wertsch%C3%B6pfung%20neu%20gedacht%20-%20KI%20&utm\\_source=hs\\_automation&utm\\_medium=email&utm\\_content=62078628&\\_hsenc=p2ANqtz-8GevShd-21Em8Xr2LYZnvnvZTcs6b3d8x47ACk-4GC7-1dtWvvDaxshFmLdaoL7b9JP9n0ZRAR4ix0dxVG9ljXAZe\\_qFxxuM4kxLhJ5NjUB6vO2pM&\\_hsmi=62078628](https://hub.kpmg.de/hubfs/LandingPages-PDF/KPMG_Studie_Artif_Intelligence_April_2017_BF_SEC.pdf?utm_campaign=Studie%20-%20Wertsch%C3%B6pfung%20neu%20gedacht%20-%20KI%20&utm_source=hs_automation&utm_medium=email&utm_content=62078628&_hsenc=p2ANqtz-8GevShd-21Em8Xr2LYZnvnvZTcs6b3d8x47ACk-4GC7-1dtWvvDaxshFmLdaoL7b9JP9n0ZRAR4ix0dxVG9ljXAZe_qFxxuM4kxLhJ5NjUB6vO2pM&_hsmi=62078628).

Köster, S. (18. Dezember 2017). Automatisierung von Geschäftsprozessen mit Chatbots. Aufgerufen am 30.11.2019 von <https://www.koester-econsulting.com/post/prozesse-Chatbots> [Web Log Eintrag].

Lotze, N. (2016). Chatbots. Eine linguistische Analyse. Frankfurt a.M: Peter Lang.

Lücke, B. (13. Dezember 2018): Die Zukunft von Chatbots: Wie lassen sie sich wirklich sinnvoll

einsetzen? Aufgerufen am 30.11.2019 von <https://borisluecke.de/2018/12/13/die-zukunft-von-chatbots-wie-lassen-sie-sich-wirklich-sinnvoll-einsetzen/> [Web Log Eintrag].

Maggi Chatbot Kim (2017). Aufgerufen am 30.11.2019 von <https://www.maggi.de/artikel/chatbot-kim>.

Meggel, G. (Hrsg.). (1993). *Handlung, Kommunikation, Bedeutung*. Mit einem Anhang zur Taschenbuchausgabe 1993. Frankfurt am Main: Suhrkamp Verlag.

Munz, M. (2018). CATS: Forschungsprojekt zu Chatbot-Systemen im Personalrecruiting startet. Aufgerufen am 30.11.2019 von <https://www.hs-rm.de/de/hochschule/aktuelles/details/artikel/cats-forschungsprojekt-zu-chatbot-systemen-im-personalrecruiting-startet/>.

Olsen, P. (2019). Banks Are Promoting 'Female' Chatbots To Help Customers, Raising Concerns Of Stereotyping. *Forbes Magazine Online*. Aufgerufen am 9.12.2019 von <https://www.forbes.com/sites/parmyolson/2019/02/27/banks-are-promoting-female-chatbots-to-help-customers-raising-concerns-of-stereotyping/#6f31027c5a8f>.

Personalmarketing 2Null (10. Juli 2017). Mit Chatbots im Recruiting zur perfekten Candidate Experience. Aufgerufen am 30.11.2019 von <https://personalmarketing2null.de/2017/07/10/chatbots-recruiting-candidate-experience/> [Web Log Eintrag].

Petry, T., Jäger, W. (Hrsg.). (2018). *Digital HR. Smarte und agile Systeme, Prozesse und Strukturen im Personalmanagement*. Freiburg/München/Stuttgart: Haufe.

Piccolo, L. S. G., Mensio, M. & Alani, H. (2019). Chasing the Chatbots. In: Bodrunova, S. S. et al. (Hrsg.). *Internet Science. INSCI 2018 International Workshops, St. Petersburg, Russia, October 24-26, 2018, Revised selected papers*. Cham, Switzerland, S. 157– 169.

Picot, A. & Rohrbach P. (1995). In: *Information Management* 10, 1995, 1, S. 28–35.

Porter, M. E. (1998). *Competitive advantage. Creating and sustaining superior performance; with a new introduction*. New York: The Free Press.

Reeves, B. & Nass, C. (1996). The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Pla. In: Bibliovault OAI Repository, the University of Chicago Press, 1996.

Sparkasse Chatbot Linda (2018). Neu: Chatbot Linda in der Chat Suite. Aufgerufen am 30.11.2019 von [https://www.sparkassen-finanzportal.de/news/Chatbot-Linda-in-der-Chat-Suite\\_.html](https://www.sparkassen-finanzportal.de/news/Chatbot-Linda-in-der-Chat-Suite_.html).

Spierling, Ulrike; Luderschmidt, Johannes (2018). Chatbots und mediengestützte Konversation. In: Christoph Kochhan und Alexander Moutchnik (Hrsg.): Media Management. Ein interdisziplinäres Kompendium. Wiesbaden: Springer Gabler, S. 387–408.

Storp, M. (2002). Chatbots: Möglichkeiten und Grenzen der maschinellen Verarbeitung natürlicher Sprache. In: Networx, Nr. 25. Hannover. Aufgerufen am 30.11.2019 von <http://www.medien-sprache.net/networx/networx-25.pdf>.

Stucki, T., D'Onofrio, S. & Portmann, E. (2018). Chatbot - Der digitale Helfer im Unternehmen: Praxisbeispiele der Schweizerischen Post. In: HMD Praxis der Wirtschaftsinformatik 55, 2018, 4, S. 725–747.

Svanidze, D. (22. März 2019). Was kostet ein Chatbot?: Entwicklung und Kosten eines Chatbots. Aufgerufen am 30.11.2019 von <https://app3null.com/kostet-ein-chatbot-entwicklung-und-kosten-eines-chatbots/> [Web Log Eintrag].

Vom Hofe (23. Februar 2017). Fragen Sie Tinka. Aufgerufen am 30.11.2019 von <https://www.telekom.com/de/blog/konzern/artikel/fragen-sie-chatbot-tinka-485982> [Web Log Eintrag].

Weizenbaum, J. (1976). Die Macht der Computer und die Ohnmacht der Vernunft. Frankfurt am Main: Suhrkamp.

Wilson, H., Daugherty, P.R. & Bianzino, N.M. (2017). When AI Becomes the New Face of Your Brand. Harvard Business Review Online. Aufgerufen am 30.11.2019 von <https://hbr.org/2017/06/when-ai-becomes-the-new-face-of-your-brand>.

## 9. Political Bots im Wahlkampf

Vortragstag und Gruppenname:	11. Juli 15:00 Uhr, Political Bots im Wahlkampf
Titel der Ausarbeitung:	Political Bots im Wahlkampf
Verfasser/-innen:	Ansgar Kuhn, Janine Lollert, Torben Wiedenhaupt
Betreuerin:	PD Dr. Sandra Nuy

### Übersicht über die Struktur der Ausarbeitung

1. Einleitung
  2. Grundlagen
    - 2.1 Social Bots/Political Bots
    - 2.2 Twitter
  3. Akteure
  4. Einsatzstrategien
  5. Wirkungskonstrukt
  6. Messung/Aufdeckung
  7. Anwendungsbeispiele
    - 7.1 Political Bots während der Abstimmung zum „Brexit“
    - 7.2 Political Bots während der „Krim-Krise“
  8. Political Bots in Deutschland
  9. Ergebnis und Ausblick
- Literaturverzeichnis
-



## **1. Einleitung**

Im Rahmen der diesjährigen Europawahl wurde das Phänomen der Political Bots erneut öffentlich diskutiert. Die Frage nach der Beeinflussung politischer Meinungsbildung steht wiederholt im Raum (Tagesschau.de, 2019). Spätestens seit dem US-Präsidentenwahlkampf 2016, bei dem Political Bots beobachtet wurden, wird auch in Deutschland vor dem Auftreten sogenannter Meinungsroboter im Wahlkampf gewarnt und der Begriff gelangte in den öffentlichen Diskurs (Neudert, 2017, S. 5).

Inwieweit Deutschland von politischen Bots in sozialen Netzwerken – die Forschung bezieht sich bislang fast ausschließlich auf Twitter – betroffen war und ist, soll hier durch die Auswertung verschiedener Studien und mit vergleichenden Beispielen aus anderen Ländern dargestellt werden. Ziel ist hierbei, die länderspezifischen Unterschiede und Erklärungsansätze für den Status in Deutschland herauszuarbeiten. Die zu untersuchenden Fragen lauten dabei: Welche Ausprägungen von Political Bots lassen sich in einzelnen Ländern beispielhaft beobachten und wie ist, auch im Vergleich dazu, die Situation in Deutschland momentan und zukünftig einzuschätzen?

Dafür werden zunächst als Grundlagen die Begrifflichkeit der Political Bots – in der Literatur oft auch Social Bots genannt – erläutert sowie die untersuchte Plattform Twitter und die beteiligten Akteure vorgestellt. Darüber hinaus werden als theoretische Hinleitung die Einsatzstrategien, die Wirkung und die Messung bzw. Aufdeckungsmöglichkeiten von Political Bots dargelegt, bevor schlussendlich die Anwendungsfälle betrachtet werden.

## **2. Grundlagen**

### **2.1. Social Bots/Political Bots**

Bei Social Bots handelt es sich um Imitationen von menschlichen Akteuren, genauer gesagt imitieren sie Nutzer sozialer Medien und deren Kommunikationsverhalten, vorzugsweise auf Twitter. Sie wollen die öffentliche Richtung und die Deutungsrichtung von Themen beeinflussen (Graber & Lindemann, 2018, S. 52). Um dies möglichst unerkannt zu tun, nutzen bzw. missbrauchen Social Bots neueste Techniken und Algorithmen. Menschen orientieren sich bei ihren eigenen Stellungnahmen, Bewertungen und Gewichtungen von Natur aus an Informationen und den Meinungen ihrer Mitmenschen, um mit ihrer Meinung nicht isoliert zu sein. Dieses reale menschliche Verhalten überträgt sich auch auf ihr Verhalten im Internet. Somit kommt es auch hier zu manipulierten Mehrheitsmeinungen (Graber & Lindemann, 2018, S. 52). Gelingt den Social Bots dies, so

haben sie ihr Ziel, die Manipulation von Einstellungen und Verhaltensweisen, erreicht. Social Bots täuschen anderen Internetnutzern vor, reale Menschen zu sein. Dadurch können sie die Nutzer unterschwellig beeinflussen, denn diese vertrauen darauf, dass der vermeintlich „echte“ Nutzer – der Social Bot – in Kontakt mit wirklichen, realen Menschen steht (Graber & Lindemann, 2018, S. 53).

Ist die Rede von Social Bots, sind intelligente Computerprogramme gemeint. Diese „(...) besitzen keine eigene Meinung, sondern folgen einer vorgegebenen Agenda“ (Graber & Lindemann, 2018, S. 57). Sie sind mit Profilinginformationen ausgestattet und zielen darauf ab, einen virtuellen Freundeskreis aufzubauen. Dieser virtuelle Freundeskreis soll realen Nutzern einen solchen vortäuschen, um sich anschließend mit den realen Nutzern erfolgreich zu vernetzen. Somit handelt es sich bei Social Bots lediglich um Computersysteme, die menschliches Verhalten vortäuschen (Graber & Lindemann, 2018, S. 57). Das gesamte Prinzip folgt einem sogenannten Schneeballprinzip. Reagiert das intelligente Programm auf einen geteilten Inhalt mit einem Like, Kommentar o. Ä., so wird dies dem realen Nutzer angezeigt. Möglicherweise verbreitet dieser reale Nutzer den Beitrag ebenfalls, dadurch wird der Beitrag dessen gesamten Freundeskreis angezeigt, wodurch diese eventuell auf das Profil des Bots aufmerksam werden. Kommen nun also mehrere Social Bots mit derselben Zielsetzung zum Einsatz, so können ihrerseits Meinungstendenzen entstehen. Somit entsteht ein fälschlicher Eindruck großer Zustimmung. Um die Authentizität des Social Bots möglichst vielen echten Nutzern vorzutäuschen, wird dieser mit einem glaubwürdigen Profil ausgestattet (Leistert, 2017, S. 224). Je mehr solcher Bots vorhanden sind, desto effektiver ist die Verbreitung von Inhalten. Der Vorteil des Ganzen ist es, dass sich die Betreiber von Social Bots zum einen die Konstruierbarkeit von Medienwahrheiten und zum anderen das Prinzip der sozialen Bewährtheit zu Nutze machen können (Graber & Lindemann, 2018, S. 57).

Werden Social Bots im politischen Wahlkampf eingesetzt, nennt man sie auch Political Bots. In wieweit diese in Deutschland vorhanden sind, wird in Kapitel 8 näher erläutert. Eine große Öffentlichkeit erfuhr das Thema der Political Bots vor allem durch den Wahlkampf des jetzigen US-Präsidenten Donald Trump im Jahr 2016 (Bessi & Ferrara). Deshalb wird hier eine erneute Betrachtung ausgespart und in Kapitel 7 auf weitere Beispiele geschaut. Seit dem US-Wahlkampf sind viele Internetnutzer verunsichert, denn sie wissen nicht, wie sie Social Bots bzw. Political Bots erkennen können oder wie sie auf die Gefahr der daraus resultierenden Manipulation reagieren sollen (Graber & Lindemann, 2018, S. 53). Zuletzt ist noch zu klären, ob Social Bots mit Propaganda gleichzusetzen sind. Bei Propaganda handelt es sich um die „[...] Absicht der

systematischen Änderung von Einstellungen und Meinungen durch ein instrumentalisiertes (Massen-)Medium“ (Graber & Lindemann, 2018, S. 54). Vergleichen man dies nun mit dem Verhalten und dem Einsatz von Social Bots, so sind ähnliche Züge zu erkennen, jedoch sind Social Bots auch die Folge des technischen Wandels auf dem Gebiet der digitalen Kommunikation und somit noch nicht allzu lange Teil von propagandistischen Unternehmungen. Propaganda allgemein existiert schon deutlich länger (Graber & Lindemann, 2018, S. 56-57). Social Bots können also als Teil von Propaganda mit dieser teilweise gleichgesetzt werden. Auch wird in späteren Kapiteln in diesem Zusammenhang von Propaganda gesprochen.

## **2.2. Twitter**

Bei Twitter handelt es sich um einen Mikro-Bloggingdienst und ein soziales Netzwerk, welches im März 2006 von Jack Dorsey gegründet wurde und insbesondere seit Donald Trump als prominentem Nutzer allgemeine Bekanntheit erlangte (Pfaffenberger, Adrian & Heinrich, 2019, S. 97). Es kann kostenlos genutzt werden. Twitter ist ein Begriff aus dem Englischen und kann im deutschen als „Gezwitscher“ übersetzt werden, was sich im Logo des Unternehmens widerspiegelt, welches einen stilisierten Vogel darstellt. Die User haben die Möglichkeit, über das Netzwerk Kurznachrichten, sogenannte Tweets – in der Literatur analog zu anderen Netzwerken gelegentlich auch Posts genannt – abzusetzen. Diese sind jedoch auf 280 Zeichen (bis 2017 nur 140 Zeichen) begrenzt. Die Zeichenobergrenze der Tweets ist charakteristisch für die Plattform und somit ein Alleinstellungsmerkmal von Twitter. Neben der Möglichkeit, zeichenbegrenzt Kurznachrichten zu schreiben, somit zu „twittern“, kamen im Laufe der Jahre weitere Funktionen hinzu. Charakteristisch für Twitter ist der sogenannte Hashtag, welcher mittlerweile auch von anderen Plattformen übernommen wurde. Es handelt sich um eine Zeichenfolge mit unmittelbar vorangestelltem Raute-Zeichen (#Hashtag), mit dem der eigene Tweet verschlagwortet und von anderen Nutzern wiedergefunden werden kann. Follower werden Nutzer genannt, die einer Person folgen, also somit deren Tweets abonniert haben und diese im eigenen Feed wiederfinden. Zudem gibt es noch eine sogenannte Retweet-Funktion. Hier können Tweets von anderen Nutzern kommentiert oder unkommentiert mit den eigenen Followern geteilt werden (Twitter Inc., 2019).

