



Hochschule Neubrandenburg  
University of Applied Sciences

# **Nutzen und Chancen von schriftlicher Onlineberatung mit Kindern und Jugendlichen**

## **Potenziale und Risiken bei der Einbindung von Künstlicher Intelligenz**

Masterthesis zur Erlangung des akademischen Grades

**„Master of Arts“ (M.A.)**

im Studiengang

**Beratung – Psychosoziale Beratung in den Handlungsfeldern Soziale Arbeit/Sozialpädagogik,**

**Bildung und Erziehung**

des Fachbereichs

**Soziale Arbeit, Bildung und Erziehung**

an der

**Hochschule Neubrandenburg – University of Applied Sciences**

**Verfasser:** Fabian Binder  
**Erstbetreuerin:** Prof.in Dr.in phil. habil. Barbara Bräutigam  
**Zweitbetreuerin:** Prof.in Dr.in Miriam Burzlaff  
**URN-Nummer:** urn:nbn:de:gbv:519-thesis2024-0439-1

*Nobody phrases it this way,  
but I think that artificial intelligence is almost a humanities discipline.  
It's really an attempt to understand human intelligence and human cognition.*

**Sebastian Thrun (2013)**

*Moral goodness might be more like a precious metal  
than an abundant element in human nature,  
and even after the ore has been processed and refined  
in accordance with the prescriptions of the CEV proposal,  
who knows whether the principal outcome will be  
shining virtue, indifferent slag, or toxic sludge?*

**Nick Bostrom (2014)**

*People always overestimate the change that will occur in the next two years,  
and underestimate the change that will occur in the next ten.*

**Bill Gates (1996)**

# Inhaltsverzeichnis

1	Einleitung.....	1
2	Grundlagen schriftlicher Onlineberatung mit Kindern und Jugendlichen .....	5
2.1	Schriftliche Onlineberatung – eine Begriffsbestimmung .....	6
2.2	Instrumente der schriftlichen Onlineberatung .....	11
2.2.1	Mailberatung.....	11
2.2.2	Chatberatung.....	12
2.2.3	Messengerberatung .....	13
2.3	Haltung und Methoden.....	14
2.4	Wirk- und Einflussfaktoren.....	18
2.5	Mediennutzungsverhalten und dessen Bedeutung für die Onlineberatung .....	23
3	Grundlagen Künstlicher Intelligenz .....	28
3.1	Geschichte .....	28
3.2	Definition.....	44
3.2.1	Symbolische KI.....	47
3.2.2	Subsymbolische KI.....	49
3.2.2.1	Künstliche neuronale Netze .....	50
3.2.2.2	Trainingsverfahren .....	56
3.2.3	Hybride KI.....	58
4	Künstliche Intelligenz in der schriftlichen Onlineberatung .....	60
4.1	Chatbots .....	64
4.2	Assistive Tools .....	78
4.3	Einsatz in der Ausbildung .....	84
5	Fazit .....	88
	Literaturverzeichnis.....	93
	Eidesstattliche Erklärung.....	114

# 1 Einleitung

Im letzten Jahr formulierte die Onlineberaterin und Professorin für Digitale Transformation in Sozialen Handlungsfeldern und Gesellschaft Emily Engelhardt angesichts des „gefühlte ‚plötzlichen‘ Auftauchen[s]“ (Engelhardt 2023a, S. 406) von ChatGPT und generativer KI stellvertretend für ihre Profession die Sorge, dass das Feld der Beratung nach gerade überstandener Pandemie vor der nächsten Disruption stehe (vgl. ebd.). Generative KI, so ihre Vermutung, könne „das Selbstverständnis von Beratung und der Profession der Beratenden durchaus erschüttern“ (ebd.). Damit gehört Beratung einer Reihe von Berufsfeldern an, deren Zugehörigen seit der Veröffentlichung von ChatGPT am Ende des Jahres 2022 die Sorgenfalten ins Gesicht geschrieben stehen, aufgrund der Befürchtung, bald gar nicht mehr oder nur noch teilweise benötigt zu werden (vgl. Zinkula 2023). Angesichts freudestrahlender Verkündungen von Geschäftsführer:innen, die sich erhoffen, Einsparungen erzielen zu können, indem sie beispielsweise ein Drittel ihrer Belegschaft durch Künstliche Intelligenz ersetzen können, verwundert dies nicht (vgl. Kelly 2023). Möglich machen soll dies ChatGPT, Vertreter einer neuerlichen Form Künstlicher Intelligenz namens generativer KI, die neben einigen anderen wundersamen Dingen vordergründig in der Lage ist, natürliche Sprache scheinbar verstehen und sinnhaft generieren zu können (vgl. Nazir & Wang 2023, S. 1). Damit stellt ChatGPT die neueste Errungenschaft eines interdisziplinären Forschungsfeldes dar, das sich in den 1950er Jahren aufmachte, menschliche Intelligenz versteh- und künstlich replizierbar zu machen (vgl. Sheikh et al. 2023, S. 1). Gerade in den vergangenen Jahren hat die Entwicklung von KI noch einmal rasch an Fahrt aufgenommen und immer leistungsstärkere Produkte hervorgebracht (vgl. Vaske 2023). In der Folge hat der Gebrauch von KI auch mehr und mehr den Weg in die Alltäglichkeit gefunden: „Der Geist ist aus der Flasche“ (Gravelmann 2024b, S. 5). Somit scheint klar: „Whatever field you are in, if it uses language, it is about to be transformed“ (Marche 2021). Demnach stellt sich aktuell also auch für Beratende in klassischen Settings und stärker noch in der Onlineberatung – etwas drastisch formuliert – die Frage, ob es bald überhaupt noch eine Verwendung für sie geben wird (vgl. Engelhardt 2023a, S. 406).

In Anbetracht dieser Entwicklung kann die gesamtgesellschaftliche Aufgabe also nicht lauten, das Aufkommen dieser neuen Technologie ungeschehen zu machen oder sie zu ignorieren, sondern es muss darum gehen, einen Umgang mit ihr zu finden – auch mittels gesetzlicher Regulierungen (vgl. Gravelmann 2024b, S. 4). Ein erster logischer Schritt auf diesem Weg scheint die ausführliche Auseinandersetzung mit dem Thema ‚Künstliche Intelligenz‘ zu sein – auch, weil eine höhere Informiertheit zu einer Demystifizierung von KI beitragen kann und eine bessere

Einordnung des Phänomens ermöglicht (vgl. Sheikh et al. 2023, S. 9 f.). Dadurch könnte KI womöglich etwas vom beschriebenen ersten Schrecken verlieren. Gleichzeitig erlaubt Aufgeklärtheit auch eine kompetentere Positionierung bezüglich KI, die gesellschaftliche Mitgestaltungsprozesse und individuelle Handlungsfähigkeit möglich werden lässt (vgl. ebd., S. 10 f.) Für (Online-)Beratende in der psychosozialen Arbeit könnte sich so die Frage, „Werde ich durch KI ersetzt?“, über, „Wie wird sich KI auf meine Berufspraxis auswirken?“, hin zu, „Wie kann ich mir KI für meine Berufspraxis verantwortungsvoll zunutze machen, um diese effizienter und qualitativ hochwertiger zu gestalten?“, wandeln (vgl. Abrams 2024b, S. 47).

Mit der Frage, wie die vorliegende Situation bestmöglich zu bewältigen sei, beschäftigten sich Beratende zuletzt gezwungenermaßen in der Covid-19-Pandemie, die sich hinsichtlich der Digitalisierung als Katalysator erwiesen hat und dazu führte, „dass das Digitalisierungsthema im Alltag der Sozialen Arbeit auf breiter Ebene angekommen ist“ (Gravelmann 2024a, S. 22). Die in dieser Zeit gemachten Erfahrungen können der Beratungsprofession möglicherweise zugutekommen, um die kommenden, durch KI hervorgerufenen Herausforderungen bewältigen und die sich darbietenden Chancen wahrnehmen zu können. Gerade im Hinblick auf Kinder und Jugendliche erscheint dies von hoher Relevanz, da vor allem sie gerne niedrigschwellige Onlineberatungsangebote in Anspruch nehmen und Hemmungen zur Inanspruchnahme geringer ausfallen im Vergleich zu Beratungsangeboten, die in Präsenz aufzusuchen wären (vgl. ebd., S. 114). Kinder und Jugendliche sind, was neue Technologien angeht, häufig ‚Early Adopter‘, die sich oftmals scheinbar mühelos neue Technologien aneignen und diese zu nutzen wissen (vgl. Quenzel & Hurrelmann 2022, S. 47). Es ist davon auszugehen, dass die Technologie ‚Künstliche Intelligenz‘ in dieser Hinsicht keine Ausnahme darstellen wird, zumal KI bereits jetzt in vielen lebensweltlichen Dingen verankert ist, beispielsweise in Form von Chatbots in von Kindern und Jugendlichen genutzten Social-Media Apps (vgl. Gravelmann 2024b, S. 9; vgl. Weinhardt 2024, S. 65). Die Soziale Arbeit orientiert sich stets an der Lebenswelt ihrer Klient:innen, mitsamt den einhergehenden Bedürfnissen an soziale Dienstleistungsprozesse (vgl. Klein & Pulver 2020, S. 190 f.). Deshalb wird sie sich auf KI fraglos einstellen müssen. Ein weiterer Faktor geht von Fachkräften der Sozialen Arbeit aus, die im Sinne der Digitalität KI wie selbstverständlich in Arbeitsprozesse hineintragen werden (vgl. Weinhardt 2023a, S. 43 f.).

Seelische Leiden, die einen großen Anteil der weltweiten gesundheitlichen Probleme ausmachen, betreffen auch häufig Kinder und Jugendliche, denn das erste Einsetzen von psychischen Problemen tritt häufig schon im Kindes- oder Jugendalter zutage (vgl. Solmi et al. 2021, S. 286; vgl. World Health Organization 2022, S. 35 ff.). So gerät beispielsweise zunehmend auch der Stressfaktor Einsamkeit in den Blick der Öffentlichkeit, von dem laut Bertelsmann Stiftung jeder

zweite junge Mensch zumindest teilweise betroffen ist (vgl. Steinmayr et al. 2024, S. 17). Dies zeigt, dass der aktuelle Bedarf an psychosozialer Versorgung für junge Menschen unverändert hoch ist.

Es wird also interessant sein, herauszufinden, wie KI sich in der schriftlichen Onlineberatung mit Kindern und Jugendlichen einsetzen lassen könnte und welche Potenziale und Risiken sich hieraus für Ratsuchende und Beratende ergeben. Ein Blick soll auch darauf geworfen werden, welche Implikationen daraus für die Praxis der schriftlichen Onlineberatung entstehen können. Vorliegende Masterarbeit setzt sich zum Ziel, einen Beitrag zum derzeitigen wissenschaftlichen Diskurs zu leisten, indem eine Bestandsaufnahme aktuellster Forschung zu der Thematik durchgeführt wird.

Es ergibt sich folgende Forschungsfrage: *Inwiefern kann sich die schriftliche Onlineberatung mit Kindern und Jugendlichen Künstliche Intelligenz unter Berücksichtigung auftretender Potenziale und Risiken, sowohl für Ratsuchende als auch Beratende, zunutze machen?*

Eingangs werden in Kapitel 2 Grundlagen schriftlicher Onlineberatung mit Kindern und Jugendlichen erläutert, die für den Rahmen dieser Arbeit von Relevanz sind. Daher soll zu Beginn eine begriffliche Einordnung von schriftlicher Onlineberatung erfolgen (vgl. Kap. 2.1). Dabei sollen auch die Besonderheiten des Onlinesettings herausgearbeitet werden. Anschließend werden gängige Verfahren (vgl. Kap. 2.2) sowie Haltung und Methoden, die in der schriftlichen Onlineberatung eine zentrale Rolle einnehmen, näher beleuchtet (vgl. Kap 2.3) und deren Wirk- und Einflussfaktoren erläutert (vgl. Kap 2.4). Zum Schluss werden in Kapitel 2.5 Kinder und Jugendliche unter Berücksichtigung ihres Mediennutzungsverhaltens als Klientel von Onlineberatung betrachtet und es wird herausgearbeitet, welche Implikationen hierdurch für die schriftliche Onlineberatung entstehen.

Kapitel 3 der Arbeit soll einen Einblick in die Grundlagen Künstlicher Intelligenz gewähren. Die Wissensgrundlage, die in diesem Kapitel gelegt wird, ist für das weitere Verständnis der Arbeit unumgänglich. Um die Technologie KI zu ergründen, soll eingangs die nicht immer lineare Geschichte Künstlicher Intelligenz näher beleuchtet werden. Dabei findet nicht nur eine bloße Aufzählung und Nennung der maßgeblich Beteiligten statt, sondern es sollen auch historische Umstände, Mythen, Dichtungen, philosophische Überlegungen, sowie Treiber und Hemmer der Entwicklung von KI Erwähnung finden, die zum heutigen Status Quo des Entwicklungsstandes von KI geführt haben (vgl. Kap. 3.1). Daran anschließend werden die zuvor eingeführten Termini

Künstlicher Intelligenz aufgegriffen und, so wie sie sich in ihren verschiedenen Formen und Funktionsweisen konstituieren, in Kapitel 3.2 näher erläutert.

In Kapitel 4 findet eine thematische Zusammenführung der beiden zuvor behandelten Themenfelder ‚schriftliche Onlineberatung mit Kindern und Jugendlichen‘ und ‚Künstliche Intelligenz‘ statt. Dabei soll untersucht werden, inwiefern Künstliche Intelligenz in der schriftlichen Onlineberatung bereits zur Anwendung kommt und in welchen Anwendungsfeldern KI nutzbar gemacht werden kann. Einleitend wird hierbei ein kurzer Blick auf den Status Quo der Einbindung von KI in der Sozialen Arbeit erfolgen. Zudem soll untersucht werden, welche gewinnbringenden Anwendungsmöglichkeiten zur Implementierung von KI in der Onlineberatung bereits existieren beziehungsweise sich derzeit für den Einsatz in naher Zukunft andeuten. Hierbei werden zum einen Chatbots in den Blick genommen (vgl. Kap. 4.1) und zum anderen assistive KI-Tools beleuchtet (vgl. Kap. 4.2). Außerdem wird geprüft, inwiefern KI in der Ausbildung von Onlineberater:innen zur Anwendung kommen kann (vgl. Kap. 4.3). Demgemäß soll insbesondere eine kritische Einordnung etwaiger Potenziale und Risiken stattfinden, die bei der Benutzung von KI in den beschriebenen Anwendungsfeldern auftreten können.

Zum Abschluss erfolgt ein Resümee, das in besonderem Maße die für die Forschungsfrage relevanten Erkenntnisse darstellt. Unter Berücksichtigung der erarbeiteten Ergebnisse dieser Thesis soll ein kurzer Ausblick auf die Zukunft sowie mögliche weitere Forschungsmöglichkeiten geworfen werden.

## 2 Grundlagen schriftlicher Onlineberatung mit Kindern und Jugendlichen

Die stetig zunehmende Mediatisierung und Digitalisierung unseres Alltags macht auch vor dem Feld der psychosozialen Beratung keinen Halt und wirkt sich mitunter stark auf dieses aus (vgl. Kutscher et al. 2020, S. 12). So wird das Internet nicht nur genutzt, um Informationen bereitzustellen und Beratungen auf den Weg zu bringen, sondern Beratungen werden auch über die Kommunikationsmöglichkeiten, die das Internet bereithält, durchgeführt (vgl. Engelhardt 2021, S. 11). Eine große Chance der fortschreitenden Digitalisierung sieht Wenzel (2019) vor allem in der Schaffung zusätzlicher Zugangswege zur Beratung für Klient:innen (vgl. S. 222). Zu beachten ist jedoch, dass ein großflächigeres und breiter gefächertes Angebot an Beratungsmöglichkeiten, die auch online von Klient:innen genutzt werden können, nicht automatisch gleichbedeutend damit ist, dass diese neuen Zugänge zu Unterstützungsangeboten auch unbedingt effektiv besser in Anspruch genommen werden können (vgl. Klein & Pulver 2020, S. 192). Psychosoziale Onlineberatung<sup>1</sup> findet seit Mitte der 1990er-Jahre statt. Genau genommen war es im Jahre 1995 die Telefonseelsorge, die durch Beratung via Telefon schon Erfahrung in medial vermittelter Beratung besaß und das erste Onlineberatungsangebot in Deutschland offerierte (vgl. Justen-Horsten & Paschen 2016, S. 19). Im Laufe der Jahre wurde „der Zugang zu Onlineberatung [...] durch technologische Fortschritte immer niedrigschwelliger, insbesondere da die Popularität mobiler Endgeräte und die Vielfalt an internetfähigen Geräten und Anwendungen zunahm“ (Lippert et al. 2024, S. 43). Ausgelöst durch die Coronapandemie und den damit einhergehenden temporären Kontaktbeschränkungsmaßnahmen gewann Onlineberatung spätestens im Jahr 2020 deutlich an Relevanz (vgl. Engelhardt 2021, S. 9). In Deutschland wird Onlineberatung hauptsächlich durch Beratungsstellen angeboten und von ihnen durchgeführt (vgl. ebd., S. 20). So gibt es auch Onlineberatungsstellen, die sich speziell an Kinder und Jugendliche richten, wie zum Beispiel die ‚JugendNotmail‘ oder der ‚krisenchat‘.

---

<sup>1</sup> Wenn im Kontext dieser Arbeit auf Onlineberatung verwiesen wird, ist damit immer die Art der Beratung gemeint, die im psychosozialen Kontext stattfindet. Nicht von Bedeutung sind dabei Onlineberatungen, die im Zusammenhang mit Vertragsabschlüssen oder Käufen bestehen und von Seiten dieser Geschäftspartner angeboten werden, um beim Kauf oder Vertragsabschluss behilflich zu sein.



## 2.1 Schriftliche Onlineberatung – eine Begriffsbestimmung

Nestmann & Sickendiek (2018) definieren Beratung wie folgt:

Beratung (engl. counselling) ist eine spezifische Form der Kommunikation: Eine Person ist einer anderen Person dabei behilflich, Anforderungen und Belastungen des Alltags oder schwierigere Probleme und Krisen zu bewältigen. Beraten können sich ebenso Gruppen und Organisationen, oder die beratende Seite kann ein Softwareprogramm sein. Beratung leistet Beistand bei der kognitiven und emotionalen Orientierung in widersprüchlichen und unübersehbaren Situationen und Lebenslagen. Sie unterstützt Ratsuchende dabei, Wahlmöglichkeiten abzuwägen, sich zwischen Alternativen zu entscheiden oder aber Optionen bewusst offenzuhalten. Beratung fördert Zukunftsüberlegungen und Pläne, die aus neu gewonnenen Zielrichtungen und Entscheidungen resultieren, sie hilft Ratsuchenden die Planungsschritte zu realisieren und begleitet erste Handlungsversuche mit Reflexionsangeboten (S. 110).

In Beratung erfolgt eine Fokussierung auf Themenkomplexe und Problembereiche, die mittels fachlicher Beratungsmethoden bearbeitet werden. Diese kann Ratsuchenden dabei behilflich sein, „mit der Fülle von Eindrücken umzugehen, die Komplexität des Alltags angemessen zu reduzieren und bei Bedarf Prioritäten in der Problembearbeitung zu setzen“ (Wenzel 2019, S. 220). Psychosoziale Beratung, in professioneller Form ausgeübt, gilt als „interaktiver Prozess“ (Engelhardt 2023a, S. 404), der darauf abzielt, „bei der Gestaltung der Lebensführung und im Umgang mit Lebensbedingungen und Herausforderungen Orientierungs-, Planungs-, Entscheidungs-, Handlungs- und Reflexionshilfe zu bieten“ (Schubert et al. 2019, S. 28). Beraten werden kann (therapie-)schulenübergreifend mittels verschiedener Beratungsansätze, die auf Grundlage eines interdisziplinären Wissensverständnisses fußen (vgl. Deutsche Gesellschaft für Beratung e.V. 2020, S. 3). Dieses Wissen entstammt der Psychologie, Pädagogik, Psychotherapie, Medizin, Sozialarbeit, Soziologie, Rechtswissenschaft und bisweilen der Theologie, wobei die Beratungsansätze je nach Anforderung und Umfeld an psychoanalytische, individualpsychologische, gesprächspsychotherapeutische und verhaltenstherapeutische Herangehensweisen angelehnt sein können (vgl. Hoff & Zwicker-Pelzer 2015, S. 45). Beratung ist auf Klient:innen fast aller Alters-, Bevölkerungs- und Kulturgruppen ausgerichtet und schließt damit also jegliche Menschen der „Mehrheitsbevölkerung“, aber auch [...] sozialer[r] und gesundheitsbezogene[r], Risiko- und Problemgruppen“ (Nestmann & Sickendiek 2018, S. 110) mit ein.

Unter dem Begriff Onlineberatung wird vor allem textgebundene Beratung, die via E-Mail, Chat oder Messenger durchgeführt wird, verstanden (vgl. Eichenberg & Kühne 2014, S. 43). Nicht zuletzt wegen der Umstände der zurückliegenden Coronapandemie spielten in den letzten Jahren auch audio-visuelle durchgeführte Beratungen via Videokonferenzsoftware eine immer größere Rolle (vgl. Hörmann et al. 2023, S. 14). Engelhardt & Storch (2013) resümieren daher,

Onlineberatung subsumiere sämtliche Formen von Beratung, die medial durchgeführt über das Internet ablaufen würden, wozu sie sowohl textgebundene Verfahren wie E-Mail-Beratung oder Chats als auch textungebundene Verfahren wie Videochats oder Internettelefonie zählen (vgl. S. 4). Darüber hinaus seien auch Mischformen möglich, die ein oder mehrere Merkmale der verschiedenen Beratungsformen kombinieren würden (vgl. ebd.). Wie in Abbildung 1 ersichtlich wird, kann Onlineberatung sowohl synchron, das heißt in Echtzeit ablaufend, als auch asynchron, also zeitlich zerdehnt, erfolgen (vgl. Kühne & Hintenberger 2020, S. 36 f.).

	<b>Textbasiert</b>	<b>Nicht-textbasiert</b>
<b>Synchron</b>	Chat, Messenger	Video, Internettelefonie, Avatare
<b>Asynchron</b>	Mail, Forum	Sprachnachrichten, Videonachrichten

*Abbildung 1: Medien der Onlineberatung (Engelhardt 2021, S. 18)*

In dieser Arbeit soll ausschließlich die textbasierte Onlineberatung thematisiert werden, weshalb Videotelefonie und weitere schriftungebundene Beratungsformen in dieser Ausarbeitung ausgeklammert werden. Daher referiert im weiteren Verlauf der Arbeit der Begriff Onlineberatung auf textgebundene Onlineberatung, die via E-Mail, Chat, Messenger oder vergleichbar getartet erfolgt.

Mittlerweile hat sich Onlineberatung als feste Größe etabliert. Reindl (2018) postuliert diesbezüglich, „dass Onlineberatung nach mehr als 20 Jahren Praxis einen Stand erreicht hat, der sie neben der Präsenzberatung und der Telefonberatung als ‚dritte Kraft‘ in der psychosozialen Beratungslandschaft ausweist“ (S. 23). Bezogen auf das Institutionalierungsmodell von Tolbert und Zucker (1996, zit. n. Eichenberg & Kühne 2014, S. 33) wird die Onlineberatung momentan in der zweiten von drei Phasen, der Phase der annähernden Institutionalisation, verortet (vgl. Eichenberg & Kühne 2014, S. 38). Eichenberg & Kühne (2014) stellen in Aussicht, diese Phase sei noch mitten im Prozess und ein Fortschreiten zur nächsten Phase stünde noch nicht unmittelbar bevor (vgl. ebd.). Sie begann zeitgleich mit dem Aufkommen des Web 2.0 im Jahre 2003, also dem Internet, das nicht nur durch passives Konsumieren gekennzeichnet ist, sondern sich durch die interaktive Mitgestaltung seiner Nutzenden hervorhebt (vgl. Engel 2019, S. 12). Diese Phase zeichnet sich insbesondere durch eine erhöhte wissenschaftliche Reflexion des Arbeitsgebietes aus, sodass ein größeres Aufkommen an Fachöffentlichkeit und objektiviertem Wissen entsteht, das Qualitätsmerkmale von Onlineberatung herausarbeitet, anhand derer Qualifizierungs- und Fortbildungsmaßnahmen für Akteur:innen der Sozialen Arbeit ermöglicht werden (vgl. Reindl 2018, S. 18). Die vorangegangene sogenannte Pionierphase stach demgegenüber eher durch

einen Ansatz heraus, der aufgrund der damals neuartigen Möglichkeiten, die das Internet hervorbrachte, mit Learning by Doing beschrieben werden kann, da zu dieser Zeit ein wissenschaftlicher und theoretischer Unterbau in Bezug auf Onlineberatung noch fehlte und die Technik wenig ausgereift war (vgl. Schubert et al. 2019, S. 229). Damit die Onlineberatung in die letzte Phase der vollständigen Institutionalisierung eintreten kann, benötigt es allgemein gültige Standards und Normen und ein großes theoretisches Fundament, das qualitativ hochwertige professionelle und standardisierte Onlineberatung ermöglicht. Einmal dort angekommen, sehen Eichenberg & Kühne (2014) es zudem auch als unumgänglich an, kontinuierlich nah am Zahn der Zeit des Mediums Internets zu bleiben, da nicht abzusehen sei, ob und inwiefern Veränderungen am und im Internet eintreten könnten, die eine grundsätzliche erneute Adaptation an das Arbeitsfeld der Onlineberatung nötig machen würden (vgl. S. 38 f.). Obwohl sich Onlineberatung auch angetrieben durch die Entwicklungen der letzten Jahre längst etabliert hat und die Chancen digitalisierter Beratung hinlänglich bekannt und nachgewiesen sind (vgl. Kap. 2.4), bemängelt Wenzel (2019), dass „es [...] nach wie vor keine flächendeckende Regelfinanzierung von Onlineberatung (gibt)“ (S. 226).

Die Grenzen zwischen Therapie und Beratung sind mitunter nicht immer klar zu definieren. In Bezug auf Onlinetherapie und Onlineberatung wird dies anhand der Definition von Barak & Grohol (2011) deutlich, die Onlinetherapie als „therapeutic intervention that is administered individually or in groups by a trained mental health professional, and may occur via e-mail, chat, forums, and audio or video (e.g., webcam)“ (S. 160) definieren. Wie zuvor in diesem Kapitel herausgearbeitet wurde, können fast alle in diesem Zitat benannten Merkmale auch auf Onlineberatung zutreffen. Ein signifikanter Unterschied zwischen Onlinetherapie und -beratung besteht allerdings hinsichtlich Finanzierung und Durchführung: Zwar dürfen Psychotherapeut:innen mit Eintritt des EBM-Beschlusses vom ersten Juli 2022 30 Prozent ihrer Gesamtstunden per Videobehandlung durchführen, jedoch haben einige Abläufe wie probatorische Sitzungen, Diagnostik und Indikationsstellung weiterhin ausschließlich im persönlichen Kontakt zu erfolgen, sodass sie von der Krankenkasse übernommen werden können (vgl. Bundespsychotherapeutenkammer 2022, S. 5; vgl. Kassenärztliche Bundesvereinigung 2022). Demgegenüber kann Onlineberatung ohne vorherigen in Präsenz stattgefundenen Kontakt in Anspruch genommen werden und ist losgelöst von der Finanzierung der Krankenkasse, die lediglich für Gesundheitsleistungen, denen Psychotherapie zuzuordnen ist, aufkommt, wohingegen Beratung als Sozialleistung definiert ist, die von anderer Stelle finanziert wird (vgl. Engelhardt 2021, S. 20). Dass diese strikte Trennung beider Arbeitsfelder mitunter ungünstig ist, arbeitet Bräutigam (2022) heraus, indem sie dafür plädiert, die psychotherapeutische Praxis täte gut daran, neben der Perspektive, die vor

allem auf die Behandlung psychischer Störungen zielt, soziokulturelle Faktoren stärker in Therapie miteinzubeziehen (vgl. S. 133). Gleichzeitig sei Beratung gefordert, über die Lösungsfindung in verschiedenen gelagerten herausfordernden Lebenssituationen hinaus, ein größeres Bewusstsein für klinisch-pathologische Sichtweisen entwickeln, da bei Ratsuchenden teils schwere psychische Störungen vorliegen können (vgl. ebd.).

Das Besondere an der Onlineberatung ist nicht etwa, dass sie eine zur Präsenzberatung abweichende Methode darstellt, sondern, dass sie in einem unterschiedlichen Setting stattfindet, da auch „online sowohl systemisch, behavioral, psycho-analytisch usw. beraten werden kann“ (Thiery 2018, S. 7). Die Losgelöstheit von zeitlichen und räumlichen Bedingungen kann für Ratsuchende sowohl einen einladenden als auch ausladenden Charakter mit sich führen, weshalb die Merkmale von Präsenz- und Onlineberatung sich ergänzen und nicht gegenseitig ausschließen, da die Beschränkungen des einen Settings die Möglichkeiten des anderen hervorheben und vice versa (vgl. ebd.). Onlineberatung bietet vor allem die Chance zusätzliche Zielgruppen anzusprechen, die Themen vorbringen können, wegen derer sie in der Präsenzberatung möglicherweise nicht vorstellig werden würden (vgl. Kap. 2.4) oder aufgrund anderer Barrieren Beratungsstellen vor Ort nicht ansteuern können (vgl. Stüwe & Ermel 2019, S. 125). Im Setting der Onlineberatung herrschen im Vergleich zur Face-to-Face-Beratung auch Unterschiede hinsichtlich der Kommunikation. So stehen sowohl Ratsuchenden als auch Beratenden lediglich Schrift beziehungsweise Text zur Verfügung, um sich mitzuteilen (vgl. Knatz 2008, S. 2). Thiery (2008) hebt hervor, „dass Offline-Beratung prototypisch im Medium der gesprochenen Sprache (Stimme) stattfindet, [während] [...] die Schrift als prototypisches Medium der Online-Beratung [gilt]“ (S. 8). Durch das Ausbleiben der körperlichen Kopräsenz aller an der Beratung beteiligten Personen entfallen andere Eindrücke, die durch die Sinnesorgane Auge, Ohr und Nase entstehen könnten (vgl. Knatz 2009b, S. 61). Diese beschriebene Kanalreduktion in der Kommunikation kann aus einer defizitorientierten Perspektive betrachtet werden – klar ist allerdings auch, dass diese für die Onlineberatung positive Implikationen haben kann, wie sich in Kapitel 2.4 noch zeigen lassen wird (vgl. Hintenberger 2006, S. 2). Eine anonyme Beratung, die via eines frei gewählten Nicknames stattfindet, beugt zusätzlich einer möglichen Voreingenommenheit vor, da ein Rückschluss auf die (soziale) Herkunft entfällt, wenngleich anzumerken ist, dass die Wahl des Nicknames das Potenzial birgt ganz eigentümliche Assoziationen auf Seiten der Beratenden hervorzurufen (vgl. Ploil 2009, S. 17). Auch wenn Onlineberatung in der Praxis seitens der Ratsuchenden häufig anonym in Anspruch genommen wird, ist dies aber je nach Kontext nicht zwingend notwendig und kann generell auch nicht-anonym erfolgen (vgl. Hörmann et al. 2023, S. 14). Die durch die körperliche Absenz ausgelöste Spärlichkeit an zur Verfügung stehenden Informationen auf

körperlicher, sozialer und psychologischer Ebene können aber Projektionen auf den:die Gesprächspartner:in begünstigen und eine Idealisierung des Gegenübers zur Folge haben (vgl. Brunner 2009, S. 39).