## **3. Akteure**

In Zusammenhang mit der Debatte um Political Bots im Wahlkampf sind verschiedene Akteure tätig. An erster Stelle stehen die Propagandisten. Sie sind die Akteure, die sogenannten Betreiber oder Programmierer der Political Bots. Diese geben sie entweder in Auftrag oder aktivieren sie

sogar selbst (Graber & Lindemann, 2018, S. 63-64). Im direkten Gegenzug zu den Propagandisten gibt es allerdings auch einen Gegenstrom. Dieser versucht anhand von Algorithmen, Systeme zur Erkennung von Social Bots zu entwickeln, um diese aufzudecken. Zudem versucht dieser Gegenstrom die Allgemeinheit zur Erkennung von Social Bots zu mobilisieren (Graber & Lindemann, 2018, S. 64).

Des Weiteren zu nennen sind die Journalisten bzw. die Medien allgemein. Sie sind sowohl für die objektive als auch für die subjektive Berichterstattung verantwortlich (Graber & Lindemann, 2018, S. 64). Die Wähler sind auf diese vermittelten Informationen angewiesen, denn ein direkter Informationsaustausch wie früher ist nicht mehr möglich, bzw. reicht nicht aus. Durch die Entwicklung des Internets läuft der Wahlkampf heute unter Vielkanalbedingungen ab und wird somit für die Wähler immer unübersichtlicher (Haller, 2017, S. 13).

Weitere wichtige Akteure sind die Politiker und Parteien. Diese müssen zwischen relevanten und unwichtigen Kommunikationskanälen unterscheiden. Zudem müssen sie während des Wahlkampfes aktiv sein, um ihre Position zu verdeutlichen und die Bürger von dieser zu überzeugen (Graber & Lindemann, 2018, S. 65). Abschließend sind die Betreiber der sozialen Netzwerke zu nennen, welche versuchen, die Glaubwürdigkeit ihres Unternehmens aufrecht zu erhalten. Das Offenlegen eines Vorgehens gegen Social Bots kann sich negativ bei den Werbekunden auswirken, weshalb auch keine genauen Zahlen zu Bots veröffentlicht werden (Egli & Rechsteiner, 2017, S. 256).

#### **4. Einsatzstrategien**

Social Bots agieren nach bestimmten Mustern, aus denen sich drei Einsatzstrategien ableiten lassen können. Die erste Strategie ist das sogenannte Astroturfing. Der Begriff steht für eine künstliche Graswurzelbewegung, also eine Bürgerbewegung, welche aus der gesellschaftlichen Basis – von unten – entsteht und nicht von oben oktroyiert wird (Brachten, Stieglitz, Hofeditz, Kloppenborg & Reimann, 2017, S. 3). Abgeleitet wird der Name von einer amerikanischen Marke für Kunstrasen, „Astroturf“, welcher im Englischen zum allgemeinen Synonym für Kunstrasen geworden ist (Lobbypedia, 2017, Absatz 2). Beim Astroturfing durch Social Bots wird in Form von zentral gesteuerten Tweets und Retweets künstlich simuliert, dass eine bestimmte politische Meinung durch viele angebliche Einzelmeinungen eine breitere Unterstützung erhält als es in Wirklichkeit der Fall ist (Brachten et al., 2017, S. 3). Dieser Agenda gehen die Bots auf diversen Wegen nach, indem sie nach Schlüsselbegriffen auf Twitter suchen und vorgefertigte Kommentare zu

entsprechenden Tweets veröffentlichen. Besonders smarte Bots sind dialogfähig, „[...] wandeln dabei ihre Sprache, variieren die Veröffentlichungsfrequenz, wechseln zwischen Kommentar, Like und eigenem Post“ (Puscher, 2017, S. 28).

Die zweite Strategie ist das sogenannte Smoke Screening bzw. Misdirecting. Diese eigentlich separaten Intentionen werden zu einer Strategie zusammengefasst, da sie in Bezug auf Twitter kaum zu trennen sind (Brachten et al., 2017, S. 7). Hier wird versucht, ein unter einem bestimmten Hashtag diskutiertes Thema in eine andere Richtung zu lenken und somit unwichtig erscheinen zu lassen. Dabei soll verborgen werden, worum es bei dem Hashtag wirklich geht und Tweets abzusetzen, die nicht dem Hashtag entsprechen, aber dem Thema im Groben zuzuordnen sind (Brachten et al., 2017, S. 3).

Als dritte Strategie gilt das auch aus anderen Kontexten bekannte Spamming, In der Regel wird hier durch massenhafte Accounts Werbung verbreitet, die mit populären oder aktuellen Hashtags versehen sind und diese damit stilllegen können (Brachten et al., 2017, S. 3). Hier handelt es sich im Allgemeinen um unpolitische Inhalte, welche aber auch mit politischen Hashtags versehen werden. Um Spamming einzudämmen, beschloss Twitter im April 2019 die Zahl der Accounts, denen am Tag neu gefolgt werden darf, auf 400 zu begrenzen (OnlineMarketing.de, 2019, Absatz 1).

Möglich ist auch eine Kombination von Strategien, je nach Intention der Propagandisten. Daneben werden einige weitere, teilweise länderspezifische Aktionen von Social Bots beobachtet, welche aber keiner dieser Strategien direkt zuzuordnen und individuell zu betrachten sind, worauf in Kapitel 7 näher eingegangen wird.

## **5. Wirkungskonstrukt**

Soziale Netzwerke zielen auf das Mitteilen ab, denn Medienplattformen werden i.d.R. genutzt, um sich anderen mitzuteilen. Es gilt also, je mehr Unterstützung ein Beitrag hat, desto größer ist seine Reichweite. Die Betreiber der Social Bots nutzen genau dieses Grundverständnis der Plattformnutzung für ihr Vorhaben aus (Graber & Lindemann, 2018, S. 62). Der Wirkungserfolg eines Bots ist nicht allein durch die Menge seiner Interaktionen gegeben, viel wichtiger ist die persönliche Identifikation der Nutzer mit den Tweets der Bots. Ohne diese hat die Quantität der Aktivität keinen Erfolg. Hat ein Tweet also keinen oder nur einen sehr geringen Bezug zu den Usern, so wird er trotz des hohen Bot-Einsatzes nur eine geringe Reichweite erzielen (Graber & Lindemann, 2018,

S. 62).

Hinzu kommt, dass eine Medienwahrheit keine Überprüfung verlangt. So können es beispielsweise Bilder ohne jegliche Aussagekraft dennoch schaffen, in Verbindung mit einer weitläufigen Meinungsäußerung gebracht zu werden (Graber & Lindemann, 2018, S. 63). So entstehen schnell polarisierende Meinungen. Hierbei ist es egal, was die Bilder oder Texte wirklich aussagen, ob sie erfunden oder einfach ein Fantasiekonstrukt eines Einzelnen sind. Polarisierende Nachrichten werden leicht von Anderen übernommen und weiterverbreitet, auch wenn diese nicht einmal wissen, ob es sich bei der Nachricht um die Wahrheit handelt. Somit gibt es die Befürchtung, dass Social-Media-Propaganda einen immer größeren Einflussbereich erlangt, weil die Social Bots nicht schnell genug erkannt werden (Graber & Lindemann, 2018, S. 65).

Zum Thema Wirkungskonstrukt gehört auch die Medienwirkung, welche in vier verschiedenen Ausprägungen Einflüsse auf den Rezipienten hat (Haller, 2017, S. 15):

Als Erstes zu nennen ist eine Veränderung des Wissens von Mediennutzern. Dieses tritt beispielsweise durch den Erwerb von politischem Wissen ein, z. B. erfahren die Rezipienten durch die Lokalzeitung, das Lokalradio oder auch durch das Internet, was die kommunale Politik aktuell plant. Über überregionale Entwicklungen, wie z. B. nationale oder internationale Wahlen, können sich die Rezipienten über das Fernsehen oder auch über überregionale Zeitungen informieren.

An zweiter Stelle kann es Auswirkungen auf den emotionalen Zustand von Mediennutzern geben. Hier verändert sich das Erregungsniveau eines Rezipienten. Dies kann sowohl kurzfristig als auch langfristig geschehen. Solche Veränderungen können verschiedene emotionale Zustände hervorrufen wie beispielsweise Aggressionen, Furcht, Freude oder auch positive Entspannung.

Die dritte Veränderung zielt auf die Veränderung der Meinung bzw. Einstellung zu bestimmten Sachverhalten ab. Diese Veränderungen benötigen jedoch eine vorangegangene Änderung auf der Wissens- und Emotionsebene, denn es muss eine Wissensaneignung über den Sachverhalt gegeben sein. Die Veränderung der Meinung ist vor allem auch über Propaganda möglich.

Als vierte und letzte Medienwirkung gibt es die Verhaltensveränderung der Mediennutzer. Hier wird z. B. das Ziel verfolgt, eine Stimmabgabe für die eigene Partei im Wahlkampf zu erreichen. Eine Verhaltensveränderung ist jedoch keine zwingende Folge der Mediennutzung.

Entscheidende Faktoren sind bei dieser Veränderung jedoch das Vorwissen, die gegebene Einstellung und die Beeinflussung der Mitmenschen. Somit haben die Medien nicht zwangsläufig eine Wirkung auf die Einstellung bzw. die Wahrnehmung der Menschen.

## **6. Messung/Aufdeckung**

Zu Beginn dieses Abschnittes ist es wichtig zu unterstreichen, dass Kommunikation zwischen zwei Seiten über die sozialen Netzwerke und somit auch Plattformen wie z. B. Facebook und Twitter immer indirekt abläuft (Graber & Lindemann, 2018, S. 63). Sie verläuft zu jedem Zeitpunkt unpersönlich sowie zeitlich und räumlich versetzt. Auch die Art und Weise, wie miteinander kommuniziert wird, ist meist auf wenige Aktionen beschränkt. Darauf stützen sich die Betreiber von Political Bots, denn der Nutzer kann zu keinem Zeitpunkt vollständig und korrekt einschätzen, ob sein fiktives Gegenüber in den Netzwerken ein reales ist (Graber & Lindemann, 2018, S. 64). Die Forschung ist in diesem Bereich noch nicht so weit fortgeschritten. Als ein erster Lösungsansatz könnte hier ein Algorithmus genannt werden, der in einer Zusammenarbeit zwischen der Indiana University und der University of Southern California entstanden ist. Dieser soll helfen, Political Bots auf der Kommunikationsplattform Twitter zu identifizieren. Das Forscherteam griff dabei auf insgesamt sechs Kategorien zurück, die in den Algorithmus miteingeflossen sind: User Meta-Data, Sentiment, Friends, Network, Content und Timing (Varol, Ferrara, Davis, Menczer & Flammini, 2017, S. 2).

Der Terminus User Meta-Data beschreibt die generelle Seriosität des zu untersuchenden Twitter-Accounts. Political Bots sind vor allem darauf ausgelegt, Inhalte zu verbreiten. Deshalb sind ihre Profilinformationen häufig nicht vollständig oder sinnvoll eingepflegt. Gleichzeitig nehmen diese Profile auch nur selten eigenständigen Kontakt zu anderen, genuinen Nutzern auf (Varol et al., 2017, S. 2). Daran anschließend wird unter der Kategorie Sentiment das Reaktionsschema eines Profils auf einen Tweet oder einen Kommentar verstanden. Political Bots antworten mit einer programmierten Aktion bzw. mit einem abgestimmten Schema, das auf festgelegte Schlagwörter reagiert, wenn sie in einem Tweet bzw. Kommentar erscheinen. Die Absicht der Political Bots-Algorithmen ist es, gezielt Gefühle bei echten Nutzern zu erzeugen. Im Rückkehrschluss fallen Bots durch immer wieder gleiche Reaktionen auf. Bei realen Nutzern lassen sich diversifizierte Handlungsmuster beobachten (Varol et al., S. 2).

Die Kategorien Friends und Network beziehen sich auf die zwischenmenschliche Kommunikation eines Nutzers. Es geht darum, was ein bestimmter Account teilt, wie oft es von anderen Nutzern

erwähnt wird und welche Hashtags ausgesucht wurden. Auch hier weisen echte Profile einen diversen Umgang mit den unterschiedlichsten Themen und Menschen auf. Political Bots beschränken sich für gewöhnlich auf ein bestimmtes Themengebiet, verteilen vorgegebene Inhalte und kommunizieren nahezu allein über bestimmte Kanäle. Letzteres bedeutet, dass Freunde und das gesamte Netzwerk eines Political Bot-Accounts meist um ein und dasselbe Thema herum aufgebaut sind und keine natürliche Verteilung zu finden ist (Varol et al., S. 3).

Die Kategorie Content bezeichnet die verwendete Sprache, der untersuchten Political Bot-Accounts. Es geht hier um die sogenannte „Entropie der Sprache“ (Varol et al., S. 3). Kriterien sind hierbei u.a. Satzbau, Länge oder die Diversität des Wörteregebrauchs. Verwendet der zu untersuchende Account in verschiedenen Meldungen beispielsweise immer wieder dieselbe Grammatik, dieselben Wörter oder kopiert den Wortlaut anderer Nutzer, liegt auch die Schlussfolgerung nahe, dass es sich um einen Political Bot handelt (Varol et al., 2017, S. 3).

Schließlich prüft die Kategorie Timing, in welcher Frequenz Interaktionen stattfinden, und wie die sogenannte Reaktivität, also das Feedback eines Accounts, beurteilt werden kann. Political Bots können wesentlich schneller auf neue Inhalte reagieren. So gibt es beispielsweise Profile, die auf bestimmte Themeninhalte innerhalb von Sekunden reagieren (Varol et al., 2017, S. 4).

Die Studie legt nahe, dass sich hinter Accounts, die mehr als 50 Tweets am Tag versenden und diese Zahl kontinuierlich beibehalten, in den meisten Fällen ein Bot zu finden ist. Diese Angabe ist allerdings kritisch zu betrachten, insofern eine solche Zahl normale Nutzer, die sich als sogenannte „Heavy User“ permanent in sozialen Netzwerken aufhalten, diese Zahl durchaus erreichen und überschreiten (Varol et al, 2017, S. 3; vgl. Graber & Lindemann, 2018, S. 64). Auch ein Algorithmus kann keine vollständige Erkennung von Bots gewährleisten:

„Im selben Maße, wie sich Gegenmaßnahmen weiterentwickeln, entwickeln sich auch die Political Bots. Es gilt der Leitsatz: Je diverser das Verhalten der Programme wird, desto schwerer wird es, sie als solche zu entlarven“ (Graber & Lindemann, 2018, S. 64).

## **7. Anwendungsbeispiele**

Im Jahr 2016 berichtete das Medienmagazin „Zapp“ des Norddeutschen Rundfunks (NDR), dass Social Bots als Instrument in den sozialen Netzwerken eingesetzt werden, um gezielt politische Stimmung für oder gegen eine bestimmte Thematik im Sinne der Betreiber dieser Bots zu



entwickeln (Drehkopf, 2016). Bei der Recherche nach entsprechenden Studien, welche nicht nur den US-Wahlkampf 2016 betreffen, stellten sich zwei europäische Beispiele aus der jüngeren Vergangenheit als geeignetes Anschauungsmaterial heraus, welches hier vorgestellt werden soll. Zum einen konzentrieren sich die folgenden Kapitel auf Social Bots in Großbritannien während des Wahlkampfs zum Referendum im Juni 2016, bei dem über den sogenannten „Brexit“, dem Austritt des Landes aus der Europäischen Union abgestimmt wurde (Schuppen, J., Forßmann, C., Berndt, G. & Bogazliyan, S., 2017). Zum anderen geht es um Social Bots in der Ukraine während des Konflikts mit Russland um die Halbinsel Krim im Schwarzen Meer. Im Februar und März 2014 leisteten sich Russland und die Ukraine zum Teil gewalttätige und bewaffnete Auseinandersetzungen. Die Krim, die bis dato zur Ukraine gehörte, wurde im Verlauf des Konfliktes durch russische Kämpfer zum größten Teil eingenommen und nach einem umstrittenen Referendum im März 2014 schließlich komplett annektiert (Bundeszentrale für politische Bildung, 2019).