Knatz (2008) schlussfolgert, es sei alleine der Text, der in der Onlineberatung eine Wirkung entfalten würde, weshalb Beratende mittels des Textes im Dialog mit sich selber stünden, da Eindrücke und hervorgerufene Emotionen stets im Kontext der eigenen Biografie und auf Grundlage des eigenen, auch fachlichen, Wissens zu betrachten sind (vgl. S. 2). Methodisch wird Ratsuchenden also in Form von Texten begegnet und nicht als anwesende Person im unmittelbaren Kontakt, auch wenn es sich bei Texten von Klient:innen um deren Produkte handelt, mit denen sie ihren Anliegen, Ansichten und Gefühlszuständen Ausdruck verleihen (vgl. Brunner 2006, S. 3). Aus diesem Grund ist die kritische Reflexion für Beratende in der Onlineberatung dahingehend wichtig, dass mögliche Erfahrungsschätze und Kenntnisse, die aus der Beratung in Präsenz herrühren und dort funktionieren, sich nicht unbedingt unverändert auf das Setting der Onlineberatung transferieren lassen (vgl. Knatz 2009b, S. 60). Aufgrund dieser Unterschiede sind Sozialarbeitende nicht automatisch für Onlineberatung qualifiziert, da sie neben technischen Kenntnissen, die sich auf die Medien der Onlineberatung beziehen, auch spezifisches fachliches Wissen und auf Onlineberatung zugeschnittene Beratungskompetenzen benötigen (vgl. Stüwe & Ermel 2019, S. 125). Gerade aufgrund des Ausbleibens von Sinneseindrücken hebt Knatz (2008) bei der Arbeit mit Texten in der Onlineberatung die Bedeutung der Fähigkeit, zwischen den Zeilen lesen zu können, hervor: „Es ist ein Zeichen von emotionaler Kompetenz, hinter vordergründiger Information noch andere Dimensionen zu erkennen und darauf zu reagieren“ (S. 2). Dabei kann die Kompetenz, empathische und pointierte auch auf die Textsprache des Ratsuchenden angepasste Texte verfassen zu können, nicht hoch genug bewertet werden, da die „hohe Kunst des Lesens und Schreibens [in der Onlineberatung] ein bestimmendes Gewicht hat“ (Lang 2015, S. 99). Neben auf schriftsprachliche Beratung bezogene Fähigkeiten, tiefergehendes Wissen um Textanalyse und -produktion und Methodenkompetenzen, die der medial gestützten Form der Kommunikation zugrunde liegen, benötigen Onlineberatende Grundkenntnisse in der Beherrschung mit der Technik, deren Tools und Hilfsmitteln sowie ein sensibles Bewusstsein für Datenschutz, sodass Ratsuchende in dieser Hinsicht abgesichert sind und sich zu öffnen vermögen (vgl. Blackstein 2023, S. 29 f.; vgl. Engelhardt 2021, S. 33; vgl. Thiery 2018, S. 8). Da diese Fähigkeiten nicht unbedingt zur Grundausbildung Sozialarbeitender gehören, plädieren Stüwe & Ermel (2019) dafür, Onlineberatung in der Hochschullehre fest zu verankern und Fort- und Weiterbildungsangebote anzubieten, damit in der Onlineberatung Tätige den Ansprüchen des Settings der Onlineberatung hinreichend gerecht werden können (vgl. S. 125).

## 2.2 Instrumente der schriftlichen Onlineberatung

Die textgebundene Onlineberatung kann mittels unterschiedlicher Instrumente umgesetzt werden, die im Folgenden kurz, die wichtigsten Merkmale betonend, beschrieben werden sollen. Die Forenberatung soll an dieser Stelle unberücksichtigt bleiben, da für das Thema dieser Arbeit lediglich eins-zu-eins Beratung relevant ist, die in einem geschützten, nicht öffentlich einsehbareren virtuellen Raum abläuft. Nicht betrachtet verbleibt auch die Peer-Beratung, die oft mit der Forenberatung einhergeht, weil dort keine Beratung zwischen einer professionellen Fachkraft und einer ratsuchenden Person stattfindet. Ausgeklammert werden soll im Rahmen dieser Arbeit auch Blended Counseling, dessen Vorteile durch die Kombination der jeweiligen Kommunikationskanäle aber unbestreitbar sind, da „gerade in längeren Prozessen mit eventuellen Wechselschritten zwischen online und offline [...] schriftlich die Themen geäußert werden können, die im direkten Kontakt nicht oder noch nicht angesprochen werden können, auf die es aber auf diese Weise dann doch wieder möglich ist, Bezug zu nehmen“ (Brunner 2009, S. 37).

### 2.2.1 Mailberatung

Die Mailberatung ist die gängigste Form der textbasierten Onlineberatung (vgl. Stüwe & Ermel 2019, S. 123). Sie war, wie eingangs beschrieben wurde, auch die erste Variante der durchgeführten Onlineberatung (vgl. Kühne 2009, S. 4). Anders als vielleicht angenommen werden kann, findet diese Form zumeist über plattforminterne und SSL-verschlüsselte Nachrichtensysteme statt, um einen datenschutzsicheren Rahmen zu bieten, und nicht über herkömmliche E-Mail-Klienten (vgl. Eichenberg & Kühne 2014, S. 88). Rein aus technischer Perspektive betrachtet, kann Mailberatung aber auch über diese erfolgen. Mailberatung ist eine asynchrone Form der Onlineberatung, was bedeutet, dass diese Art der Kommunikation zeitversetzt passiert und ortsunabhängig umgesetzt werden kann (vgl. Engelhardt 2021, S. 64). Aufgrund der Zeitunabhängigkeit erstreckt sich die Produktion des Textes über einen gedehnten Zeitraum, der „das Planen, das Formulieren, das Niederschreiben und das Überarbeiten“ (Knatz 2009b, S. 67) sowohl auf Seiten der Ratsuchenden als auch der Beratenden beinhaltet und erlaubt. Auf Grundlage der wechselseitigen Mailkommunikation entwickelt sich ein Beratungsprozess (vgl. Stüwe & Ermel 2019, S. 123). Dieser findet in aller Regel im direkten Kontakt zwischen zwei Personen statt und kann sich über einen längeren Zeitraum erstrecken, was einen persönlichen Dialog begünstigen kann (vgl. Engelhardt 2021, S. 65). Eichenberg & Kühne (2014) merken diesbezüglich aber an, Mailberatung eigne sich aufgrund der zeitlichen Abstände zwischen den Mails nur in eingeschränktem Maße für akute Krisen (vgl. S. 89). Trotz dessen gilt die Mailberatung als geeignete Form der Onlineberatung beispielsweise in der Suizidprävention bei jungen Menschen (vgl.

Oswald 2018, S. 9). In dieser Hinsicht ist für Ratsuchende der zuverlässige und regelmäßige Austausch der Mailberatung bedeutsamer als vergleichsweise schnellere und kürzere Kontakte, wie sie in der Chatberatung geschehen können (vgl. ebd., S. 8). Weinhardt (2010) betrachtet die Mailberatung „als universelles Medium, [um] unterschiedliche Beratungsanlässe zu bearbeiten“ (S. 8), und schlussfolgert, sie sei für ein breites Spektrum an Typen von Nutzenden geeignet, um eine große Palette an unterschiedlichen Beratungsanlässen zu bearbeiten, und daher für Ratsuchende aller sozialen Schichten als funktional zu erachten (vgl. ebd.).

### 2.2.2 Chatberatung

Die Chatberatung wird als synchrones Beratungssetting ausgemacht, in dem Ratsuchende und Beratende zur gleichen Zeit online sind und prinzipiell in Echtzeit miteinander in Kommunikation treten (vgl. Eichenberg & Kühne 2014, S. 100). Bei genauerer Betrachtung wird die Chatberatung als ‚quasi-synchrones‘ Setting definiert, da Chatbeiträge nicht unmittelbar Buchstabe für Buchstabe wie in einem Schreibprogramm auf dem Monitor des Gegenübers erscheinen, sondern in mal mehr oder wenig langen Sequenzen erst nach Betätigung der Senden-Taste an die am Chat Beteiligten gesendet werden (vgl. Heider 2008, S. 1). Dennoch bleibt festzuhalten: Im Diskurs rund um Onlineberatung üblicherweise gebräuchlich sind dennoch hauptsächlich nur die Begriffe synchron und asynchron (vgl. Engelhardt 2021, S. 66). Als synchrones Medium unterliegen Teilnehmende im Vergleich zur asynchronen Mailberatung einer deutlichen Beschleunigung. Trotz der Gleichzeitigkeit und Kopräsenz der am Chat beteiligten Personen im virtuellen Raum kann im Vergleich zur Präsenzberatung nur ein geringeres Maß an Informationen transportiert werden (vgl. Hintenberger 2009, S. 69). Chatberatung kann aus Sicht der Beratenden im Einzelkontakt mit einer ratsuchenden Person stattfinden oder im Gruppenchat mit mehreren Klient:innen, wodurch je nach Setting andere Anforderungen an Beratende gestellt werden (vgl. Engelhardt 2021, S. 66). So können in Gruppenchats Gesprächsstränge von mehreren Teilnehmenden ineinander verwoben sein und mehrere davon gleichzeitig ablaufen, was das Folgen von Wortbeiträgen erschweren und die Kommunikation generell herausfordernder machen kann (vgl. Stüwe & Ermel 2019, S. 124). Der ‚Expertenchat‘ stellt eine besondere Variante des Gruppenchats dar, in dem sich jemand Fachkundiges oft zu einem vorher vereinbarten Thema möglichen Fragen von interessierten Ratsuchenden stellt, was beispielsweise häufig in der Prävention Anwendung findet (vgl. Hintenberger 2009, S. 77). Verglichen mit der Mailberatung betonen Stüwe & Ermel (2019), „die Kommunikation im Chatraum [ist] weniger ausführlich, tief-schürfend und reflektiert, sie ist eindeutig spontaner und emotionaler“ (S. 124), weshalb Chatten vor allem bei jüngeren Menschen Zuspruch genieße. Gerade von jungen Ratsuchenden können im Chat, anders als in der Mailberatung, Codes in Form von Akronymen, Emojis oder

Umgangssprache zum Einsatz kommen, weshalb Beratenden auch in dieser Hinsicht Kenntnisse behilflich sein können, um missverständlicher und holpriger Kommunikation vorbeugen zu können (vgl. Knatz 2022, S. 24; vgl. Hintenberger 2009, S. 71). Ein daran angrenzendes, in der Onlineberatung – allen voran in der Chatberatung – zu beobachtendes Phänomen wird Oraliteralität genannt und beschreibt den Umstand, dass die verschriftlichten Sätze stark der gesprochenen Sprache angeglichen werden (vgl. Hintenberger 2010, S. 4). Der Einsatz von Oraliteralität stellt eine Möglichkeit dar, die „mit der Kanalreduktion einhergehende Schwierigkeit, Gefühle, Emotionen und Ausdruck zu übermitteln“ (Engelhardt 2021, S. 56), zu kompensieren.

### 2.2.3 Messengerberatung

Die Messengerberatung ist noch eine sehr junge Form der Onlineberatung, die laut Engelhardt & Piekorz (2022) mutmaßlich während der Coronapandemie an Relevanz hinzugewann (vgl. S. 19). Vor allem junge Menschen nutzen ihr Smartphone vordergründig für das Internet und damit einhergehend für Messengerdienste (vgl. Kap. 2.5), wodurch ein grundsätzlicher positiver Effekt der Messengerberatung darin auszumachen ist, dass sie eine sehr lebensweltnahe Form der Onlineberatung darstellt (vgl. Schubert et al. 2019, S. 228). Daher ist auch davon auszugehen, das Nutzungsverhalten könnte eine der Hauptantriebsfedern für die Entwicklung der Messengerberatung gewesen sein (vgl. Engelhardt & Piekorz 2022, S. 24).

Obwohl das Naturell der Messengerberatung, die auf mobilen Endgeräten über geläufige Apps wie WhatsApp, aber auch auf Desktop-PCs über Web-Anwendungen stattfinden kann, in seiner Ausgestaltung der Chatberatung prinzipiell sehr ähnelt, gibt es in der Praxis dennoch einige gravierende zu beobachtende Unterschiede (vgl. Blackstein 2023, S. 108; vgl. Engelhardt 2021, S. 70). Hervorzuheben ist vor allem die Möglichkeit via Messengern nicht nur über Textnachrichten kommunizieren zu können, sondern auch Funktionen zur Hand zu haben, die es Nutzenden erlauben, Video- und Sprachnachrichten zu verschicken, sowie Video- und Sprachanrufe zu veranstalten (vgl. Engelhardt & Piekorz 2022, S. 19). Sprachnachrichten können Barrieren abbauen und kommen Ratsuchenden entgegen, die sich schriftlich entweder nicht ausdrücken können oder wollen (vgl. Knatz 2022, S. 16). Beratenden kommt zu Gute, dass Emotionen aus der Stimme herauszuhören sind und daher weniger Missverständnisse in der Kommunikation auftreten (vgl. ebd., S. 17). Darüber hinaus kommt Ratsuchenden der entlastende Effekt entgegen, den das Aufsprechen einer Nachricht haben kann, der bei schriftlicher Kommunikation auch zu beobachten ist (vgl. Kap. 2.4), nur dass eine Sprachnachricht im Vergleich zum Verfassen einer Mail spontaner aufgenommen werden kann (vgl. Knatz 2022, S. 17). Gängige Messenger ermöglichen es zudem, Dateien aller Art auszutauschen (vgl. Blackstein 2023, S. 108). Durch die



Funktionsvielfalt von Messengern können sich unterschiedlichste Beratungssettings ergeben, die mal asynchron, mal synchron verlaufen können (vgl. Hörmann et al. 2023, S. 32). Die Kommunikation via Messenger – temporäre Audio- oder Videoanrufe ausgeklammert – kann demnach als eine Art ‚Stream‘ bezeichnet werden, in dem Dialoge oftmals nicht klar durch Beginn und Abschluss gerahmt sind, weshalb übliche Gesprächsformalitäten weniger und seltener zur Anwendung kommen (vgl. Engelhardt 2021, S. 59). Engelhardt & Piekorz (2022) sehen das größte Potenzial der Messengerberatung darin, multimediale Inhalte über die vorliegenden multifunktionalen Werkzeuge in die Beratung einzubetten, um Nutzende multimodal anzusprechen zu können, was gerade im Bereich von Psychoedukation immense Vorteile bieten kann, da mediale Inhalte unmittelbar in den Kommunikationsverlauf eingebunden werden können (vgl. S. 21 ff.).

Die vorgestellte vorzufindende Vielseitigkeit von Messengerberatung nährt allerdings die Befürchtung, ein fachlicher Qualitätsverlust könne die Folge sein, da sich gängige fokussierende Beratungsprozesse unter Umständen nur beschwerlich realisieren lassen können (vgl. Engelhardt 2021, S. 70). Ein weiterer Nachteil der Messengerberatung besteht hinsichtlich des Datenschutzes. Zwar versenden heutzutage fast alle auf dem Markt befindlichen Messengerdienste ihre Nachrichten und Dateien in verschlüsselter Form, jedoch erheben Konzerne wie Meta, denen der geläufigste Messenger WhatsApp zugehörig ist, Metadaten über das Nutzungs- und Kommunikationsverhalten seiner Nutzenden (vgl. Blackstein 2023, S. 110 f.). Ratsuchende sollten daher über diesen Umstand in Kenntnis gesetzt werden, bei gleichzeitigem Verweis auf mögliche bestehende sichere Anlaufstellen, die betreffende Beratungsstellen anbieten sollten (vgl. ebd., S. 111). Abschließend sei festgestellt: Diese noch recht junge Form der Onlineberatung wurde bislang wenig untersucht, weshalb sie weiterer systematischer Forschung bedarf (vgl. Engelhardt & Piekorz 2022, S. 25).

### **2.3 Haltung und Methoden**

Haltungen, die in der Präsenzberatung eingenommen werden, und Methoden beziehungsweise Interventionen, die bei der Face-to-Face-Beratung zur Anwendung kommen, lassen sich unter Berücksichtigung der in Kapitel 2.1 beschriebenen Eigen- und Besonderheiten des Settings Onlineberatung adaptiert auf dieses übertragen. Brunner (2006) postuliert diesbezüglich, Techniken „wie aktives Zuhören, zirkuläres Fragen [und] Rollenspiele“ (S. 2) und Methoden, die „mit Beratungskonzepten, Beratungstheorien, Menschenbildern und ähnlichem zu tun [haben]“ (ebd.), seien losgelöst vom Setting zu betrachten und Ausdruck der Haltung des Beratenden gegenüber Ratsuchenden „im Kontext von Beratungshandeln“ (ebd.), weshalb vor allem die Haltung hervorzuheben sei. Bündig zusammengefasst beschreibt Ploil (2009) die in der

psychosozialen Onlineberatung gängig vertretene Haltung als „personenzentriert, lebenswelt- und ressourcenorientiert, geschlechtersensibel und mit systemischer Perspektive sowie dem Prinzip ‚Hilfe zur Selbsthilfe‘“ (S. 39). Zu der Ansicht, dieser von Ploil (2009) beschriebene polyeklektische Ansatz fände Anwendung, gelangte auch Mode-Scheibel (2014), die in einer empirischen Studie untersuchte, welche Beratungsansätze unter Interventionen in der Onlineberatung zum Einsatz kommen würden. Dabei gaben 98 Prozent der Befragten an, einen polyeklektischen Ansatz bei ihrer Tätigkeit in der Onlineberatung zu bevorzugen, bei dem therapie- beziehungsweise theorieschulenübergreifende Methoden und Interventionen angewendet werden (vgl. Mode-Scheibel 2014, S. 93). Neben dem Umstand, dass Beratungsansätze verschiedener auf das Individuum ausgerichteter Therapieschulen häufig im Studium und in Fort- und Weiterbildungen an Beratende vermittelt werden, scheint die Präferenz eines integrativen Ansatzes auch dem breiten Aufgaben- und Tätigkeitsfeld geschuldet zu sein, das in der beraterischen Praxis abgedeckt wird (vgl. Schubert et al. 2019, S. 63).

Anhand der in Abbildung 2 dargestellten Ergebnisse lässt sich die Wichtigkeit der drei Dimensionen personenzentrierter Beratung beziehungsweise Therapie nach Carl Rogers<sup>2</sup> erkennen, nämlich Empathie, Wertschätzung und Kongruenz, die laut Nußbeck (2019), „die heute allgemein akzeptierten und als Grundlage jeglicher Berater-Klient-Beziehung angesehenen Variablen sind“ (S. 112).

---

<sup>2</sup> Zur weiteren Vertiefung: Rogers, C. R. (1999). *Therapeut und Klient* (14. Aufl.). Frankfurt a. M.: Fischer.; Rogers, C. R. (2007). *Die nicht-direktive Beratung* (12. Aufl.). Frankfurt a. M.: Fischer; Rogers, C. R. (2016). *Die klientenzentrierte Gesprächspsychotherapie* (20. Auflage). Frankfurt a. M.: Fischer.

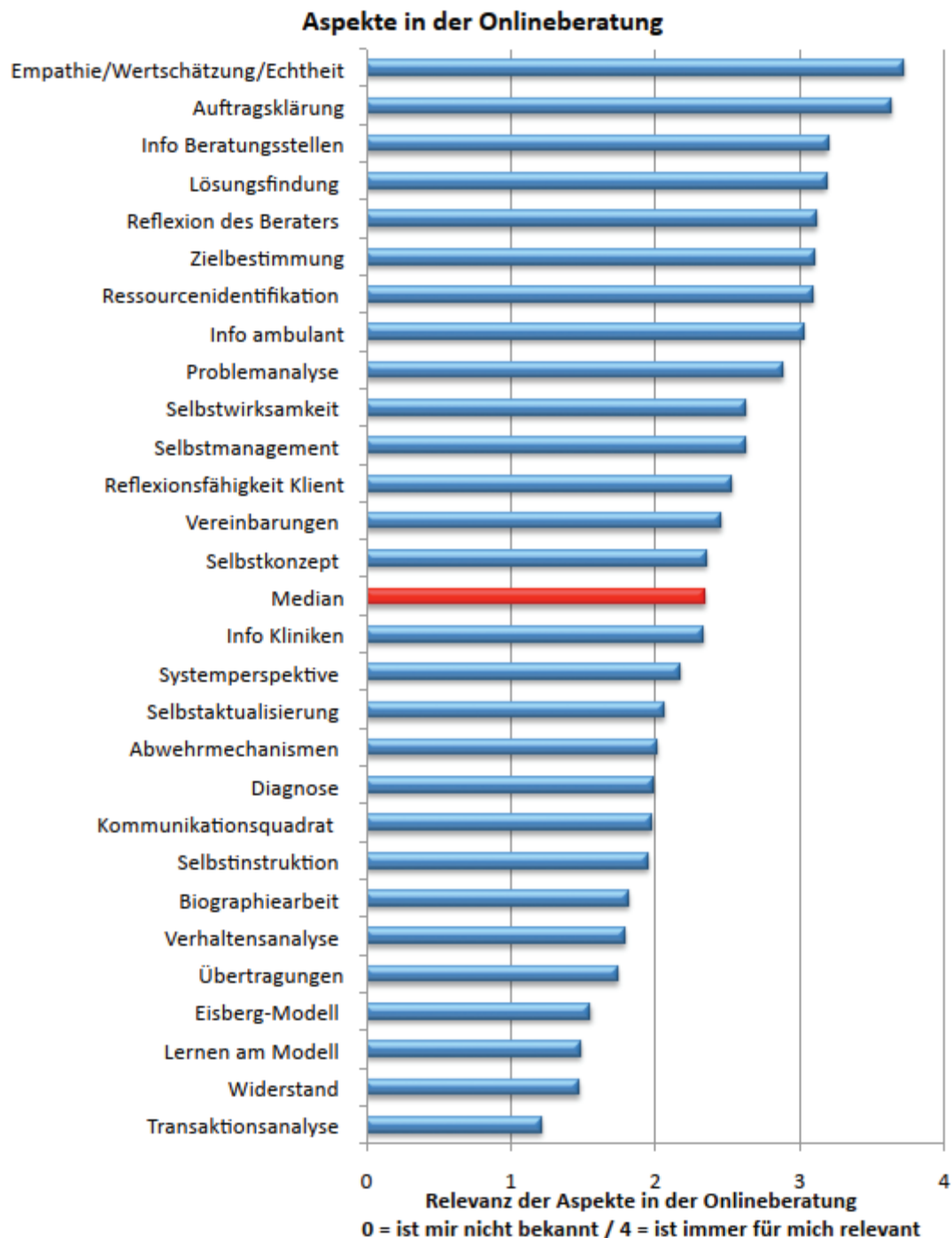


Abbildung 2: Aspekte der Onlineberatung (Mode-Scheibel 2014, S. 95)

Personenzentrierte Beratung, die der humanistischen Psychologie zuzuordnen ist, kann ein wenig wie eine Grundhaltung verstanden werden, die sich eher als Menschenbild beziehungsweise grundsätzliche Herangehensweise an das Gegenüber versteht, denn als Technik oder Methode (vgl. Schubert 2019, S. 81). Neben der für Beratung im Allgemeinen immanent bedeutsamen Auftragsklärung scheint auch die ressourcen- und lösungsorientierte Beratung einen gewichtigen Platz bei Onlineberatenden in der Praxis einzunehmen, wie sich an der hoch eingestuftem

Lösungsfindung und Ressourcenidentifikation in Abbildung 2 und der als am relevantesten empfundenen Lösungsorientierung in Abbildung 3 erkennen lässt.

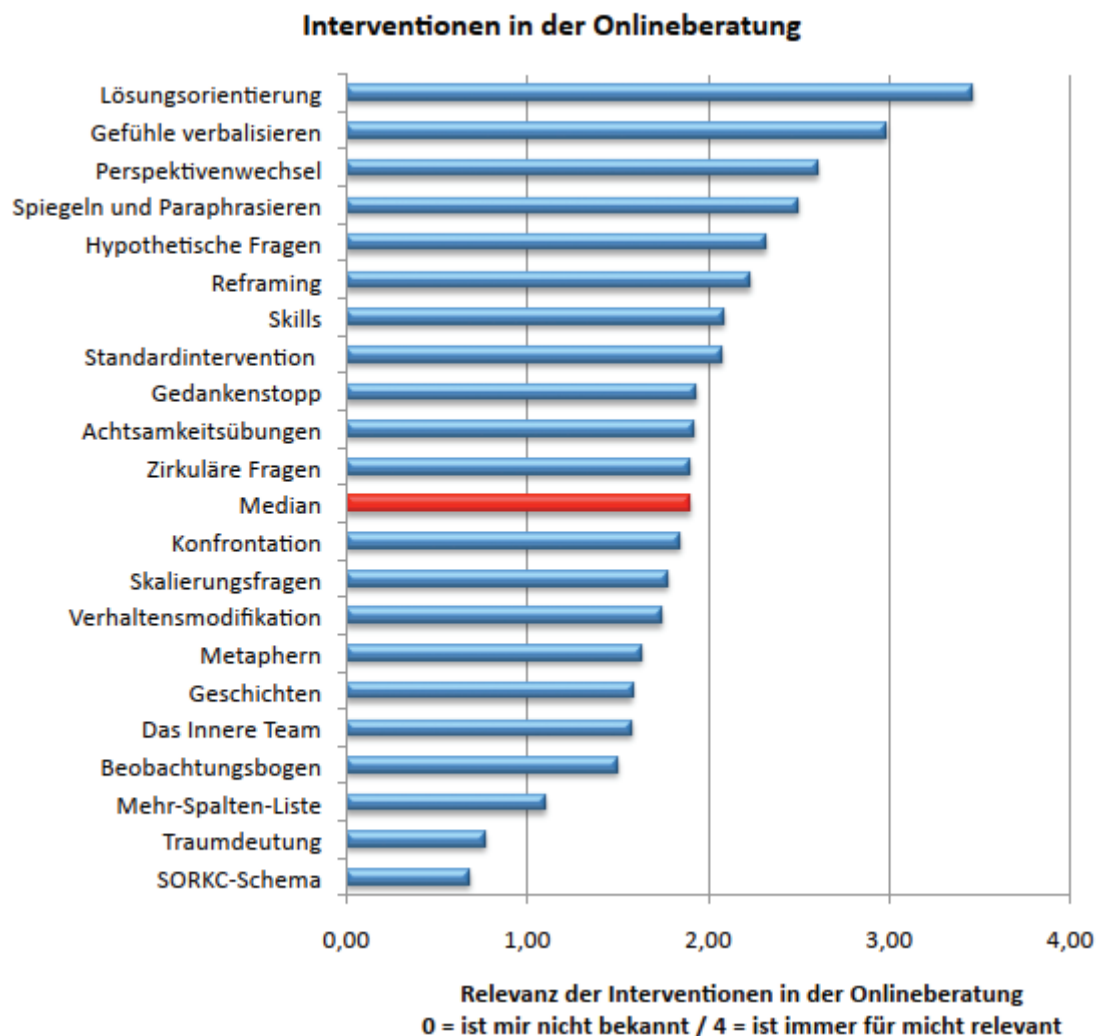


Abbildung 3: Interventionen der Onlineberatung (Mode-Scheibel 2014, S. 97)

In lösungs- und ressourcenorientierter Beratung wird versucht, bereits Funktionierendes auf Seiten der Ratsuchenden aufzutun oder mögliche neue Quellen von Ressourcen zu identifizieren, deren Potenziale dann genutzt werden können, um an der Umsetzung der Lösung zu arbeiten (vgl. Nußbeck 2019, S. 73). Aus diesem Grund wird in der lösungs- und ressourcenorientierten Beratung der Fokus weg von Problemen hin zu funktionierenden Dingen gerichtet und es gilt herauszufinden, unter welchen Umständen Probleme nicht zu Tage treten (vgl. ebd.). Diese Haltung wird auch in der systemischen Beratung vertreten (vgl. von Schlippe & Schweitzer 2016, S. 209). Für sie typisch sind die in Abbildung 3 überdurchschnittlich oft als relevant benannten Interventionen Perspektivenwechsel, Reframing und hypothetische sowie zirkuläre Fragen (vgl. Lindemann 2023, Kap. 2.4). Bei Betrachtung der beiden Abbildungen ist auch auffällig, wie bedeutsam für Onlineberatende das Verbalisieren von Gefühlen und das Spiegeln und

Paraphrasieren gegenüber Ratsuchenden ist. Ein Erklärungsansatz bietet die bereits thematisierte Kanalreduktion, die es den Onlineberatenden unmöglich macht, parasprachliche Mittel oder Gestik und Mimik dafür zu nutzen, um Verständnis und Empathie zu signalisieren oder anderweitig zu kommunizieren (vgl. Kap. 2.1). Das Verbalisieren von Gefühlen ist in dieser Hinsicht eine gute Möglichkeit, um dem Gegenüber seine emotionale Anteilnahme zu signalisieren, und das Spiegeln und Paraphrasieren des Geschriebenen der Ratsuchenden eignet sich, um sich der faktischen und emotionalen Wirklichkeit anzunähern, indem kommuniziert wird, was auf sachlicher und emotionaler Ebene verstanden wurde (vgl. Knatz 2009a, S. 106). In der Onlineberatung müssen Beratende daher „Ausdrucksformen wie Aktionswörter, Emoticons und ganz generell die Beschreibung und Umschreibung dessen forcieren, was in der Interaktion unter Abwesenden eben nicht unmittelbar wahrnehmbar ist“ (Brunner 2009, S. 33). Als überraschend erachtet werden kann in den Ergebnissen von Mode-Scheibels Erhebung, dass das 4-Folien-Konzept nach Knatz & Dodier (2003) vielen Onlineberatenden gänzlich unbekannt war beziehungsweise in der Beratungspraxis keine Anwendung findet (vgl. Mode-Scheibel 2014, S. 94). Das 4-Folien-Konzept gilt als verbreitete und durchdachte Strukturierungshilfe, die speziell für die Onlineberatung konzipiert wurde und insbesondere in der Mailberatung gut genutzt werden kann. Bei Anwendung des Konzeptes liest man eine Mail mehrfach und setzt bei jedem Lesen einen unterschiedlichen Fokus, bevor die Antwort im letzten Schritt verfasst wird, sodass die verschiedenen Ebenen der Nachricht des Ratsuchenden möglichst gut mitberücksichtigt werden können (vgl. Knatz 2009a, S. 108 ff.).