Die Auswertung der ausgewählten Beispiele basiert dabei auf zwei groß angelegten Studien u.a. von der University of Oxford zum Brexit-Wahlkampf, sowie einer Forschung des Teams rund um den Wissenschaftler Simon Hegelich vom Forschungsprojekt „Social Media Forensics“ zum Krim-Konflikt im Rahmen des Forschungskollegs „Zukunft menschlich gestalten“ (FoKoS) der Universität Siegen. Beide bieten ein vollständiges Bild, sodass im Rahmen dieser Arbeit ein anschaulicher Überblick gelingt, wie Social Bots genau funktionieren. Zudem lassen sich die Studien in manchen Punkten gut vergleichen. Beispielsweise sind die Ansatzpunkte in beiden Fällen z. B. Schlagworte, die in den betroffenen Tweets auftauchen:

„Wir sind auf die Bots gestoßen, als wir angefangen haben, Daten von Twitter mit dem Hashtag #Ukraine zu analysieren. Dabei ist uns aufgefallen, dass es sehr viele Meldungen gab, wo die gleiche Meldung von unterschiedlichen Nutzern gesendet worden ist. Es war aber nicht als ein Retweet gekennzeichnet. Das ist ein ungewöhnliches Verhalten und jetzt wühlen wir sozusagen das gesamte Netzwerk auf.“ (Simon Hegelich im NDR ZAPP-Bericht: Drehkopf, 2016, ab Min. 1:29)

### **7.1. Political Bots während der Abstimmung zum „Brexit“**

Die Wissenschaftler Bastos und Mercea, beide von der Universität zu London, kommen zu dem Schluss, dass Political Bots als politische Stimmungsmacher im Vorfeld des Brexit-Referendums im Jahr 2016 zum Einsatz kamen. Die Studie analysiert die im Wahlkampf veröffentlichten Inhalte der verschiedenen Parteien unter Berücksichtigung mehrerer ausschlaggebender Faktoren wie

z.B. Umfang, Intensität und Reichweite der Tweets, was ein Novum darstellt (Bastos & Mercea, 2017, S. 1). Zentral für die Forscher waren hier die benutzten Hashtags in den Tweets. Gut zehn Millionen Hashtags in über 790.000 Twitter-Profilen wurden im untersuchten Zeitraum von Anfang Juni bis Anfang Juli 2016 untersucht. Ihre Studie legt dar, dass gut 13.500 dieser Accounts die in Kapitel 6 erläuterten Merkmale zu Erkennung von Political Bots aufweisen (Bastos & Mercea, 2017, S. 44).

Beispielsweise bildeten die Bots ein Netzwerk von sogenannten „Zombie-Agenten“ (Bastos & Mercea, 2017, S. 44). Damit betrieben sie gezielt „Meinungsmache“ und verschwanden kurz nach der Entscheidung wieder. Es war zu erkennen, dass es insgesamt eine stärkere Unterstützung des „Leave“-Lagers gab. Ein Drittel der analysierten Profile enthielt das Schlüsselwort „Leave“ (Bastos & Mercea, 2017, S. 44). Daraus lässt sich erkennen, dass automatisierte Bot-Programme beispielsweise auf Twitter als „beaufsichtigtes Netzwerk von Zombie-Agenten“ (Billo, 2017, Absatz 4) fungierten, so die Wissenschaftler. Sie unterstützen bzw. verstärken den Meinungs-austausch zugunsten der einen oder der anderen Seite. Allerdings erklären die Forscher ausdrücklich auch:

"Wir haben keine Beweise gefunden, dass Bots gefälschte Nachrichten verbreitet haben. Stattdessen wurden sie zum Einspeisen und Wiederholen benutzer-kuratierter, parteiischer und polarisierender Informationen genutzt" (Bastos in: Billo, 2017, Absatz 5).

Im Nachgang der Wahl lässt sich erkennen, wie sehr sogenannte Bot-Programme und -netzwerke in sozialen Medien strategisch eingesetzt werden, um bestimmte Positionen gezielt zu unterstützen bzw. zu verfälschen. Hier zeigt sich vor allem die Strategie des Astroturfings erkennbar. Im Wahlkampf um den Brexit zeigt sich im Anschluss der Entscheidung, dass etliche Accounts ausgeschaltet oder gänzlich gelöscht wurden bzw. die Benutzernamen hin zur vollständigen Unkenntlichkeit abgeändert wurden. Dies geschah bei insgesamt über 40.000 Accounts (Billo, 2017, Absatz 7). Diese Streichungen oder Veränderungen der Konten in den Wochen nach dem Referendum zeigen, dass die hinter den Aktivitäten stehenden "Strippenzieher" leicht verborgen bleiben können (Billo, 2017, Absatz 6). Insgesamt gibt die Studie ein ernüchterndes Bild ab: Sie zeigt, wie schwer es tatsächlich ist, Political Bots zu entdecken, richtig auszuwerten und gegen sie vorzugehen. Ob die Verwendung von Social Bots im Wahlkampf trotz der übermäßigen Anzahl von Pro-Brexit-Bots letztendlich auch zu der gewünschten Verschiebung des Wahlergebnisses seitens der Auftraggeber geführt hat, kann ebenfalls nicht ausreichend festgestellt werden (Erleben, 2016). Zu diesem Ergebnis kommen auch die ebenfalls zum Brexit forschenden Wissenschaftler Howard

und Kollanyi: „It is difficult to say how much public opinion is shaped by political discourse on this topic over social media or what the influence of bots is on public sentiment. Nevertheless, we can identify the role that bot algorithms play in political communication around StrongerIn-Brexit issues. We find that political bots have a small but strategic role in the referendum conversations: (1) the family of hashtags associated with the argument for leaving the EU dominates, (2) different perspectives on the issue utilize different levels of automation, and (3) less than 1 percent of sampled accounts generate almost a third of all the messages“ (Howard & Kollanyi, 2016, S. 5).

## **7.2. Political Bots während der „Krim-Krise“**

Der Konflikt zwischen Russland und der Ukraine um die Halbinsel Krim fand seinen Höhepunkt im Februar und März 2014. Nach bewaffneten Auseinandersetzungen und einem international nicht anerkannten Referendum auf der Krim über den zukünftigen Status dieses Landesteils annectierte Russland die Insel im Schwarzen Meer und verletzte dabei u.a. internationales Recht (Bundeszentrale für politische Bildung (2019). Im Nachgang zu diesem Konflikt begann ein Forschungsteam rund um Simon Hegelich mit der Analyse des Konfliktes in den sozialen Netzwerken. Dabei stießen sie auf ein großes, undurchsichtiges Bot-Netzwerk auf Twitter. Rund 15.000 Accounts waren mit Beendigung der Studie aktiv (Hegelich, 2016, S. 5). Durchschnittlich veröffentlichten diese über 60.000 Meldungen am Tag. Die Tweets richteten sich mehrheitlich an die Interessen junger männlicher Einwohner der Ukraine (Hegelich, 2016, S. 5). Ihre Inhalte reichen von Fußball, sexistischen Witzen und Links zum illegalen Download aktueller amerikanischer Kinofilme. Zwischen diesen Meldungen flossen immer wieder gezielt Propaganda-Nachrichten des „Rechten Sektors“, einer ultranationalistischen ukrainischen Vereinigung mit paramilitärischem Ableger, ein (Hegelich, 2016, S. 5-6). Die Forscher konnten hier verschiedene Strategien der Manipulation identifizieren:

„Zum einen geht es auch hier um die Verfälschung von Trends, indem bestimmte Hashtags besonders populär gemacht werden. Darüber hinaus verknüpfen die Bots aber bewusst Schlagworte [...] offenbar, um die Algorithmen von Twitter dazu zu bringen, Nutzern [...] auch Inhalte des „Rechten Sektors“ zu präsentieren. Als weitere Strategie wird die Verbreitung von Falschinformationen hier deutlich: So verbreitete das Botnetz die Nachricht, die Separatisten hätten von Russland Raketen erhalten und würden nun auf Kiew schießen. Außerdem folgen die Bots gezielt ukrainischen Politikern, um ihre eigene Reichweite zu erhöhen“ (Hegelich, 2016, S. 5).

Hier wird also teilweise die Strategie Astroturfing mit Misdirecting kombiniert, soweit die Bots als solche erkennbar waren. Insgesamt verfügen die Political Bots im Krimkonflikt über verschiedene Möglichkeiten, wie sie klassischen Bot-Erkennungsalgorithmen ausweichen können:

„Sie folgen sich gegenseitig und haben dadurch ein ausgewogenes Verhältnis von Freunden und Followern, sie posten nach einem zeitlichen Muster, das Pausen und Schlafenszeiten simuliert, aber trotzdem zufällig ist [...].“ (Hegelich, 2016, S. 5)

Dabei besitzen sie aber auch die Fähigkeit, Nachrichten geringfügig abzuwandeln, sodass die Aussage zwar identisch bleibt, automatische Erkennungsprogramme diese Texte aber nicht als identische Inhalte erkennen können (Hegelich, 2016, S. 6).

Ob Social Bots auch in diesem politischen Entscheidungsprozess als ausschlaggebender Impuls gewertet werden können, kann auch hier nicht ausreichend beurteilt werden (Erleben, 2016). Trotz groß angelegter Forschungen ist es nach wie vor nicht überprüfbar, ob Social Bots bei Wahlen mitentscheidend sind, so Simon Hegelich:

„Ich glaube nach wie vor nicht, dass Bots ein geeignetes Instrument sind, um Massen umzudrehen. Aber in Situationen, wo wenige Stimmen entscheidend sein können, haben Bots möglicherweise einen Einfluss. Es wird allerdings unmöglich sein, das jemals nachzuweisen – weil man nie nachweisen kann, ob die Leute nicht ohne die Bots das Gleiche gewählt hätten. Generell muss man allerdings sagen: Social Bots spielen in Wahlkampagnen im Moment noch keine Rolle. Aber das kann ja noch kommen“ (Simon Hegelich im taz-Interview: Laaff, 2016, Antwort zu Frage 12).

Hinzu kommt die bereits erwähnte und ununterbrochene technische Weiterentwicklung der Bots, die es für die Wissenschaftler schwer macht, diese programmierten Anwendungen zu erfassen und in einem stetig aktualisierten Rahmen zu untersuchen.

## **8. Political Bots in Deutschland**

Untersuchungen zu Political Bots in Deutschland fanden vor allem im Superwahljahr 2017, ein Jahr nach Trump-Wahl und Brexit-Abstimmung statt. Hervorzuheben sind hierbei die Bundespräsidentenwahl im Februar, die Landtagswahl in Nordrhein-Westfalen im Mai und die Bundestagswahl im September. Studien zu diesen drei Wahlen kommen zu dem Ergebnis, dass durch Political Bots ein kaum messbarer Einfluss auf das Wahlergebnis festzustellen ist, wobei der

Bundespräsident nicht durch das Volk, sondern durch die Bundesversammlung gewählt wird (Bundestag.de, 2019). Diese Wahl ist also nur in Teilen mit den anderen Wahlen zu vergleichen. Nichtsdestotrotz waren hier 22 Bots aktiv, die 6000 Tweets in drei Tagen abgesetzt haben und in der Mehrheit rechte Propaganda, in Form von negativen Kommentaren zu Immigration sowie xenophobischen und AfD-befürwortenden Kommentaren, verbreitet haben. Darüber hinaus waren Attacken gegen die beiden damaligen Kanzlerkandidaten Angela Merkel und Martin Schulz zu beobachten (Neudert, 2017, S. 12). Im Vorfeld der Landtagswahl in NRW waren 61 Political Bots beobachtet worden, wobei der Begriff hier zu hinterfragen ist, da es hauptsächlich Bots ohne politische Strategie und mit fast ausschließlich Spam abseits politischer Themen als Inhalt waren (Brachten et al., 2017, S. 10).

Bei der Untersuchung der Political Bots vor und während des Wahlkampfes zur Bundestagswahl wird zwischen existierenden Bots und aktiven Bots unterschieden. Während der Anteil von allen Bots von 7,1 % auf 9,9 % gestiegen ist, ging der Anteil der aktiven Bots von 2% auf 1,4 % sogar zurück, was am Ende 314 aktive Bots im Wahlkampf bedeutete (Keller & Klinger, 2019, S. 19). Fast keiner dieser Bots agierte politisch, auch hier ist vor allen Dingen Spamming zu beobachten. Nur einer der von den aktiven Bots genutzten Hashtags war auf die deutsche Politik bezogen: #AfD. Die AfD ist von allen Parteien diejenige mit dem geringsten Anteil an Social Bots unter ihren Followern, aber unter diesen Bots diejenige mit dem größten Anteil an aktiven Bots. (Keller & Klinger, 2019, S. 20).

In Deutschland ist also noch keine relevante Anzahl von Meinungsrobotern auf Twitter festzustellen, welche wirklich als Political Bots zu bezeichnen wären. Vielmehr handelt es sich um hochaktive, automatisierte Accounts und Fake Follower mit mehr Retweets statt eigenen Inhalten (Neudert, 2017, S. 11). Der niedrige Status der Political Bots in Deutschland kann auf verschiedene Gründe zurückgeführt werden. Twitter als das für Bots anfälligste soziale Netzwerk ist in Deutschland kein Massenmedium, sondern ein Elitemedium, welches hauptsächlich von Politikern, Journalisten und Intellektuellen genutzt wird und somit für politisches Agenda-Setting eine nachrangige Rolle spielt (Pfaffenberger, Adrian & Heinrich, 2019, S. 120; vgl. Puscher, 2017, S. 29). Die Parteien in Deutschland sind im Vergleich z.B. zu den USA mit weniger Budget für Online-Wahlkampf und weniger Daten über die Wahlberechtigten ausgestattet. Auch erschweren die festgefahrenen Strukturen und Prozesse der alteingesessenen, deutschen Parteien die Einbindung digitaler Technologie. Neue Parteien können die Digitalstrategie von Beginn an mit einbinden (Jung-herr, 2017, S. 97).

Zuletzt stellt sich die Frage, ob das deutsche Wahlsystem mit der indirekten Wahl und dem damit verbundenen, untergeordneten Personenwahlkampf sowie der fehlenden dichotomen Entscheidung wie in den USA, beim Brexit oder dem Krim-Referendum, für Political Bots überhaupt interessant ist, da die mögliche Beeinflussung eines Ergebnisses zu gering ist.

## **9. Ergebnis und Ausblick**

Es kann festgehalten werden, dass versuchte Meinungsmanipulation durch Political Bots unterschiedlich stark ausgeprägt ist, wobei generell ein Anstieg in Ländern zu verzeichnen ist, in denen Wahlkampf geführt wird. Die anfangs gestellte Frage nach besonderen Ausprägungen von Political Bots, konnten die Beispiele aus Großbritannien und der Ukraine, jeweils im Vorfeld des jeweiligen Referendums, exemplarisch beantworten. In beiden Beispielen wurden starke Manipulationsversuche beobachtet, wobei im zweiten Beispiel die strategisch unauffälligere Beeinflussung eine automatische Erkennung deutlich erschwerte.

Der sogenannte Cyber Troops Atlas, welcher Länder weltweit vergleicht, weist darüber hinaus die USA, Russland, Israel und China als Länder mit dem größten Anteil an versuchter Meinungsmanipulation aus. Deutschland wird hierbei eine geringe Aktivität zugewiesen (Bradshaw & Howard, 2018, S. 17). Das deckt sich mit den bisher geführten und hier dargelegten Studien, die besagen, dass der Einfluss auf Wahlen in Deutschland marginal ist, die Sensibilisierung hingegen hoch. Die Beeinflussung ist so gut wie nicht vorhanden und lange nicht so groß, wie er von den Medien gemacht wurde (Jungherr, 2017, S. 98). Damit kann auch der zweite Teil der Ausgangsfrage nach der Situation in Deutschland beantwortet werden. Eine relevante Zahl von Political Bots mit entsprechendem Einfluss – vor allem im Vergleich zu vorangegangenen Beispielen – ist in Deutschland also bislang nicht zu verzeichnen und die Aktivitäten sind vergleichsweise gering. Als mögliche Gründe dafür wurden der Status von Twitter, die Parteistrukturen und das Wahlsystem in Deutschland diskutiert.