Davon abgesehen gilt es laut Boeger (2024) außerdem zu beachten, dass Phänomene wie Übertragung und Gegenübertragung in Beratungen jeglichen Kontextes und Settings auftreten können, weshalb es diese zu berücksichtigen gilt und es in der Beratung angeraten sein kann, diese zu thematisieren, um möglichen Kontaktabbrüchen und Konflikten vorbeugen zu können (vgl. S. 22). Welche theorie- und therapieschulenspezifische beziehungsweise -übergreifende Herangehensweise an den Onlineberatungsprozess Onlineberatende auch wählen, der Beratende „muss in der Lage sein, ein Beziehungsangebot schriftlich auszudrücken“ (Knatz 2009a, S. 107), denn eine gelingende Beziehung zwischen Beratenden und Ratsuchenden ist, wie sich im folgenden Unterpunkt zeigen wird, auch in der Onlineberatung einer der größten Wirkfaktoren von gelingender Beratung.

## **2.4 Wirk- und Einflussfaktoren**

Die allgemeinen Wirkfaktoren von Psychotherapie, die therapieschulenübergreifend eine besonders gute Wirkung erzielen, wurden von Grawe (2005) in der Ressourcenaktivierung, der

Problemaktualisierung, der Problembewältigung, der motivationalen Klärung und der Therapiebeziehung ausgemacht (vgl. S. 7). Fundamental wichtig ist hierbei vor allem „die Qualität der Therapiebeziehung, [die] signifikant zu einem besseren oder schlechteren Therapieergebnis [beiträgt]“ (ebd.). Auch Brunner (2009) weist darauf hin, dass die „Beziehung von zahlreichen Beratungs- und Therapieansätzen als das wesentliche Medium der Beratungs- und Therapiearbeit gesehen [wird]“ (S. 40), auf deren Basis in der Beratung „eine spezifische, reflexiv-fragende Form der Kommunikation [...] [stattfindet]“ (Weinhardt 2024, S. 62). Die in Kapitel 2.3 dargelegte zumeist polyeklektische Herangehensweise der Onlineberatenden deckt sich mit Grawes Plädoyer, „allgemeine Wirkfaktoren müssen patientenspezifisch verwirklicht werden“ (Grawe 2005, S. 8), da so die patienten- und problemspezifischen Bedürfnisse und Aspekte Berücksichtigung finden können. Berger et al. (2013) konkludieren hinsichtlich der therapeutischen Allianz, diese sei auch im Setting der textgebundenen Onlineberatung gut herzustellen, und schätzen diese vergleichbar gut mit Beratungen und Therapien in Präsenz ein (vgl. S. 10). Zu einem vergleichbaren Ergebnis kommt auch Eichenberg (2007), die im Zuge einer Studie die Wirksamkeit und Wirkweise der ‚sextra-Online-Beratung‘ der pro familia untersuchte. So postuliert sie,

dass eine virtuell vermittelte positive Beziehung zum Beratenden wie auch eine dem Medium adäquate Antwort auf das Beratungsgesuch ausschlaggebend für die Effektivität der Beratung ist, [wodurch] die allgemein bekannten Wirkfaktoren für hilfreiche psychotherapeutische Interventionen auch für die Online-Beratung [gelten] und internetbasierte Beratungsangebote durchaus fruchtbar sein [können], wenn die medien-spezifischen Besonderheiten im Beratungsprozess berücksichtigt werden (Eichenberg 2007, S. 261).

Auch Barak & Grohol (2011), die die Wirksamkeit von Onlineberatung und -therapie anhand von mehreren vorliegenden Studien untersucht haben, sehen hinsichtlich der Effektstärke von Onlinetherapie und Face-to-Face-Therapie kaum signifikante Unterschiede und bescheinigen gerade Interventionen, die über Mail oder Chat ablaufen, eine gute Wirksamkeit (vgl. S. 160).

Wie Wenzel (2019) hervorhebt, kann die Wirksamkeit von Onlineberatung also als begründet und erbracht erachtet werden (vgl. S. 226). „Online-Beratung [...] kann [...] eine hohe Wirksamkeit nicht nur in den virtuellen, sondern auch in den realen Dimensionen der Lebenswelten ihrer Nutzerinnen entfalten“ (Brunner 2009, S. 43). Darüber hinaus stellt Onlineberatung durch ihre spezifischen, dem Setting zugrundeliegenden Merkmale sogar eine günstige Ergänzung für Ratsuchende dar (vgl. Wenzel 2019, S. 225). Hervorzuheben sei anfangs die Anonymität der Ratsuchenden, die zwar nicht obligatorisch vorkommt, jedoch in der Praxis häufig gegeben ist, da Onlineberatungsangebote häufig keine Klardaten der Ratsuchenden verlangen beziehungsweise nicht auf diese angewiesen sind, um Onlineberatung gewährleisten zu können (vgl. Hörmann et al. 2023, S. 14). Die Anonymität gepaart mit der bereits beschriebenen Kanalreduktion kann bei

vielen Ratsuchenden in der Onlineberatung den Effekt haben, sich leichter dem nicht direkt sicht- und greifbaren Gegenüber zu öffnen (vgl. Knatz 2009b, S. 59). Dieser enthemmende Effekt, der sich generell in der Onlinekommunikation beobachten lässt, wird „Disinhibition Effect“ (Suler 2004, S. 321) genannt. Laut Suler sorgen Anonymität und das Nicht-gesehen-werden-können bei Nutzenden sowie etwaige verzögerte Reaktionen des Gegenübers in der Kommunikation für den enthemmenden Effekt (vgl. ebd., S. 322 f.). Auch spielen intrapersonelle und soziale Faktoren eine geringere Rolle, wie Sule es in der Minimierung von Status und Autorität beschreibt, die Onlinekommunizierende weniger beeinflussen, was zusätzlich eine Öffnung in der Onlinekommunikation begünstigt (vgl. ebd. S. 324). Dieses beschriebene Phänomen verringert die Schwellenangst und trägt bei Ratsuchenden dazu bei, in Anbetracht von schambehafteten beziehungsweise stigmatisierenden Beratungsanlässen und -themen, leichter eine Onlineberatung in Anspruch nehmen zu können (vgl. Hörman et al. 2023, S. 13). So kann das Selbstwertgefühl unverletzt bleiben und die Würde aufrechterhalten werden (vgl. Schultze 2007, S. 4). Nicht nur kann dieser Umstand dazu führen, dass sich Klient:innen generell in der Lage sehen, sich öffnen zu können, er kann auch zur Folge haben, dass Ratsuchende in der schriftlichen Onlineberatung schneller den Kern des Problems oder des Beratungsanlasses benennen können (vgl. Brunner 2009, S. 37). „Das geschriebene Problem steht im Vordergrund, nicht die Person in ihrer Erscheinung“, (Eichenberg & Kühne 2014, S. 83), weshalb textbasierte Onlineberatung – trotz Möglichkeiten zur Video- oder Audiotelefonie über das Internet – unverändert eine hohe Relevanz besitzt (vgl. ebd.). Petzold (2006) resümiert daher, kanalreduzierte Medien zur Kommunikation wie textgebundene Onlineberatung würden sich je nach Aufgabe und Anforderung in mancherlei Hinsicht besser eignen als Kontakte, die in Präsenz oder per Videotelefonie erfolgen (vgl. S. 5). Die Anonymität und Nichtsichtbarkeit des Gegenübers kann aber nicht nur eine Chance für Ratsuchende darstellen: Da Beratende aufgrund fehlender optischer Eindrücke in der Onlineberatung weniger dazu neigen, ihr Gegenüber in eine Schublade einzuordnen, können auch entstehende bewusste und unbewusste oberflächliche Bewertungen und mögliche Fehleinschätzungen darüber, wer der/die Ratsuchende sei und an was es ihm/ihr mangle, verringert werden (vgl. Blackstein 2023, S. 35; vgl. Lang 2002, zit. n. Engelhardt 2021, S. 28).

Anonymität trägt auch zur Niedrigschwelligkeit von Onlineberatung bei, da Ratsuchende schnell, unkompliziert und ohne größeres Aufsehen Beratung in Anspruch nehmen können und mögliche themen- und personenbezogene Skrupel weniger wirken (vgl. Reindl 2018, S. 23; vgl. Risau 2009, S. 207). Niedrigschwelligkeit stellt in Bezug auf Beratung generell ein Qualitätskriterium dar und ist bei Onlineberatung in besonderem Maße gegeben, da diese orts- und bei asynchroner Kommunikation zeitunabhängig erfolgen kann (vgl. Hörmann 2023, S. 35). Da Ratsuchende, wie

zuvor dargelegt, leichter in den Kontakt kommen können, „kann Onlineberatung eine hilfreiche Vorstufe zu intensiveren Kontakten am Telefon oder in der persönlichen Beratung sein und zu diesen hinführen“ (Knatz 2009b, S. 66). Gleichwohl ist dieser Schritt nicht als zwingende Konsequenz, sondern als Angebot zu verstehen, da Klient:innen nicht immer den geschützten Rahmen der Onlineberatung verlassen möchten. Zugleich werden im Onlineberatungsprozess vonseiten der Ratsuchenden mitunter Idealbilder auf die Onlineberatenden projiziert, die förderlich auf den Beratungsverlauf wirken können und sich bei einem Face-to-Face-Kontakt auflösen oder sogar umkehren könnten (vgl. Schultze 2007, S. 4 f.). Daher schreiben Beratende einem Wechsel von Online- zu Präsenzberatung nicht nur förderliche Potenziale zu, wie die Auswertung eines Projektes von Hörmann et al. (2019), das unter anderem den Wechsel von Online- zur Präsenzberatung untersuchte, zeigen konnte (vgl. S. 53).

Ein weiterer Einflussfaktor in der Onlineberatung ist die erhöhte Autonomie, die Ratsuchenden „ein hohes Maß an Kontrolle in der Selbstdarstellung und der Selbstenthüllung“ (Knatz 2009b, S. 59) gibt und ihnen oftmals erst ermöglicht, den Kontakt zu suchen, da dieser ohne Sorge vor einem drohenden Gesichtsverlust oder anderweitiger Konsequenzen jederzeit eingestellt werden kann (vgl. Ploil 2009, S. 18; vgl. Stüwe & Ermel 2019, S. 125). In dieser Hinsicht gibt Brunner (2009) jedoch auch zu bedenken, dass diese Art der Beziehung im Vergleich zum Kontakt in Präsenz unter Umständen auf einem weniger stabilen Fundament fußt, was schneller Kontaktabbrüche seitens der Ratsuchenden zur Folge haben kann, da Missverständnisse, Fehler oder inadäquate Verhaltensweisen von Beratenden, anders als in der Face-to-Face-Beratung, nicht unmittelbar geklärt und aufgefangen werden können (vgl. S. 37). Dem stellt Schultze (2007) gegenüber, Antworten der Beratenden würden bei Ratsuchenden auf eine geringere psychische Gegenwehr stoßen, wodurch die Tragfähigkeit der Beziehung zwischen Beratenden und Ratsuchenden in Konsequenz belastbarer sei (vgl. S. 6). Durch die stärkere Autonomie können Ratsuchende einen erhöhten Einfluss auf das Beratungsgeschehen und die Kontaktsteuerung in der Onlineberatung nehmen, indem sie bestimmen, „wohin die Reise geht, was auf dem Weg passiert und wann das Ziel erreicht wird oder ob überhaupt ein Ziel erreicht werden muss oder soll“ (Blackstein 2023, S. 32). Dadurch sind Klient:innen der Onlineberatung umfänglich dagegen gefeit, sozialen Erwartungen gerecht werden zu müssen, mit Peinlichkeiten konfrontiert zu werden oder sich Situationen aussetzen zu müssen, in denen sie durch äußerliche Einflussfaktoren unter Umständen einen Kontrollverlust erleiden könnten (vgl. Knatz 2009b, S. 62). So greift Onlineberatung durch ihr ortloses Setting des virtuellen Raumes auch stärker möglichen Ängsten vor, die beispielsweise hinsichtlich drohender Zwangsmaßnahmen bestehen könnten (vgl. Stüwe & Ermel 2019, S. 125).



Ein weiterer Einflussfaktor von textgebundener Onlineberatung ist die Schriftlichkeit, da sie mehrere positive Effekte hat. „Wer schreibt, schaut sich selbst beim Denken zu“ (Knatz 2009a, S. 106). Knatz' Feststellung deutet auf den selbstreflexiven und Gedanken ordnenden Effekt hin, den der Schreibprozess für Ratsuchende in der Onlineberatung haben kann und der im eher entschleunigten Prozess des Schreibens eine tiefergehende und strukturierte Auseinandersetzung mit den Themen, die Ratsuchende beschäftigen, erlaubt (vgl. Heimes 2023, S. 18). So berichten viele Ratsuchende davon, dass bereits das Verfassen der Eingangsmail ihnen eine erste Entlastung und Erleichterung verschaffen konnte (vgl. Knatz 2009b, S. 65). Zusätzlich kann Schreiben, was von Nutzenden oft als bedeutsam und nützlich empfunden wird, auch zur Reduktion von Symptomen und zu einem verbesserten allgemeinen Wohlbefühl führen (vgl. Zagorscak & Knaevelsrud 2021, S. 238). Insbesondere die Verschriftlichung der Gedanken kann Ratsuchenden helfen, Reflexionsprozesse anzuschieben, was einen großen Anteil am wirksamen Effekt der textgebundenen Onlineberatung ausmacht, sowie zur Klärung beizutragen und Gedanken zu strukturieren (vgl. Engelhardt & Storch 2013, S. 4). Nicht selten erleben Ratsuchende hierbei ein verstärktes Gefühl von Selbstwirksamkeit, da sie „durch ihr Schreiben einen aktiven Beitrag zur Verbesserung ihrer Gesundheit [...] leisten, was wiederum ihr Vertrauen in sich selbst stärkt“ (Heimes 2023, S. 29).

Der verschriftlichte Dialog in der Onlineberatung ermöglicht auch das, was Knatz (2009b) als „scheinbar paradoxe Situation einer Nähe durch Distanz“ (S. 64) beschreibt: Trotz körperlicher Absenz und dem Ausbleiben von wahrzunehmenden und zu deutenden Körpersignalen wie Gestik und Mimik kann psychologische Nähe aufgebaut werden (vgl. Döring 2003, S. 367). Dieses Phänomen ermöglicht es Ratsuchenden selbst über stark schambehaftete beziehungsweise gesellschaftlich stigmatisierende Themen in Kommunikation zu treten (vgl. Knatz 2009b, S. 64). Onlineberatung bietet Klient:innen daher in dieser Hinsicht die Option, sich in einer sicheren Umgebung über das Geschriebene in den Austausch zu begeben, was sie in gesprochener Form vielleicht nicht tun würden. „Es gibt Gedanken, Gefühle und Probleme, die man keinem Menschen sagen, aber auch nicht für sich behalten möchte“ (ebd., S. 59) und so kann Onlineberatung Ratsuchenden ermöglichen, darüber zu schreiben, was nicht ausgesprochen werden kann (vgl. Blackstein 2023, S. 82).

Durch Schreiben als kreativer Prozess kann ein Zugang zum inneren Erleben hergestellt werden, was eine stärkere und fokussierte Auseinandersetzung mit der eigenen Gefühlswelt erlaubt (vgl. Heimes 2023, S. 18). Es kann auch dabei helfen, bisher unterdrückte belastende Erlebnisse und Gefühle zu verarbeiten und diese zu integrieren (vgl. ebd., S. 24). Gleichzeitig kann eine Distanz zu den eigenen Erfahrungen, Umständen und Erlebnissen hergestellt werden, indem diese über

das Schreiben nach außen transportiert werden, was die Chance bietet, neue Sichtweisen und Perspektiven für die Lösungsfindung zu entwickeln und somit (wieder) Handlungsfähigkeit zu erlangen (vgl. Knatz 2009b, S. 66). Das Verschriftlichen der Gedanken kann Ratsuchenden in dieser Hinsicht dabei behilflich sein, verschiedene Handlungsweisen auszuloten und sich detailliert mit ihnen zu beschäftigen, bevor sie anfangen diese in der Realität umzusetzen (vgl. Heimes 2023, S. 18).

Der verschriftlichte Beratungsdialo g kann ferner förderlich auf Ratsuchende wirken, da sie diesen im Beratungsverlauf und auch nach Abschluss erneut beliebig oft nachlesen können, weshalb dieser einen nachhaltigen Effekt hat und eine wiederholende Selbstreflexion begünstigt wird (vgl. Schultze 2007, S. 7). Zudem bescheinigen Justen-Horsten & Paschen (2016) konstruktiven, systemisch orientierten Fragen eine hohe Nachhaltigkeit, da diese „beim wiederholten Lesen sehr kraftvoll sein [können]“ (S. 128) und durch ihren Nachhall beim Ratsuchenden längerfristige Wirkung haben können, wodurch Entwicklungs- und Veränderungsprozesse begünstigt werden (vgl. ebd.). Durch die Nachhaltigkeit der Schriftlichkeit können Inhalte und Aspekte, die in der Beratung thematisiert wurden, im Laufe der Zeit verarbeitet werden und dadurch auch allmählich neue Erkenntnisse gewonnen werden, was einen erhöhten Lerntransfer zur Folge haben kann (vgl. Schultze 2007, S. 7 f.).

Abschließend betrachtet kann also angenommen werden, dass in der schriftlichen Onlineberatung für das Setting spezifische Wirkfaktoren und eingangs beleuchtete allgemeine Wirkfaktoren gleichermaßen bedeutsam sind und entsprechend zum Tragen kommen.

## **2.5 Mediennutzungsverhalten und dessen Bedeutung für die Onlineberatung**

Wie im vorigen Kapitel ersichtlich wurde, bietet das Internet durch seine Anonymität einen als sicher wahrgenommenen Raum, dem Menschen sich zuwenden, um Fragen und als unangenehm empfundene Themen, nicht nur in der Onlineberatung, auf den Grund gehen zu können (vgl. Ashekian 2022).

Digitale Medien haben schon vor vielen Jahren Einzug in deutsche Haushalte gehalten und sind mittlerweile als integraler Bestandteil dieser zu erachten. So nutzen rund 80 Prozent der Deutschen das Internet täglich (vgl. Beisch & Koch 2023, S. 1). Schätzungsweise 68 Millionen Deutsche besitzen ein Smartphone, wovon vor allem die jüngere Hälfte der Bevölkerung einen Großteil der Nutzenden ausmacht (vgl. Tenzer 2024). Kinder und Jugendliche zeichnen sich in der Regel durch eine hohe Affinität zu neuen und sich stets verändernden Medien aus, die sie sich

zumeist wie selbstverständlich erschließen und einen natürlichen Umgang mit ihnen pflegen (vgl. Quenzel & Hurrelmann 2022, S. 175).

Das aktuelle Mediennutzungsverhalten und die Ausstattung hinsichtlich der ‚neuen Medien‘ von Kindern und Jugendlichen in Deutschland lassen sich vornehmlich anhand der aktuellsten KIM-Studie von 2022, die seit 1999 zweijährlich erscheint, und der jüngsten JIM-Studie von 2023, die seit 25 Jahren jährlich herausgebracht wird, untersuchen. Für die KIM-Studie 2022 wurden 1.219 Kinder im Alter zwischen sechs bis dreizehn Jahren sowie das haupterziehende Elternteil befragt (vgl. Medienpädagogischer Forschungsverbund Südwest 2022, S. 2). An der JIM-Studie 2023 nahmen 1.200 Jugendliche im Alter von zwölf bis neunzehn Jahren teil (vgl. Medienpädagogischer Forschungsverbund Südwest 2023, S. 2).

In fast ausschließlich allen Haushalten der befragten Kinder und Jugendlichen bestehen internetfähige Geräte wie Smartphones, Laptops, Computer oder Tablets, wobei sogar in etwas mehr als der Hälfte der Haushalte Wearables wie etwa Smartwatches vorhanden sind (vgl. Medienpädagogischer Forschungsverbund Südwest 2022, S. 4; vgl. Medienpädagogischer Forschungsverbund Südwest 2023, S. 5). Mit 96 Prozent der Jugendlichen besitzen fast alle ein eigenes Smartphone; ungefähr drei Viertel sind entweder im Besitz eines Laptops oder Computers; 56 Prozent der Jugendlichen haben ein eigenes Tablet; gut ein Drittel der Jugendlichen haben eigene Wearables; 23 Prozent verfügen über einen eigenen Smartspeaker (vgl. Medienpädagogischer Forschungsverbund Südwest 2023, S. 6). 44 Prozent der Kinder besitzen ein eigenes Smartphone, wobei der Anteil steigend mit dem Alter zunimmt (vgl. Medienpädagogischer Forschungsverbund Südwest 2022, S. 5). Besitzen lediglich 9 Prozent der Sechs- bis Siebenjährigen ein Smartphone, so haben 81 Prozent der zwölf- bis dreizehnjährigen Kinder eines (vgl. ebd.). Ein ähnlicher Altersverlauf lässt sich auch in Bezug auf weitere internetfähige Endgeräte wie Computer, Laptops, Tablets beobachten (vgl. ebd., S. 6). Vergleichbar sieht es beim Umgang mit Smartphones und dem Internet aus: Benutzen lediglich 28 Prozent der Sechs- bis Siebenjährigen ein- oder mehrmals pro Woche ein Mobiltelefon und nur 21 Prozent dieser Altersgruppe im beschriebenen Zeitrahmen das Internet, so tun dies 96 Prozent der Zwölf- bis Dreizehnjährigen mindestens einmal in der Woche und 92 Prozent dieser Altersgruppe nutzen das Internet im gegebenen Zeitrahmen (vgl. ebd., S. 9 f.). Das Smartphone ist für Jugendliche das wichtigste Gerät, das 98 Prozent von ihnen mindestens einmal wöchentlich nutzen, wobei 95 Prozent der Jugendlichen mehrmals wöchentlich bis täglich online sind (vgl. Medienpädagogischer Forschungsverbund Südwest 2023, S. 12 ff.). Knapp die Hälfte aller befragten Jugendlichen nutzt regelmäßig Smartspeaker (vgl. ebd., S. 15).

Für 70 Prozent der Kinder, die das Internet regelmäßig nutzen, ist vor allem das Nachrichten versenden und empfangen über WhatsApp eine beliebte Tätigkeit und liegt damit noch vor der auf Videoclips und -streams fokussierte App TikTok, die für gut jedes zweite internetnutzende Kind von Bedeutung ist (vgl. Medienpädagogischer Forschungsverbund Südwest 2022, S. 83). Auch bei den befragten Jugendlichen ist WhatsApp mit Abstand die beliebteste App, sodass ganze 94 Prozent diese täglich beziehungsweise in aller Regelmäßigkeit nutzen, um Text- und Sprachnachrichten zu verschicken und zu empfangen (vgl. Medienpädagogischer Forschungsverbund Südwest 2023, S. 76). Leven & Utzmann (2019) bestätigen auf Grundlage der aktuellsten Shell-Jugendstudie von 2019 den Status von WhatsApp, die sich als wichtigste Kommunikationsplattform für junge Menschen herausstellt, und konstatieren: „Alle befragten Jugendlichen haben auf ihren Smartphones die App mit dem grünen Icon installiert. Und niemand kennt jemanden in seinem Freundes- oder Bekanntenkreis, der WhatsApp oder etwas Vergleichbares nicht verwendet“ (S. 272). Auf Platz zwei liegt Instagram, das von 62 Prozent genutzt wird, und auf Platz drei die App TikTok, die 59 Prozent nutzen (vgl. Medienpädagogischer Forschungsverbund Südwest 2023, S. 76). Es wird deutlich, dass Messengerdienste und soziale Medien für junge Menschen die wichtigsten Aktivitäten im Internet darstellen (vgl. Wolfert & Leven 2019, S. 227). Ungefähr 224 Minuten am Tag verbringen Jugendliche im Schnitt im Internet (vgl. Medienpädagogischer Forschungsverbund Südwest 2023, S. 76). Kinder hingegen kommen im Durchschnitt auf 43 Minuten Onlinezeit pro Tag, wobei die Dauer bei sechs- bis siebenjährigen Kindern mit 17 Minuten eher gering ausfällt und bei Zwölf- bis Dreizehnjährigen mit 74 Minuten eher höher bemessen ist (vgl. Medienpädagogischer Forschungsverbund Südwest 2022, S. 83). Obwohl ChatGPT erst Ende des Jahres 2022, also ungefähr ein halbes Jahr vor der stattgefundenen Befragung, veröffentlicht wurde, wussten bereits drei Viertel der Jugendlichen, was es mit der Künstlichen Intelligenz auf sich hat und fast 40 Prozent der Befragten haben den Chatbot (vgl. Kap. 4.1) sogar schon einmal benutzt (vgl. Medienpädagogischer Forschungsverbund Südwest 2023, S. 31). Es ist davon auszugehen, dass diese Zahlen im Verlauf der nächsten Jahre weiter ansteigen werden.

Viele Menschen sind mit ihren internetfähigen Endgeräten quasi daueronline und wählen sich nicht mehr wie dereinst bewusst ins Internet ein, weshalb die reale und die digitale Welt als ineinander verflochten begriffen werden müssen und eine strikte Trennung dieser Bereiche nicht mehr sinnig scheint (vgl. Engelhardt & Storch 2013, S. 9). Damit geht auch ein verändertes Kommunikationsverhalten einher, das durch digitale Medien in vielfacher Ausgestaltung realisiert werden kann, wie dieses Kapitel bereits gezeigt hat. Die Art und Weise, wie kommuniziert wird, hat im Laufe der Zeit vor allem in jüngeren Generationen einen Wandel durchlaufen, was

sich gesamtgesellschaftlich auf die Kommunikationsroutinen auswirkt (vgl. Reindl & Engelhardt 2021, S. 118). Wie die aktuelle ARD/ZDF-Onlinestudie, die jedes Jahr erhoben wird, zeigt, kommen Deutsche der Altersgruppe der Jugendlichen und jungen Erwachsenen im Alter von 14-29 Jahren auf die höchste Internetnutzungsdauer von etwas mehr als vier Stunden pro Tag, wobei allein schon knapp 90 Minuten zum Chatten, Nachrichten verschicken und Kommunizieren über Messengerdienste investiert werden (vgl. Beisch & Koch 2023, S. 5). Klein & Pulver (2020) konstatieren bezüglich dieser Entwicklung, dass „der weitreichende Mediatisierungsschub der Alltagswelten, der auf die Digitalisierung folgte, bedeutet, dass kommunikatives und soziales Handeln zunehmend auch mediales Handeln ist, respektive dass das eine das andere nicht (mehr) ausschließt“ (S. 190). Durch diese „mediatisierten Veränderungen gesellschaftlicher Kommunikationsprozesse und Lebenswelten“ (ebd.) ergeben sich für die Soziale Arbeit zwangsläufig neue Anforderungen. Insbesondere Kinder und Jugendliche sind von der Mediatisierung der Lebenswelt stark betroffen, weshalb für sie die sogenannten ‚neuen Medien‘ eine große Rolle spielen und diese dadurch gleichzeitig auch für die Soziale Arbeit an Relevanz gewinnen (vgl. von Wensierski 2018, S. 981). Das Angebot an Anlauf- und Beratungsstellen, die im Internet zu finden sind und zu denen auch unseriöse und unpassende Angebote gehören, kann Ratsuchende schnell überfordern (vgl. Klein & Pulver 2020, S. 192). Die Soziale Arbeit ist daher gefordert, diesen Umstand aufzugreifen und Präsenz im medialen Raum zu zeigen, der „als faktischer sozialer Raum zu verstehen [ist]“ (ebd. S. 191).

Schriftliche Onlineberatung ist aufgrund der Lebensweltnähe und der alltagsvertrauten Kommunikation, wie die in diesem Kapitel dargelegten Nutzungszahlen und die Beliebtheit von WhatsApp bei Kindern und Jugendlichen gezeigt haben, eine von vielen Möglichkeiten, diesen neuen Anforderungen zu begegnen und ihnen gerecht zu werden. Onlineberatung bietet Jugendlichen, die sich ihrer selbst häufig noch etwas unsicher und unerfahren sind, die Möglichkeit sich in den Austausch zu begeben, ohne sich mit ihrer ganzen Person präsentieren zu müssen, sodass sich für sie auch ein sicherer Raum bietet, der sie einlädt, sich etwas nach vorne zu wagen und sich auszuprobieren (vgl. Petzold 2006, S. 10). Onlinemedien sind für junge Menschen ein Sozialisationsinstrument und erfüllen unter anderem eine Ratgeberfunktion, die zur Hand genommen werden kann, wenn sie Hilfe benötigen (vgl. Fleischer 2018, S. 289; vgl. Rösch 2022, S. 142). Daraus ergeben sich Implikationen für die sozialarbeiterische Praxis.

Soziale Arbeit ist aufgefordert, im digitalen Raum auf mögliche Klient:innen zuzugehen und ihnen Zugangswege zu eröffnen, da „der Einstieg und Zugang zu professionellen Beratungsangeboten im Internet [...] vielfach nicht [...] über selbständiges Suchen im Internet, sondern direkt über soziale Netzwerke [erfolgt]“ (Klein & Pulver 2020, S. 196). Angebote der Sozialen Arbeit

werden also tendenziell nicht gefunden und frequentiert, sofern sie nicht auch Präsenz in den Sozialen Medien zeigen und dort im Sinne einer digitalen Streetwork auf sich aufmerksam machen (Klein & Pulver 2020, S. 196). So konnte die Bundeskonferenz für Erziehungsberatung, die seit über 10 Jahren mittels Onlineberatung Erziehungsberatung für Kinder, Jugendliche und Eltern anbietet, die Anzahl an generierter Reichweite, die sich in der Anzahl an Abonnent:innen und Follower:innen messen lässt, über eine erhöhte Social-Media-Aktivität deutlich erhöhen (vgl. Gravelmann 2024a, S. 112). Eichenberg et al. (2016) geben in dieser Hinsicht jedoch zu bedenken, dass Präsenzen in sozialen Medien sich auf Informationsquellen und Hinweise auf Kontaktangebote beschränken sollten, um den Schutz von Ratsuchenden in datenrechtlicher Hinsicht gewährleisten zu können, „gerade weil insbesondere jüngere Patienten immer mehr dazu neigen, ihre Leidens- aber auch Genesungsgeschichte öffentlich zu machen“ (S. 44). Überdies hat eine Befragung von Ratsuchenden aus dem Jahr 2020 der Bundeskonferenz für Erziehungsberatung ergeben, dass ein Großteil der Jugendlichen über Suchmaschinen auf das Angebot der bke-Onlineberatung aufmerksam geworden ist (vgl. Hoghe et al. 2020, S. 28). Eine in der Befragung zu Tage geförderte Erkenntnis lautete, 87 Prozent der Jugendlichen gaben die Empfehlung, es sei sinnvoll über Fachkräfte, die ihnen in Präsenz begegnen, auf mögliche Onlineberatungsangebote aufmerksam gemacht zu werden (vgl. ebd.). Hoghe et al. (2020) kommen daher zu dem Schluss, dass der Informationsweitergabe in der präsenten Praxis im Moment ein noch zu niedriger Stellenwert beigemessen würde, da „F[ace]-2-F[ace]-Information durch geeignete Multiplikator:innen wie z. B. Fachkräfte aus dem schulischen Kontext [...] eine sehr geeignete Möglichkeit darstellen [könnte], um Onlineberatungsangebote Sozialer Dienste bekannt zu machen“ (S. 28). Sozialarbeitende, die nicht im Onlinebereich tätig sind, scheinen daher Verbesserungspotenziale hinsichtlich der Zugangswege zu Onlineberatung zu bergen, von denen gerade Kinder und Jugendliche profitieren können.

### 3 Grundlagen Künstlicher Intelligenz

Künstliche Intelligenz (KI) als Technologie ist seit einigen Jahren in aller Munde. Deren Wichtigkeit und möglichen Einfluss auf die menschliche Geschichte betonte Google-Geschäftsführer Sundar Pichai einst, indem er der Erforschung von KI eine größere Bedeutung zuordnete als der Entdeckung des Feuers oder der Elektrizität (vgl. Clifford 2018). Zugleich kann vermerkt werden, dass seit einigen Jahren ein immenser Hype um KI entbrannt ist und falsche Versprechen in diesem Feld keine Seltenheit sind. So konnten Studien zeigen, dass viele Startups, die sich mit KI-Entwicklung brüsteten, um vom momentan ausgebrochenen KI-Hype profitieren zu können, diese jedoch gar nicht benutzten oder an ihr arbeiteten (vgl. Olson 2019). Trotz der mittlerweile bestehenden Omnipräsenz von KI im öffentlichen Diskurs scheint vielen Menschen nicht in Gänze klar zu sein, was genau sich hinter dem Begriff verbirgt und welche Technologien dieser Sammelbegriff umfasst. Dieses Unwissen mag zu großen Teilen zur Mystifizierung von KI beitragen. Diesen Umstand beschrieb Rodney Brooks, emeritierter Professor der Robotik am Massachusetts Institute of Technology und Gründungsmitglied der Nonprofit-Organisation Association for the Advancement of Artificial Intelligence, die sich mit der Erforschung und Anwendung von Künstlicher Intelligenz beschäftigt, treffend einmal mit folgenden Worten: „Every time we figure out a piece of it, it stops being magical; we say, ‚Oh, that's just a computation‘“ (Kahn 2002).

#### 3.1 Geschichte

Erste Überlegungen zu Künstlicher Intelligenz und Träume von künstlichen Menschen beziehungsweise intelligenten Maschinen reichen bis weit in die Antike zurück und haben die Forschung und Entwicklung, wie sich zeigen wird, bisweilen maßgeblich beeinflusst (vgl. Sheikh et al. 2023, S. 20 f.). So lassen sich im Verlauf der Geschichte allerlei Zeugnisse in Form von Aufzeichnungen, Niederschriften, Bildnissen, Skulpturen und Malereien finden, die das Streben des Menschen nach autonom und intelligent agierenden Maschinen und Geschöpfen bezeugen – Wesen, die durch Menschen- oder Götterhand erschaffen wurden und denen sich eine Art von künstlicher Intelligenz unterstellen ließe (vgl. Nilsson 2010, S. 3). Der Mensch als Schöpfer von intelligenten Maschinen schien die Gedanken der Menschen also schon umzutreiben, lange bevor es eine Möglichkeit zur technologischen Umsetzung gegeben hätte: „AI is one of humanity's oldest dreams“ (Wooldridge 2023).

Vor allem im antiken Griechenland lassen sich einige dieser Ideen wiederfinden und es scheint, als seien es vor allem die griechischen Götter gewesen, die einen unbändigen Drang danach verspürten, künstliches Leben zu erzeugen (vgl. McCorduck 2004, S. 4). Eine der frühesten

Aufzeichnungen über das künstliche Einhauchen von Leben geht auf die Sage von Pygmalion zurück, der als Bildhauer die Skulptur einer Frau erschuf, in die er sich verliebte, woraufhin Göttin Aphrodite erbarmte sich seiner und erweckte sie daraufhin zum Leben (vgl. O'Regan 2021, S. 295). Homer wusste von Hephaistos, Gott des Feuers und der Schmiedekunst, zu erzählen, der in seiner unterirdischen Schmiede goldene Helferinnen beschäftigte, die intelligent, stark und der Sprache mächtig waren und denen ihr Geschick durch Götterhand gegeben war (vgl. Sheikh et al. 2023, S. 22). Hephaistos war es auch, der Pandora aus Lehm erschuf, welche später mit der namensgebenden Büchse ausgestattet, alles Schlechte auf die Welt entlassen sollte (vgl. McCorduck 2004, S. 5). Vermutlich augenzwinkernd geben Sheikh et al. (2023) auf vorangegangene Dichtungen zu Protokoll, dass im antiken Griechenland schon kontemporäre Wunderdinge der Technik, namentlich Killer- und Sex-Roboter sowie mechanische Assistenz, erdacht wurden (vgl. S. 22). Aber nicht nur die sagemuwobenen Götter, sondern auch die Gelehrten des antiken Griechenlands beschäftigten sich mit intelligenten Maschinen. Als in dieser Hinsicht bedeutsam kann vor allem Heron, in der Bibliothek von Alexandria ansässiger Lehrer und Forscher, hervorgehoben werden, der zahlreiche sogenannte ‚Automata‘ entwarf (vgl. Strick 2013). Je nach Definition versteht man unter ‚Automata‘ Maschinen, deren Wortbedeutung „*self-moving* or *self-acting*“ (Geiszler 2023, S. 25) inne liegt. Zu nennen wäre in dieser Hinsicht besonders das aus Herons Feder stammende mechanische Theater, dessen Figuren sich, angetrieben durch Dampf- und Wasserdruck, selbstständig bewegen konnten und Zuschauende in Staunen versetzten (vgl. Strick 2013).

Auch anderorts lassen sich Gedanken zu künstlichen Intelligenzen nachvollziehen. So wurde in der jüdischen Tradition überliefert, Judah Löw, seines Zeichens Rabbi und Philosoph des 16. Jahrhunderts, habe den Golem, einen Diener aus Lehm, in Prag erschaffen (vgl. McCorduck 2004, S. 526). Im arabischen Raum kursieren Erzählungen von Dschinn, Wesen, die Wünsche erfüllen können; in Rom hergestellte gefährliche mechanisierte Wachen beschützen buddhistische Reliquien; die nordische Mythologie weiß vom Riesen Hrungnir zu berichten, der ein Herz aus Stein besessen haben soll und vom Donnergott Thor gerichtet wurde; im antiken China wird die Geschichte vom Handwerker Yan Shi überliefert, der einen ‚Automaton‘ bestehend aus Holz und Leder gebaut haben soll; in der estländischen Mythologie wird die Legende vom Kratt erzählt, einer magischen Kreatur, die sich aus Heu und Haushaltsgegenständen konstituiert und Besitzenden zur Hand geht, ihnen jedoch auch gefährlich werden kann, sofern sie nicht ausreichend beschäftigt wird (vgl. Sheikh et al. 2023, S. 22). Letztgenanntes Beispiel besitzt sogar einen direkten Bezug zur Gegenwart: Das sogenannten ‚Kratt Law‘ in Estland regelt Haftung und



Verantwortlichkeit in Bezug auf den Gebrauch von Algorithmen (vgl. Kap. 3.2.1) und steht damit in direkter Verbindung zur heutzutage gebräuchlichen Künstlichen Intelligenz (vgl. e-estonia 2020).

Auch im weiteren Verlauf der Geschichte beflügelten Ideen intelligenter Maschinen und künstlich erschaffener Geschöpfe immer wieder die Gedanken von Literat:innen. Prominente Beispiele solcher Werke des 19. Jahrhunderts sind beispielsweise Goethes „Zauberlehrling“, in der ein zum Leben erweckter Besen außer Kontrolle gerät, Goethes „Faust (2. Teil)“, in dessen Verlauf ein künstlicher Mensch, der Homunkulus, kreierte wird, und Mary Shelleys „Frankenstein“, in dem eine zum Leben erweckte Kreatur sich letztendlich gegen seinen Erschaffer wendet (vgl. Buchanan 2005, S. 53; vgl. Ekmekci & Arda 2020, S. 2; vgl. Sheikh et al. 2023, S. 23 f.).

Im 20. Jahrhundert gerieten Roboter, also menschenähnliche Maschinen, in den Fokus der Öffentlichkeit. Das Wort ‚Roboter‘ ist eine originär aus dem Theaterstück Rossum’s Universal Robots (R.U.R.) stammende Wortschöpfung Karel Čapeks, die auf das tschechische Wort ‚robota‘ zurückzuführen ist, was sinnhaft übersetzt ‚zur Arbeit gezwungen‘ bedeutet (vgl. Jordan 2019). Darauf aufbauend machte Isaac Asimov, der einer der einflussreichsten Science-Fiction-Autoren seiner Zeit war und im Zuge seines Schaffens etliche Robotergeschichten verfasste, mit seinen drei Gesetzen der Robotik auf sich aufmerksam. Diese besagen, dass ein Roboter nie einem Menschen schaden dürfe, er dem Menschen gehorchen müsse und er sein eigenes Überleben sicherzustellen habe, solange es nicht in Konflikt mit dem ersten Gesetz gerate (vgl. Boden 2018, S. 171). Asimovs drei Gesetze der Robotik waren auch im wissenschaftlichen Diskurs um Roboter fester Bestandteil und übten Einfluss auf diesen aus. (vgl. Decker 2019, S. 345 ff.). Isaac Asimov war es auch, der das geflügelte Wort „Frankenstein Komplex“ (The Royal Society 2018, S. 8) ersann, um eine Trope zu beschreiben, die, seitdem Künstliche Intelligenzen oder intelligente Maschinen das erste Mal die Bühne in Fiktion betreten haben, beobachtet werden kann: Das Aufbegehren der unnatürlichen beziehungsweise künstlichen Kreationen gegen ihre menschlichen Erschaffer (vgl. ebd.). Neben dem großen Aufkommen dystopischer literarischer Werke soll an dieser Stelle nicht unerwähnt bleiben, dass utopische Narrative auch existieren. Laut Hermann (2018) fungieren KI-Tropen in Fiktion als eine Art Lupe, die die Menschheit mit ihren Hoffnungen, Ängsten und (sozio-politischen) Themen in philosophischer, kultureller und psychologischer Hinsicht in den Fokus rückt und so einen Raum eröffnet, in dem bisher noch nicht formuliert gebliebene und unausgesprochene Probleme ausgehandelt werden können (vgl. S. 224). Auch traten KI-Berater und -Therapeuten in Science-Fiction-Literatur recht früh in Erscheinung wie zum Beispiel in Philipp K. Dicks 1964 erschienener Kurzgeschichte „Ach, als Blobbel hat man’s schwer!“. In dieser therapiert ein intelligenter Robotertherapeut, der per Münzeinwurf

bezahlt wird, den Protagonisten der Geschichte (vgl. Dick 2008, S. 590). Dick kam damit in der Literatur dem echten Chatbot namens ELIZA zwei Jahre zuvor, der von Joseph Weizenbaum entwickelt wurde, an die Grundhaltungen klientenzentrierter Gesprächstherapie nach Rogers angelehnt war und im weiteren Verlauf des Kapitels noch näher erläutert werden wird (vgl. Weizenbaum 1966, S. 36).

Nachdem vorab insbesondere Mythen, Überlieferungen und Fiktion in Bezug auf Künstliche Intelligenz im Fokus standen, sollen im Folgenden Meilensteine in der Entwicklung Künstlicher Intelligenz beleuchtet werden, die ganz konkret zum Vorankommen Künstlicher Intelligenz beigetragen haben. René Descartes, der im 17. Jahrhundert lebte, war Vertreter des Rationalismus, dem logisches und deduktives Denken bei der Erlangung von Wissen inhärent ist. Seine philosophischen Prinzipien und Methoden haben indirekt dazu beigetragen, Grundlagen zu schaffen, die später in der KI-Entwicklung Anwendung finden sollten. Zu nennen sei hierbei vor allem der Leib-Seele-Dualismus, der postuliert, Körper (*res extensa*) und Geist (*res cogitas*) seien voneinander getrennt (vgl. O'Regan 2021, S. 298). Dies lässt sich analog auf den Aufbau eines Computers übertragen, dessen Körper die Hardware und dessen Geist die Software darstellt. Descartes zufolge erlangt der Geist Wissen vor allem über Beobachtung, sowie mittels logischen und deduktiven Denkens (vgl. ebd.). Wie sich im weiteren Verlauf der Arbeit zeigen lassen wird, sollte dieser Ansatz von Wissensverarbeitung vor allem im Bereich symbolischer KI und in Expertensystemen zum Einsatz kommen (vgl. Kap. 3.2.1). Descartes vertrat überdies eine mechanistische Weltansicht, in der sich das Verhalten von Tieren und Menschen in einzelne mechanistische Prozesse differenzieren und hierüber erklären ließe (vgl. O'Regan 2021, S. 298 f.).

Ebenfalls großen Einfluss auf die Entwicklung Künstlicher Intelligenz hatten die Arbeiten von Gottfried Wilhelm Leibniz, der danach strebte, Wahrheiten, die in der Arithmetik oder Algebra geschlossen werden konnten, auch in anderen Bereichen schlussfolgern zu können (vgl. Teich 2020, S. 276). Durch die Arbeiten Blaise Pascals inspiriert, der zuvor mechanische Rechenmaschinen entworfen hatte, vollendete Leibniz im Jahre 1694 die sogenannte Leibnizsche Rechenmaschine, die als erste Maschine dazu in der Lage war die Grundrechenarten durchzuführen (vgl. O'Regan 2021, S. 36). Auch arbeitete er Zeit seines Lebens daran, eine universell gültige Sprache zu entwickeln, mittels derer sich das gesamte menschliche Wissen formulieren ließe (vgl. Nilsson 2010, S. 11). Dem zugrunde lag die Ansicht, dass Wissen sich aus kleineren, primitiveren Bausteinen konstituiert, in etwa so wie Wörter sich aus einzelnen Buchstaben zusammensetzen (vgl. ebd.). Damit legte er den Grundstein von KI-Logik, die unter Verwendung allgemeiner Symbole und mechanischer Abläufe funktioniert (vgl. Sheikh et al. 2023, S. 23). Seinen

Einfluss auf Künstliche Intelligenz unterstreicht auch das auf Leibniz zurückzuführende Binärsystem, auf dessen Grundlage heutige Computer rechnen und operieren (vgl. O'Regan 2021, S. 38).

Im 18. Jahrhundert machte eine Maschine namens ‚mechanischer Türke‘ (auch ‚Schachtürke‘ genannt) auf sich aufmerksam, die Wolfgang von Kempelen für die Monarchin Maria Theresia anfertigen ließ (vgl. Geisler 2023, S. 25). Diese Maschine wurde Gegenspieler:innen und Zusehenden als intelligente Maschine verkauft, die selbstgesteuert gegen sie antrat (vgl. Buchanan 2005, S. 54). Was diese nicht wussten: In ihrem Innern nahm ein Mensch Platz, der über einen Mechanismus den Schachtürken steuern konnte (vgl. Sheikh et al. 2023, S. 23). Dem Vernehmen nach spielte der Schachtürke viele Partien auf europäischem Boden und soll sogar prominente Gegner wie Napoleon und Benjamin Franklin bezwungen haben (vgl. ebd.). Wenngleich diese Maschine als Schummelei enttarnt werden sollte, die weder intelligent noch selbstgesteuert war, stellte sie doch einen direkten Vorläufer der heute geläufigen Schachcomputer dar, die diese beiden Kriterien erfüllen und im weiteren Verlauf des Kapitels noch behandelt werden sollen (vgl. Buchanan 2005, S. 54). Eine zumindest namentliche Hommage an von Kempelens mechanischen Türken kann in Amazons gleichnamigem Marktplatz erkannt werden, auf dem Dienstleistungen ausgelagert werden können (vgl. Sheikh et al. 2023, S. 23).

In den 1840er Jahren sollte die Zusammenarbeit der heute als Pioniere der Computergeschichte anerkannten Charles Babbage und Ada Lovelace große Fußabdrücke hinsichtlich der Entwicklung Künstlicher Intelligenz hinterlassen. Der Mathematiker, Philosoph und Maschinenbauer Babbage arbeitete über 10 Jahre am Konzept der sogenannten ‚Analytical Engine‘ (vgl. Jaeger 2023, S. 86). Diese sollte über Lochkarten, die das 0-oder-1-Prinzip anwendeten, gesteuert werden und war in der Lage, Zahlen mit über 200 Stellen zu verarbeiten (vgl. ebd.). Während Charles Babbage vor allem den Nutzen der Maschine darin sah, große Zahlen fehlerfrei berechnen zu können, bewies seine Forschungspartnerin geradezu bahnbrechend visionären Weitblick, indem sie konstatierte, die analytische Maschine sei, so sie denn eines Tages konstruiert werden würde, auch in der Lage sequenzielle Operationen hinsichtlich Musik, Bildern und Buchstaben durchzuführen – im Prinzip also das, was moderne generative Künstliche Intelligenz imstande zu leisten ist (vgl. Ekmekci & Arda 2020, S. 3). Die grundsätzliche Fähigkeit der Maschine, allgemeine Informationen zu verarbeiten, solange diese sich mathematisch übersetzen ließen, wie Lovelace es herausarbeitete, bildet das Herzstück heutiger Informatik (vgl. Jaeger 2023, S. 88). Insbesondere folgende zwei Sätze stehen dabei im Fokus:

1. die Transformation komplexer Informationen wie Sprache oder Musik in Mathematik,
2. der Einsatz von Algorithmen, also endlich vielen, eindeutigen Verarbeitungsinstruktionen für die eingegebene Information. (ebd., S. 88)

In Lovelaces Niederschriften finden sich Aufzeichnungen über einen Algorithmus, mit dem die analytische Maschine in der Lage gewesen wäre, die Bernoulli-Zahlenfolge automatisch zu berechnen. Damit verewigte sie sich aus historischer Perspektive als erste Programmiererin und Erfinderin von Software im Allgemeinen (vgl. Jaeger 2023, S. 88).

Der britische Mathematiker Alan Turing war sich über Lovelaces und Babbages Arbeiten bewusst und bewunderte diese (vgl. McCorduck 2004, S. 63). In seinem Aufsatz von 1936 ‚On Computable Numbers, With an Application to the Entscheidungsproblem‘ knüpfte er mit seiner Forschung im Prinzip nahtlos an die Ausführungen Lovelaces an (vgl. Turing 1936). Diese Schrift legte einen weiteren Grundstein für moderne Computer und künstliche Intelligenz, weshalb auch Turing als einer der Vorväter von Computern und Künstlicher Intelligenz gilt. Hier präzisierte er das Konzept des Algorithmus und skizzierte, durch Informationsverarbeitung und Befolgung grundsätzlicher Berechnungen könne selbst eine recht simple Maschine, die später nach ihm benannte Turingmaschine und Archetypus des modernen Computers, komplexe Berechnungen vollführen (vgl. Muhle 2023, S. 16); Lediglich vorausgesetzt sei, „dass die Rechenregel bekannt ist“ (ebd.). In Bezug auf Künstliche Intelligenz ist die zentrale Erkenntnis also, dass eine derartige Maschine in der Lage wäre, sämtliche Handlungen zu vollführen, solange es möglich ist, dies mittels einer Formel darzustellen und berechnen zu lassen. Turing galt als mathematisches Ausnahmetalent und bewerkstelligte es im zweiten Weltkrieg mit Kolleg:innen auch, die Enigma, das deutsche Nachrichtenverschlüsselungssystem, zu dechiffrieren (vgl. O’Regan, S. 62). Als aus heutiger Sicht bahnbrechend kann auch Turings unveröffentlichter Aufsatz ‚Intelligent Machinery‘ von 1948 erachtet werden. Dort beschäftigte er sich mit der Frage, wie Maschinen durch Adaptation lernen und intelligentes Verhalten zeigen könnten und stellte eine Analogie zum menschlichen Gehirn her (vgl. Turing 1948). Damit leistete Turing einen wegweisenden Vorgriff auf die heutzutage als maschinell lernende bekannte Künstliche Intelligenz (vgl. Kap. 3.2.2). Turing veröffentlichte auch ein weiteres Paper namens ‚Computing Machinery and Intelligence‘, das für großes Aufsehen in der Wissenschaftsgemeinschaft sorgte (vgl. Turing 1950). Ihm wurde nachgesagt, er sei der Ansicht gewesen, dass die Möglichkeiten des durch Computer Berechenbaren einen viel größeren Horizont umfassen müssten als all jenes, was durch dezidierte Anweisungen erfasst werden konnte (vgl. Nilsson 2010, S. 37). Demzufolge könnten Maschinen mit ausreichender Komplexität die Fähigkeit besitzen, Verhaltensweisen zu entwickeln, die diese nicht durch explizite Programmierung erworben hatten, womit sie sogar befähigt sein könnten, menschliche Kreativität und Originalität nachzuahmen (vgl. ebd.). Es verwundert daher nicht, dass er den Vorschlag unterbreitete, „to consider the question, ‚Can machines think?‘“ (Turing 1950, S. 433). Um das herauszufinden, schlug er das Imitation Game vor, das später nach ihm benannt als

‚Turingtest‘ bekannt werden sollte (vgl. Teich 2020, S. 277). In diesem Spiel kommuniziert ein Mensch mittels Tastatur mit einem weiteren Menschen und einer Maschine. Aus Sicht Turings würde eine Maschine dann als intelligent gelten, sofern der Mensch nicht unterscheiden könnte, wer von seinen beiden Gesprächspartnern die Maschine sei (vgl. Turing 1950, S. 433 f.). Zudem greift Turing über etliche Seiten hinweg Argumente auf, die eine mögliche Maschinenintelligenz in Abrede stellen könnten und versucht diese theoretisch zu widerlegen (vgl. ebd., S. 443 ff.). Seinen Aufsatz schließt er mit der Hoffnung, „machines will eventually compete with men in all purely intellectual fields“ (ebd., S. 460). Um dies erreichen zu können, böte sich an, einen Computer so zu programmieren, dass er dem Bewusstsein eines Kindes ähnelte, damit ihm dann alle weiteren Fähigkeiten beigebracht werden könnten (vgl. ebd., S. 456).

Ein weiterer bedeutsamer theoretischer Beitrag hinsichtlich der Entwicklung von Künstlicher Intelligenz wurde vom Psychiater und Neurologen Warren McCulloch und dem Mathematiker Walter Pitts geleistet. In ihrem 1943 erschienenen Beitrag ‚A Logical Calculus of the Ideas Immanent in Nervous Activity‘ stellten sie ein theoretisches Modell für neuronale Netzwerke vor (vgl. McCulloch & Pitts 1943). Ihr Modell, das auf der binären Logik (an oder aus, 0 oder 1) fußte, zeigte, wie einfache neuronale Elemente miteinander verbunden werden können, um komplexe Funktionen auszuführen (vgl. Nilsson 2010, S. 17). Diese Arbeit sollte sich im weiteren Verlauf der Geschichte als wegweisend für die Entwicklung neuronaler Netzwerke (vgl. Kap. 3.2.2.1) im Forschungsfeld der Künstlichen Intelligenz erweisen, da sie ein grundlegendes Verständnis dafür lieferte, wie biologische Neuronen Informationen verarbeiten und wie ähnliche Prinzipien auf Formen der Künstlichen Intelligenz, vorrangig dem maschinellen Lernen und dem sogenannten Deep Learning, angewendet werden können (vgl. Sheikh et al. 2023, S. 26). In Konsequenz bedeutete es, „one and the same theoretical approach – namely, Turing computation – could be applied to human and machine intelligence“ (Boden 2018, S. 30).

Die Geburtsstunde Künstlicher Intelligenz als akademisches Forschungsfeld wird gemeinhin auf den Sommer 1956 datiert. Ausschlaggebend hierfür war das ‚Dartmouth Summer Research Project on Artificial Intelligence‘, das von John McCarthy, Marvin Minsky, Nathaniel Rochester und Claude Shannon ein Jahr zuvor mittels eines Förderantrags initiiert wurde und 1956 am Dartmouth College in Hanover, New Hampshire, über zwei Monate hinweg abgehalten wurde (vgl. McCarthy et al. 2006). Im Zuge dieses Workshops wurde der Terminus ‚Artificial Intelligence‘ zum ersten Mal offiziell benutzt und geprägt (vgl. Haenlein & Kaplan 2019, S. 3). Wie das Kapitel gezeigt hat, haben aber bereits Jahre zuvor Diskussionen über maschinelle Intelligenz stattgefunden (vgl. Muhle 2023, S. 13). Teilnehmende des Workshops würden über den Zeitraum der nächsten beiden Jahrzehnte die KI-Forschung maßgeblich bestimmen und auf diesem Feld

entscheidende Fortschritte erzielen (vgl. Russell & Norvig 2010, S. 18). Ausgerufenes Ziel war es, Forschende aus unterschiedlichen Forschungsfeldern zusammenzubringen, um Grundlagenforschung zu Maschinen zu betreiben, die in Konsequenz in der Lage sein sollten, menschliche Intelligenz imitieren zu können (vgl. Kaplan 2022, S. 13). Dem zugrunde lag die Annahme, intelligentes menschliches Verhalten würde aus kleinteiligen Prozessen bestehen, die sich formalisieren ließen, sodass diese für eine Maschine reproduzierbar gemacht werden könnten (vgl. Dick 2019, S. 2). Das Vorhaben für das ‚Dartmouth Summer Research Project on Artificial Intelligence‘ wurde 1955 wie folgt formuliert:

The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer. (McCarthy et al. 2006, S. 12).

Trotz der Bedeutung des ‚Dartmouth Summer Research Project on Artificial Intelligence‘ für den weiteren Verlauf der Geschichte Künstlicher Intelligenz gab es in dieser Hinsicht enttäuschter John McCarthy im Nachgang des Projektes einmal preis, der angedachte Workshop sei von Teilnehmenden zeitweise nur mäßig und unregelmäßig besucht worden, weshalb in dieser Form ein ursprünglich angedachter dauerhafter und ineinandergreifender multidisziplinärer Arbeitsprozess und Wissensaustausch kaum stattfinden konnte (vgl. McCorduck 2004, S. 114).

In der anfänglichen Forschung von KI ging es vordergründig darum, menschliche Intelligenz so gut es geht zu verstehen, um diese anschließend simulieren zu können (vgl. Dick 2019, S. 2). Heutzutage ist eine Abkehr davon erkennbar: „Most researchers want to design automated systems that perform well in complex problem domains by any means, rather than by human-like means“ (ebd.).

Hinsichtlich der KI-Entwicklung ist häufiger von Jahreszeiten oder zwei Wellen die Rede, um die Phasen des Auftriebs (Frühling) oder der Regression (Winter) zu beschreiben (vgl. Abb. 4).