Auch ist bei den Political Bots in Deutschland in den meisten Fällen eher von menschlich programmierten oder automatisierten Accounts ohne künstliche Intelligenz im Sinne maschinellen Lernens oder gar künstlicher neuronaler Netze zu sprechen (Mayer, 2018, S. 31; vgl. Reuter, 2016). Das kann sich allerdings schnell ändern, denn die Weiterentwicklung von Bots schreitet teilweise schneller voran als die Gegenmaßnahmen (Keller & Klingern, 2019, S. 24). Auch gibt es

stille Botnetze, welche bislang unauffällig waren, aber jederzeit aktiviert werden können (Neudert, 2017, S. 13). Sollten sich Bots überproportional vermehren, könnten sie ganze Plattformen lahmlegen, wenn es für echte Nutzer sinnlos erscheint, möglicherweise nur noch mit Bots als Gesprächspartnern zu agieren (Hegelich, 2016, S. 1). Wenn Social Bots allerdings intelligenter werden als es heute der Fall ist, bleibt zukünftig zu untersuchen, wie Nutzer mit Ihnen agieren und vor allem wie Politiker und Journalisten diese behandeln. Denn sollten Social Bots solche einflussreichen Personen durchdringen, bekommen sie die Reichweite und die Glaubwürdigkeit, die sich ihre Auftraggeber erhoffen. Schon heute können Bots auf Twitter durch verschiedene Strategien Trends entstehen bzw. manipulieren lassen. Diese Trends wiederum können in politische Entscheidungsprozesse einfließen, wenn Politiker sich an diesen orientieren, was eine Gefahr der demokratischen Meinungsbildung darstellt (Hegelich, 2016, S. 3). Insofern wird weiter zu beobachten sein, wie sich Political Bots entwickeln und welche Regulierungsmaßnahmen möglicherweise zu treffen sein werden.

---

## Literaturverzeichnis

Bastos, M. T. & Mercea, D. (2017). The Brexit Botnet and User Generated Hyperpartisan. *Social Science Computer Review*. Abgerufen von <https://journals.sagepub.com/doi/pdf/10.1177/0894439317734157> (Stand: 14.09.2019).

Bessi, A. & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday* 21 (11). doi:10.5210/fm.v21i11.7090

Billo, T. (2017). Wenn Maschinen Stimmung machen. Brexit: Forscher kommen der Twitterbot-Armee auf die Spur. *Datacenter Inside*. Abgerufen von <https://www.datacenter-insider.de/brexit-forscher-kommen-der-twitterbot-armee-auf-die-spur-a-661027/> (Stand: 14.09.2019).

Brachten, F., Stieglitz, S., Hofeditz, L., Kloppenborg, K. & Reimann, A. (2017). Strategies and influence of social bots in a 2017 German state election – A case study on Twitter. In: *Proceedings of the Australasian Conference on Information Systems*, Paper 207.

Bradshaw, S. & Howard, P. N. (2018): *Challenging Truth and Trust: A Global Inventory of*

Organized Social Media Manipulation, Working paper, 1. Computational Propaganda Research Project. University of Oxford.

Bundestag.de (2019): Die Wahl des Bundespräsidenten. Abgerufen von <https://www.bundestag.de/bundesversammlung> (Stand: 11.10.2019).

Bundeszentrale für politische Bildung (2019): Vor fünf Jahren: Russlands Annexion der Krim. Abgerufen von <https://www.bpb.de/politik/hintergrund-aktuell/287565/krim-annexion#footnode3-3> (Stand 20.10.2019).

Drehkopf, K. (2016): Netzangriff: Bots feuern gegen Themen [Fernsehbeitrag]. In: ZAPP. Hamburg, Norddeutscher Rundfunk. Abgerufen von <https://www.youtube.com/watch?v=hTgp5nWfyEo> (Stand 20.10.2019).

Egli, P. & Rechsteiner, D. (2017). Social Bots und Meinungsbildung in der Demokratie. Aktuelle Juristische Praxis (AJP/PJA), 2, 249-259.

Erxleben, C. (2016): Negativer Einfluss. Wie Twitter-Bots Brexit-Wähler betrogen haben. In: Internet World Business. Abgerufen von <https://www.internetworld.de/technik/bots/twitter-bots-brexit-waehler-betrogen-1111324.html> (Stand: 20.10.2019).

Graber, R. & Lindemann, T. (2018). Erkennbarkeit von Social Bots und mögliche Gegenmaßnahmen. In: K. Sachs-Hombach & B. Zywiets (Hrsg.): Fake News, Hashtags & Social Bots. Neue Methoden populistischer Propaganda (S. 51-69). Wiesbaden: Springer VS.

Haller, A. (2017). Der Wahlkampf im Netz. Twitter, Facebook, Social Bots, Fake News und die Folgen. Politische Studien, 474 (Vol. 68), 12-21.

Hegelich, S. (2016). Invasion der Meinungs-Roboter. Analysen & Argumente, 221. Bonn: Konrad Adenauer Stiftung.

Howard, P. N. & Kollanyi, B. (2016): Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum. Oxford University. Abgerufen von <https://arxiv.org/ftp/arxiv/papers/1606/1606.06356.pdf> (Stand: 20.10.2019).



Jungherr, A. (2017). Einsatz digitaler Technologie im Wahlkampf. In: H. Gapski, M. Oberle & W. Staufer (Hrsg.): Medienkompetenz. Herausforderung für Politik, politische Bildung und Medienbildung (S. 92-101). Bonn: Bundeszentrale für politische Bildung.

Keller, T. R. / Klinger, U. (2019). Social bots in election campaigns: theoretical, empirical, and methodological implications. *Political Communication*, 36 (1). 171-189.

Laaff, M. (2016): Datenexperte über Social Bots. „Manipulation ist nicht so einfach“.

Die Tageszeitung (taz). Abgerufen von <https://taz.de/Datenexperte-ueber-Social-Bots/!5337164/>. (Stand: 20.10.2019).

Leistert, O. (2017). Social Bots als algorithmische Piraten und als Boten einer techno-enviromentalen Handlungskraft. In: R. Seyfert, R. & J. Roberge (Hrsg.): *Algorithmenkulturen. Über die rechnerische Konstruktion der Wirklichkeit* (S. 215-234). Bielefeld: Transcript.

Lobbypedia (2017). Astroturfing. Abgerufen von <https://lobbypedia.de/wiki/Astroturfing> (Stand: 14.09.2019).

Mayer, C. (2018). Künstliche Intelligenz und maschinelles Lernen: Hintergrund, Anwendungsfälle und Chancen für Medienunternehmen. *MedienWirtschaft*, 3, 30-35.

Neudert, L.-M. (2017). Computational Propaganda in Germany: A cautionary tale, Working paper, 1. Computational Propaganda Research Project. University of Oxford.

OnlineMarketing.de (2019). Twitter wehrt sich gegen Spammer. Abgerufen von <https://onlinemarketing.de/news/twitter-spam-400-follower-limit-neu> (Stand: 14.09.2019).

Pfaffenberger, F., Adrian, C. & Heinrich, P. (2019). Was bin ich – und wenn ja, wie viele? Identifikation und Analyse von Political Bots während des Bundestagswahlkampfes 2017 auf Twitter. In: C. Holtz-Bacha (Hrsg.): *Die (Massen-) Medien im Wahlkampf. Die Bundestagswahl 2017* (S. 97-124). Wiesbaden: Springer VS.

Puscher, F. (2017). Mein Bot, der Wahlkämpfer. *Absatzwirtschaft*, 6, S. 28-29.

Reuter, M. (2016). Fake News, Bots und Sockenpuppen – eine Begriffsklärung. Abgerufen von <https://netzpolitik.org/2016/fakenews-social-bots-sockenpuppen-begriffsklaerung/#maschinelle-fakeaccounts> (Stand: 12.10.2019).

Schuppen, J., Forßmann, C., Berndt, G. & Bogazliyan, S. (2017): EU-Austritt Großbritannien. Eine Chronik der Brexit-Ereignisse. Wirtschaftswoche. Abgerufen von <https://www.wiwo.de/politik/europa/eu-austritt-grossbritanniens-eine-chronik-der-brexit-ereignisse/24155158.html> (Stand: 20.10.2019).

Tagesschau.de (2019). Manipulationsversuche vor der Europawahl. Abgerufen von: <https://www.tagesschau.de/ausland/europawahl-manipulation-101.html> (Stand: 20.10.2019).

Twitter Inc. (2019). Glossar. Abgerufen von <https://help.twitter.com/de/glossary> (Stand: 14.09.2019).

Varol, O., Ferrara, E., Davis, C. A., Menczer, F. & Flammini, A. (2017). Online Human-Bot Interactions. Detection, Estimation, and Characterization. Abgerufen von <https://arxiv.org/pdf/1703.03107.pdf?ref=il> (Stand: 14.09.2019).

## 10. Diskurs über KI in den Massenmedien

Vortragstag und Gruppenname:	11. Juli 15:45 Uhr, Diskurs über KI in den Massenmedien
Titel der Ausarbeitung:	Diskurs über KI in den Massenmedien
Verfasser:	Lars Goeritz, Julian Ruppelt, Boran Özer
Betreuerin:	PD Dr. Sandra Nuy (Politikwissenschaften)

### Übersicht über die Struktur der Ausarbeitung

1. Einleitung
2. Theoretische Fundierung
  - 2.1 Narrative über intelligente Menschen
  - 2.2 Hoffnungen und Ängste
    - 2.2.1 Hoffnungen
    - 2.2.2 Ängste
3. Methodisches Vorgehen
4. Ergebnisse
  - 4.1 Wettbewerb
    - 4.1.1 Anzahl und Verteilung der Artikel
    - 4.1.2 Apokalyptische und euphorische Berichterstattung
  - 4.2 Sicherheit
    - 4.2.1 Anzahl und Verteilung der Artikel
    - 4.2.2 Apokalyptische und euphorische Berichterstattung
  - 4.3 Mensch-Maschinen-Verhältnis
    - 4.3.1 Anzahl und Verteilung der Artikel
    - 4.3.2 Apokalyptische und euphorische Berichterstattung
  - 4.4. Arbeitsplätze

#### 4.4.1 Anzahl und Verteilung der Artikel

#### 4.4.2 Apokalyptische und euphorische Berichterstattung

### 5. Fazit

#### Literaturverzeichnis

---

## 1. Problemstellung

Mit der zunehmenden Digitalisierung von Wirtschaft wird die Bedeutung künstlicher Intelligenz immer deutlicher. Hinsichtlich dieser Entwicklungen spielt auch zunehmend die Art und Weise, wie künstliche Intelligenz in der Öffentlichkeit wahrgenommen wird, in ihrer Entwicklung und ihrem Einsatz eine besondere Rolle. Aber auch die Darstellungen und die Narrative über künstliche Intelligenz sind von großer Bedeutung in der Entwicklung und Förderung. Vorherrschende KI-Narrative teilen dominante Merkmale und tendieren zu teilweise utopischen oder dystopischen Extremen. So neigen gängige Darstellungen zu einer starken Polung von extrem optimistischen und extrem pessimistischen Einstellungen gegenüber Entwicklungen von KI. Durch die rapide Entwicklung sind uns Maschinen auf vielen Feldern, die Menschen vorbehalten schienen, bereits überlegen. Besonders in den letzten fünf Jahren hat es auf dem Gebiet der künstlichen Intelligenz mehr Fortschritte gegeben als in den 50 Jahren zuvor (Eberl 2018: 8). Aber auch Algorithmen steuern immer mehr Arbeits- und Lebensbereiche und werfen Fragen auf, wie sich diese technologischen Fortschritte entwickeln werden: Was ist von Nutzen und Vorteil für den Menschen jenseits der technischen Machbarkeit? Wie verändern sich Wirtschaft, Arbeit und Alltag? Wie lassen sich Risiken steuern und inwieweit spielt Sicherheit eine Rolle? Die Antworten auf diese Fragen werden neben weiteren Einflussfaktoren im besonderen Maße von Narrativen über künstliche Intelligenz geformt, die in unserem alltäglichen Leben präsent sind. Dabei wird deutlich, dass das Denken über intelligente Maschinen kein neuer Trend ist, sondern historisch aufgearbeitet werden kann. So soll im Rahmen dieser Arbeit darüber berichtet werden, wie die Tendenzen der Berichterstattung in den drei größten Tageszeitungen Deutschlands rund um das Thema der künstlichen Intelligenz ausfallen und welche Narrative sich wiederfinden lassen.

Basis unseres Forschungsinteresses ist die Frage, in welchem Maße diese Narrative den Diskurs prägen und wie im Allgemeinen über KI berichtet wird. Dazu soll zunächst die theoretische Fundierung dieser Arbeit erläutert werden. Dabei geht es vor allem darum, wie die Narrative rund um das Thema der künstlichen Intelligenz einzuordnen sind. Im Anschluss wird das methodische Vorgehen vorgestellt, um die darauffolgenden Ergebnisse vorstellen zu können

## **2. Theoretische Fundierung**

Basis unseres Forschungsinteresses ist die Frage, wie die Tendenzen der Berichterstattung rund um das Thema der künstlichen Intelligenz in den drei größten Tageszeitungen ausfallen und inwieweit sich populäre Narrative wiederfinden lassen. Um den besonderen Faktor der Erzählungen über Künstliche Intelligenz fassen zu können, soll zunächst die Frage beantwortet werden, wie Narrative im Kontext von Künstlicher Intelligenz zu verstehen sind und welchen Einflussfaktor sie haben können. Dazu wird auf die historische Entwicklung des Denkens über Künstliche Intelligenz eingegangen, um die Merkmale, die bei der Entwicklung Künstlicher Intelligenz von Bedeutung waren herauszustellen. Diese Erzählungen sind besonders durch Hoffnungen und Ängste geprägt, die sich auch heute auf unterschiedliche Weise auswirken und den Verlauf der Technologien im besonderen Maße prägen können. Anschließend werden die für unsere Analyse wichtigen Begriffe, wie Apokalyptiker und Euphoriker definiert, welche die zuvor besprochenen Hoffnungen und Ängste als Gegensätze innerhalb der Narrative darstellen.

### **2.1 Narrative über intelligente Menschen**

Ein Narrativ ist vor allem mit dem Akt der Erzählung verbunden und wird von Monika Fludernik (2006) wie folgt beschrieben:

“A narrative is associated above all with the act of Narration and is to be found wherever someone tells us something: a newsreader on the radio, a teacher at school, a school friend in the playground, a fellow passenger on a train, a news-agent, one’s partner over the evening meal, a television reporter, a newspaper columnist or the narrator in the novel that we enjoy reading before going to bed“ (Fludernik 2006: 1).

Ob fiktiv oder nicht-fiktiv, Narrative funktionieren in der realen Welt, betreffen Einzelpersonen und kollektive und beeinflussen menschliches Handeln, Denken und soziale Auswirkungen. Dabei gibt es mindestens drei Möglichkeiten, wie diese Narrative die Technologien und ihre Auswirkungen gestalten können. Erstens könnten sie die Ziele der KI-Entwickler beeinflussen, indem Erzählungen die bereits laufende Forschung beeinflussen und weiterentwickeln können. Somit könnten diese Narrative als Anreize für neue Wege der Forschung verstanden werden (Cave 2019: S. 74). Zweitens können Narrative den öffentlichen Einfluss und besonders die Akzeptanz von KI-Systemen beeinflussen. Drittens können Narrative zukünftige Regulierungen beeinflussen, da diese sowohl politische Entscheidungsträger als auch ihre Wähler prägen und beeinflussen (ebd.).

Angesichts dieser Einflussfaktoren ist es von großer Bedeutung, dass die Erzählungen über Künstliche Intelligenz den aktuellen Stand und die Möglichkeiten der Technik weitgehend realistisch widerspiegeln. Obwohl das Konzept intelligenter Maschinen derzeit einer enormen Zunahme an Popularität ausgesetzt ist, handelt es sich um einen alten Gedanken. Die älteste bekannte Geschichte von einer Art Künstlicher Intelligenz kann in Homers Ilias aus dem 8. Jahrhundert v. Chr. gefunden werden. Dieser Gedanke wird im folgenden Zitat deutlich:

„Hephaistos hinkte sodann aus der Tür und Jungfrauen stützten den Herrscher, goldene, Lebenden gleich, mit jugendlich reizender Bildung; diese haben Verstand in der Brust und redende Stimme, haben Kraft und lernten auch Kunstarbeit von den Göttern. Schräge vor ihrem Herrn hineilten sie, er, der nachwankend, nahte, wo Thetis saß und ruht' auf schimmerndem Sessel“ (Homer, Ilias, 18. 417-21).