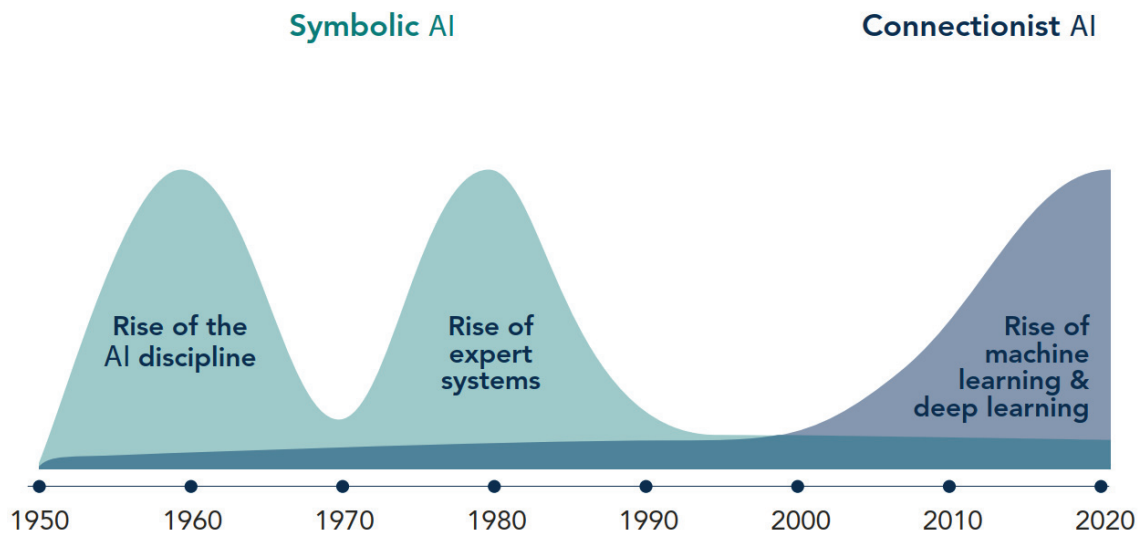


Abbildung 4: KI-Wellen beziehungsweise Jahreszeiten (Sheikh et al. 2023, S. 36)

Das ‚Dartmouth Summer Research Project on Artificial Intelligence‘ kann als Beginn eines solchen ersten Frühlings erachtet werden, der von Optimismus hinsichtlich der Entwicklung Künstlicher Intelligenzen geprägt war und in dem große Erwartungen an mögliche Fortschritte geschürt wurden. Eine hoffnungsvolle Künstliche Intelligenz stellten Herbert Simon und Allen Newell, die auch am ‚Dartmouth Summer Research Project on Artificial Intelligence‘ teilnahmen, mit dem ‚Logic Theorist‘ vor, einem symbolverarbeitenden Programm, das in der Lage war, mathematische Theoreme zu beweisen (vgl. McCorduck 2004, S. 167). Es funktionierte so gut, dass es sogar eines von Whitehead und Russells aufgestellten Theorems zu verbessern wusste (vgl. Sheikh et al. 2023, S. 29). 1959 präsentierten Newell und Simon den ‚General Problem Solver‘, der eine Art Weiterentwicklung des ‚Logic Theorist‘ darstellte und sich aus ihrer Sicht menschlicher Entscheidungsfindung bediente (vgl. Nilsson 2010, S. 87). Er sollte in der Lage sein, eine große Palette an Problemen lösen zu können, solange es sich in kleinere Ziele, Operatoren und zu vollführende Handlungen herunterbrechen ließe, die zuvor per Programmierung definiert sein mussten (vgl. Sheikh et al. 2023, S. 29). In der Realität bedeutete dies, dass er nur einige gut zu beschreibende Probleme, Puzzle, Rätsel und Logikfragen lösen konnte – dies jedoch sehr gut und zuverlässig (vgl. Taulli 2022, S. 9). Beide Programme, die sich symbolische KI (vgl. Kap. 3.2.1) zunutze machten, gelten heute als Wegbereiter von Expertensystemen, auch wenn dieser Begriff erst im weiteren Verlauf der Geschichte erdacht und gebräuchlich werden sollte. Durch die jüngsten Erfolge auf dem Feld der Künstlichen Intelligenz etwas zu euphorisiert, verkündete Herbert Simon im Jahre 1965, „Machines will be capable, within twenty years, of doing any work a man can do“ (Simon 1965, zit. n. Wennker 2020, S. 3). 60 Jahre später muss konstatiert werden, dass der von ihm skizzierte Fall bisher noch nicht eingetreten ist und eine allgemeine Künstliche Intelligenz (vgl. Kap. 3.2) noch nicht am Horizont abzusehen ist. Doch auch eine andere Form der

KI erblickte in dieser Zeit das Licht der Welt, – die auf Grundlage eines neuronalen Netzes, wie McCulloch und Pitts es 1943 beschrieben hatten, basierende KI. An dieser Technik, auf der generative Künstliche Intelligenzen beruhen (vgl. Kap. 3.2.2.1), die gegenwärtig für gehöriges Aufsehen sorgen, wurde durch Frank Rosenblatt gearbeitet (vgl. Rosenblatt 1958). Das auf den Namen Perzeptron hörende Programm musste trainiert werden, sodass es Verbindungen zwischen den künstlichen Neuronen herstellen konnte (vgl. Lenzen 2020, S. 22). Danach war es in Versuchen in der Lage, mithilfe von Fotozellen simple Formen identifizieren zu können (vgl. Taulli 2022, S. 11). Aufgrund der damaligen Technikbeschränkungen konnte das Perzeptron allerdings nicht alle Funktionen abdecken, die symbolische KI zu der Zeit bewältigen konnte (vgl. Lenzen 2020, S. 22).

Ein weiteres, besonders bezogen auf das Thema dieser Arbeit, relevantes Ereignis war der Chatbot ELIZA<sup>3</sup>. Dieser im Jahr 1966 von Joseph Weizenbaum präsentierte Chatbot war zwar nicht der erste seiner Art, jedoch war er der erste Chatbot, der an Carl Rogers Grundhaltungen personenzentrierter Gesprächsführung angelehnt war und eine Art von Psychotherapeutengespräch führen konnte (vgl. ebd.). Wie Weizenbaum (1966) untenstehend näher erläutert, war die Funktionsweise ELIZAs aus einer Haltung des Nichtwissens essenziell:

This mode of conversation was chosen because the psychiatric interview is one of the few examples of categorized dyadic natural language communication in which one of the participating pair is free to assume the pose of knowing almost nothing of the real world. If, for example, one were to tell a psychiatrist ‚I went for a long boat ride‘ and he responded ‚tell me about boats‘, one would not assume that he knew nothing about boats, but that he had some purpose in so directing the subsequent conversation. It is important to note that this assumption is one made by the speaker. Whether it is realistic or not is an altogether separate question. In any case, it has a crucial psychological utility in that it serves the speaker to maintain his sense of being heard and understood. (S. 42)

Ursprüngliches Ziel Weizenbaums war es, die Kommunikation zwischen Menschen und Computer mittels natürlicher Sprache zu erforschen (vgl. Weizenbaum 1966, S. 36). ELIZA erzeugte allerdings eine ziemlich große Resonanz und Weizenbaum verfolgte argwöhnisch, wie auskunftsfreudig Menschen dem Chatbot einen Blick in ihr Innerstes offerierten, wohlwissend, dass dieser kein echtes Verständnis für seine Gesprächspartner haben könnte (vgl. Teich 2020, S. 277). ELIZA generierte Antworten lediglich auf Grundlage vordefinierter Regeln, die vor allem über das Ausfindigmachen von Schlüsselwörtern funktionierten (vgl. Weizenbaum 1966, S. 37). Weizenbaum

---

<sup>3</sup> ELIZA kann heute noch unter folgender Webadresse ausprobiert werden: <https://www.masswerk.at/elizabot/>



empfand seine Arbeit als missverstanden und entwickelte sich im Laufe der Jahre zu einem Gesellschaftskritiker, dem blindes Vertrauen in die Informatik sowie zu euphorische Erwartungshaltungen an Computertechnik widerstrebte (vgl. Muhle 2023, S. 23). Namensgebend war er auch für das Weizenbaum-Institut, das „dazu beitragen [will], dass Digitalisierung und vernetzte Gesellschaft nicht nur besser verstanden, sondern auch nachhaltig, selbstbestimmt und verantwortungsvoll gestaltet werden“ (Weizenbaum Institut o. D.), weshalb dort zu diesen Themen unabhängige und interdisziplinäre Forschung betrieben wird (vgl. ebd.).

Trotz der für diese Zeit spannenden Entwicklungen und Innovationen im Bereich der Künstlichen Intelligenz war man insgesamt hinter den allzu euphorischen Erwartungen zurückgeblieben. Es kam infolgedessen zum ersten KI-Winter (vgl. Taulli 2022, S. 12). Neben den offensichtlichen Hardware-Limitierungen früherer Computer, die hochwertigere Künstliche Intelligenzen nicht realisieren konnten, waren hierfür insbesondere zwei Dinge ausschlaggebend: Zwei Berichte, die damals den Stand gegenwärtiger KI-Entwicklung untersuchten und den Regierungen Großbritanniens und der USA vorgelegt wurden, konkludierten, intelligente Maschinen seien Zukunftsmusik und gemachte Versprechungen seien nicht umzusetzen gewesen (vgl. Sheikh et al. 2023, S. 30). Daraufhin kam es Anfang der 1970er-Jahre zu Kürzungen der zuvor bereitgestellten Gelder und einem Rückgang betriebener Forschung (vgl. Haenlein & Kaplan 2019, S. 3).

Knapp zehn Jahre später erlebte Künstliche Intelligenz in den 1980er-Jahren einen neuen Aufschwung und die zweite Welle beziehungsweise der zweite Frühling brach an. Auslöser war eine langfristige und großangelegte Investition Japans mittels des 5th Gen. Computer Systems Project in die eigene Wirtschaft, die damals weltweit geachtet und respektiert war (vgl. Taulli 2022, S. 15). Dieser Umstand zwang westliche Länder wie Großbritannien und die USA und dort ansässige Unternehmen dazu, mit eigenen Projekten und Investitionen nachzuziehen, weshalb wieder Gelder in die Forschung und Entwicklung von KI flossen (vgl. Sheikh et al. 2023, S. 32 f.). In dieser Zeit deuteten auch einige auf symbolischer KI basierende Expertensysteme auf ihr vielversprechendes Potenzial hin und konnten kommerzielle Erfolge verbuchen. Zu nennen wären hier beispielsweise ‚MYCIN‘ und ‚R1/XCON‘. MYCIN, das an der Universität in Stanford entwickelt wurde, konnte Ärztinnen und Ärzten bei der Diagnose von Infektionskrankheiten behilflich sein und lieferte im Labor zufriedenstellende Ergebnisse, die bei der Behandlung zu Rate gezogen werden konnten (vgl. Ekmekci & Arda 2020, S. 11). Rechtliche Bedenken verhinderten jedoch einen Einsatz in der medizinischen Praxis. Das Expertensystem ‚R1/XCON‘ hingegen war in der Lage, kostspielige und fehleranfällige Bestellungen von Computerhardware in der Praxis deutlich zu reduzieren, was dessen Entwickler DEC, Hersteller von Computersystemen, Einsparungen in zweistelliger Millionenhöhe bescherte (vgl. Teich 2020, S. 281). Beide Systeme waren erfolgreich,

indem sie menschliches Expertenwissen modellierten und Nutzenden bei der Entscheidungsfindung behilflich waren (vgl. Wennker 2020, S. 4).

Doch auch im Bereich künstlicher neuronaler Netze konnten neue Erfolge vorgewiesen werden. Das 1986 vorgestellte Programm ‚NETTalk‘ von Sejnowski & Rosenberg beruhte auf genau einem solchen und trat damit in die Fußstapfen, die einst das Perzeptron hinterlassen hatte. Indem eine begrenzte Anzahl geschriebener Wörter zusammen mit ihrer Aussprache als Phonemsequenzen eingegeben wurde, war es NETTalks künstlichem neuronalen Netz beispielsweise möglich, das korrekte Aussprechen englischer Wörter zu erlernen und dieses neu erworbene Wissen auf für NETTalk unbekannte Wörter zu übertragen (vgl. Sejnowski & Rosenberg 1986, S. 663).

Thematisierter Fortschritte und erster kommerzieller Erfolge zum Trotz kam es letztendlich doch erneut zu einem Einbruch an Investitionen, was vorrangig an den zu großen Ambitionen lag, die sowohl in Japan als auch in den westlichen Ländern nie erfüllt werden konnten – aber auch die Fehleranfälligkeit der zum Einsatz kommenden KI-Systeme, die technisch noch nicht gänzlich ausgereift waren und unter limitierter Hardware litten, spielten eine Rolle (vgl. Sheikh et al. 2023, S. 34). Der anbrechende KI-Winter traf die Branche hart und war Zeugnis dessen, wie tief das allgemeine Vertrauen in die Versprechungen Künstlicher Intelligenz gesunken war (vgl. Taulli 2022, S. 15). Die Ausmaße waren so groß, dass an der Forschung und Entwicklung von KI beteiligte Menschen es zeitweise vermieden den Begriff Künstliche Intelligenz zu verwenden und stattdessen Begrifflichkeiten wie ‚Wissensverarbeitung‘ beziehungsweise ‚Knowledge-Engineering‘ nutzten, um ihr Tätigkeitsfeld zu beschreiben (vgl. Teich 2020, S. 278). Die Befürchtung, nicht ernstgenommen zu werden oder Investoren zu verschrecken, war teilweise groß (vgl. Markoff 2005).

Ende der 1990er Jahre erlangte KI wieder an Bedeutung. Dieser Zeitpunkt markiert den allmählichen Beginn des erneuten Auftriebs und der dritten und aktuell noch anhaltenden KI-Welle, die spätestens seit den in den 2010er-Jahren gemachten Fortschritten im maschinellen Lernen deutlich an Aufwind gewann (vgl. Lenzen 2020, S. 24). Obwohl diese Ära ganz im Zeichen konnektionistischer Künstlicher Intelligenz steht (vgl. Abb. 4), also KI, die über künstliche neuronale Netze beziehungsweise über maschinelles Lernen, wie zum Beispiel Deep Learning, funktioniert (vgl. Kap. 3.2.2.1), machte zuerst der Erfolg einer symbolischen KI auf sich aufmerksam. IBMs Schachcomputer ‚Deep Blue‘ gelang es 1997 den damaligen Schachweltmeister Gari Kasparov zu besiegen (vgl. Teich 2020, S. 278). Dies gelang Deep Blue, indem er auf einen großen Wissensschatz an Positionen, Spielzügen und Ausgangslagen zurückgriff und so die besten Entscheidungen für sich treffen konnte (vgl. Ekmekci & Arda 2020, S. 14). Für das Forschungsfeld

der Künstlichen Intelligenz war das ein durchschlagender Erfolg. Zwar hatte es schon Jahrzehnte zuvor erste Schachcomputer gegeben – die auf Amateurniveau spielende Schach-KI ‚Mac HACK VI‘ war 1967 sogar zum Ehrenmitglied des US-amerikanischen Schachverbandes ernannt worden –, jedoch hatte noch nie zuvor eine Schach-KI existiert, die befähigt war, den amtierenden Weltmeister zu schlagen (vgl. Nilsson 2010, S. 194).

Ein weiterer durchschlagender Erfolg war die Künstliche Intelligenz ‚Watson‘, die als sogenannte hybride KI (vgl. Kap. 3.2.3) größtenteils auch auf Basis einer wissensbasierten und symbolischen KI (vgl. Kap 3.2.1) operierte und 2011 in der Quizshow ‚Jeopardy!‘ gegen die beiden amtierenden Meister antrat und gewann (vgl. Ertel 2020, S. 23). ‚Watson‘ beruhte auf mehreren ineinander verzahnten und im Kollektiv zusammenarbeitenden KI-Systemen, die über einen zentralen Server koordiniert waren, da sie neben dem schnellen Zugriff auf einen großen Wissensfundus zum Beispiel auch in der Lage sein musste, menschliche Sprache zu verarbeiten (vgl. ebd.). ‚Watson‘ hatte allerdings kein tiefgreifendes Verständnis von komplexer menschlicher Sprache, da sich das formelhafte Frage-Antwort-Spiel von ‚Jeopardy!‘ leicht auf Wissensdatenbanken wie Wikipedia anwenden ließ (vgl. Sheikh et al. 2023, S. 35). Zu einer scheinbar offenen Unterhaltung, zu der beispielsweise heutzutage ChatGPT fähig ist, wäre sie nicht imstande gewesen (vgl. Engelhardt 2023c, S. 17). Für diese KI-Welle typisch, machten Künstliche Intelligenzen immer schnellere Fortschritte. Ein gutes Beispiel, das diesen Umstand illustriert, war die durch die dem US-amerikanischen Verteidigungsministerium zugehörige Firma DARPA ausgerufene Herausforderung für selbstfahrende Autos, die 2004 stattfand (vgl. Taulli, S. 183). Diese sollten eine vorher festgelegte Strecke von Kalifornien nach Nevada absolvieren. Von den ursprünglich 106 teilnehmenden Teams traten letztendlich nur 15 zum Start an, wovon jedoch keines der Autos in der Lage war, das Ziel zu erreichen (vgl. Ekmekci & Arda 2020, S. 13f.). Bei einer Wiederholung nur ein Jahr später gelang es bereits fünf von 23 angetretenen Teams, den Kurs erfolgreich zu bestehen, da durch längere und bessere Trainings viele der zuvor gemachten Fehler präventiv behoben werden konnten (vgl. Buchanan 2005, S. 54).

Im Jahre 2012 ebnete die KI ‚AlexNet‘ die Siegesstraße für das maschinelle Lernverfahren Deep Learning (vgl. Krizhevsky et al. 2012). Bei der ‚ImageNet Large-Scale Visual Recognition Challenge‘, die eine jährlich stattfindende Olympiade im Bereich der Computer Vision darstellt, bei der KIs gefordert sind, Bilder korrekt zu erkennen und zu klassifizieren, gelang es ‚AlexNet‘, den ersten Platz zu erringen (vgl. Deshpande 2016). Das Ergebnis sorgte für Begeisterung und Staunen, da ‚AlexNet‘ eine sehr geringe Fehlerrate aufwies und mit großem Abstand auf den zweiten Platz siegte (vgl. ebd.). ‚AlexNet‘ bescherte Deep Learning einen großen Popularitätsschub, indem es zeigte, welche präzisen und herausragenden Ergebnisse mit einem mehrschichtigen,

tiefen, künstlichen neuronalen Netzwerk (vgl. Kap. 3.2.2.1) erzielt werden können. Nur zwei Jahre später nutzten alle topplatzierten KIs bei der ‚ImageNet Large-Scale Visual Recognition Challenge‘ Deep Learning Verfahren (vgl. Gershgorin 2017). Matthew Zeiler, der mit seiner KI die Challenge 2013 gewann, ist sich sicher: „This Imagenet 2012 event was definitely what triggered the big explosion of AI today“ (ebd.). Auf diesen Durchbrüchen baute eine weitere Deep Learning basierte KI auf: Im Jahre 2016 gelang es ‚AlphaGo‘, so der Name der Künstlichen Intelligenz, den amtierenden Champion Lee Sedol im Brettspiel Go zu schlagen (vgl. Teich 2020, S. 279). Der Erfolg war insofern bahnbrechend, da Go in Bezug auf seine Komplexität deutlich höher als Schach anzusiedeln ist, weshalb lange Zeit der Glaube bestand, eine KI würde niemals in der Lage sein, einen Menschen darin zu schlagen (vgl. Haenlein & Kaplan 2019, S. 4). Ein Jahr darauf gelang es ‚AlphaGo‘, der Nachfolger-KI von ‚AlphaGo Zero‘, nach einem dreitägigen Training gegen sich selbst (vgl. Kap. 3.2.2.2), ‚AlphaGo‘ in 100 Spielen 100-mal zu besiegen (vgl. Mockenhaupt & Schlaghauf 2024, S. 73). Hingegen nur vier Stunden Training gegen sich selbst benötigte ‚AlphaGo Zero‘, um Schach zu meistern: In 100 Spielen behielt sie 28-mal gegen die bisher beste Schach-KI ‚Stockfish‘ die Oberhand bei 72 Unentschieden (vgl. Boden 2018, S. 67). Damals wie heute nutzen KI-Entwickler Spiele und Wettbewerbe gerne, um die neue Leistungsfähigkeit ihrer KIs der Öffentlichkeit unter Beweis zu stellen.

Ein weiterer Meilenstein war 2017 die Entwicklung einer Deep-Learning-Architektur, die sich ‚Transformer‘ (vgl. Kap. 3.2.2.1) nennt und als eine der größten bisherigen Errungenschaften in der Entwicklung Künstlicher Intelligenz überhaupt gilt (vgl. Olson 2023). Vaswani et al. (2017) entwickelten den Transformer, um die Computerlinguistik (auch Natural Language Processing, kurz NLP, genannt und nicht zu verwechseln mit neurolinguistischer Programmierung) voranzutreiben und bessere Übersetzung zwischen verschiedenen Sprachen ermöglichen zu können. Ihr bescheidenes Vorhaben revolutionierte letztendlich die Art und Weise, wie Künstliche Intelligenzen natürliche Sprache verarbeiten, da ihr Transformer die Konkurrenz der damaligen Zeit sowohl im Ergebnis als auch in der Trainingseffizienz deutlich übertraf (vgl. ebd., S. 8). Viele von den generativen KI-Tools, die momentan für Begeisterung und Gesprächsstoff sorgen, basieren mindestens zu Teilen auf dieser Technologie, beispielsweise ChatGPT, die Bildgenerierungssoftware DALL-E, Sora und Midjourney, aber auch die Proteinanalyse-KI AlphaFold, mit der man sich erhofft, unheilbare Krankheiten behandeln zu können (vgl. Olson 2023). Der Einfluss dieser Arbeit lässt sich nur schwer leugnen: „If modern artificial intelligence has a founding document, a sacred text, it is Google’s 2017 research paper ‚Attention Is All You Need‘“ (Toews 2023).

Anfang der 2010er-Jahre sorgten die Smartspeaker-KIs Siri (Apple), Alexa (Amazon) und Co. noch für Erstaunen, da sie gesprochene Sprache der Nutzenden einigermaßen präzise verarbeiten

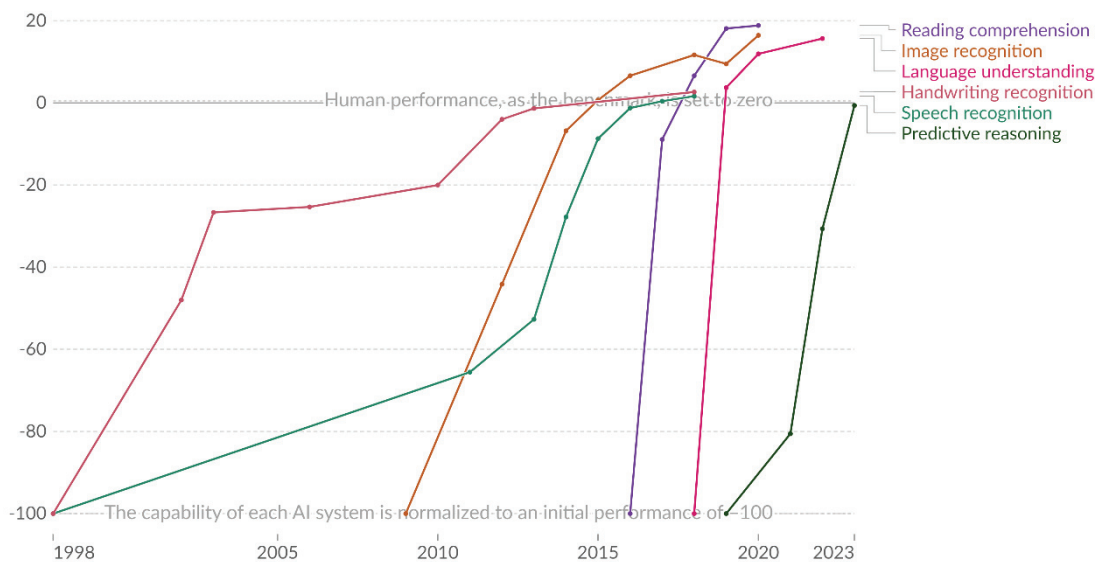
konnten. Die Präsentation von der KI Google Duplex 2018, bei der die Sprach-KI täuschend menschlich einen Frisörtermin vereinbarte, ohne dass das Gegenüber die KI als eine solche entlarvte, hatte eine enorme Resonanz zur Folge, die aufgrund der Lebensechtheit der KI-Stimme zwischen Erstaunen und Entsetzen schwankte (vgl. Herbig 2018). Nur 6 Jahre später sollte eine Reihe von Videos, in denen OpenAI ihr neuestes ChatGPT-Modell 4o am 13.05.2024 auf ihrem Youtube-Kanal präsentierte, alle bisher dagewesenen Sprach-KIs deutlich in den Schatten stellen. Nicht nur zeigte das Modell, dass lückenlose Unterhaltungen mit ihm möglich sind, sondern es verblüffte auch dadurch, dass es künstliche Emotionen transportieren konnte und bei menschlichen Gegenübern via Kameraschnittstelle in der Lage war, diese zu lesen (vgl. OpenAI, o. D.). Damit setzte OpenAI, was große Sprachmodelle (auch Large Language Models, kurz LLM) angeht und im weiteren Verlauf der Arbeit noch näher erläutert werden sollen (vgl. Kap. 3.2.2.1), zum wiederholten Male neue Maßstäbe.

Merkmal der beschriebenen KI-Welle ist, dass Erfolge immer schneller erzielt werden – insbesondere dann, wenn es sich um differenzierte und klar definierte Teilleistungen handelt (vgl. Abb. 5).

### Test scores of AI systems on various capabilities relative to human performance



Within each domain, the initial performance of the AI is set to -100. Human performance is used as a baseline, set to zero. When the AI's performance crosses the zero line, it scored more points than humans.



Data source: Kiela et al. (2023)

OurWorldInData.org/artificial-intelligence | CC BY

Note: For each capability, the first year always shows a baseline of -100, even if better performance was recorded later that year.

Abbildung 5: KI-Leistung im Vergleich zum Menschen (Kiela et al. 2023, zit. n. Roser 2022)

Vorrangig kann dies an drei Faktoren festgemacht werden, nämlich wissenschaftlichem Fortschritt, fortschrittlicherer Hardware inklusive besserer Anbindung an die Software und einer Fülle verarbeitbarer Daten, auch Big Data genannt, die maschinell lernende Künstliche Intelligenzen benötigen, um ihre Wissensdatenbanken zu füllen und anhand dieser zu lernen (vgl. Sheikh et al. 2023, S. 39). Einige der wissenschaftlichen Fortschritte der letzten Jahre wurden auf den vorherigen Seiten näher erläutert. Auch stärkere Hardware, die deutlich bessere Ergebnisse über effizientere Trainings maschinell lernender KI ermöglicht, hat einen großen Anteil am jüngsten KI-Boom. Veranschaulicht werden kann dies gut anhand des börsennotierten Unternehmens NVidia, das weltweit führender Hersteller von Grafikkarten und Grafikchips ist, die den Großteil der Rechenleistung beim Training von maschinell lernender KI übernehmen. Alleine im letzten Finanzjahr stieg der Wert der Aktie um 239 Prozent an, wodurch NVidia zum drittgrößten Unternehmen weltweit aufstieg (vgl. Statista Research Department 2024). Verdanken kann NVidia dies hauptsächlich dem enormen Hype, der KI umgibt, und der deutlich erhöhten Nachfrage nach Grafikchips, auf die alle namhaften Techkonzerne angewiesen sind (vgl. Bowman 2024). Die große Nachfrage wird auch daran sichtbar, dass mehreren Medienberichten zufolge kürzlich bekannt wurde, Elon Musk plane für sein KI-Startup ‚xAI‘ den weltgrößten Supercomputer im Wert von 300 Millionen Euro zu bauen, der im Verbund aus 30.000 Einzelsystemen der Firma Nvidia bestehen soll (vgl. Kremp 2024b). Letztlich sorgte auch Big Data dafür, dass KI immer besser wurde. Anfang der 2010er-Jahre kam es zu einem Paradigmenwechsel: Anstatt, wie zuvor geschehen, wissensbasierte Expertensysteme individuell mit Regeln zu programmieren, rückte maschinelles Lernen immer mehr in den Fokus. Diese Systeme können mittels Daten auf verschiedene Weisen trainiert werden (vgl. Kap. 3.2.2.2), wofür sie allerdings riesige Datensätze benötigen (vgl. Kaplan 2022, S. 15). Das Aufkommen einer enormen Fülle an Datensätzen steht insbesondere in direkter Korrelation mit der flächendeckenden Nutzung des Internets weltweit durch fast alle Menschen. Dies hat zur Folge, dass jede Menge Daten frei zur Verfügung stehen, die sich maschinell lernende Künstliche Intelligenzen für Trainings zu Nutze machen können (vgl. Sheikh et al. 2023, S. 38).

In den letzten Jahren ist zudem die Arbeit an Quantencomputern immer weiter fortgeschritten. Diese sind erst am Anfang ihrer Entwicklung und stehen aktuell noch nicht so sehr im Fokus, jedoch ruhen große Hoffnungen in ihrem Potenzial, das vor allem in deutlich erhöhter Rechenleistung begründet wird (vgl. Lenzen 2020, S. 28). Ein Durchbruch in diesem Bereich könnte der Forschung einen ganz neuen Schub verleihen und weiterhin dafür sorgen, dass die Rasanz der Fortschritte bei der Entwicklung Künstlicher Intelligenzen deutlich zunimmt. Wie an Abbildung 5 abzulesen war, ist bereits gegenwärtig die Leistung Künstlicher Intelligenzen beachtlich und

sie „dringt in kognitive Bereiche vor, die bisher dem Menschen vorbehalten waren“ (Wennker 2020, S. 8). Dass KI mittlerweile in der Mitte der Gesellschaft angekommen ist und ein regelrechter Hype um KI ausgebrochen ist, zeigen die Nutzendenzahlen von ChatPGT. Wie Hu (2023) berichtet, erreichte ChatGPT bereits zwei Monate nach seiner Veröffentlichung am Ende des Jahres 2022 eine Anzahl von Nutzenden, die die 100-Millionen-Marke überschritt. Damit ist der Chatbot Rekordhalter, was die Geschwindigkeit angeht, diese Marke überschritten zu haben. Zum Vergleich: Die derzeit beliebteste Social-Media-Plattform TikTok benötigte für diese Nutzeranzahl neun Monate und die Social-Media-Instanz Instagram mehr als zwei Jahre (vgl. ebd.).

Vielen Menschen erscheinen die Dinge, die sich mit generativer KI heutzutage anstellen lassen wie interessante aber letztlich noch unnütze und nicht ganz ausgereifte Spielereien. Der Wirtschaftsprofessor Ethan Mollick (2024) weist in seinem Blog darauf hin, KI würde sich nicht von anderen wichtigen Technologien unterscheiden und daher durch fortlaufende Entwicklung immer besser werden. Allerdings sei die Geschwindigkeit, in der im Feld der Künstlichen Intelligenz Fortschritte erzielt würden, vergleichsweise deutlich größer. Abschließend prophezeit er, KI werde in naher Zukunft die Schwelle zur Nützlichkeit überschreiten, dann nicht mehr wegzudenken sein und eventuelle Konkurrenzprodukte oder -technologien vom Markt drängen, so wie es einst die Digitalkamera mit der Fotokamera getan hat (vgl. Mollick 2024). Unstrittig ist also, dass KI als Technologie im Laufe der Zeit immer ausgefeiltere Varianten hervorbringen wird. Da sich diese auch zum Nudging, zur Überwachung, zur Manipulation und zur Kontrolle der Menschen einsetzen lassen, ist es wichtig, fortlaufend einen gesellschaftlichen Diskurs darüber zu führen, wie sich der vielgestaltete Einsatz von KI unter Berücksichtigung ethischer und datenschutzrechtlicher Bedenken vereinbaren lässt, sodass das Recht auf Selbstbestimmung des Menschen unversehrt bleibt (vgl. Abrams 2024a; vgl. Beranek et al. 2021, S. 15; vgl. Greef 2023, S. 91 ff.).