In diesem Zitat wird deutlich, dass diese Vorstellung intelligenter Maschinen besonders auf dem Gedanken von dienenden Maschinen aufbaut und den heutigen Vorstellungen intelligenter Maschinen sehr ähnlich ist. Dem Menschen und in diesem Falle dem Schmiedegott Hepahistos soll die Arbeit durch Maschinen abgenommen werden, die von ihren Göttern gelernt haben, wie sie diese einzusetzen haben. Andere Legenden schrieben Hephaistos ähnliche technologische Wunder zu. Besonders interessant ist Talos, eine riesige menschenähnliche Maschine aus Bronze, die an den Ufern Kretas patrouilliert und Steine auf Piraten und Eindringlinge wirft (Cave, Dihal 2019: 7). An dieser Stelle wird deutlich, dass die Vorstellung von Maschinen, die zur Herstellung von Sicherheit aber auch zur Ausübung militärischer Macht kein neuer Gedanke ist. Talon könnte also als der erste Killerroboter gesehen werden. Die größte Dichte an fiktiven Erzählungen, die sich mit künstlicher Intelligenz beschäftigen, lässt sich nach der Prägung des Begriffs Artificial Intelligence im Jahre 1955 beobachten. Diese Erzählungen bieten eine Fülle von phantasievollen Denkansätzen in Bezug auf eine Reihe von Themen. Viele diese Erzählungen sind dystopisch, einige utopisch und andere beinhalten Elemente aus beiden (Cave, Dihal 2019: 8). Durch die folgenden Ausführungen wird jedoch deutlich, dass derzeit viele Hoffnungen und Ängste, die mit Künstlicher Intelligenz verbunden werden, durch die Realität aus dem Gleichgewicht geraten.

## **2.2 Hoffnungen und Ängste**

Die britische Studie „Scary Robots“: Examining Public Responses to AI“ von Steven Cave und Kanta Dihal des Leverhulme Centre for the Future of Intelligence der Cambridge Universität und Kate Coughlan der BBC analysiert, wie Entwicklungen der Künstlichen Intelligenz an die breite

Öffentlichkeit kommuniziert und von ihr wahrgenommen werden. Um das Bewusstsein für gängige Narrative über Künstliche Intelligenz in der britischen Allgemeinbevölkerung zu analysieren, wurde online eine quantitative Umfrage unter 1078 Befragten durchgeführt. In Bezug auf Einstellungen gegenüber Technologien und Gerätebesitz im Allgemeinen stellte die Stichprobe ein typisches Abbild der britischen Bevölkerung dar, wobei 78% der Befragten ein Smartphone, 70% einen Laptop, 65% ein Tablet und 42% einen Desktop-Computer besaßen. Lediglich 1% besaß keines der aufgeführten Geräte. Bezüglich der Rolle, die Technologie in ihrem Leben spielt, waren 39% der Meinung, dass die Aussage „Ich kann nicht ohne sie leben, aber manchmal brauche ich eine kleine Pause“ ihre Situation am besten beschreibt. 1% waren der Meinung, dass die Aussage „Technologie macht mir Angst“ sie am besten repräsentiert (Cave, Dihal 2019: o.S). Sieben Fragen der Umfrage konzentrierten sich gesondert auf Künstliche Intelligenz. Dabei wurden die Befragten gebeten, acht Behauptungen über KI zu bewerten. Diese Behauptungen spiegelten Hoffnungen und Ängste wider, die dem phantasievollen Denken über Intelligente Maschinen zugrunde liegen. So argumentieren Cave und Dihal (2019), dass die Reaktionen auf KI in einen Rahmen von vier Dichotomien gestellt werden können - also vier Hoffnungen und vier parallele Ängste. Diese Dichotomien der Hoffnungen und Ängste lassen sich in der Berichterstattung der drei größten Tageszeitungen Deutschlands wiederfinden und werden in dieser Arbeit aufgenommen und bilden somit die Kategorien der Analyse. Dabei stehen die Hoffnungen für euphorische und die Ängste für apokalyptische Berichterstattungen.

### **2.2.1 Hoffnungen**

Die vier Hoffnungen, Unsterblichkeit, Mühelosigkeit, Genugtuung und Dominanz werden mit Narrativen in Verbindung gebracht, in denen intelligente Maschinen die Gesellschaft transformativ positiv beeinflussen. Unsterblichkeit spricht für den grundlegenden Antrieb, durch KI am Leben zu bleiben und gesund zu leben (Cave, Dihal 2019: 74). Das Streben nach Gesundheit und Langlebigkeit ist ein grundlegender menschlicher Antrieb. Daher ist es nicht verwunderlich, dass eine der großen Hoffnungen für die KI darin besteht, dass sie genau das tun wird. So werden der KI revolutionäre Möglichkeiten wie beispielsweise bessere Diagnosen, personalisierte Medizin und weniger medizinische Fehler im Gesundheitswesen zugeschrieben. Cave stellt an dieser Stelle zwei unterschiedliche Arten der Unsterblichkeit durch KI fest: Transformation und Transparenz. So werden in Transformationsprozessen immer ausgeklügeltere Vorbeugemaßnahmen, Medikamente und Prothesen eingesetzt, um den Körper immun gegen Alterung und Krankheiten zu machen. Im Extremfall ist dies ein Prozess der Cyborgisierung, bei dem die unzuverlässigen organischen Teile des Körpers durch langlebigere Maschinenteile ersetzt werden. Im Gegensatz dazu

beinhaltet der Prozess der Transzendenz das völlige Verlassen des menschlichen Körpers und das Hochladen des Geistes auf eine Maschine, die per Definition KI sein muss (Cave, Dihal 2019: 75). Mühelosigkeit bezieht sich auf den Wunsch, frei von Mühen zu sein und die Hoffnung, dass KI zunehmend viele Aufgaben erfüllen wird, die Menschen nicht tun wollen. Wie zuvor in Homers Zitat über Hephaistos maschinelle goldene Dienstmädchen deutlich wurde, ist dies eine sehr alte Hoffnung, die durch intelligente Maschinen erfüllt werden soll. Genugtuung bezieht sich auf die Art und Weise, wie man die freie Zeit nutzt und welche Rolle KI bei der Erfüllung dieser Wünsche spielen könnte. Sobald die KI die Hoffnungen auf ein längeres Leben und Bequemlichkeit erfüllt hat, besteht die Hoffnung darin, die verbleibende Zeit mit dem zu füllen, was uns Menschen Freude bereitet. So könnte die KI beispielsweise der perfekte Begleiter sein: immer da, immer bereit zuzuhören und nie etwas dafür zu verlangen. Schließlich kann Dominanz über andere als Mittel gesehen werden, indem KI zu mächtigen neuen Verteidigung- und Sicherheitsmitteln beiträgt (ebd.). Wie zuvor im Beispiel von Talos erwähnt ist auch die Vorstellung autonomer Waffen eine sehr alte Überlegung.

### **2.2.2 Ängste**

Mit jeder dieser Hoffnungen sind parallele Ängste verbunden. So birgt die Hoffnung auf Unsterblichkeit die Gefahr des Unmenschlichen, bei der der Mensch auf dem Weg zu einer immer längeren Lebensdauer riskiert, seine Identität zu verlieren und mehr Maschine als Mensch zu werden. Die zentrale Frage ist, ob es einem Individuum möglich ist, seine Identität durch eine radikale Verwandlung zu bewahren, die erforderlich ist, um einen gewöhnlichen Sterblichen in etwas Unsterbliches zu verwandeln. So kann dieser Verlust der Menschlichkeit so etwas wie den Verlust von menschlichen Werten und Emotionen bedeuten (Cave, Dihal 2019: 76). Der Hoffnung der Leichtigkeit steht die Angst der Obsoleszenz gegenüber. Der Wunsch, frei von Arbeit zu sein, wird also zur Angst, nicht mehr gebraucht zu werden, da Maschinen jegliche Tätigkeiten übernehmen könnten. Es scheint eine Grenze zu geben, wie viel Freizeit Menschen tolerieren können, bevor die Angst obsolet zu werden, einsetzt. Besonders, weil Arbeit dem Menschen nicht nur die Möglichkeit zur Erhaltung eines Lebensstandards durch Einkommen ermöglicht, sondern auch mit einer Rolle innerhalb der Gesellschaft ausstattet (ebd.). Die Genugtuung birgt das Risiko der Entfremdung, wenn Menschen in ihrem Wunsch nach künstlich perfekten Wechselwirkungen entfremdet werden und lieber mit Maschinen als mit Menschen agieren. Aber auch die Angst, dass etwas Unmenschliches eine so große Rolle in menschlichen Beziehungen einnehmen könnte, spielt eine besondere Rolle. Während diese Angst darauf beruht, dass KI nicht menschlich genug ist, gibt es auch Befürchtungen, dass KI in Zukunft besser als Menschen sein könnte, sodass



Menschen durch die Erfüllung aller Wünsche durch KI Systeme füreinander redundant werden könnten. Das Streben nach mehr Dominanz ruft Ängste vor einem Aufstand hervor, bei der die KI-fähige Macht eines Volkes gegen sie wendet (Cave, Dihal 2019: 77). Diese Erzählung lässt sich in Karel Čapek's Stück R.U.R (Rossum's Universal Robot) wiederfinden und ist seitdem in fiktiven aber auch nichtfiktiven Vorstellungen ein anhaltendes Thema.

### **3. Methodisches Vorgehen**

Nachdem der Forschungsstand über Künstliche Intelligenz zu den für uns relevanten Fragenstellungen vorgestellt wurde, werden wir nun unser methodisches Vorgehen für unsere eigene Untersuchung vorstellen. Für unsere Analyse haben wir uns für eine qualitative Inhaltsanalyse nach Mayring (2015) entschieden. Diese Untersuchungsform erlaubt es – anders als etwa quantitative Inhaltsanalysen – auch Einzelfälle stärker in den Blick zu nehmen, weshalb sie sich anbietet, um das inhaltlich breite Spektrum des KI-Diskurses nachzuzeichnen. Die Forschung dieser Arbeit folgte Mayrings inhaltsanalytischen Ablaufmodell (siehe Anhang S. 30, Abbildung 5). Als Untersuchungsgegenstand dienten uns die Online-Artikel der drei größten deutschen überregionalen Tageszeitungen im Zeitraum vom 7. März bis zum 5. April 2019. Die drei Zeitungen – FAZ, SZ und BILD – wurden ausgewählt, weil sie als Leitmedien gelten und somit großen Einfluss auf den Diskurs über KI haben (u.a. Pfanner 2011, Weischenberg/Malik/Scholl 2006 S. 346 ff.). Aus arbeitsökonomischen Gründen haben wir uns für Online-Artikel entschieden. Alle Online-Artikel der drei Zeitungen können nach Stichworten durchsucht werden, wodurch wir schneller feststellen konnten, ob sich ein Artikel inhaltlich mit dem Thema KI beschäftigt. Für unsere Untersuchung haben wir die Datenbanken der drei Zeitungen nach dem Schlagwort „künstliche Intelligenz“ durchsucht und alle Artikel berücksichtigt, die im Zeitraum vom 7. März bis zum 5. April von den Datenbanken gefiltert wurden. So kamen wir auf 114 auszuwertende Artikel. Von den 114 Artikel entfielen 66 Artikel auf die Süddeutsche Zeitung, 36 Artikel auf die FAZ und 12 Artikel auf die BILD-Zeitung. Die Süddeutsche Zeitung sticht hier klar hervor, weil sie im Untersuchungszeitraum eine eigene Serie zum Thema KI hatte und dadurch dieses Thema wohl auch verstärkt auf ihrer medialen Agenda hatte. Der Untersuchungszeitraum ist rund um den EU-Gipfel am 21. und 22. März gewählt worden, der das Thema KI auf der Tagesordnung hatte (Aamann/Laas 2019). Unsere Vorannahme dabei war, dass die politische Agenda maßgeblich auch die mediale Agenda der Tageszeitungen beeinflusst. Von der Wahl des Zeitraumes erhofften wir uns also eine höhere Artikelanzahl. Denn wir gehen mit dem Agenda-Building-Ansatz von Lang/Lang (1981) davon aus, dass politische Akteure ihre Themen auf die mediale Agenda setzen. Ein Nachteil der Wahl des Untersuchungszeitraums könnte sein, dass die Artikel thematisch stärker politisch gefärbt sind

und andere Blickweisen auf das Thema unter den Tisch fallen. Jedoch haben wir in der Berichterstattung ein sehr diverses Themenfeld vorgefunden, sodass der derzeitige KI-Diskurs weitgehend abgebildet sein sollte.

Für die Kategorienbildung unserer Inhaltsanalyse haben wir uns für ein induktives Vorgehen entschieden, was bedeutet, dass wir unsere Kategorien aus dem Material gebildet haben. Diese Kategorien wurden nach jedem Materialdurchlauf weiter spezifiziert und angepasst. Auch wenn wir über Vorkenntnisse über den KI-Diskurs (siehe Kapitel 2) verfügten, flossen diese nicht direkt in das Kategoriensystem ein. Die Kategorien wurden also autonom und ausschließlich aus den analysierten Zeitungsartikeln erstellt. Somit können unsere Untersuchungsergebnisse nun mit anderen Untersuchungen zum KI-Diskurs abgeglichen werden. Folgende Kategorien haben wir gebildet: Sicherheit, Wettbewerb, Mensch-Maschinen-Verhältnis und Arbeitsplätze (vollständiges Kategoriensystem: siehe Anhang S. 30 ff., Abbildung 6). Jede dieser Kategorien haben wir in Apokalyptiker und Euphoriker unterteilt, um die unterschiedliche Sichtweisen auf KI deutlich herauszustellen. Die Unterscheidung zwischen Apokalyptikern und Euphorikern geht weitgehend auf Umberto Eco (1984) zurück, der allerdings anstelle des Wortes Euphoriker das Wort Integrierte wählte. Nach seinem Ansatz gibt es in der Reaktion auf neue Technologien bzw. Medien immer zwei Gruppen. Die Apokalyptiker – bestehen meist aus Vertretern des alten Mediums – verbinden mit dem neuen Medium den Untergang (Eco 1984, S. 16 ff.). Die Integrierten – bestehen meist aus Technikern und Ingenieuren – überhöhen das neue Medium geradezu und sehen in diese eine Art Erlösung (ebd.).

Eco plädierte für einen dritten Weg, einen Weg der Mitte, der diese beiden Sichtweisen miteinander vereint (ebd. S. 38 ff.). Einen solchen intrigierenden Weg haben wir in der Berichterstattung allerdings nicht gefunden, weshalb wir uns auf die Opposition Apokalyptiker und Euphoriker beschränkt haben.

Folgende Analyseeinheiten haben wir für die Untersuchung gewählt:

Kleinste Kodiereinheit: ein Wort

Größte Kontexteinheit: der ganze Zeitungsartikel eines Autors

Auswertungseinheit: Fundstellen des Erhebungszeitpunktes

Die Kontext- und Kodiereinheiten wurden großzügig gewählt, damit die Forschung nicht eingeschränkt werden muss.

## 4. Ergebnisse

Die im Zeitraum vom 7. März 2019 bis zum 5. April 2019 erhobenen 114 Artikel sind unsere Grundgesamtheit. Eine apokalyptische Sichtweise konnte in 62 Artikeln festgestellt werden. Demgegenüber stehen 91 Artikel, in denen euphorisch über das Thema KI berichtet wird. Mit Blick auf die Verteilung der drei ausgewählten Zeitungen wird festgestellt, dass die SZ 30 apokalyptische und 42 euphorische Artikel, die FAZ 31 apokalyptische und 38 euphorische Artikel und die Bild-Zeitung 1 apokalyptischen und 11 euphorischen Artikel stellt. Die nachstehenden Abbildungen 1 bis 4 zeigen jeweils eine Tabelle, in der die Ergebnisse der 4 Kategorien dargestellt sind. Bevor die apokalyptische und euphorische Berichterstattung einer jeden Kategorie vorgestellt wird, soll zu Beginn kurz auf die Anzahl und Verteilung der Artikel der jeweiligen Kategorie eingegangen werden.

### 4.1 Wettbewerb

	Süddeutsche	FAZ	BILD
<b>Apokalyptiker: Anzahl Artikel</b>	12 Artikel	9 Artikel	1 Artikel
<b>Dominantes Narrativ 1</b>	Investitionen in KI notwendig, um wettbewerbsfähig zu bleiben → ohne Investitionen sind Arbeitsplätze und Wohlstand bedroht	Großer Nachholbedarf Deutschlands im KI-Bereich → hohe Investitionen notwendig, um Rückstand aufzuholen	
<b>Dominantes Narrativ 2</b>	Konkurrenzsituation mit China und Amerika	Krieg um KI	China soll als Rivale im KI-Bereich angesehen werden
<b>Euphoriker: Anzahl Artikel</b>	18 Artikel	13 Artikel	5 Artikel
<b>Dominantes Narrativ 1</b>	Wer in KI investiert, ist international konkurrenzfähig	Durch Investitionen international wettbewerbsfähig werden	Investitionen in KI → nächstes großes Thema
<b>Dominantes Narrativ 2</b>	Produktivitätssteigerungen durch KI	Produktivitätssteigerungen durch KI	

**Abbildung 1: Kategorie Wettbewerb**

#### 4.1.1 Anzahl und Verteilung der Artikel

Die erste Kategorie Wettbewerb ist mit 58 Artikeln auch die größte Kategorie. Alle 58 Artikel beschäftigten sich mit dem Einfluss von KI auf den Wettbewerb von Unternehmen und Staaten. 22 Artikel haben eine apokalyptische Sichtweise, 36 Artikel hingegen eine euphorische. SZ (30 Artikel) berichtet am häufigsten, dicht gefolgt von der FAZ (23 Artikel). Die Bild-Zeitung berichtet eher spärlich über die Thematik (6 Artikel).

#### 4.1.2 Apokalyptische und euphorische Berichterstattung

In den 22 Artikeln, die den Apokalyptikern der Kategorie Wettbewerb zugeordnet werden können, liegt der Fokus auf Nachteilen, die Unternehmen und Staaten durch mangelnde Investitionen im KI-Bereich haben oder haben könnten.

Die SZ berichtet in insgesamt 12 Artikeln vor allem darüber, dass Politik und Unternehmen dringend in KI investieren müssten. Andernfalls drohe der Verlust der Wettbewerbsfähigkeit, die wiederum zur Gefährdung des Wohlstands und zum Verlust von Arbeitsplätzen führe. In einigen Artikeln wird ferner auch die Konkurrenzsituation mit China und Amerika betrachtet (u.a. Kolb 2019; Kornelius 2019 und Fromm 2019). Gerade China gelte – aufgrund von größerer Population und protektionistischer Marktpolitik – als weitaus größere Gefahr für Europa. Aus diesem Grund appelliert die SZ in mehreren Artikeln, dass Europa im Hinblick auf KI-Investitionen mehr zusammenarbeiten solle (u.a. Dostert/Kutsche 2019). Letztlich seien Investitionen – gerade in Wissenschaft und Forschung – notwendig, um international konkurrenzfähig zu sein. In diesem Zusammenhang sieht auch die FAZ Deutschland unter Zugzwang. So wird in einem der neun FAZ-Artikeln eine Umfrage des Verbandes deutscher Ingenieure zitiert, nach der etwa zwei Drittel der befragten Ingenieure angaben, dass Deutschland die Kompetenzen fehlen, um KI-Technologie effizient einsetzen zu können. Dies sei eine bedenkliche Einschätzung der deutschen Fähigkeiten und Perspektiven in einer Zeit, in der künstliche Intelligenz auch die industrielle Produktion übernehmen könnte. Ähnlich wie in der SZ werden auch in den Artikeln der FAZ China und Amerika als führende KI-Nationen genannt (u.a. Pennekamp 2019, Kafsack 2019 und Armbruster 2019). Deutschland hingegen sei ein schwächelnder KI-Standort und spiele keine Rolle mehr, was sich besonders im Bereich des Deep Learning zeigt. In einem Artikel (Was ist künstliche Intelligenz?, 26.03.2019) prophezeit der Historiker Niall Ferguson zudem gar einen Krieg um künstliche Intelligenz, der noch Jahrzehnte andauern wird. Deutschland und die EU müssten daher mehr in KI investieren. Die EU (und damit auch Deutschland) soll mit Blick auf den Artikel (Chinapolitik im Wandel, 30.03.2019) in der Bild-Zeitung China weniger als Partner, sondern vielmehr als Rivale im KI-Bereich ansehen. Der Autor Alexander von Schönburg befürwortet daher vor allem die geplante staatliche Investitionskontrolle für chinesische Investitionen in Europa.

Der euphorische Blick auf KI ist in dieser Kategorie in 36 Artikeln zu finden. In allen 36 Artikeln werden Wettbewerbsvorteile von Staaten und Unternehmen durch Investitionen in KI thematisiert. Dabei werden vor allem die Narrative der Mühelosigkeit und der Genugtuung aufgegriffen.

In den 18 Artikeln der SZ heißt es überwiegend, dass durch Investitionen in künstliche Intelligenz die internationale Wettbewerbsfähigkeit erhalten bleibt. Ein entscheidender Wettbewerbsvorteil dabei: KI steigert die Produktivität. Vor allem im Bereich der Kundenbetreuung kann KI eingesetzt werden, um Mitarbeiter zu entlasten und mögliche neue Kunden (für das Unternehmen und den Staat) zu gewinnen (u.a. Hahn 2019, Högler/Jung 2019 und Kuhn 2019). Zudem könne KI – gerade im Bereich der Industrie – dafür sorgen, dass Maschinen rechtzeitig gewartet werden oder gegebenenfalls Schrauben schneller produziert werden (u.a. Beenen 2019). Auch in den 13 Artikeln der FAZ wird über Produktivitätssteigerungen, die KI ermöglichen kann, berichtet. Grundsätzlich sei künstliche Intelligenz in nahezu allen Sektoren einsetzbar. Gerade in der Industrie verspricht man sich große Effizienzsteigerungen und Produktionsgewinne. Besonders hervorgehoben werden die Strukturierung großer Datenmengen und die Verarbeitung von Prognosen durch künstliche Intelligenz in der Finanzwelt (u.a. Frühauf 2019, Germis 2019 und Giersberg 2019). In einem Artikel (Turing-Award für die drei Pioniere der künstlichen Intelligenz, 27.03.2019) werden der künstlichen Intelligenz generell wirtschaftliche und gesellschaftliche Auswirkungen wie einst bei der Erfindung der Elektrizität zugesprochen.

Der Tenor in den FAZ-Artikeln ist zudem: Deutschland steht in der KI-Forschung besser da als vermutet. Bei der Entwicklung der jüngsten und bisher erfolgreichsten Generation von künstlicher Intelligenz stünden deutsche Wissenschaftler nämlich an führender Stelle (u.a. Armbruster 2019 und Harhoff 2019). Allgemein verfügten die Länder der EU zusammengenommen über eine gute bis sehr gute Ausgangsposition. Um auch zukünftig international wettbewerbsfähig bleiben zu können, will die Bundesregierung daher bis zum Jahr 2023 etwa 500 Millionen investieren und führender KI-Standort werden (Ankenbrand 2019). Auch in den 5 Artikeln der Bild-Zeitung werden Investitionen in KI begrüßt, damit die Unternehmen und Staaten wettbewerbsfähig bleiben. In einem Artikel (Leipzig erforscht jetzt künstliche Intelligenz, 02.04.2019) wird KI beispielsweise nach der Digitalisierung als „DAS große Thema“ genannt, welches Firmen nicht verpassen dürfen um konkurrenzfähig bleiben zu können.

## 4.2 Sicherheit

	Süddeutsche	FAZ	BILD
Apokalyptiker: Anzahl Artikel	3 Artikel	9 Artikel	Kein Artikel in dieser Kategorie
Dominantes Narrativ 1	Große Datenschutzbedenken durch den Einsatz von KI → Eingriffe in Privatsphäre werden kritisch gesehen	Permanente Überwachung mithilfe von KI → politische Kontrolle durch KI	
Dominantes Narrativ 2		KI muss reguliert werden, um die Bürger zu schützen	
Euphoriker: Anzahl Artikel	6 Artikel	5 Artikel	1 Artikel
Dominantes Narrativ 1	Ausblendung von Inhalten durch KI machen das Internet zu einem besseren Ort	Sicherheit im Netz wird durch KI erhöht	KI kann unseriöse Immobilienangebote herausfiltern
Dominantes Narrativ 2	Bevölkerung in China: Kameraüberwachung mithilfe von KI nicht so schlimm	Förderung von Mitarbeitern mithilfe von KI	

Abbildung 2: Kategorie Sicherheit

### 4.2.1 Anzahl und Verteilung der Artikel

Auf die Kategorie Sicherheit entfallen 24 Artikel. Alle Artikel haben den Einfluss von KI auf die Sicherheit in den Bereichen Datenschutz oder Privatsphäre thematisiert. Die Verteilung der Artikel ist dabei ausgeglichen: zwölf Artikel weisen eine apokalyptische Sichtweise auf, wogegen in den anderen zwölf Artikeln euphorisch über die Thematik berichtet wird. Die FAZ stellt die meisten Artikel (14 Artikel), dicht gefolgt von der SZ, die in neun Artikeln berichtet. Die Bild-Zeitung stellt hingegen lediglich einen Artikel.

### 4.2.2 Apokalyptische und euphorische Berichterstattung

Mit Blick auf die apokalyptische Sichtweise kann festgestellt werden, dass alle zwölf Artikel die Risiken von KI die den Datenschutz oder die Privatsphäre betreffen, thematisieren. In allen Artikeln wird zudem die Angst vor dem Aufstand als Narrativ verwendet.

In den drei Artikeln der SZ werden vor allem große Datenschutzbedenken geäußert, die durch den Einsatz von KI entstehen (u.a. Chinas große App, 20.03.2019). Grundsätzlich werden die Eingriffe in die Privatsphäre kritisch gesehen, weshalb Marc Zuckerberg in einem weiteren Artikel andere Unternehmen – und auch Facebook selbst – in die Pflicht nimmt, den Datenschutz bei KI zu erhöhen (Hurtz 2019). Die Angst davor, dass sich zukünftig die KI selbst oder ein KI-fähiger Staat gegen die Menschen wendet wird auch in den neun Artikeln der FAZ aufgegriffen. So wird

darüber berichtet, dass durch die Überwachung und Massenspeicherung von Daten ein Bonitäts-system geschaffen werden soll. Damit soll unter anderem das Verhalten der Bevölkerung in allen Lebensbereichen kontrolliert und bewertet werden (Eppelsheim / Kusicke 2019). In einem weiteren Artikel spricht auch ein KI-Roboter selbst darüber, dass KI dazu missbraucht werden könne, um der Welt und der Menschheit zu schaden (Scheer 2019). Die Microsoft-Deutschland-Chefin Sabine Bendiak mahnt mit Blick auf den Einsatz von KI daher, dass auch über den richtigen gesetzlichen Rahmen gesprochen werden muss. Insbesondere die Gesichtserkennung müsse zwingend reguliert werden, um die Bürger zu schützen (Armbruster 2019).

Die zwölf euphorisch zugeordneten Artikel beschäftigen sich vor allem mit den Chancen von KI im Hinblick auf den Datenschutz oder die Privatsphäre. In den Artikeln wird hauptsächlich das Narrativ der Dominanz verwendet. Die SZ und die Bild-Zeitung greifen in ihrer Berichterstattung zudem auf das Narrativ der Mühelosigkeit zurück.

In den sechs Artikeln der SZ wird darüber berichtet, dass KI in Zukunft viele Aufgaben erfüllen könne, die Menschen nicht beziehungsweise nur ungerne übernehmen würden. Explizit geht es um das Ausblenden und Löschen von rassistischen und terroristischen Inhalten im Internet (u.a. Ratzesberger 2019; Moorstedt 2019; Brühl 2019). In einem anderen Artikel wird ferner auch auf die Kameraüberwachung in China eingegangen. Ein Großteil der Chinesen beurteilt diese – durch KI ausgeführte – Überwachung positiv, da die Kriminalität gesunken sei (Mayr 2019). Deutschland hingegen sei – um nun einen Blick auf die fünf Artikel der FAZ zu werfen – noch nicht in der Lage zu einer flächendeckenden Überwachung (Kuner 2019). Auch die FAZ berichtet in ihren Artikeln über die Erhöhung der Sicherheit im Internet und greift hiermit im Speziellen das Narrativ der Dominanz auf: so wurden Apps entwickelt, die Kinder beim Chatten im Netz vor Gefahren schützen sollen (u.a. Ludwig 2019). Die Bild-Zeitung berichtet in ihrem Artikel („Eine dubiose Firma inserierte Immobilie – offenbar eine Masche. Im Internet wird mein Haus verkauft“, 11. März 2019) darüber, dass Immobilien-Scout24 durch den Einsatz von KI unseriöse Immobilienangebote herausfiltern kann. Unklar ist jedoch, wie genau das geschieht.

### 4.3 Mensch-Maschinen-Verhältnis

	Süddeutsche	FAZ	BILD
Apokalyptiker: Anzahl Artikel	9 Artikel	10 Artikel	Keine Artikel in dieser Kategorie
Dominantes Narrativ 1	Mensch nicht mehr bestimmende Instanz / menschliche Werte nicht mehr entscheidend	Intelligente Maschinen steuern die Menschen	/
Dominantes Narrativ 2	Diskriminierung gesellschaftlicher Gruppen durch KI	Menschen dürfen durch KI nicht diskriminiert werden	/
Dominantes Narrativ 3	KI wird von Menschen als Machtinstrument missbraucht	KI als Gefahr für die Demokratie	/
Euphoriker: Anzahl Artikel	14 Artikel	14 Artikel	4 Artikel
Dominantes Narrativ 1	Künstliche Intelligenz ersetzt nicht den Menschen	Maschine kann menschliche Komplexität noch nicht abbilden	
Dominantes Narrativ 2	Den Menschen kann durch KI im Alltag geholfen werden	KI als Unterstützer des Menschen	Funktionierende Mensch-Maschine-Interaktion

Abbildung 3: Kategorie Mensch-Maschinen-Verhältnis

#### 4.3.1 Anzahl und Verteilung der Artikel

Der Einfluss von KI auf das Verhältnis von Mensch und Maschine wurde in 51 Artikeln thematisiert. Davon weisen 19 Artikel eine apokalyptische Sichtweise auf. Demgegenüber stehen 32 Artikel in denen euphorisch über die Thematik berichtet wird. Die FAZ (24 Artikel) und die SZ (23 Artikel) berichten gemessen an der Anzahl der Artikel relativ ausgeglichen. Die Bild-Zeitung setzt sich in vier Artikeln mit dem Thema auseinander.

#### 4.3.2 Apokalyptische und euphorische Berichterstattung

In den 19 apokalyptischen Artikeln werden die Gefahren die für den Menschen durch den Einsatz von KI entstehen angesprochen.



In den neun Artikeln der SZ wird häufig die Angst der Obsoleszenz als Narrativ aufgegriffen. So wird befürchtet, dass der Mensch durch den Einsatz von KI nicht mehr die bestimmende Instanz ist. In Zukunft seien die menschlichen Werte somit nicht mehr entscheidend, da die Maschinen über den Menschen herrschen würden. Die Angst der Obsoleszenz ist somit eng verbunden mit der Angst vor dem Aufstand. Ferner heißt es auch, dass Menschen die KI als Machtinstrument nutzen können, um zum Beispiel Kontrolle über andere Menschen zu erlangen. Damit einher geht auch die Diskriminierung gesellschaftlicher Gruppen durch KI (u.a. Dennett C. 2019; Spiekermann 2019). Die FAZ warnt in ihren zehn Artikeln ebenfalls vor dem Aufstand und einer möglichen Diskriminierung durch KI. In einem Artikel („Kritik an Zusammensetzung. Google löst KI-Ethikrat nach einer Woche wieder auf“, 5. April 2019) spricht beispielsweise der Google-Chef Sundar Pichai, dass sichergestellt werden müsse, dass Menschen nicht aufgrund von Hautfarbe oder Geschlecht benachteiligt werden. In einem weiteren Artikel heißt es, dass ohne Transparenz und Zuverlässigkeit intelligente Maschinen irgendwann den Menschen steuern werden. Die KI würde damit eine Gefahr für die Demokratie darstellen, da keine politischen Wahlen mehr stattfinden müssten. Denn die Präferenzen können dann – ähnlich wie die politische Handlungsstrategie – einfach errechnet werden. Dies würde Parlamente überflüssig machen. In diesem Zusammenhang wird auch von einem digitalen Imperialismus gesprochen, bei dem die Überwachung und Herrschaft zukünftig durch intelligente Maschinen ausgeführt wird (Eppelsheim/Kusicke 2019).

Diese düsteren Zukunftsaussichten werden in den meisten Artikeln hingegen nicht behandelt. Im Gegenteil: in den 32 euphorisch zugeordneten Artikeln stehen die Verbesserungen und Chancen für den Menschen im Fokus. Um diese – neben der Überlegenheit der menschlichen Intelligenz – zu verdeutlichen, nutzen die Zeitungen in ihrer Berichterstattung die Narrative der Unsterblichkeit und der Mühelosigkeit.