### **3.2 Definition**

Den Begriff Künstliche Intelligenz zu definieren, ist ein nicht unkompliziertes Unterfangen. Über die Jahre hinweg betrachtet, erblickten zahllose unterschiedliche Definitionen das Licht der Welt. Die Gründe hierfür sind vor allem zweierlei. Weil sich die KI-Entwicklung, wie im vorigen Unterkapitel ersichtlich wurde, fortlaufend in regem Wandel befindet und eine hohe von Durchbrüchen und Innovationen geprägte Dynamik aufweist, verändert sich das Verständnis von KI entsprechend laufend mit (vgl. Mockenhaupt & Schlagenhaut 2024, S. 57). Mit fortschreitender Zeit verschieben sich die Grenzen dessen, was als KI und was als grundlegende, ins alltägliche Leben eingebettete und für normal befundene Technologie gewertet wird ganz wie von selbst

(vgl. Kaplan 2022, S. 7). Darüber hinaus zeichnet sich das Forschungs- und Arbeitsfeld von KI-Entwicklung seit jeher durch seine Multi- beziehungsweise Interdisziplinarität aus, weshalb eine einheitliche Definition schwerfällt (vgl. Kap 3.1). Die verschiedenen an der Entwicklung von KI beteiligten Disziplinen warten teils mit stark unterschiedlichen Herangehensweisen, Paradigmen und Methoden auf und betrachten den Gegenstand Künstliche Intelligenz bisweilen aus voneinander abweichenden Perspektiven (vgl. Ekmekci & Arda 2020, S. 19).

Überdies erscheint es nicht überraschend, Künstliche Intelligenz nicht leicht einheitlich definieren zu können, wenn miteinbezogen wird, dass KI je nach Modell unterschiedlich stark eine Simulation beziehungsweise Emulation menschlicher Intelligenz ist, die bis heute von Psycholog:innen, Behaviourist:innen und Neurowissenschaftler:innen nicht gänzlich durchdrungen und einvernehmlich erklärbar gemacht werden kann (vgl. Mockenhaupt & Schlagenhauf 2024, S. 53 ff.; vgl. Pontes 2014). Auch die Perspektive auf Intelligenz und den Intelligenzbegriff befindet sich in stetem Wandel, wie die Historie zu zeigen weiß. War man 1950 noch der Meinung, das erfolgreiche Programmieren einer Schach-KI könne einen großen Beitrag zur Entschlüsselung menschlicher Intelligenz sein, weiß man heute, dass dem nicht so ist (vgl. Sheikh et al. 2023, S. 16). Konträr dazu war es den Forschenden lange Zeit ein blinder Fleck, welche kognitiven Höchstleistungen bereits Kleinkinder schier mühelos vollbringen und dass diese als ein Teil von menschlicher Intelligenz zu betrachten sind, wie McCorduck (2004) zu berichten weiß:

Intelligence, Brooks writes, was for his colleagues 'best characterized as the things that highly educated male scientists found challenging' — chess, symbolic integration, proving mathematical theorems, and solving complicated word algebra problems. 'The things that children of four or five years could do effortlessly, such as visually distinguishing between a coffee cup and a chair, or walking around on two legs, or finding their way from their bedroom to the living room were not thought of as activities requiring intelligence. Nor were any aesthetic judgments included in the repertoire of intelligence-based skills.' (S. 456)

Dieser Umstand ist heute als Moravecsches Paradox bekannt. Es beschreibt, dass Aufgaben wie Schachspielen oder fortgeschrittene Kalkulationen, die für Menschen schwierig zu bewältigen sind, von Computern vergleichsweise einfach gelöst werden können – umgekehrt fällt es Künstlichen Intelligenzen ungleich schwerer, Dinge zu erlernen, die schon Einjährigen natürlich von der Hand gehen (vgl. Moravec 1988, S. 15). An diesem Beispiel wird auch deutlich, wie unterschiedlich die Maßstäbe und Anforderungen an Künstliche Intelligenzen sein können. Fokussieren sich beispielsweise manche Künstliche Intelligenzen nur auf die konkrete Lösung anwendungsbezogener Probleme, für die sie speziell programmiert oder trainiert wurden, wird sie auch als schwache oder enge KI bezeichnet (vgl. Mockenhaupt & Schlagenhauf 2024, S. 60). Trotz der etwas despektierlich anmutenden Beschreibung der KI mittels des Wortes ‚schwach‘



sollte nicht der Fehler begangen werden, diese zu unterschätzen. Gerade dann, wenn beispielsweise schnelle und fehlerlose Datenverarbeitung vonnöten ist, übertrifft schwache oder schmale KI sämtliches, was ein Mensch vergleichsweise zu leisten im Stande wäre (vgl. Rainsberger 2021, S. 10). Sollen wiederum andere KI-Systeme die gleichen intellektuellen Fähigkeiten wie Menschen erreichen und einer Vielzahl von Problemen und Situationen, auf die sie nicht speziell vorbereitet wurden, die Stirn bieten oder sogar übertreffen können, spricht man von starker oder weiter KI (vgl. Mockenhaupt & Schlagenhaut 2024, S. 60). Starke KI „would equal and outperform human beings in many domains“ (Kaplan 2022, S. 15) und sie wäre dazu imstande, eigenständig Probleme aufzudecken, diese zu umreißen und klar zu benennen, um selbstständig Lösungen für sie zu finden und zu kreieren (vgl. Poltermann et al. 2024, S. 22). In diesem Zusammenhang werden auch oft die Begriffe ‚Artificial Narrow Intelligence‘ (ANI) und ‚Artificial General Intelligence‘ (AGI) genannt, wobei ersterer für schmale und zweiterer für starke KI steht (vgl. Kaplan 2022, S. 15). Eine starke KI, die auch allgemeine KI bezeichnet wird, ist zum jetzigen Zeitpunkt noch nicht existent und wird es auch auf viele Jahre hinaus nicht sein, auch wenn einige Stimmen, wie zum Beispiel die von Elon Musk in Bezug auf das neueste ChatGPT-4-Modell, etwas anderes behaupten, was jedoch vermutlich auf wirtschaftliches Eigeninteresse und Konkurrenzgefahren zurückzuführen ist (vgl. Kemper 2024). Reiner Bestand menschlicher Vorstellungskraft ist bisher auch die ‚Artificial Super Intelligence‘ (ASI), die auch als Singularität bezeichnet wird und, mit echtem Bewusstsein ausgestattet, menschliche Fähigkeiten in jeder erdenklichen Hinsicht übertreffen würde (vgl. Kaplan 2022, S. 15).

Dass diese genannten Beispiele nicht so einfach konsolidiert werden können, leuchtet ein. Für diese Arbeit soll daher eine Definition von KI zitiert werden, die eine unabhängige, hochrangige von der europäischen Kommission eingesetzte Expertengruppe für Künstliche Intelligenz im Jahr 2019 herausgearbeitet hat und die sämtliche momentan existierende Künstliche Intelligenzen mitberücksichtigt:

Systeme der künstlichen Intelligenz (KI-Systeme) sind vom Menschen entwickelte Softwaresysteme (und gegebenenfalls auch Hardwaresysteme), die in Bezug auf ein komplexes Ziel auf physischer oder digitaler Ebene handeln, indem sie ihre Umgebung durch Datenerfassung wahrnehmen, die gesammelten strukturierten oder unstrukturierten Daten interpretieren, Schlussfolgerungen daraus ziehen oder die aus diesen Daten abgeleiteten Informationen verarbeiten, und über das bestmögliche Handeln zur Erreichung des vorgegebenen Ziels entscheiden. KI-Systeme können entweder symbolische Regeln verwenden oder ein numerisches Modell erlernen,[sic!] und sind auch in der Lage, die Auswirkungen ihrer früheren Handlungen auf die Umgebung zu analysieren und ihr Verhalten entsprechend anzupassen.

Als wissenschaftliche Disziplin umfasst die KI mehrere Ansätze und Techniken wie z. B. maschinelles Lernen (Beispiele dafür sind „Deep Learning“ und bestärkendes Lernen), maschinelles Denken (es

umfasst Planung, Terminierung, Wissensrepräsentation und Schlussfolgerung, Suche und Optimierung) und die Robotik (sie umfasst Steuerung, Wahrnehmung, Sensoren und Aktoren sowie die Einbeziehung aller anderen Techniken in cyber-physische Systeme). (Europäische Kommission 2019, S. 6)

Kerntätigkeit Künstlicher Intelligenzen ist zumeist, mit großen Mengen an Daten zu arbeiten, um aus diesen Erkenntnisse abzuleiten (vgl. Kreuzer 2023, S. 23). Auf Grundlage dessen kann ein KI-System dann ganz spezifisch Aufgaben lösen und – je nach Sinn und Zweck – Voraussagungen treffen, Entscheidungen ableiten oder gar gänzlich neue Inhalte kreieren, die auf Wahrscheinlichkeitsrechnungen beruhen und nicht *echter* Kreativität entspringen (vgl. Kap. 3.2.2.1). Der Aufbau eines KI-Systems und dessen Funktionsweise und Einbettung in die Welt kann in Abbildung 6 nachvollzogen werden.

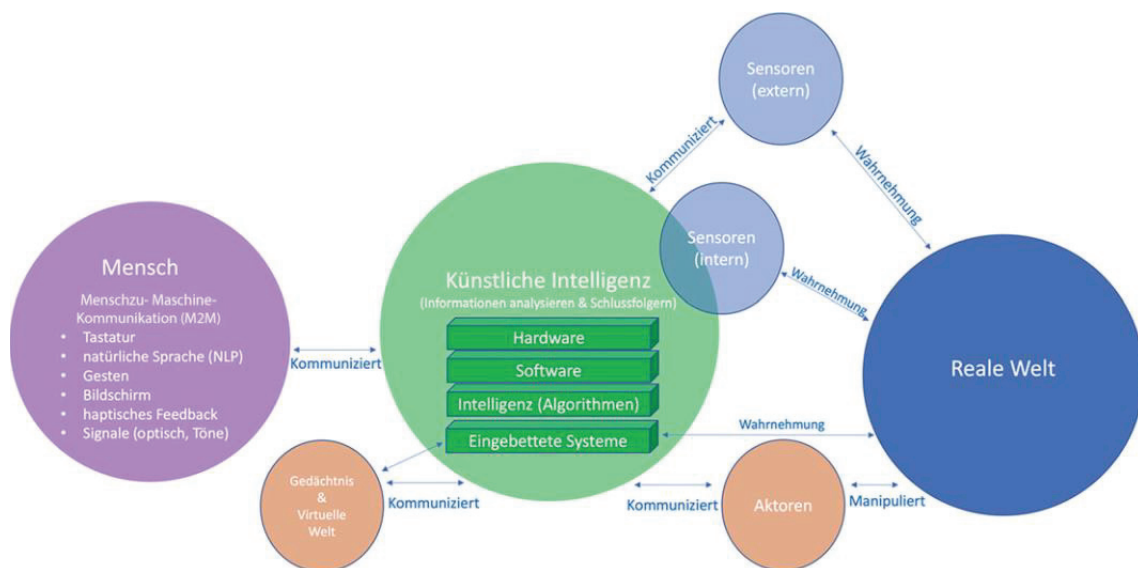


Abbildung 6: Aufbau eines KI-Systems (Mockenhaupt & Schlagenhaut 2024, S. 59)

Angemerkt werden soll an dieser Stelle, dass diese Abbildung verschiedene Möglichkeiten eines Aufbaus demonstriert. Abgesehen von der internen Konstitution einer KI (Hardware, Software, Algorithmen und eingebettete Systeme) sind die verschiedenen Merkmale, Funktionsweisen und Ausstattungen als optional zu betrachten und nicht jedem KI-System inhärent. KI-Systeme können also durchaus große Unterschiede zueinander aufweisen, wie sich im Folgenden zeigen wird.

### 3.2.1 Symbolische KI

Symbolische Künstliche Intelligenz wird heutzutage etwas herablassend auch als ‚Good-old-fashioned AI‘ (GOFAI) bezeichnet, obwohl zahllose auf symbolischer KI beruhende Systeme bis heute im praktischen Einsatz sind (vgl. Rainsberger 2021, S. 11). Die symbolische KI hat ihren

Namen daher, da sie sich „ursprünglich [...] an Regeln und Formeln der symbolischen Logik“ (Mainzer & Kahle 2022, S. 2) orientierte und Computerprogramme auf diese Weise in der Lage sind, diese zu verarbeiten (vgl. Taulli 2022, S. 10 f.). Für symbolische KI typisch sind wissensbasierte Expertensysteme, die individuell mit in formale Regeln übersetztem Fachwissen gespeist werden, sodass dieses für einen bestimmten Zweck und eine vordefinierte Aufgabe gebräuchlich gemacht werden kann, wie zum Beispiel die in Kapitel 3.1 angesprochenen KI-Systeme MYCIN oder XCON gezeigt haben (vgl. Mainzer & Kahle 2022, S. 2). Symbolischer KI immanent ist, dass sie einen Top-Down-Ansatz verfolgt, was bedeutet, dass sie die Daten, die sie bekommt, eindeutig nachvollziehbar anhand eines vorher definierten Algorithmus verarbeitet, woraus im Ergebnis eine Schlussfolgerung errechnet wird (vgl. Mockenhaupt & Schlagenhaut 2024, S. 65). Mockenhaupt & Schlagenhaut (2024) erklären einen Algorithmus wie folgt:

Ein Algorithmus ist eine präzise, endliche Vorgehensweise zur Lösung einer Aufgabe. Er enthält eindeutig definierte schrittweise Anweisungen zur Erreichung eines bestimmten Ziels. Bei Wahlmöglichkeiten muss es definierte Antwortwege geben. (S. 183)

Algorithmen dienen dem Menschen durch errechnete Ergebnisse zur Entscheidungsunterstützung, „oder [sie] oder entscheiden selbst – in dem Sinn, dass sie auf Basis ihres Ergebnisses eine Folgeaktion auslösen“ (Greef 2023, S. 94). Kaplan (2022) weist darauf hin, ein Nachteil der symbolischen KI sei, sie könne nur in einem vorher bestimmten Rahmen funktionieren und sei daher starr, das heißt, sie lerne im Laufe der Zeit nicht (vgl. S. 10). Sollten Veränderungen am Algorithmus gewünscht sein, müssen diese händisch programmiert werden und Wissen, das sich über Zeit aktualisiert hat, muss manuell in die Datenbank eingepflegt werden und eventuell nicht mehr gültige Informationen händisch entfernt werden (vgl. ebd.). Zudem ist es für eine symbolische KI wichtig, dass sie über eine möglichst lückenlose Datenbasis verfügt, da sie nur sehr bedingt imstande ist, mit Datenunsicherheit umgehen zu können (vgl. Mockenhaupt & Schlagenhaut 2024, S. 65).

Doch die symbolische KI hat auch Vorteile: Durch den Top-Down-Ansatz ist zu jeder Zeit nachvollziehbar, aus welchem Grund die symbolische KI eine Entscheidung getroffen hat und wie sie zu diesem Ergebnis gekommen ist, wodurch sie vollends berechenbar ist (vgl. Sheikh et al. 2023, S. 76). Symbolische KI ist in der Praxis daher insbesondere in Umgebungen von Vorteil, in denen

sich anhand von vorab klar definierten Regeln konkrete Probleme in statischen Umgebungen lösen lassen (vgl. Mockenhaupt & Schlagenhauf 2024, S. 65).

### 3.2.2 Subsymbolische KI

Im Gegensatz zur symbolischen KI beruht subsymbolische KI auf der Idee eines Systems, das sich selbständig auf den Weg der Problemlösung begeben soll (vgl. Lenzen 2020, S. 24). Die Strukturen, die hierbei gebildet werden, können nicht durch Symbole erfasst werden, sondern sind sehr viel abstrakter, weshalb der Begriff ‚subsymbolisch‘ beim Beschreiben dieser Form der KI gewählt wurde (vgl. ebd.). Ein anderer Begriff, der häufig in diesem Zusammenhang fällt, um diese Form der Künstlichen Intelligenz zu beschreiben, ist das sogenannte maschinelle Lernen (vgl. Sheikh et al. 2023, S. 28). Ein maschinell lernendes KI-System wird mit Datensätzen gespeist und mittels dieser Trainings unterzogen, in Folge derer der Algorithmus, der dem jeweiligen KI-Modell zugehörig ist, Muster, Auffälligkeiten, Unterschiede und Gemeinsamkeiten im vorliegenden Datensatz erkennt und auf diese Art und Weise eine Lösung entwickelt, anstatt das dies, wie bei der zuvor vorgestellten symbolischen KI, durch eine explizite Programmierung geschieht (vgl. Rainsberger 2021, S. 11). Die KI erzeugt also *Wissen* aus Daten, ohne hierfür ausdrücklich programmiert worden zu sein und kann in Folge dessen das neu gewonnene Wissen auf neue unbekannte Daten beziehungsweise die Umwelt anwenden (vgl. Wennker 2020, S. 12). Bevor dem KI-Modell Daten gegeben werden, die es verarbeiten soll, findet das sogenannte Data-Pre-Processing statt, was das Aufbereiten der Daten beschreibt (vgl. Mockenhaupt & Schlagenhauf 2024, S. 145). Dieser Schritt ist sehr wichtig, da minderwertige Datensätze, die von doppelten, fehlerhaften oder fehlenden Daten geprägt sind, schnell zu unbrauchbaren Ergebnissen führen können, weshalb dieser Prozess sehr arbeitsaufwändig ist und bis zu 80 Prozent der Gesamtarbeit in der Entwicklung einer KI ausmachen kann (vgl. Wennker 2020, S. 10).

Subsymbolische Künstliche Intelligenz folgt im Gegensatz zur symbolischen KI einem Bottom-Up-Ansatz. Die Regeln, nach denen sie zu verfahren hat, werden ihr nicht von einem Menschen aufoktroiert, sondern sie formuliert diese selber (vgl. Lenzen 2020, S. 24). Im Laufe des Trainings erweitert sie ihr *Wissen* und ihre *Intelligenz* stetig, weshalb sie im Vergleich zur symbolischen KI zu weitaus komplexeren Aufgaben eingesetzt werden kann und in dynamischen Umgebungen bestehen kann (vgl. Kaplan 2022, S. 10). Der Nachteil hieran ist die sogenannte Black-Box-Problematik, die den problematischen Umstand beschreibt, dass die Entscheidungsfindung, nach der die subsymbolische KI agiert, nicht en Detail transparent zu machen ist und nicht klar ist, auf welcher Grundlage der zum Einsatz kommende Algorithmus zu seiner Entscheidung gefunden hat (vgl. Mockenhaupt & Schlagenhauf 2024, S. 84). Da die maschinellen Lernverfahren

eine große Menge an Daten benötigen, um aus ihnen zu lernen und so komplexe Zusammenhänge errechnen zu können, stellte sich Big Data als große Triebfeder in der Entwicklung subsymbolischer KI heraus (vgl. Gutwald et al. 2021, S. 5). Ein Problem an Daten ist, dass sie veralten, weshalb andauernd neue Daten in großen Mengen benötigt werden (vgl. Lenzen 2020, S. 77). „Sie müssen gesammelt, gespeichert, aufbereitet und für den Prozess des überwachten Lernens annotiert werden, bevor man sie verwenden kann“ (ebd.). Deshalb kann die Abhängigkeit von Daten sich auch als Problem entpuppen, wenn keine mehr übrig sind, anhand derer sich eine KI noch trainieren ließe (vgl. Jürgens 2024). Es zeigt sich, dass symbolische und subsymbolische KI einige Verschiedenheiten aufweisen, wodurch sie sich für unterschiedliche Anwendungsfelder eignen (vgl. Abb. 7).

Schwache KI	Symbolische KI	Beispiel: Assistenzsystem beim Auto
	<ul style="list-style-type: none"> <li>- (sichere) Daten</li> <li>- Bekannte Ereignisse &amp; Zusammenhänge</li> <li>- Wissens- &amp; regelbasierte Verfahren</li> <li>- Transparente Entscheidung</li> </ul>	<ul style="list-style-type: none"> <li>- Zielvorgabe, z.B. rückwärts einparken</li> <li>- Abgleich Umgebungsdaten der Sensorik</li> <li>- Prognose aufgrund von (Verkehrs-)Regeln</li> <li>- Nachvollziehbare, wenig störanfällige Verarbeitung</li> </ul>
Starke KI	Subsymbolische KI	Beispiel: Autonomes Fahren
	<ul style="list-style-type: none"> <li>- Große Datenmenge</li> <li>- Daten ggf. unsicher oder nicht vollständig</li> <li>- Maschinelles Lernen &amp; neuronale Netze</li> <li>- Robuste (fehlertolerante) Verarbeitung</li> <li>- Entscheidung für Menschen nicht immer transparent bzw. nachvollziehbar</li> </ul>	<ul style="list-style-type: none"> <li>- Komplexe Verkehrssituation</li> <li>- Nicht alle Teilnehmer halten sich an Regeln</li> <li>- Schnelle, z. T. unlogische Veränderungen</li> <li>- Nachahmung der menschl. unbewussten, sensomotorischen Reaktionsweise</li> <li>- Entscheidung bewegt sich innerhalb von Leitplanken</li> </ul>
	Zukunft	Beispiele:
	<ul style="list-style-type: none"> <li>- Lernen aus kleinen Datenmengen</li> <li>- Rückschlüsse aus 1 auf <math>\infty</math></li> <li>- Regeltolerante Systeme</li> <li>- Chaostheoretischer Ansatz</li> </ul>	<ul style="list-style-type: none"> <li>- Kind lernt aus einer Katzenzeichnung und erkennt sicher alle Katzen</li> <li>- Zulassen kreativer Lösungen, die Regeln ausweiten („Learn the rules and then break them“ nach: Pablo Picasso )</li> </ul>

Abbildung 7: Symbolische und subsymbolische KI (Mockenhaupt & Schlagenhaut 2024, S. 66)

### 3.2.2.1 Künstliche neuronale Netze

Ein neuronales Netz wird im Fachgebiet der Künstlichen Intelligenz auch als künstliches neuronales Netz (KNN) bezeichnet (vgl. Kreuzer 2023, S. 11). Es stellt ein Rechenmodell dar, das von der Funktionsweise des menschlichen Gehirns inspiriert ist und dessen Aufbau aus einer Vielzahl von miteinander verbundenen Knoten besteht – in etwa wie Synapsen im menschlichen Gehirn miteinander verknüpft sind (vgl. Wennker 2020, S. 18). Diese werden als Neuronen bezeichnet

und organisieren sich in mehreren Schichten (vgl. Kreutzer 2023, S. 11). Die Rohdaten, anhand derer KI-Modelle, die auf künstlichen neuronalen Netzen basieren, lernen, werden über die Eingangsschicht (Input Layer) eingepflegt (vgl. Lenzen 2020, S. 38 f.). Anschließend finden in den verborgenen Schichten (Hidden Layers) die eigentlichen Berechnungen statt, woraus sich schlussendlich das Ergebnis der Berechnungen ergibt, welches über die Ausgangsschicht (Output Layer) ausgegeben wird (vgl. Abb. 8) (vgl. Kreutzer 2023, S. 11).

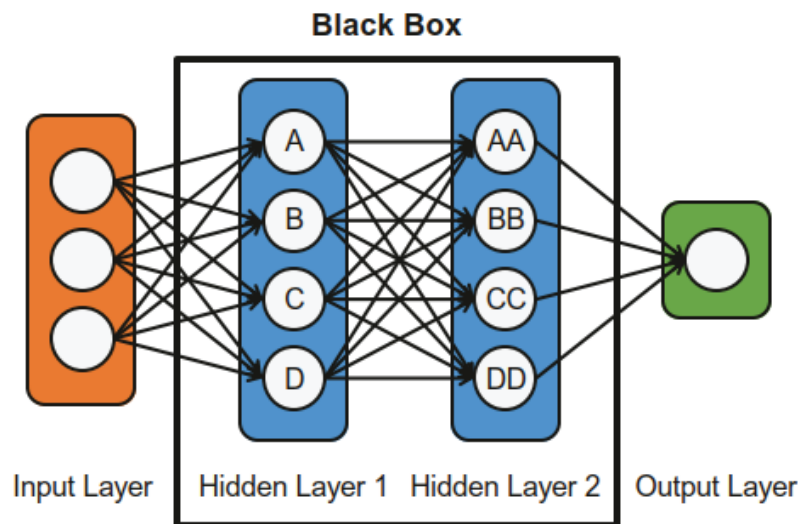


Abbildung 8: Schichten eines neuronalen Netzes (Kreutzer 2023, S. 11)

Terminologisch werden künstliche neuronale Netze und neuronale Netzwerke oft synonym verwendet (vgl. Taulli 2022, S. 83). Sie beschreiben also künstliche neuronale Netze, die eine oder mehrere verborgene Schichten, wie sie für Deep-Learning-KIs typisch sind, besitzen (vgl. Lenzen 2020, S. 37). Allerdings kann ein neuronales Netzwerk auch die Gesamtheit des KI-Systems beschreiben, wobei dann, wie in Kapitel 3.2 dargelegt wurde, die Hard- und Softwareinfrastruktur miteinzubeziehen wäre (vgl. Kreutzer 2023, S. 11).

Mit zunehmender Komplexität und Tiefe des neuronalen Netzes verstärkt sich die Black-Box-Problematik (vgl. Kap. 3.2.2) und es kann nicht mehr nachvollzogen werden, wie das KI-Modell zu seinem Ergebnis gekommen ist (vgl. Lenzen 2020, S. 40). Kommen mehrere Schichten zum Einsatz, wird das künstliche neuronale Netz tiefer und leistungsfähiger, weshalb diese Art der KI auch den Namen Deep Learning trägt (vgl. Rainsberger 2021, S. 12). Herkömmliche künstliche neuronale Netze leiten Eingangssignale normalerweise unidirektional von der Eingangsschicht über die verborgene Schicht zur Ausgangsschicht und verfügen zumeist nur über bis zu drei Schichten, wohingegen die Architektur von Deep-Learning-Netzen oftmals aus zehn oder weit

mehr Schichten besteht und über einen komplexeren Signalfluss verfügen kann (vgl. Kreuzer 2023, S. 13; vgl. Wennker 2020, S. 26). Das populäre und sehr komplexe KI-Modell von ChatGPT-3 setzt sich beispielsweise aus 96 verborgenen Schichten und 175 Milliarden Neuronen zusammen (vgl. Knapp 2023). Durch die tiefen und komplexen Netze sind Deep-Learning-KI-Modelle in der Lage, „kontinuierlich zu lernen und Erlerntes bei weiteren Interaktionen mit der Welt anzupassen und zu modifizieren“ (Wennker 2020, S. 9). Moring (2023) zufolge ist die Anzahl an verborgenen Schichten, die ein neuronales Netzwerk haben kann, in der Theorie unbegrenzt (vgl. S. 15). Allerdings erhöht sich die benötigte Rechenleistung mit jeder neuen verborgenen Schicht, weshalb die Anzahl verborgener Schichten, die in der Praxis bei neuronalen Netzwerken zum Einsatz kommen, nicht ins Unermessliche wachsen (vgl. ebd.). Es gibt eine Reihe unterschiedlicher KI-Architekturen, die dem Deep Learning zuzuordnen sind, wovon die prominentesten im Folgenden kurz erläutert werden sollen.

Das Convolutional Neural Network (CNN, zu Deutsch etwa: faltendes neuronales Netzwerk) sticht vor allem in der Bild-, Text- und Audioverarbeitung sowie bei der Gesichtserkennung hervor und ist in seiner Funktionsweise dem Teil des Gehirns nachempfunden, der Bilder erfasst und verarbeitet (vgl. Alzubaidi et al. 2021, S. 13; vgl. Kreuzer 2023, S. 13). In den vergangenen Jahren haben Convolutional Neural Networks (CNN) rasch an Popularität gewonnen und Künstliche Intelligenzen, die auf der CNN-Architektur beruhen, decken ein breites Spektrum an Einsatzmöglichkeiten ab (vgl. Raitoharju 2022, S. 35). So kommen sie in der historischen Analyse, Medizin, Landwirtschaft, Insektenkunde, Fernerkundung und Robotik zum Einsatz und sind fester Bestandteil selbstfahrender Autos (vgl. ebd.).

Der Hauptunterschied von Convolutional Neural Networks zu einfachen neuronalen Netzen besteht in der Architektur und Art und Weise, wie Informationen verarbeitet werden. Ein typisches CNN setzt sich in der Regel aus vier Schichten zusammen, die aber in mehrfacher Ausführung innerhalb eines CNNs vorkommen können: Convolutional Layers, Pooling Layers, Fully Connected Layers und einem Output Layer (vgl. ebd., S. 37). Anders als bei simpleren künstlichen neuronalen Netzen sind die im vorangegangenen Satz benannten Schichten des CNN bis auf die Ausgabeschicht nicht vollends miteinander vernetzt, sondern nur mit wenigen Neuronen der darunter liegenden Schicht, was vor allem den Grund hat, Rechenleistung sparen zu wollen (vgl. Ertel 2020, S. 325). Mittels sogenannter Filter beziehungsweise Kernels, die sich in den Convolutional Layers befinden, erfasst das CNN Informationen aus dem vorliegenden Bild (vgl. Wennker 2020, S. 28). Das Verfahren, wie eine CNN arbeitet, kann verbildlicht werden, wenn man sich eine Kamera vorstellt, die eingangs Schritt für Schritt das gesamte Bild abtastet (vgl. Moring 2023, S. 17). Auf die Convolutional-Schicht folgt stets eine Pooling-Schicht. Die Pooling-Schicht

reduziert die Dimensionen des empfangenen Bildes, um weitere Rechenleistungseinsparungen gewährleisten zu können (vgl. Ertel 2020, S. 325). Zunächst werden nur grobe Merkmale des Bildes erkannt, aber mit jeder weiteren Schicht, die das Bild verarbeitet, wird die Mustererkennung präziser und die erkannten Strukturen feiner (vgl. Derry et al. 2023, S. 1270). Die Daten werden in den Convolutional- und Pooling-Schichten immer wieder neu abgetastet und gefiltert bis in einem letzten Schritt Ergebnisse entstehen, die sich „aus den charakteristischen Merkmalen der vorgegebenen Klassen, die erkannt werden sollen“ (Moring 2023, S. 17), ergeben.

Eine weitere KI-Architektur ist im rekurrenten neuronalen Netzwerk (RNN) auszumachen. Ein RNN ist eine Architektur künstlicher neuronaler Netze, die speziell für die Verarbeitung von Sequenzen entwickelt wurde, wie zum Beispiel Texte oder gesprochene Sprache (vgl. Ertel 2020, S. 331). Besonderes Merkmal eines rekurrenten neuronalen Netzwerkes ist es, Informationen aus vorherigen Schritten einer Sequenz zu speichern und sie in aktuelle Berechnungsvorgänge miteinzubeziehen (vgl. Lenzen 2020, S. 38). Dies geschieht, indem die Neuronen innerhalb einer Schicht verknüpft sind und mit Neuronen vorangegangener Schichten verbunden werden (vgl. Wennker 2020, S. 26). Dadurch können RNNs Kontext und Abhängigkeiten in Sequenzen erkennen und darauf basierend Vorhersagen treffen oder Aufgaben wie Analysen, Übersetzungen sowie Handschrift- und Spracherkennung übernehmen (vgl. Moring 2023, S. 16). Während des Lernens verwenden rekurrente neuronale Netzwerke eine Technik, die sich Backpropagation nennt. Mit dieser lassen sich Häufung und Schwere gemachter Fehler reduzieren, indem der von der Ausgabeschicht produzierte Wert mit dem Zielwert abgeglichen wird (vgl. Wennker 2020, S. 27). In diesem sich wiederholenden Prozess, „erfolgt somit ein Lernen durch Rückkopplung in Feedback-Schleifen, in denen der Algorithmus systematisch die (initial zufälligen) Parameter des Netzwerks anpasst und neu kalibriert, um die Erkennungsleistung des nächsten Durchgangs zu verbessern“ (Greef 2023, S. 96).

Die Erfindung der Transformer-Architektur war bahnbrechend für große Sprachmodelle (LLMs) und die natürliche Sprachverarbeitung (NLP) beziehungsweise Computerlinguistik von Künstlicher Intelligenz (vgl. Kap. 3.1). Gegenwärtig bildet diese Architektur das Rückgrat aller großen Sprachmodelle wie ChatGPT, BERT oder Googles ‚Gemini‘ (früher Bard genannt) (vgl. Atallah et al. 2023, S. 609). Als große Sprachmodelle werden KI-Modelle bezeichnet, die auf die Interaktion zwischen Menschen und Computer über menschliche Sprache spezialisiert sind (vgl. Kreuzer 2023, S. 245). Ihre Fähigkeiten bestehen vordergründig darin, geschriebenes Wort und gesprochene Sprache sinnhaft analysieren und in Zusammenhang setzen zu können (vgl. Atallah et al. 2023, S. 609). So wie sie natürliche Sprache verarbeiten können, sind sie gleichsam auch in der Lage, Sprache generieren zu können – sowohl in geschriebener als auch in gesprochener Form



(vgl. Kreutzer 2023, S. 245). Die besten großen Sprachmodelle wie ChatGPT-4o sind dabei nicht nur befähigt, eine sinnhafte und kohärente Unterhaltung mit einer menschlichen nutzenden Person zu führen, indem sie auf ihre Sätze reagieren und ihr Auskunft erteilen, sie sind beispielsweise auch fähig, Programmcodes zu schreiben, über Noten ein Musikstück zu komponieren oder Kochrezepte zu verfassen (vgl. Atallah et al. 2023, S. 609). Transformer werden genutzt, um die NLP-Forschung voranzutreiben, deren gegenwärtige „Ziele [...] u.a. das Erkennen von Stimmungen in Texten, das automatisierte Generieren von Sprache, die nicht oder kaum von menschlich verfassten Texten zu unterscheiden ist (z.B. zur Anwendung in Chatbots) und das Erfassen des Inhalts sowie des Kontexts von Sprache“ (Linnemann et al. 2023a, S. 200). Darüber hinaus nutzen auch sogenannte generative Künstliche Intelligenzen, wie zum Beispiel die Bildergeneratoren ‚Midjourney‘ und ‚DALL-E‘, unter anderem große Sprachmodelle, um aus einer Textaufforderung Bilder zu generieren (vgl. Oppenlaender 2022, S. 200 f.). Bevor Transformer das Feld der großen Sprachmodelle für sich vereinnahmten, beruhten die meisten LLMs auf rekurrenten neuronalen Netzwerken (vgl. Hutchins et al. 2022, S. 1). Allerdings haben Transformer gegenüber RNNs vor allem drei Vorteile: Transformer-Modelle sind in der Lage Texte parallel zu verarbeiten, was deutlich schnellere Trainings und Textverarbeitung ermöglicht; Transformer können viel leichter auf Daten vergangener Zeitsequenzen zugreifen als RNN und leiden deshalb nicht so stark wie RNN unter Speicherlimitierungen; die Selbstaufmerksamkeitsfunktion von Transformern ermöglicht ihnen, die Bedeutungen von Wörtern in Sätzen besser verstehen zu können, was eine bessere Erfassung von Langzeitabhängigkeiten zur Folge hat, wodurch sie fähig sind, Gesprächszusammenhänge von Wörtern, Sätzen und deren Bedeutungen besser und länger aufrechterhalten zu können, als RNN es tun (vgl. ebd.).

Ein statistisches Sprachmodell, das auf enormen menschlichen Textmengen beruht, bildet das grundsätzliche Fundament eines Transformer-Modells (vgl. Kreutzer 2023, S. 245). Das Training eines Transformers ist zweigeteilt. Zuerst lernt er via unüberwachtem Lernen (vgl. Kap. 3.2.2.2). Dafür erhält der Transformer Trainingstexte anhand derer er lernt, die richtigen nächsten Worte in Sätzen vorauszusagen – ähnlich einer Autovervollständigung beim Nachrichtentippen am Smartphone (vgl. Ertel 2020, S. 399). Eine Besonderheit stellt dabei dar, dass der Transformer Worten innerhalb von Sätzen eine unterschiedliche Gewichtung anhand ihrer Relevanz erteilt, was dem Transformer dabei hilft, die Wahrscheinlichkeit zu erhöhen, das nächste richtige Wort passend vorauszusagen (vgl. Atallah et al. 2023, S. 610). Nach Beendigung des unüberwachten Lernens kann ein Transformer auch mit überwachtem Lernen austrainiert werden (vgl. Ertel 2020, S. 399). Das eröffnet auch die Möglichkeit, einen Transformer für einen ganz bestimmten Zweck zu trainieren oder Übersetzungen in andere Sprachen zu ermöglichen (vgl. ebd.). Hieraus

ergibt auch das Akronym ‚GPT‘ in ChatGPTs Namen Sinn, das für ‚Generative Pretrained Transformer‘ steht und auf das vorangegangene Training hindeutet (vgl. Stavridis & Wacker 2023, S. 1). Bei der Nutzung eines großen Sprachmodells wie ChatGPT-4o kann schnell der Eindruck entstehen, der Chatbot würde über ein tiefgreifendes Verständnis über die Sprach- oder Gesprächsinhalte verfügen. Allerdings hat ein auf einer Transformer-Architektur basierender Chatbot kein profundes Verständnis von weltlichen Konzepten (vgl. Taulli 2022, S. 99). Seine durchaus beeindruckende Leistung verdankt er lediglich einem riesigen Wissensfundament und einer gut trainierten Aufmerksamkeit, wichtige Wörter zu entschlüsseln und Zusammenhänge zu erkennen, um daraufhin Texte zu produzieren, die mit sehr hoher Wahrscheinlichkeit eine für die menschliche nutzende Person passende Antwort darstellen, denn daraufhin wurde er ausgebildet (vgl. Atallah et al. 2023, S. 610).

Große Sprachmodelle können eine große Vielfalt an verschiedenen Texten produzieren weswegen sie zur generativen KI zu zählen sind. Generative KI beschreibt KI-Modelle, die in der Lage sind, neue Inhalte auf Grundlage der Daten zu erzeugen, anhand derer sie trainiert wurden (vgl. Poltermann et al. 2024, S. 22). Solche Inhalte reichen von Texten über (Bewegt-)Bilder bis hin zu Musik und weiteren Datenformen (vgl. Mockenhaupt & Schlagenhaut 2024, S. 38). Wie die Transformer werden auch Generative Adversarial Networks (GAN) via unüberwachtem Lernen trainiert (vgl. Lenzen 2020, S. 41). GANs bestehen aus zwei künstlichen neuronalen Netzwerken, die als Kontrahenten gegenüberstehen (vgl. Taulli 2022, S. 89). Das eine Netzwerk ist dabei in der Rolle des Generators und versucht Daten zu produzieren, die vergleichbare Eigenschaften besitzen wie die Trainingsdaten (vgl. Lenzen 2020, S. 41). Dem gegenüber steht das andere Netzwerk, der Diskriminator, den man als eine Art Kunstkritiker bezeichnen könnte (vgl. Taulli 2022, S. 89). Dieses Netzwerk unterzieht die produzierten Daten des Generators einer Bewertung und gleicht sie mit den Trainingsdaten ab (vgl. Lenzen 2020, S. 41). Das geschieht, indem der Diskriminator herauszufinden versucht, ob er die Daten des Generators von den originär vorliegenden Trainingsdaten unterscheiden kann (vgl. ebd.). Umgekehrt versucht der Generator Daten zu produzieren, die der Diskriminator als echt, also als von den Trainingsdaten nicht zu unterscheiden, deklariert (vgl. Ertel 2020, S. 337 f.). Mittels Backpropagation, also dem rückgeleiteten Signal, erhält der Generator die Rückmeldung des Diskriminators und lernt so mit fortschreitender Trainingsdauer, Daten zu erstellen, die den realen Trainingsdaten immer näherkommen (vgl. Taulli

2022, S. 89). Diese Schleife durchläuft viele Iterationen, bis der Generator Ergebnisse liefert, die bisweilen von echten Daten nicht mehr zu unterscheiden sind<sup>4</sup> (vgl. Wennker 2020, S. 31).

### 3.2.2.2 Trainingsverfahren

Es gibt eine Vielzahl unterschiedlicher Trainingsverfahren, mit denen KI-Modelle trainiert werden können. Aufgrund der Begrenztheit dieser Arbeit können an dieser Stelle nicht alle Erwähnung und Behandlung finden. Wie zuvor bei den KI-Architekturen geschehen, sollen die populärsten Trainingsformen allerdings fortfolgend erläutert werden (vgl. Abb. 9).

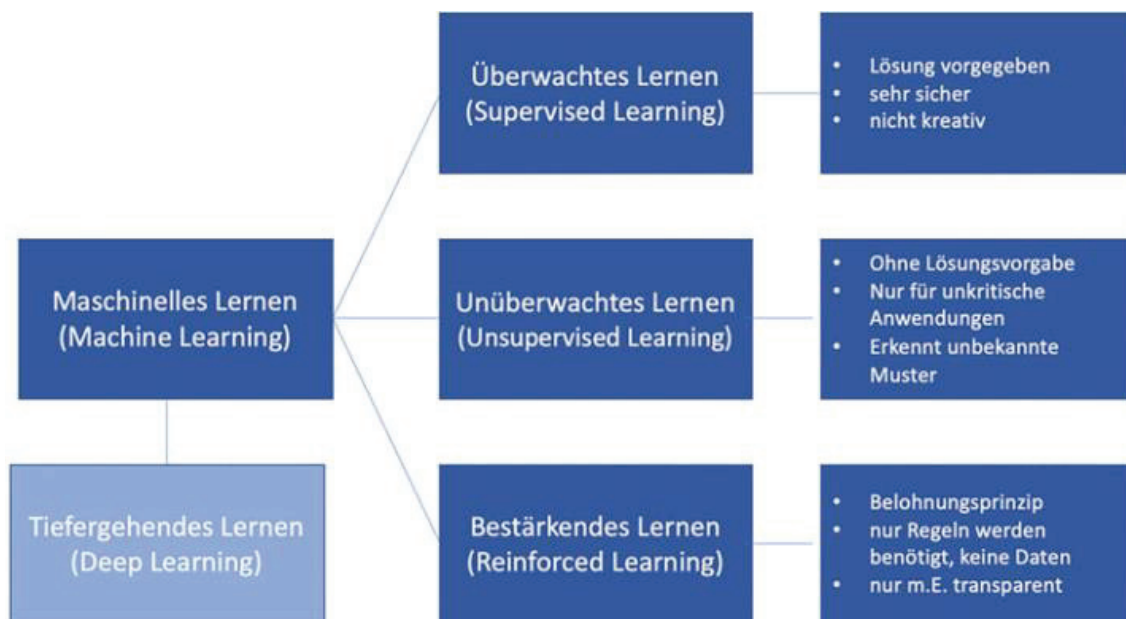


Abbildung 9: Methoden des Maschinellen Lernens (Mockenhaupt & Schlagenhauf, S. 175)

Die unterschiedlichen Trainingsverfahren können sowohl beim herkömmlichen maschinellen Lernen als auch bei künstlichen neuronalen Netzen und dem dazugehörigen Deep Learning eingesetzt werden (vgl. Lenzen 2020, S. 33 ff.). Wichtig anzumerken ist auch, dass beim Training von KI-Modellen mehrere unterschiedliche Trainingsverfahren kombiniert werden können, wie

---

<sup>4</sup> Um zu sehen, zu welchem imposanten Resultaten GANs imstande sind, können auf der Internetseite [www.thispersondoesnotexist.com](http://www.thispersondoesnotexist.com) eine schier unerschöpfliche Anzahl an Bildern von menschlichen Gesichtern eingesehen werden, die, wie die namensgebende Internetadresse andeutet, in der Realität nicht existieren, obwohl dies beim ersten Blick (und in manchen Fällen auch beim zweiten und dritten) den Anschein macht. Um ein neues Gesicht anzuzeigen zu lassen, muss die Seite einfach erneut aufgerufen beziehungsweise neu geladen werden.

es beispielsweise bei Transformer-Modellen in der Regel der Fall ist, wodurch beim Einsatz in der Praxis oftmals bessere Ergebnisse erzielt werden können (vgl. Ertel 2020, S. 399).

Überwachtes Lernen beginnt typischerweise mit einem bereits etablierten Datensatz und einem gewissen Verständnis darüber, wie diese Daten zu klassifizieren sind (vgl. Lenzen 2020, S. 51). Ziel ist es nun, Muster in den Daten zu finden, damit diese für analytische Prozesse in Bezug auf neue Daten genutzt werden können (vgl. Kreutzer 2023, S. 15). Anhand zuvor erarbeiteter Beispiele werden die Algorithmen trainiert und ihre Leistung kann dementsprechend bewertet werden (vgl. Hurwitz & Kirsch 2018, S. 15). Problematisch kann das sogenannte ‚Overfitting‘ werden, wenn das trainierte KI-Modell sich genau auf die Trainingsdaten abgestimmt hat, aber sich nur schlecht oder fehlerbehaftet auf unbekannte Daten anwenden lässt (vgl. Taulli 2022, S. 88). Um ‚Overfitting‘ zu vermeiden, müssen Modelle mit unbekannt beschrifteten Daten getestet werden, sodass die Genauigkeit des Modells bei der Vorhersage von Ergebnissen evaluiert werden kann (vgl. Sheikh et al. 2023, S. 37).

Beim unüberwachten Lernen wird der KI ein großer Datensatz vorgesetzt, der keinerlei Klassifikationen oder Beschreibungen besitzt (vgl. Lenzen 2020, S. 40 f.). Aufgabe der Algorithmen ist es nun, innerhalb dieses Datensatzes eigenständig Muster zu entdecken (vgl. Lehmann 2024, S. 166). Der Vorteil an der fehlenden Zielvorgabe dieses Lernverfahrens ist es, dass die KI möglicherweise verborgene und unvorhergesehene Muster zu Tage fördern kann, die unter Umständen als kreativ beziehungsweise neuartig bezeichnet werden könnten (vgl. Lenzen 2020, S. 41). Ein Problem, was bei allen Lernformen auftreten kann – insbesondere beim unüberwachten Lernen durch die fehlende Überwachung –, ist die sogenannte ‚Simulation Gap‘, die den „Schritt von der idealistischen Simulation über Trainingsdaten in die weniger ideale ‚schmutzige‘ Realität mit allen möglichen Nebenbeeinflussungen“ (Mockenhaupt & Schlagenhaut 2024, S. 177) beschreibt, wobei es zu fehlerhaften Interpretationen seitens der KI beim Einsatz in der Praxis kommen kann.

Das bestärkende Lernen ist eine Lernmethode, die auf behavioristischem Lernen fußt (vgl. Hurwitz & Kirsch 2018, S. 15). Hierbei wird der KI zwar ein Ziel vorgegeben, jedoch wird sie weder angeleitet noch wird ihr ein idealer Lösungsweg vorgegeben und einen Datensatz, anhand dessen die KI lernen kann, gibt es auch nicht (vgl. Lenzen 2020, S. 41). Die KI probiert eigenständig verschiedene Lösungswege innerhalb eines sich wiederholenden Trial-and-Error-Prozesses aus und sammelt hierüber Erfahrung, die sie sich für weitere Versuche nutzbar macht (vgl. Sheikh et al. 2023, S. 37). Gesteuert wird dieser Prozess durch häufig angewendete Belohnung und Bestrafung, sodass im Trainingsverlauf Lösungsideen, die zu unerwünschten Ergebnissen führen,

verworfen werden, während auf den erfolgreichen Lösungsansätzen aufgebaut wird (vgl. Lenzen 2020, S. 41 f.). Durch dieses Verfahren lernt die KI ihr Verhalten aufgrund der gemachten Erfahrung zu verfeinern (vgl. Moring 2023, S. 57). Bestärkendes Lernen eignet sich besonders, wenn nur wenige Trainingsdaten vorhanden sind oder das ideale Ergebnis nicht klar definiert werden kann (vgl. Mockenhaupt & Schlagenhaut 2024, S. 177). Auch kann es gut zur Anwendung kommen, wenn das Lernen erst durch Interaktion mit der Umwelt möglich ist, wie es beispielsweise bei über KI-Systeme gesteuerten Robotern der Fall ist (vgl. Kreutzer 2023, S. 19). Wie in Kapitel 3.1. besprochen wurde, konnte die das Brettspiel Go spielende KI ‚AlphaGoZero‘ eine überragende Leistung erlangen. Dies gelang ihr, indem sie unter anderem über bestärkendes Lernen in Abermillionen Spielen gegen sich selbst antrat und so immer besser werden konnte (vgl. Lenzen 2020, S. 42). Schließlich gelang es ihr, weltbeste Ergebnisse zu erzielen, ohne dass hierfür ein Datensatz als Grundlage oder menschliche Eingriffe während des Trainings vonnöten gewesen wären (vgl. Kreutzer 2023, S. 20).

### 3.2.3 Hybride KI

Unter hybriden KI-Systemen versteht man einen holistischen Ansatz, der symbolische und subsymbolische KI-Verfahren vereint (vgl. Ertel 2020, S. 11). Aus epistemologischer Sicht entspricht dieses System in etwa der Idee des menschlichen Organismus, der Zugang zu Weltwissen und logischem Schließen hat und dies mit seiner Wahrnehmung und der Verarbeitung der Informationen, die daraus gewonnen werden können, verbinden kann (vgl. Mainzer & Kahle 2022, S. 6). Deshalb wird dieser Art der KI im Vergleich zu rein symbolischen oder subsymbolischen Verfahren auch der größte Grad an Intelligenz beigemessen (vgl. Abb. 10) (vgl. ebd.).

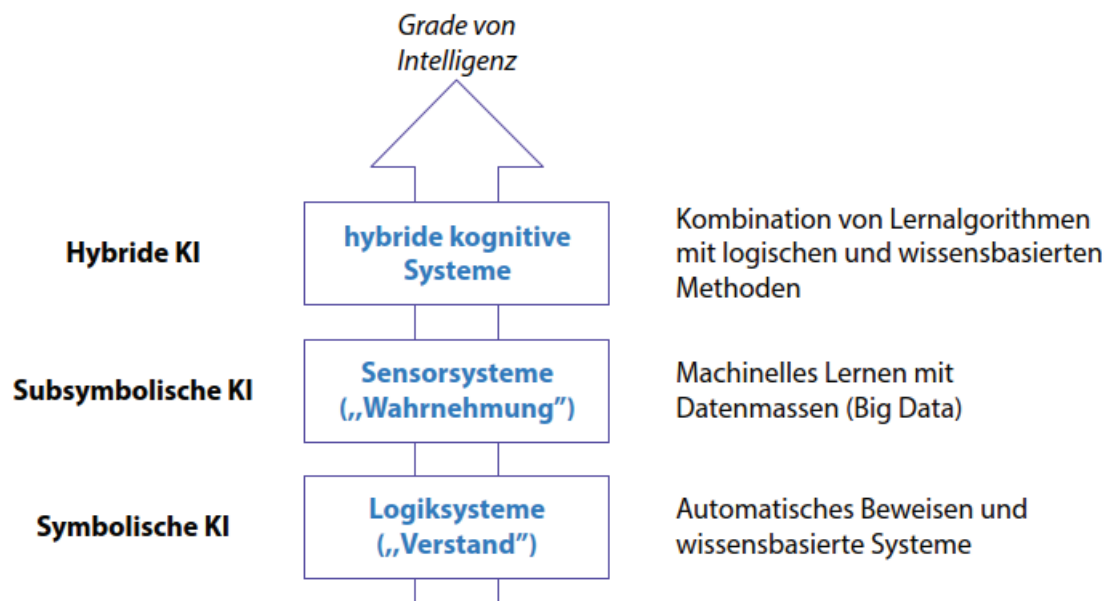


Abbildung 10: Von der symbolischen zur subsymbolischen zur hybriden KI (Mainzer & Kahle 2022, S. 5)

Die Idee dahinter ist, das Beste beider Welten zusammenbringen zu können, indem die symbolische KI das Spielfeld absteckt, auf dem sich die subsymbolische KI bewegt, um dort ihre Leistung zu verfeinern und ihre Tätigkeit zu verrichten (vgl. Lenzen 2020, S. 49). Die Vorteile, die sich daraus ergeben, lassen sich wie folgt beschreiben: Durch den Einsatz hybrider KI erhofft man sich eine erhöhte Erklärbarkeit und Transparenz, da die maschinell lernenden Verfahren beispielsweise auf einen Datensatz zugreifen, der durch eine komplett einsehbare symbolische KI bereitgestellt wird (vgl. Mockenhaupt & Schlagenhaut 2024, S. 64). Mittels symbolischer KI können so Grenzen gesetzt werden, die subsymbolische KI zu berücksichtigen hat. Dies könnte beispielsweise Geschwindigkeiten betreffen, die nicht überschritten werden dürfen, Bereiche, die nicht betreten werden sollen, oder Bewegungen und Aktionen, die zu unterlassen sind (vgl. Lenzen 2020, S. 49 f.). Auch das KI-System ‚Watson‘, was in Kapitel 3.1 erwähnt wurde, war ein hybrides KI-System, das den Wissensschatz einer symbolischen Expertensystem-KI mit maschinellem Lernen verknüpfte (vgl. Ertel 2020, S. 23).

Der Begriff hybride KI ist nicht immer ganz trennscharf abgegrenzt. So kann hybride KI neben der Verbindung symbolischer und subsymbolischer KI auch KI-Modelle beschreiben, die mehrere unterschiedliche maschinell lernende Verfahren in Kombination nutzen (vgl. Kreutzer 2023, S. 246).

## 4 Künstliche Intelligenz in der schriftlichen Onlineberatung

Seit Jahren gewinnt die Digitalisierung für die Soziale Arbeit immer mehr an Bedeutung, wobei nicht mehr nur die Lebenswelten von Kindern, Jugendlichen und Familien sowie die daraus entstehenden medien- und sozialpädagogischen Herausforderungen im Fokus stehen, sondern auch die Implikationen für und Auswirkungen auf Arbeits- und Organisationsprozesse zunehmend in den Vordergrund rücken (vgl. Seelmeyer & Ley 2018, S. 658). Dem professionellen Handeln in der Sozialen Arbeit ist inhärent, dass es durch berufsspezifische und gesellschaftliche Rahmenbedingungen beeinflusst wird (vgl. Campayo 2020, S. 290). Angesichts der zunehmenden Bedeutung digitaler Technologien im Alltag und in den verschiedenen Arbeitsbereichen der Sozialen Arbeit ist die Beeinflussung des professionellen Handelns durch die fortschreitende Digitalisierung beziehungsweise Mediatisierung unausweichlich, wodurch sich in Konsequenz „Anlässe, Formen und Rahmenbedingungen der Erbringung sozialer Dienste [verändern]“ (Kutscher 2019, S. 42). Gravelmann (2024a) postuliert diesbezüglich, dass „die digitalen Entwicklungen allesamt enorme Spuren in sämtlichen gesellschaftlichen Feldern sowie im Leben eines jeden Einzelnen hinterlassen [haben], [weshalb] kein Mensch und kein Arbeitsfeld, keine Institution und keine Gesellschaft [umhinkommt], sich diesen Veränderungen mit den vielfältigen Chancen und zugleich erheblichen Risiken zu stellen“ (S. 18).

Hat die digitale Transformation in der Sozialen Arbeit zu Beginn vor allem digitale Assistenzsysteme betroffen, die den Arbeitsalltag beispielsweise durch digitale Werkzeuge wie Abrechnungs- und Dokumentationssysteme verändert haben, „berührt der Einsatz digitaler Technologien in der Sozialen Arbeit jedoch zunehmend auch den Kern interaktiver Dienstleistungserbringung“ (Seelmeyer & Waag 2020, S. 180). Diesbezüglich wäre die Onlineberatung als explizites Beispiel zu nennen (vgl. Gutwald et al. 2021, S. 6). Rehme & Seelmeyer (2023) betonen in dieser Hinsicht, eine dichotome Ausdifferenzierung zwischen Online- und Offlineberatungssettings sei im Zuge des digitalen Wandels aus fachlicher Sicht nicht mehr als zeitgemäß zu erachten, da Beratung sich einer breiten Palette analoger und digitaler Werkzeuge sowie Kommunikationsformen bedienen könne, um Ratsuchenden der Situation adäquat zugeschnittene Unterstützung zukommen lassen zu können (vgl. S. 23). Wie Kapitel 2.5 zeigen konnte, hat Digitalisierung und Mediatisierung insbesondere in der Arbeit mit Kindern und Jugendlichen eine hohe Relevanz, da diese neue Technologien häufig schnell in ihren Lebensalltag integrieren und sich an neue technologische Entwicklungen schnell anpassen (vgl. Gravelmann 2024a, S. 18). Neben der

Onlineberatung in der Arbeit mit Kindern und Jugendlichen stellt auch ‚Webwork‘<sup>5</sup>, also digitale ‚Streetwork‘, eine digitale Dienstleistung dar, in der Fachkräfte sich mittels Social-Media in die digitale Lebenswelt Kinder- und Jugendlicher begeben, um dort Präsenz zu zeigen und Zugänge zur Sozialarbeit zu eröffnen (vgl. Gravelmann 2024a, S. 24). Andere Bereiche der Sozialen Arbeit hingegen haben sich erst im Laufe der Zeit mit der Digitalisierung angefreundet (vgl. Kutscher 2019, S. 42).

Generell lässt sich feststellen, dass Deutschland, was den Fortschritt der Digitalisierung betrifft, keine Vorreiterrolle einnimmt (vgl. Heister & Krings 2024, S. 41). Diesbezüglich konstatiert Gravelmann (2024b): „Die Akteure und Arbeitsfelder der Sozialen Arbeit hinken den rasanten Entwicklungen der Digitalisierung und den Handlungsbedarfen beziehungsweise Handlungsoptionen oft hinterher“ (S. 9). Allerdings hat sich nicht zuletzt die vergangene Covid-19-Pandemie hinsichtlich der Digitalisierung als Treiber erwiesen und zu hastig umgesetzten Umstrukturierungs- und Veränderungsprozessen geführt (vgl. Lehmann 2024, S. 169). Görder (2021) hebt daran anknüpfend hervor, dem Einsatz von Digitaltechnik in der Sozialen Arbeit würde häufig Skepsis entgegenschlagen und mit einer Abwehrhaltung begegnet, was einen durchaus nicht unproblematischen Umstand darstellt, da digitale Technik zum einen den Menschen stark beeinflusst und vice versa (vgl. S. 4). Aus diesem Grund seien Mensch und Technik miteinander verknüpft und nicht als getrennt zu betrachtende Bestandteile in zwei unterschiedlichen Welten zu verorten (vgl. ebd.). Andere hemmende Faktoren von Digitalisierung in der Sozialen Arbeit werden in mangelnder technologischer Infrastruktur sowie im Fachkräftemangel ausgemacht, der dazu führt, dass der Fokus auf pädagogischen Kerntätigkeiten liegen muss und Fachkräften kaum Raum und Zeit zur Aus- und Weiterbildung zur Verfügung steht (vgl. Linnemann et al. 2023b, S. 416). Gleichzeitig könnte der Einsatz von innovativer digitaler Technologie aber auch gewinnbringenden Einsatz versprechen, der in letzter Konsequenz Akteurinnen und Akteure in der Sozialen Arbeit entlasten könnte, um so dem großen Fachkräftemangel in der Sozialwirtschaft entgegenwirken zu können (vgl. Lehmann 2024, S. 164).