In den 14 Artikeln der SZ wird unter anderem darüber berichtet, dass KI für den Menschen unliebsame Aufgaben – wie etwa das Bedienen des Navigationsgerätes – übernehmen könne (Reek 2019). In einem anderen Artikel („Themen & Trends. Mehr oder weniger“, 2. April 2019) heißt es auch, dass KI den Menschen nicht vollends ersetzen könne, da der Wille ebenso wie die Entscheidung zentrale menschliche Eigenschaften bleiben. Letztlich seien es die Menschen selbst, die Maschinen – ausgestattet mit künstlicher Intelligenz – überhaupt erst lernfähig machen. Daher setzt künstliche Intelligenz menschliche Intelligenz immer voraus (Dostert / Kutsche 2019). Aus diesem Grund können Maschinen nur die Macht übernehmen, wenn der Mensch sie auch lässt. Derzeit sind Maschinen nicht in der Lage, im Kantischen Vernunftsinne zu handeln und seien

daher nur der Parasit der menschlichen Intelligenz (Dennett C. 2019). Auch die FAZ stellt 14 Artikel und betrachtet die KI ebenfalls als Unterstützer des Menschen. In den Artikeln steht primär das Narrativ der Unsterblichkeit des Menschen im Fokus. Gerade die gesamte Medizinbranche könne durch den Einsatz von KI profitieren. So könnten beispielsweise Pflegekräfte in Zukunft zunehmend entlastet werden und daneben wäre es durch KI auch möglich, die Genauigkeit von Diagnosen und Behandlungen zu steigern. Vor allem im Hinblick auf die Krebsdiagnostik und -Therapie können Tumore somit schneller erkannt werden (u.a. Winkels / Herzog 2019; Dannenberger/Herzog 2019). Und auch die Bild-Zeitung verweist in einem („Börsenkolumnistin Beate Sander. Mit Bio-Tech-Aktien Geld verdienen“, 19. März 2019) ihrer vier Artikel auf das Narrativ der Unsterblichkeit. Unter der Voraussetzung einer funktionierenden Interaktion zwischen Mensch und Maschine sei es demnach in Zukunft möglich, mithilfe von großen Datenmengen lebensbedrohlichen Krebs, einen Herzinfarkt, Alzheimer und andere schwere Krankheiten erfolgreich zu bekämpfen. Außerdem könne KI ferner auch an Bahnhöfen und Flughäfen zum Einsatz kommen, um dort Kunden bei etwaigen Fragen und Problemen zu helfen (Goedert 2019).

#### 4.4 Arbeitsplätze

	Süddeutsche	FAZ	BILD
<b>Apokalyptiker: Anzahl Artikel</b>	6 Artikel	3 Artikel	Keine Artikel in dieser Kategorie
<b>Dominantes Narrativ</b>	Durch Roboter fallen menschliche Arbeitsplätze weg (allgemein)  Bsp. BCG-Prognose: bis 2025 fallen bis zu 300 000 Stellen weg	Durch Roboter fallen menschliche Arbeitsplätze weg (konkret)  Bsp.: Übersetzer, Dolmetscher und Juristen	/
<b>Euphoriker: Anzahl Artikel</b>	4 Artikel	6 Artikel	1 Artikel
<b>Dominantes Narrativ 1</b>	Arbeitsplätze bleiben erhalten: lediglich Verlagerung in andere Bereiche	Durch KI entstehen neue Jobs → nicht zwingend weniger Arbeitsplätze	
<b>Dominantes Narrativ 2</b>	Maschinen können nicht alle menschlichen Aufgaben übernehmen		Maschinen können nicht jeden Job übernehmen

**Abbildung 4: Kategorie Arbeitsplätze**

#### 4.4.1 Anzahl und Verteilung der Artikel

In der Kategorie Arbeitsplätze haben sich insgesamt 20 Artikel mit dem Einfluss von KI auf die Arbeitsplätze beschäftigt. Euphorische und apokalyptische Sichtweisen halten sich in etwa die Waage. 9 Artikel haben eine apokalyptische Sichtweise; 11 Artikel eine euphorische. In der SZ (10 Artikel) und FAZ (9 Artikel) erschienen zudem auch fast gleich viele Artikel. Die BILD-Zeitung thematisiert das Thema der Arbeitsplätze nur in einem Artikel.

#### 4.4.2 Apokalyptische und euphorische Berichterstattung

In allen 9 Artikeln mit apokalyptischer Sichtweise auf KI wird davon ausgegangen, dass die Arbeitslosigkeit durch den Einsatz von KI steigt. Hier spielt das Narrativ der Obsoleszenz eine große Rolle.

In der **SZ** wird dieses Narrativ eher allgemein beschrieben. In einem Artikel (*Mit jeder Milliarde kommen die Roboter näher*, 31.03.2019) wird eine Prognose der BCG erwähnt, wonach bis zum Jahr 2025 in Deutschland bis zu 300.000 Stellen wegfallen. In weiteren Artikeln wird unter anderem befürchtet, dass Roboter menschliche Aufgaben übernehmen könnten. Welche das sein könnten, wird jedoch nicht näher erläutert (u.a. Dostert/Kutsche 2019, Prantl 2019 und Wagner 2019). Konkreter wird es da schon in den drei Artikeln der **FAZ**. Gerade die Branche der Übersetzer und Dolmetscher sei demnach durch den Einsatz von künstlicher Intelligenz gefährdet. In einem Artikel heißt es etwa, dass der SPD-Generalsekretär Lars Klingbeil diese Berufe bereits 2018 auf die rote Liste der gefährdeten Berufe setzte (Brummer 2019). Auch Juristen könnten durch KI zukünftig abgelöst werden (dpa/FAZ 2019). Bei den 11 Artikeln mit euphorischer Sichtweise auf das Thema KI werden in dieser Kategorie keine negativen Auswirkungen durch den Einsatz von KI auf die Arbeitsplätze erwartet. Dabei wird vereinzelt in den Artikeln auf das Narrativ der Mühelosigkeit eingegangen. In den 4 Artikeln der SZ heißt es, dass sich die Arbeitsplätze lediglich in andere Bereiche verlagern (u.a. Beenen 2019 und Fromm/Radomski 2019). In der Landwirtschaft etwa können nicht alle Aufgaben von KI übernommen werden (Hoferichter 2019). Der Tenor in den 6 Artikeln der FAZ ist, dass durch den Einsatz von künstlicher Intelligenz zukünftig andere und teilweise neue Berufsbilder entstehen. So wird in einem Artikel („Der Arbeitsmarkt für KI-Experten ist leergefegt“, 19.03.2019) darauf hingewiesen, dass insbesondere der Arbeitsmarkt für Expertinnen und Experten im maschinellen Lernen derzeit nicht gegeben ist. Diejenigen, die mit der neuen Technik umgehen können, erhalten somit zukünftig eine höhere Wertschätzung. In einem weiteren Artikel (So will Bertelsmann mit Google und Co.

mithalten, 26.03.2019) glaubt Bertelsmann-Vorstand Thomas Rabe, dass die Ansprüche in der Zukunft zwar steigen werden, es jedoch dadurch nicht zwingend weniger Arbeitsplätze geben würde.

Die **Bild-Zeitung** hat in diesem Zusammenhang lediglich einen Artikel veröffentlicht (*Hier gibt ein Roboter den Takt vor*, 14.03.2019). Sie beschwichtigt in ihrem Artikel die Bedenken der Apokalyptiker. Ein Roboter, der bei einem Konzert der Düsseldorfer Philharmoniker dirigierte, sei kein Anzeichen dafür, dass Roboter jeden Job übernehmen können.

## 5. Fazit

Nach der einhergehenden qualitativen Inhaltsanalyse kann festgehalten werden, dass die drei größten deutschen überregionalen Tageszeitungen eher euphorisch über künstliche Intelligenz berichten. Die Ergebnisse der Analyse haben gezeigt, dass eine euphorische Sichtweise in 91 Artikeln erkennbar ist. Allerdings werden auf der anderen Seite auch die Risiken von KI in den Artikeln angesprochen und erläutert. Es konnte in 62 Artikeln eine apokalyptische Berichterstattung festgestellt werden.

Bei genauerer Betrachtung der drei Zeitungen fällt zudem auf, dass die SZ (*66 Artikel*) am häufigsten über die Thematik berichtet. Die FAZ folgt mit 36 Artikeln, während die Bild-Zeitung nur 12 Artikel stellt. Als zentrale Themenschwerpunkte konnten vor allem der Wettbewerb (*58 Artikel*), das Verhältnis von Mensch und Maschine (*51 Artikel*), die Sicherheit (*24 Artikel*) und die Arbeitsplätze (*20 Artikel*) in den Zeitungen ausgemacht werden. Daneben verdeutlicht die Analyse, dass sich auch die eingangs vorgestellten Narrative der Hoffnung und Narrative der Ängste in den insgesamt 114 Artikeln wiederfinden lassen.

Insbesondere die Hoffnung der Mühelosigkeit – also der Wunsch des Menschen, dass unbeliebte Aufgaben zukünftig von der KI übernommen werden – wird in einem Großteil der Artikel aufgegriffen. So könne KI beispielsweise im Internet negative oder terroristische Kommentare in kürzester Zeit filtern und löschen. Außerdem sei KI in allen Sektoren einsetzbar und könne große Datenmengen besser strukturieren als der Mensch. Dies hat nicht nur zur Folge, dass sowohl Prognosen schneller verarbeitet werden können, als auch Produktionsgewinne eingefahren werden können. Andererseits wird auch die Angst der Obsoleszenz in den Artikeln thematisiert. Schließlich sollen Roboter zukünftig immer häufiger menschliche Aufgaben übernehmen. Dies hat zur Folge, dass immer mehr Stellen

wegfallen. Speziell die Berufe der Dolmetscher und Übersetzer seien gefährdet. Aber auch Juristen – die heutzutage teilweise schon durch KI unterstützt werden – könnten zukünftig durch diese abgelöst werden. Der Wunsch durch KI ein möglichst langes und gesundes Leben zu leben – sprich die Hoffnung auf Unsterblichkeit – wird von den Zeitungen ebenfalls angesprochen. Demnach wird gehofft, dass durch KI unter anderem lebensbedrohliche Krankheiten erfolgreich bekämpft werden können. Grundsätzlich überwiegen in den meisten Artikeln aber die Hoffnung der Dominanz und somit auch die Angst vor dem Aufstand. Durch den Einsatz von KI soll besonders die Nutzung des Internets sicherer werden. In den analysierten Artikeln werden explizit Apps genannt, die Kinder und Jugendliche vor Gefahren im Netz schützen sollen. Daneben soll auch die Kriminalität durch eine KI-gesteuerte Überwachung gesenkt werden. Die Angst vor dem Aufstand, bei der sich die KI selbst – oder ein KI-fähiger Staat – gegen die Menschen wendet wird ebenfalls intensiv in den Artikeln diskutiert. Gerade autoritäre Staaten können durch den Einsatz von KI-Technologie die eigenen Bürger kontrollieren und massiv überwachen. Diese teilweise schwerwiegenden Eingriffe in die Grundrechte seien heutzutage ohne Probleme möglich. Auch die KI selbst könne irgendwann den Menschen steuern und eine Gefahr darstellen, weshalb in dieser Hinsicht bereits von einem digitalen Imperialismus die Rede ist. Dabei werden Menschen von intelligenten Maschinen überwacht, beherrscht und diskriminiert.

Alles in allem bemühen sich die drei Zeitungen um eine objektive Berichterstattung und wägen in ihren Artikeln die Pros und Contras gegeneinander ab. Der Leser möge zudem erkennen, dass die in dieser Arbeit getroffenen Aussagen grundsätzlich nicht forschungsrelevant sein können. Dafür ist der gewählte Zeitraum vom 7. März 2019 bis zum 5. April 2019 zu klein und birgt einen hohen Unsicherheitsfaktor. Um präzisere und unserer Meinung nach forschungsrelevantere Aussagen treffen zu können, müsste der Untersuchungszeitraum deshalb erweitert werden. Dabei könnte es sehr wahrscheinlich sein, dass sich die getroffenen Aussagen in die eine wie in die andere Richtung verändern.

---

## Literaturverzeichnis

Aamann, Preben/Laas, Jüri (2019): Europäischer Rat, 21./22.03.2019. Wichtigste Ergebnisse.Europäischer Rat. Online verfügbar unter <https://www.consilium.europa.eu/de/meetings/european-council/2019/03/21-22/#> (zuletzt geprüft am 11.09.2019).

Cave, Stephen; Claire, Craig; Dihal, Kanta; Dillon, Sarah; Montgomery, Jessica; Singler, Beth; Taylor, Lindsay (2018): Portrayals and Perceptions of AI and Why The Matter. London: The Royal Society. <https://royalsociety.org/~media/policy/projects/ai-narratives/AI-narratives-workshop-findings.pdf> (zuletzt geprüft am: 14.09.2019).

Cave, Stephen; Coughlan, K.: Dihal, Kanta (2019): Scary Robots: examining public responses to AI. AIES. [http://www.aies-conference.com/wp-content/papers/main/AIES-19\\_paper\\_200.pdf](http://www.aies-conference.com/wp-content/papers/main/AIES-19_paper_200.pdf) (zuletzt geprüft am 14.09.2019).

Cave, Stephen; Dihal, Kanta (2019): Hopes and Fears for Intelligent Machines in Fiction and Reality. Nature Machine Intelligence. Nr.2 (Februar).

Eberl, Ulrich (2018): Was ist Künstliche Intelligenz - Was kann sie leisten? Aus Politik und Zeitgeschichte (APuZ), S. 8–14.

Eco, Umberto (1984): Apokalyptiker und Integrierte. Zur kritischen Kritik der Massenkultur. 2. Aufl., Frankfurt a.M.: Fischer.

Fludernik, Monika (2006): An Introduction to Narratology. Darmstadt: Wissenschaftliche Buchgesellschaft.

Lang, Gladis E.; Lang, Kurt (1981): Watergate. An Exploration of the Agenda-Building Process. In: Wilhoit, G. Cleveland/ Bock, Harold de (Hg.): Mass communication review yearbook. Sage: Beverly Hills, Calif, S. 277–290.

Mayring, Philipp (2015): Qualitative Inhaltsanalyse. Grundlagen und Techniken. 12., überarb. Aufl. Weinheim: Beltz (Beltz Pädagogik).

Pfanner, Eric (2011): Gloves Off in German Media Scramble. New York Times. Online  
Verfügbar unter <https://www.nytimes.com/2011/03/14/business/global/14bild.html?pagewanted=all> (zuletzt geprüft am 11.09.2019).

Weischenberg, Siegfried/ Malik, Maja/Scholl, Armin (2006): Journalismus in Deutschland 2005. Zentrale Befunde der aktuellen Repräsentativbefragung deutscher Journalisten. In: Media-Perspektiven, S. 346–361.

Wilhoit, G. Cleveland/ Bock, Harold de (Hg.) (1981): Mass communication review yearbook. Sage: Beverly Hills, Calif.

## 11. Fazit und Ausblick des Projekts

Die Vorträge der Tagung „Zukünfte KI - KI in Medien, Politik, Gesellschaft und Wirtschaft“ beschäftigten sich mit verschiedenen Themen rund um künstliche Intelligenz und betrachteten dabei unterschiedliche gesellschaftliche Bereiche, die von dieser Technologie berührt werden. Die Interdisziplinarität der Tagung stand dabei besonders im Vordergrund, da so zum einen der interdisziplinären Ausrichtung des Masterstudiengangs Medien und Gesellschaft, zum anderen aber auch der bereichs- und wissenschaftsübergreifenden Relevanz der Thematik „Künstliche Intelligenz“ Rechnung getragen wird. Ziel der Tagung sollte es sein, anhand von Beispielen darzulegen, wie künstliche Intelligenz die Gegenwart in Bereichen der Medien, Politik, Gesellschaft und Wirtschaft gestaltet, einen Ausblick darüber zu geben, wie sich diese Bereiche und das Verhältnis von Mensch und Maschine zukünftig verändern könnten und sich im Diskurs über mögliche Zukünfte mit künstlicher Intelligenz auszutauschen.

Eröffnet wurde die Tagung durch den Keynote Speaker Dr. Ralf Winkler. Er berichtete über einen Algorithmus, der in seinem Hamburger StartUp QualiFiction eingesetzt wird, um Manuskripte innerhalb kürzester Zeit zu analysieren und einen Erfolgsscore für diese zu ermitteln. Die Vorteile dieser Technologie liegen darin, dass sie nicht nur zeitsparend und arbeitserleichternd ist, sondern auch einem „Übersehen-Werden“ von Manuskripten vorbeugt, die sonst gegebenenfalls aufgrund von Zeitmangel oder ähnlichem von Verlagen unbeachtet blieben. Herr Dr. Winkler präsentierte damit einen praktischen Anwendungsfall künstlicher Intelligenz im Redaktionswesen und ermöglichte, diese mit ihren fallspezifischen Vorteilen und Grenzen anhand eines konkreten Beispiels zu erfahren.

Der Vortrag „Ethische Problemstellungen und Grenzen der KI“ leitete anschließend in die Vortragsreihe der Tagung ein und eröffnete eine ethische Perspektive auf das Thema „Zukünfte KI“. Es wurden einige grundlegende Fragen bezüglich des Verhältnisses von Mensch und künstlicher Intelligenz aufgeworfen und Raum geboten, diese ethisch zu reflektieren und zu diskutieren. Besonders im Vordergrund standen dabei zum einen die Aspekte der Autonomie von künstlicher Intelligenz, das Machtverhältnis Mensch-Maschine und die Bedeutung, die diese beiden Aspekte für unsere Gesellschaft und unsere Identität als Mensch haben können. Zum anderen wurde auch die Frage der Verantwortung diskutiert; wer haftet, wenn Maschinen autonom handeln? Aus dem Vortrag und der anschließenden Diskussion ist deutlich geworden: Die Folgen, die eine Weiterentwicklung künstlicher Intelligenz und eine mögliche Autonomie dieser auf unsere Gesellschaft und unsere Identität als Mensch haben können, sind kaum



abschätzbar, was wiederum die Relevanz eines gesellschaftlichen und Disziplinen übergreifenden Diskurses sowie einer ethischen Betrachtung des Themas verdeutlicht.

Ethische Fragen wurden auch in dem zweiten Vortrag „Autonomie vs. Unterstützung - Künstliche Intelligenz in der Medizin“ behandelt. Dieser Vortrag untersuchte, wie der Einsatz von künstlicher Intelligenz und eine mögliche Autonomie von Maschinen in der Medizin aussehen und was dies für Patienten und Personal bedeuten könnte. Der Vortrag zeigte dabei auf, dass die Chancen, die Maschinen, wie Pflege- oder Operationsroboter, und ihre zukünftige Weiterentwicklung bieten, nicht nur ökonomischer Natur, sondern auch gesellschaftlicher und sozialer sind, insbesondere vor der Annahme eines wachsenden Pflege-Notstandes. Doch auch die Risiken, die mit einer steigenden Autonomie von Maschinen im Bereich der Medizin verbunden sind, betreffen mehr als einen gesellschaftlichen Bereich. Neben der Frage nach der Verantwortung für die Handlungen dieser Maschinen ist hier auch der Punkt der Datenspeicherung und Datenübertragung zu benennen. Die Digitalisierung von Patientendaten und ihrer Sicherheit bedarf nicht nur einer rechtlichen, sondern auch moralischen Betrachtung. Darüber hinaus sind auch die Datengenauigkeit und -richtigkeit für Big-Data-Analysen, wie sie bei Diagnosen verwendet werden, und die zielgerichtete Entwicklung künstlicher Intelligenz maßgeblich. Zu klären bleibt jedoch, wer diese Richtigkeit und Sicherheit definiert und sie garantiert.

Mit Big Data beschäftigte sich auch der dritte Vortrag „Mikroarbeit oder die Organisation des künstlichen Geistes“. Dieser betrachtete die Zerlegung von menschlicher Arbeitskraft in „Mikroarbeit“, also der Zerlegung in elementare, kleinschrittige Aufgaben, und ihrer Integration in digitale Infrastrukturen. Die Durchführung dieser „Mikroarbeit“ wirkt dabei nach außen hin, als würde sie von Maschinen durchgeführt, tatsächlich werden für diese Elementartätigkeiten jedoch Menschen eingesetzt, sie simulieren also technische Prozesse. Maschinelle Systeme können dabei auf Grundlage der Integration der Mikroarbeit, ihrer Datafizierung und Formalisierung, von den menschlichen Arbeitsprozessen lernen. Sie werden somit zum Beobachter menschlichen Handelns. Als Weiterentwicklung dieser Mikroarbeit ohne monetäre Entlohnung dient eine Integration des Nutzers durch dessen Verhaltensfeedback. Künstliche Intelligenz und die Simulation maschineller Prozesse verändern damit die Organisation von Arbeit und steuern sowohl Mikroarbeit als auch Nutzerverhalten.

Mit dem Vortrag „Künstliche Intelligenz in Literatur und Film“ und der Betrachtung von Figurenkonzepten hinsichtlich KI wurde am zweiten Tag der Tagung eine Brücke zu einer ästhetischen Betrachtung des Themas „Künstliche Intelligenz“ geschlagen. Die Entwicklung von Figurenkonzepten intelligenter Maschinen reicht in Film und Literatur vom Assistenten, Begleiter oder dem Objekt der Begierde bis hin zum Sklaven und Kontrollsystem und vom Antagonisten hin zum Protagonisten. So vielfältig wie die Figur der intelligenten Maschine in Film und Literatur sind auch die Themen der Werke, die sich mit diesen Maschinen auseinandersetzen. Sie greifen ganz unterschiedliche Ängste, Hoffnungen und Vorstellungen über künstliche Intelligenz und ihre zukünftige Gestalt auf. Ängste werden beispielsweise durch KI als emotionsloses Wesen, als Bedrohung oder als Kontrollsystem thematisiert. Dem gegenüber steht die KI als emotionales Wesen, das der Menschheit dienlich ist oder sie rettet. Doch nicht nur die Themendarstellung als Repräsentant von Ängsten und Hoffnungen lassen Rückschlüsse auf die Vorstellung von künstlicher Intelligenz und dem kulturellen Umgang mit dieser zu, auch die visuelle Figurengestaltung der KI ist hier interessant. Sie spiegelt Vorstellungen wider, die Menschen über diese Maschinen haben, und reicht vom immateriellen Wesen über eine technische Darstellung bis hin zu einem Ebenbild des Menschen selbst, einer Maschine mit Emotionen und geschlechterspezifischem Aussehen.

Die Menschlichkeit als Merkmal der Maschine, die geschlechterspezifische Kennzeichnung und viel mehr noch, die Hybridisierung von Mensch und Maschine bildeten einen zentralen Punkt des anschließenden Vortrags „Mythos Cyborg - Feministische Perspektiven auf das Verhältnis von Mensch und Technologie“. Am Beispiel von Donna Haraways „A Cyborg Manifesto“ wurde deren Vorstellung einer „Post-Gender-Welt“ und der damit einhergehenden Verwischung von genderspezifischen aber auch menschlichen Grenzen erörtert. Wie auch im Vortrag zu „KI in Film und Literatur“ gezeigt wurde, weisen intelligente Maschinen, die mit Weiblichkeit attribuiert sind, häufig bestimmte stereotypische Merkmale auf, wie Emotionalität, weniger analytisches Denken oder die Einnahme einer schwächeren Position. Zudem werden sie oft auch als Objekt der Begierde auf ihre Weiblichkeit reduziert und objektiviert. Die Vorstellung Haraways, diese stereotype Darstellung durch einen geschlechtslosen Cyborg abzulösen, verweist auf die Vorstellung eines Post-Genderismus und dessen, was künstliche Intelligenz in der Zukunft leisten kann, um die Verwischung und Auflösung geschlechtsspezifischer und humaner Grenzen zu erreichen. Auch wenn Haraways Manifesto in den 1980er Jahren veröffentlicht wurde, verlieren das Manifesto und der Vortrag mit der Veranschaulichung dieser Zukunftsvision eines genderunabhängigen Cyborgs und der Betrachtung der Rolle der

Weiblichkeit im Zusammenhang mit künstlicher Intelligenz und Technologie besonders unter Betrachtung der zeitgenössischen Diskussionen rund um das Thema Gender nicht an Aktualität und Relevanz und geben Anlass über Geschlechter, Menschlichkeit, Technologie und das Mensch-Maschine-Verhältnis zu diskutieren.

Von diesem ästhetischen und genderbezogenen Ansatz der ersten beiden Vorträge aus, verlagerte sich durch den nächsten Vortrag „Auswirkungen des Einsatzes von KI und ML in der Musikbranche, insbes. in Produktion und Vertrieb“ die Diskussion um KI in den Bereich der Musik. Zentral war hier die Frage danach, inwieweit Musik, die von künstlicher Intelligenz hervorgebracht wird, als kreative Kunst zu sehen ist, und inwieweit diese Werke von den Rezipienten akzeptiert werden. Der Vortrag zeigte anhand eines Beispiels aus Japan, dass diese Akzeptanz kulturell abhängig ist und Rückschlüsse auf den kulturellen Umgang mit KI zulässt. Zudem wurde die Frage nach dem Urheberrecht, also inwiefern künstlicher Intelligenz als Urheber Rechte zugesprochen werden können, in den Raum gestellt. Auch der ökonomische Aspekt wirft Grund zur Diskussion auf und zwar, inwiefern sich die Wertschöpfungskette der Musikproduktion durch künstliche Intelligenz verändert und wie viel Substitutions- oder auch Interaktionspotenzial in ihr steckt. Darüber hinaus wird zu beobachten sein, wie und ob sich der Fankult, der sich um Musiker, Künstler und Bands entwickelt, verändert, wenn keine realen Menschen mehr die Musik hervorbringen, sondern reine Technologien.

Mit einer anderen Form der Interaktion von künstlicher Intelligenz und Mensch beschäftigte sich der Vortrag „Chatbots in Geschäftsmodellprozessen“. Im Fokus standen hier Chatbots, die mit Hilfe von künstlicher Intelligenz auf erlerntes Wissen zurückgreifen und auf dieser Grundlage mit dem Menschen interagieren, mit besonderem Fokus ihres Einsatzes im Personal-Recruiting. Potenziale von Chatbots im Recruiting sollen vor allem in einer geringeren Diskriminierung der Bewerber liegen und in der Erleichterung des Bewerbungsverfahrens. Grenzen dieser Technologie äußern sich jedoch vor allem in Kommunikationsbarrieren zwischen Mensch und Maschine und der Anlernung der Maschine, die gegebenenfalls durch einseitiges Datenmaterial die Vorteile der geringeren Diskriminierung wieder aushebeln können. Chatbots werden aber nicht nur im Recruiting, sondern generell bei der Unterstützung von Interaktionsprozessen eingesetzt, wie beispielsweise bei der Beantwortung von Fragen zu einem Produkt oder einem Service. Offen bleibt, wie viel Chatbots noch erlernen können und wie sich hierdurch die Kommunikation zwischen Mensch und Maschine verändert. Durch diese Veränderung der Kommunikation und die Übernahme neuer Aufgaben steckt in Chatbots das Potenzial,

die zukünftige Gestalt des Arbeitsmarktes mit zu bestimmen und zu verändern.

Künstliche Intelligenz in Form von Bots thematisierte auch der anschließende Vortrag „Political Bots im Wahlkampf“. Im Vordergrund dieses Vortrags stand die manipulative Verbreitung von politischen Inhalten durch Bots in sozialen Netzwerken, die mittels Imitation menschlichen Handelns versuchen, Einfluss auf andere menschliche Plattformnutzer zu nehmen. Hierfür wurde der Einsatz von Bots in öffentlichen, vornehmlich politisch motivierten Diskursen via Twitter und die Einflussnahme dieser Bots auf die Meinung und das Verhalten anderer Plattformnutzer untersucht. Eine Problematik, die diese Untersuchung ergeben hat, ist, dass, anders als bei Chatbots, die Political Bots nicht unbedingt leicht als Maschinen zu identifizieren sind. Neben der Imitation menschlicher Akteure werden auch ihre Profile auf den Plattformen möglichst authentisch gestaltet, um somit ihre Glaubwürdigkeit und Reichweite zu steigern. Auch wenn der Einsatz solcher Bots in Deutschland bisher kaum verbreitet ist, besteht dennoch die Möglichkeit einer wachsenden zukünftigen Verwendung. Der Vortrag zeigte, dass eine Beobachtung dieses Phänomens und der Auswirkungen einer durch Political Bots ermöglichten einseitigen oder manipulativen Verbreitung einer bestimmten Meinung zukünftig an Relevanz gewinnen könnte, da die Auswirkungen des vermehrten Einsatzes von Chatbots eine breite Öffentlichkeit betreffen und deren Meinung und damit zukünftige politische Diskussionen und Ereignisse wie Wahlen beeinflussen können.

Abschließend wurden im letzten Vortrag „KI in den Massenmedien“ Narrative bezüglich der Hoffnungen und Ängste in Verbindung mit künstlicher Intelligenz aufgegriffen und darüber hinaus am Beispiel von drei deutschen Tageszeitungen untersucht, wie sich die Berichterstattung rund um das Thema „Künstliche Intelligenz“ in Massenmedien äußert. Die für diesen Zweck durchgeführte qualitative Inhaltsanalyse zeigte auf, dass in der Mehrheit der untersuchten Artikel eine euphorische Sichtweise auf künstliche Intelligenz vorherrscht, wenngleich auch apokalyptische Szenarien thematisiert werden. Die thematischen Schwerpunkte der untersuchten Artikel betreffen Berührungspunkte mit künstlicher Intelligenz, die auch in den vorangegangenen Vorträgen diskutiert wurden, wie den Wettbewerb, das Verhältnis von Mensch und Maschine, Sicherheit und Arbeitsplätze. Dieser Vortrag schlug damit eine Brücke zu den vorherigen Beiträgen und zeigte auf, dass viele bereits thematisierten Aspekte, Hoffnungen und Ängste, die mit künstlicher Intelligenz verbunden werden, auch in der Berichterstattung öffentlicher Medien zu finden sind.

Dies verdeutlicht noch einmal die Reichweite des Themas „Künstliche Intelligenz“ sowie die Relevanz seiner bereichsübergreifenden Betrachtung und Diskussion. Denn künstliche Intelligenz mit all ihren Potenzialen, Grenzen und den Konsequenzen, die bestimmte Entwicklungen mit sich ziehen können, stellt eine Herausforderung für unsere Gesellschaft dar, sie betrifft diese als Kollektiv aber vermutlich auch jedes gesellschaftliche Mitglied als Individuum. Die weitere technologische Entwicklung und insbesondere eine Autonomie von Maschinen wird, sofern sie möglich wird, vermutlich kaum überblickbare Herausforderungen und Auswirkungen mit sich bringen und das Verhältnis von Menschen und Maschine neu definieren. Es reicht daher nicht aus, diese Problematik nur aus einer ökonomischen, technischen oder rechtlichen Perspektive zu betrachten. Sie erfordert vielmehr eine umfassende Auseinandersetzung verschiedener gesellschaftlicher Bereiche und vor allem auch eine ethische Hinterfragung. Dies wird beispielsweise auch beim Aspekt der Verantwortung und Haftung deutlich. Die Zuschreibung von Verantwortlichkeit zu autonomer künstlicher Intelligenz wird mit Fortschreiten der Forschung immer relevanter und aktueller werden und steht nicht zuletzt rechtlich und ethisch zur Diskussion. Mit Hilfe der interdisziplinären Betrachtung der Thematik ist es der Tagung damit gelungen, viele verschiedene Aspekte rund um künstliche Intelligenz und ihre Potenziale und Grenzen aus unterschiedlichen Blickwinkeln zu betrachten, Verbindungen zwischen Disziplinen und gesellschaftlichen Bereichen aufzuzeigen und eine vielseitige Diskussion anzuregen.

---

## Literaturverzeichnis

Brynjolfsson, Erik und McAfee, Andrew (2017): The business of artificial intelligence. In: Harvard Business Review. <http://starlab-alliance.com/wp-content/uploads/2017/09/The-Business-of-Artificial-Intelligence.pdf> (Zugriff: 26.09.2019).

Eberl, Ulrich (2018): Was ist künstliche Intelligenz – Was kann sie leisten? In: Aus Politik und Zeitgeschichte. Künstliche Intelligenz. S. 8 – 14.

Rocke, Valentin und Osoria, Elisa (2019): Einführungsfilm der Tagung „Zukünfte KI – Künstliche Intelligenz in Medien, Politik, Gesellschaft und Wirtschaft“.

Wangler, Leo; Botthof, Alfons; Wittpahl, Volker (2019): E-Governance: Digitalisierung und KI in der öffentlichen Verwaltung. In: Wittpahl, Volker (Hrsg.): iit-Themenband. Künstliche

Intelligenz. Technologie | Anwendung | Gesellschaft. Berlin Heidelberg: Springer 2019, S. 122  
– 141.