In den hier nur sehr kurz skizzierten technologischen Wandel fügt sich KI nahtlos ein, die im Alltag bereits seit einiger Zeit in einem großen Spektrum an Einsatzbereichen, in Form von

---

<sup>5</sup> Zur weiteren Vertiefung: Stieler, M. & Zauter, S. (2022). Digital Streetwork. Aufsuchende Arbeit mit und in Sozialen Medien. *FORUM Sozialarbeit + Gesundheit*, 4. Abgerufen am 09.07.2024: <https://www.e-beratungsinstitut.de/wordpress/wp-content/uploads/2022/10/2022-04FORUMStielerZauter.pdf>; Hey, P., Schüürmann, S., Stieler, M., & Lehmann, R. (2022). Erfolgsfaktoren des Projektes Streetwork im Netz – Erste Ergebnisse des Modellprojekts zur Qualitätssicherung der webbasierten aufsuchenden Sozialarbeit von Condrops e.V. *Suchttherapie*, 23(S 01), S39\_1. <https://doi.org/10.1055/s-0042-1756092>



Smartspeakern, Suchmaschinen, Empfehlungsalgorithmen, Wortvervollständigungen, Mustererkennungsalgorithmen, Organisationsapps und Ähnlichem, Anwendung findet (vgl. Gravelmann 2024a, S. 79). Es ist anzunehmen, dass KI häufig als gängiger Teil des Alltags wahrgenommen wird und sich Nutzende daher nicht stets darüber bewusst sind, wenn sie mit einer KI in Berührung kommen (vgl. Poltermann et al. 2024, S. 22). Davon sei auch die Soziale Arbeit nicht ausgenommen, merkt Görder (2021) an, weshalb er davon ausgeht, KI und Algorithmen würden auch in diesem Bereich in den Arbeitsalltag „weitgehend unbemerkt und unreflektiert [...] Einzug [...] halten“ (S. 3). Gleichzeitig fällt bei der Recherche auf, „dass der Einsatz von KI [...] vielerorts weder in Projekten organisiert, noch (standardisiert) evaluiert wird“ (Linnemann et al. 2023b, S. 415), weshalb eine präzise und lückenlose Erhebung des Status Quo vom Einsatz von KI in der Sozialen Arbeit in Deutschland zum momentanen Zeitpunkt schwerlich realisierbar erscheint (vgl. ebd.). Neben von Fachkräften während der Arbeit privat genutzten Applikationen, die, wie am Anfang des Absatzes beschrieben wurde, KI beinhalten können, wären beispielsweise auch am Arbeitsplatz genutzte Anwendungssysteme zu nennen, die durch ihre Handhabung vor allem administrative und dokumentarische Aufgaben erleichtern und von Sozialarbeitenden, wie beschrieben, nicht unbedingt als KI ausgemacht werden (vgl. Heister & Krings 2024, S. 43).

Doch auch die bewusste Implementierung KI-gestützter digitaler Hilfesysteme, die in Kernprozesse sozialarbeiterischer Praxis eingreifen, kann beobachtet werden (vgl. Gutwald et al. 2021, S. 6). Während solche Systeme in Deutschland zum gegenwärtigen Zeitpunkt aufgrund von Datenschutzbedenken und -verordnungen nicht denkbar wären, lassen sich auf internationaler Ebene einige Beispiele aufzählen (vgl. Poltermann et al. 2024, S. 22). Görder (2021) benennt hier vor allem Systeme, die datengetrieben zur Risikovorhersage von Kindeswohlgefährdungen eingesetzt werden, aber auch zur Suizidprävention, zur Unterstützung Arbeitssuchender oder zur Überprüfung von möglicherweise erschlichenen Sozialleistungen kommt KI außerhalb von Deutschland zum Einsatz<sup>6</sup> (vgl. S. 3). Die Daten, die KI in der Sozialen Arbeit verarbeiten könnte, um Vorhersagungen zu treffen oder neue Muster und Informationen zu generieren, könnten Stammdaten, Artefakte, (handschriftliche) Notizen, Kontaktinformationen inklusive Dauer und Häufigkeit, Kosten und viele weitere sein (vgl. Linnemann et al. 2023a, 198 f.). Auch wenn derlei KI-Analyse- und Vorhersagesysteme in Deutschland derzeit nicht in der Praxis benutzt werden,

---

<sup>6</sup> Aktuelle Betrachtungen von Entscheidungsalgorithmen und Risikovorhersagesystemen im internationalen Einsatz: Burghardt & Lehmann 2023, S. 410 ff.; Gravelmann 2024a, S. 82 f.; Gravelmann 2024b, S. 12; Gutwald et al. 2021, S. 7 ff.

geben Linnemann et al. (2023b) zu bedenken, dass es vermutlich nur eine Frage der Zeit sei, bis solche oder ähnliche Systeme auch in Deutschland zur Anwendung kommen würden, alleine um der Wohlfahrtsarbeit zur Hand gehen zu können, die mit spärlichen Ressourcen zu haushalten habe (vgl. S. 416). Aufgrund der jüngsten Fortschritte im Bereich von KI erscheint es Burghardt & Lehmann (2023) zum gegenwärtigen Zeitpunkt deshalb unstrittig, dass diese KI-Systeme in naher Zukunft Veränderungen anstoßen und nachhaltigen Einfluss auf die Soziale Arbeit nehmen werden (vgl. S. 413). Neben einiger Chancen, die man sich durch einen Einsatz von KI vor allem hinsichtlich genauerer Entscheidungsfindungen, effizienterer Arbeitsprozesse und als Arbeitshilfe erhofft, wird die Überlegung der Implementierung solcher Entscheidungsalgorithmen und vergleichbarer KI-Systeme naturgemäß unter Betrachtung ethischer, professioneller und datenschutzrechtlicher Gesichtspunkte mit einem kritischen Diskurs begleitet (vgl. ebd.).

Besonders für Fachkräfte der Sozialen Arbeit, die mit Kindern und Jugendlichen arbeiten, ist KI ein Thema von großer Relevanz. Wie die Nutzungszahlen von ChatGPT gezeigt haben (vgl. Kap. 2.5), werden Kinder der aktuell heranwachsenden Generation einen spielend natürlichen Umgang mit KI erlernen, so wie es die ‚Digital Natives‘ mit digitaler Technologie zuvor getan haben. „Ein Verständnis für die Funktionsweise von KI zu entwickeln, ihre Potenziale zu fördern und den kritischen Umgang mit ihr zu üben“ (Wilke et al. 2023, S. 423), stellt daher eine gesamtgesellschaftliche Herausforderung dar, der aber insbesondere auch Fachkräfte im Umgang mit Kindern und Jugendlichen gerecht werden müssen. Einerseits sollten sich Fachkräfte der Chancen, Grenzen und Risiken von KI für ihre Arbeit bewusst sein und andererseits müssen sie befähigt sein, Kindern und Jugendlichen im Digitalisierungskontext fachkundige Unterstützung angemessen zukommen zu lassen (vgl. Gravelmann 2024b, S. 9). Linnemann et al. (2023b) betonen in dieser Hinsicht die Bedeutsamkeit der Begleitung von Kindern bei der Nutzung von KI, weil sie aus entwicklungspsychologischer Sicht erst sukzessive lernen, Realität und Fiktion voneinander zu unterscheiden (vgl. S. 421). Da Menschen ohnehin dazu tendieren, kommunikationsfähige technische Systeme als soziale Akteure zu betrachten, ist dieser Effekt bei Kindern folglich noch um einiges größer, weshalb ihnen in dieser Hinsicht besondere Aufmerksamkeit zuteilwerden sollte (vgl. ebd.).

Engelhardt (2023a) weist darauf hin, dass dem Einsatz von KI in der Beratung in den vergangenen Jahren im Diskurs bisher eine eher zu vernachlässigende Rolle beigemessen wurde (vgl. S. 404). Zwar wurden der Einsatz und ein möglicher Nutzen von Chatbots diskutiert, ansonsten sah man die Einsatzmöglichkeiten jedoch als begrenzt an (vgl. Linnemann et al. 2023a, S. 206). Allerdings haben sich, angestoßen durch die Veröffentlichung des KI-Modells ChatGPT-3.5 Ende des Jahres 2022 und der damit einhergehenden Zugänglichmachung für die breite Öffentlichkeit, auch

neue Fragen, Horizonte und Anwendungsmöglichkeiten von KI für die Profession der (Online-)Beratung ergeben (vgl. Engelhardt 2023a, S. 404). Mittlerweile zeichnet sich ein mannigfaltiges Spektrum an Einsatzmöglichkeiten von KI im Kontext der Onlineberatung ab, das weit über den bloßen Einsatz von Chatbots für Ratsuchende hinausgeht (vgl. Kap 4.1). Es deutet sich an, dass auch Beratende sich LLMs zunutze machen können. Zu nennen wären hier beispielsweise Textverarbeitungsfunktionen, aber auch als assistives Tool in der Evaluation von erfolgter oder gerade erfolgreicher Kommunikation (vgl. Kap. 4.2) kann diese Art der KI Anwendung finden (vgl. Linnemann et al. 2023a, S. 202). Erste Gehversuche deuten zudem daraufhin, dass Chatbots zukünftig in der Ausbildung von Beratenden zum Einsatz kommen könnten, wie in Kapitel 4.3 dargestellt werden wird.

#### **4.1 Chatbots**

Mit Joseph Weizenbaums ELIZA erblickte 1966 einer der ersten Chatbots das Licht der Welt (vgl. Kap. 3.1). Aufgrund seiner auf symbolischer KI beruhenden Natur konnte er in einem sehr begrenzt abgesteckten Rahmen auf Texteingaben Nutzender reagieren, um so eine, an das Psychotherapeutengespräch angelehnte, Konversation mit ihnen zu führen (vgl. Lenzen 2020, S. 22). Seit diesem ersten Auftreten eines Chatbots haben sich die Möglichkeiten, die sie bieten, erheblich weiterentwickelt (vgl. Waag et al. 2020, S. 181).

Bei Chatbots, die auch unter dem Namen ‚Conversational Agents‘ geläufig sind, handelt es sich um Programme oder KI-Modelle, mit denen Menschen über natürliche Sprache kommunizieren können (vgl. Rapp et al. 2021, S. 1). In der Praxis geschieht dies zumeist über eine Sprachschnittstelle mittels geschriebener Worte, die der Chatbot als Eingabe vom Nutzenden erhält und dem Nutzenden ebenso in selbiger Form ausgibt (vgl. Griol et al. 2013, S. 760). Wie in Abbildung 11 ersichtlich wird, organisiert der Chatbot über das Dialogmanagement vorher verarbeitete Texte und bezieht dabei den bisherigen Dialogverlauf mit ein, greift auf Wissensdatenbanken zu und lässt vordefinierte Skripte ablaufen, sofern diese während der Programmierung eingespeist wurden (vgl. Bendig et al. 2019, S. 270).

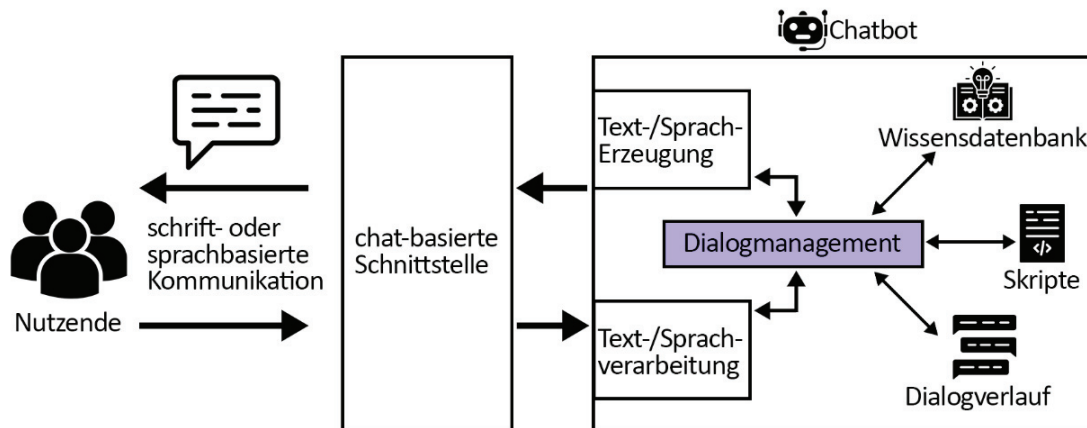


Abbildung 11: Technische Umsetzung eines Chatbots (eigene Darstellung in Anlehnung an Bendig et al. 2019, S. 270)

Manche Chatbots, wie beispielsweise ChatGPT-4o, können jedoch auch gesprochene Sprache als Ein- und Ausgabe nutzen, wenn sie befähigt sind, gesprochene Sprache transkribieren und generieren zu können (vgl. Frischholz 2024). In ihrer Funktionsweise können sie dadurch einem Smartspeaker ähneln. „Chatbots erhalten damit eine soziale Qualität, indem menschliche Kommunikation bis hin zu Beziehungsführung imitiert wird“ (Steiner & Tschopp 2022, S. 469), weswegen es Nutzenden, wie in Kapitel 3.1 anhand von Google Duplex deutlich wurde, mitunter schwerfallen kann, zu erkennen, ob es sich bei dem Gegenüber um einen Menschen handelt oder nicht (vgl. Mockenhaupt & Schlagenhaut 2024, S. 77).<sup>7</sup>

Smartspeaker wie Apples ‚Siri‘ und Amazons ‚Alexa‘ beförderten die Mensch-Computer-Interaktion im vergangenen Jahrzehnt in den Fokus der Öffentlichkeit und diese ist heute Teil der Lebensrealität vieler Menschen, wie die Anzahl Nutzender von Chatbots oder Smartspeakern zeigt (vgl. Kap. 2.5; vgl. Mornhinweg 2018, S. 144). Obwohl sich Chatbots schon seit einigen Jahren in reger Verwendung befinden, erlangten diese spätestens seit der Veröffentlichung von ChatGPT-3.5 im November 2022 einen massiven Bekanntheitsschub, sodass heutzutage mehr als die Hälfte der deutschen Bevölkerung weiß, worum es sich bei einem Chatbot handelt (vgl. Beisch & Koch 2023, S. 1). Zudem gaben in einer kürzlich durchgeführten Onlinebefragung annähernd

<sup>7</sup> Im Folgenden sollen Chatbots jedoch im Sinne von ‚Conversational Agents‘ ausschließlich auf ihre Fähigkeit, sich mit Nutzenden in eine textbasierte Interaktion begeben zu können, reduziert werden, um in diesem Bereich ihre Potenziale und Risiken für die Praxis in der Onlineberatung untersuchen zu können.

neun von zehn der 1541 Befragten an, schon einmal Kontakt mit einem Chatbot gehabt zu haben (vgl. Fokina 2024).

Heutzutage befinden sich Chatbots vor allem auf Internetseiten, in Smartphone-Apps oder in Sozialen Netzwerken als interaktiver Dienstleistungsservice in breiter Anwendung (vgl. Mornhinweg 2018, S. 144; vgl. Waag et al. 2020, S. 184). Diesen Chatbots liegt mehrheitlich symbolische KI zugrunde, weshalb sie oft nur in der Lage sind, nach erfolgter Textanalyse, auf Schlagwörter reagierend, passende Informationen für die Nutzenden per vordefinierter Textbausteine auszugeben (vgl. Waag et al. 2020, S. 183). Mornhinweg (2018) sieht eine Schwierigkeit dieser Chatbots darin, „dass sie sich aufgrund ihrer begrenzten Fähigkeiten oft verlaufen und in die Leere gehen“ (S. 147), weshalb Chatverläufe in Schleifen ablaufen können. Dies kann auf Nutzende eine unbefriedigende Wirkung haben, wenn der Gesprächsverlauf aufgrund der möglicherweise bestehenden Limitierung des Chatbots wiederkehrend an den Anfangspunkt des Chats zurückgesetzt wird (vgl. ebd.). In ihrer Steuerung können sie Nutzenden daher manchmal wie interaktive und kommunizierende Volltext-Suchmaschinen erscheinen (vgl. Bendig et al. 2019, S. 267). Gleichzeitig liegt dort aber auch eine Stärke von Chatbots, da sie für Nutzende durch ihre interaktive Steuerung und Benutzerfreundlichkeit vergleichsweise leicht zu handhaben sind, was eine niedrige Einstiegshürde bietet und somit digitale Unterstützungsprozesse einfacher angebahnt werden können (vgl. Martin & Pengel 2024, S. 25). So konnten sich Chatbots in der Selbsthilfebehandlung von Depression beispielsweise als effektiver erweisen als Bibliothekstherapie (vgl. Liu et al. 2022, S. 5).

Chatbots wie ChatGPT, die auf großen Sprachmodellen beruhen und subsymbolischer Natur sind, verfügen im Vergleich zu regelbasierten Chatbots häufig über einen größeren Funktionsumfang (vgl. Kap. 3.2.2.1), können komplexere und offenere Unterhaltungen führen und sogar Emotionen erkennen, auf das Geschriebene Nutzender umfassender eingehen, anhand des Chatverlaufs fortlaufend lernen und auf Nachfragen mittels sogenannter ‚Prompts‘ (Eingabeaufforderungen) präzisere Ausgaben produzieren (vgl. Bendig et al. 2019, S. 269; vgl. Ekin 2023, S. 3 ff.; vgl. Waag et al. 2020, S. 183). Momentan scheint ChatGPT eine deutliche Vormachtstellung eingenommen zu haben und wird außerdem fortlaufend weiterentwickelt. So hat ChatGPT-3.5 im Anschluss an seine Veröffentlichung zwei konsequente Nachfolgermodelle mit ChatGPT-4 und ChatGPT-4o erhalten, die die ursprünglich veröffentlichte Version 3.5 in seiner Performanz und Leistungsfähigkeit deutlich übertreffen (vgl. Frischholz 2024). Aber es handelt sich nicht um das einzige geläufige große Sprachmodell. Konkurrenz wäre beispielsweise in Googles Gemini oder Metas Llama auszumachen, die bei Entwickelnden sehr beliebt sind, da sie im Gegensatz zu ChatGPT über einen Open-Source Code verfügen (vgl. Cardillo 2024). Der sich abzeichnende

Trend geht deutlich weg von symbolischer KI hin zu großen Sprachmodellen, denen ein enormes Marktwachstumspotenzial im Laufe der kommenden Jahre bescheinigt wird (vgl. Dimension Market Research 2024). Ganz grundsätzlich erhofft man sich im weiteren Verlauf der Entwicklung von Chatbots weitere technische Innovation, bessere Performanz, ein vergrößertes Anwendungsspektrum sowie mehr Effizienz hinsichtlich des Trainings und eine höhere Zuverlässigkeit in Bezug auf die produzierten Textausgaben (vgl. Yahoo Finance 2024).

Es wird deutlich, dass auf symbolischer KI basierende Chatbots eher von simpler Ausgestaltung und für einen sehr speziellen Zweck programmiert sein können. Generative Chatbots hingegen können hochkomplexe KI-Modelle mit einem großen Funktionsumfang und breiteren Einsatzmöglichkeiten darstellen.

Waag et al. (2020) halten die Anwendung von Chatbots in der Onlineberatung für grundsätzlich gangbar und machen Einsatzpotenzial unter anderem in dem Bereich „Beratung als Wissensweitergabe“ (S. 181) aus. Typisch für die Beratung als Wissensweitergabe ist, dass Ratsuchende Zugang zum zur Verfügung stehenden Expertenwissen der Beratenden erhalten, Kontakte tendenziell eher kurzgehalten sind und diese zumeist nach erfolgter Wissensweitergabe enden (vgl. ebd., S. 182). Die schriftbasierte Interaktion ermöglicht es, Beratende durch Chatbots „zumindest im Bereich der Wissensweitergabe mitunter sogar [zu] ersetzen“ (ebd., S. 183), denn wie in Kapitel 2.1 anhand der Definition von Beratung ersichtlich wurde, kann „die beratende Seite [auch] [...] ein Softwareprogramm sein.“ (Nestmann & Sickendiek 2018, S. 110). Vor allem für die Wissensweitergabe zu schambehafteten Themen scheinen Chatbots prädestiniert zu sein. So lassen sich in dem unter anderem vom Bundesamt für Familie, Senioren, Frauen und Jugend geförderten Bot ‚Lovis‘, dem von der gemeinnützigen Organisation Planned Parenthood betriebenen Bot ‚Roo‘ und dem von der Nichtregierungsorganisation Population Foundation of India entwickelten Bot ‚SnehAI‘ bereits mehrere Chatbots finden, die Nutzenden die Möglichkeit zur sexuellen Aufklärung bieten (vgl. FragLovis 2024; vgl. Planned Parenthood 2024; vgl. SnehAI 2021). Führt der Einsatz von Chatbots eine größere Anzahl an möglichen Klient:innen Beratung zu, die im Falle der drei genannten Chatbots vor allem bei jüngeren Nutzenden zu verringertem risikohaften Sexualverhalten beziehungsweise sexueller Aufgeklärtheit beitragen kann, könnten sich eingesetzte Chatbots als gesamtgesellschaftlicher Gewinn herausstellen (vgl. Wang et al. 2022, S. 12 ff.).

Auch als Orientierungshilfe im sozialen Kontext könnten sich Chatbots als hilfreich erweisen. Dort könnten sie zum Beispiel in Bezug auf „Fragen zu Zuständigkeiten, Informationen zu Beratungs- und Hilfsangeboten, Unterstützung bei rechtlichen Fragestellungen oder Anleitungen

zum Ausfüllen von Formularen eine Rolle spielen“ (Gravelmann 2024b, S. 10 f.), und somit Ratsuchenden sowohl bei Fragen als auch bei Bearbeitungsaufgaben unterstützen. In dieser Hinsicht ist beispielsweise der Chatbot des Startups Limbic positiv hervorzuheben, der in Großbritannien innerhalb des Online-Services des staatlichen Gesundheitssystems NHS getestet wurde, um Überweisungen zu Gesprächstherapien zu vereinfachen (vgl. Brakemeier et al. 2024). In seiner Funktion stellt der Chatbot Nutzenden eine Reihe von Fragen, deren zugehörige Antworten dann in das System des NHS eingepflegt werden und für professionellen Fachkräften einsehbar sind (vgl. Williams 2024). Diese können Nutzende daraufhin kontaktieren (vgl. ebd.). Die Forschenden fanden heraus, dass der Chatbot im Vergleich zum ansonsten gebräuchlichen Webformular zu einer leicht erhöhten Überweisungsrate führte (vgl. Habicht et al. 2024, S. 596). Wie Habicht et al. (2024) zudem zutage förderten, profitierten insbesondere Menschen mit einer nichtbinären Geschlechtsidentität vom Chatbot, deren Überweisungsrate im Vergleich zu vorher um 179 Prozent anstieg (vgl. ebd.). Auch Menschen, die sich der Gruppe ethnischer Minderheiten zuordnen, kam der Chatbot zugute, denn deren Überweisungsrate konnte einen Anstieg von 29 Prozent verzeichnen (vgl. ebd., S. 597). Georg Schomerus, Direktor der Klinik und Poliklinik für Psychiatrie und Psychotherapie des Universitätsklinikums Leipzig, sieht einen großen Nutzen in der „vorurteilsfreie[n] Herangehensweise eines Chatbots, [weswegen] man [...] als Nutzer keine Angst haben [muss], aufgrund einer bestimmten Gruppenzugehörigkeit nicht ernst genommen zu werden“ (Brakemeier et al. 2024). Dies könnte vor allem die Intensität von psychologischen Barrieren, wie Scham, belastenden Grundannahmen und Ängsten vor Stigmatisierung mindern, die bei Menschen, die sich Minderheiten zugehörig fühlen, eine hohe Prävalenz aufweisen (vgl. ebd.). Ein solcher Einsatz von Chatbots könnte demnach vor allem Menschen begünstigen, die herkömmliche Beratungsangebote aus unterschiedlichsten Gründen nicht in Anspruch nehmen würden (vgl. Waag et al. 2020, S. 187).

Waag et al. (2020) sehen ein weiteres Einsatzpotenzial im Bereich „Beratung als Koproduktion von Lösungskompetenz“ (S. 181). In der Beratung als Koproduktion von Lösungskompetenz können Kontakte zwischen Ratsuchenden und Beratenden in einem längeren zeitlichen Rahmen und wiederkehrend stattfinden, wobei Ratsuchende als Expert:innen ihrer Wirklichkeit auftreten und Beratende ihnen bei der Lösungserarbeitung mittels beraterischer Methoden und Kompetenzen unterstützend zur Seite stehen (vgl. ebd., S. 182). Wie Filipiak (2016) hervorhebt, steht hier der Inhalt der Beratung nicht unbedingt von vornherein fest, denn ein gemeinsames koproductives Orientieren und Verständigen auf den Kern des Problems wird bereits als inhärenter Bestandteil des Beratungsprozesses erachtet und gilt oft als erster Schritt auf dem Lösungsweg (vgl. S. 6 f.). Auch Fulmer (2019) sieht in dem beschriebenen Bereich Einsatzpotenzial. An die

Definition von Beratung der American Counseling Association angelehnt, arbeitet er drei hauptsächliche Qualitäten von Beratung heraus, die er in der Beziehung, dem Empowerment und der Lösungserreichung ausmacht (vgl. Fulmer 2019, S. 811). Aus seiner Sicht sei mit Chatbots zum gegenwärtigen Zeitpunkt nur die Hilfe zur Lösungserreichung realistisch, während Empowerment ungewiss sei und Beziehungsaufbau nicht stattfinde, was aus seiner Sicht erst mit dem Aufkommen starker beziehungsweise allgemeiner KI realisierbar erscheint (vgl. ebd., S. 814).

Zu einem grundsätzlich anderen Ergebnis hinsichtlich des Beziehungsaufbaus gelangen Skjuve et al. (2021), die in ihrer Untersuchung nachweisen konnten, dass Teilnehmende mit dem Social-Chatbot ‚Replika‘, der Nutzenden als eine Art virtueller Freund dienen soll, Beziehungen aufbauen können, die Gemeinsamkeiten mit Mensch-zu-Mensch-Beziehungen aufweisen (vgl. S. 13). Diese Art von Beziehung, die von Menschen mit Maschinen oder KI eingegangen werden, wird quasisozial genannt (vgl. Linnemann et al. 2023a, S. 202). Social-Chatbots werden als Chatbots mit emotionalen Kapazitäten beschrieben, die Fähigkeiten vereinen sollen, die es ihnen ermöglichen, die Gefühle Nutzender zu erkennen, aufzugreifen und auf diese zu reagieren (vgl. Brandtzaeg et al. 2021, S. 1). „Menschen reagieren auf kommunikative Mitteilungen eines Chatbots grundsätzlich ganz ähnlich wie auf entsprechende Mitteilungen eines menschlichen Gegenübers, selbst wenn sie wissen, dass es sich um einen Bot handelt“ (Waag et al. 2020, S. 184). Es sei allerdings an dieser Stelle angemerkt, dass die von Skjuve et al. (2021) beschriebenen Beziehungen freundschaftlicher Natur waren und nicht therapeutischen Allianzen glichen, wie sie Fulmer (2019) in seiner Ausführung hinsichtlich von Beratung herausstellte. Darcy et al. (2021), die die Möglichkeit der Formung einer therapeutischen Allianz von Nutzenden mit einem Chatbot erforschten, kamen allerdings zu einem ähnlich positiven Ergebnis wie Skjuve et al. (2021). Ihre Ergebnisse zur ermittelten therapeutischen Allianz mit Chatbots waren vergleichbar hoch wie zur therapeutischen Allianz, die zwischen Menschen hergestellt werden kann (vgl. Darcy et al. 2021, S. 5). Gemäß der Überlegungen der Autor:innen könnte diesen Umstand begünstigt haben könnte, dass der Chatbot nicht versucht habe, sich als Mensch zu gerieren, um einen wie in Kapitel 3.1 beschriebenen Turing-Test zu bestehen, sondern seine maschinelle Natur den Nutzenden transparent gemacht habe (vgl. ebd.).

Skjuve et al. (2022), die in einer weiteren Studie die Beziehungen von Menschen zu Social-Chatbots untersuchten, betonen zudem, wie wichtig es für Nutzende sei, dass sie die Möglichkeit zur Selbstoffenbarung als gegeben empfinden, damit ein gelungener Beziehungsaufbau stattfinden könne (vgl. S. 8). Vor allem die gute Ausdrucksweise von generativen Chatbots kann bei Nutzenden oftmals den Eindruck von umfangreicher Kompetenz und Einfühlungsvermögen erwecken und Selbstoffenbarungseffekte weiter verstärken (vgl. Linnemann et al. 2024, S. 13). Wie in



Kapitel 2.4 gezeigt wurde, fällt es Menschen aufgrund der Kanalreduktion in der Onlineberatung leichter, sich zu öffnen, als im kopräsenten Setting. Lucas et al. (2014) konnten in ihrer Forschung zeigen, dass dieser Effekt noch einmal größer zu sein scheint, wenn das Gegenüber als nicht-menschlich wahrgenommen wird. So ergaben ihre Untersuchungen, dass Teilnehmende, die glaubten, mit einem Computer zu interagieren, von einer geringeren Furcht vor Selbstoffenbarung berichteten, was von außenstehenden Beobachtenden untermauert wurde, und zudem dazu tendierten, ihre Gefühle – in diesem Falle Traurigkeit – stärker zum Ausdruck zu bringen (vgl. Lucas et al. 2014, S. 98 f.). Damit zeigt sich beim Einsatz von Chatbots ein enormes Potenzial bei stigmatisierenden Themen, bei starker Scham oder für benachteiligte Menschen, für die das Ansprechen gewisser Themen ein Tabubruch darstellen würde oder sie hierfür Strafen beziehungsweise Verfolgung zu fürchten hätten (vgl. Haque & Rubya 2023, S. 2). Chatbots könnten also möglicherweise Zugangswege zur Beratung für Menschen eröffnen, die diese weder im kopräsenten Setting noch in Onlinesettings mit menschlichen Beratenden wahrnehmen würden.

Aus der Sicht von Skjuve et al. (2022) scheint eine gelingende Beziehung zwischen Mensch und Chatbot außerdem davon abzuhängen, ob ein Chatbot in der Lage ist, variierende Unterhaltungen zu führen, und beim Gegenüber sowohl Einsamkeit lindern kann als auch Selbstreflexion zu fördern weiß (vgl. S. 13). Gelingt dies, kann ein Chatbot für Nutzende ein emotionaler Rückhalt sein, der mit nützlichen Rat- und Vorschlägen aufwartet und sie unterstützt, sich selbst besser einzuschätzen und sich positiver zu betrachten (vgl. Ta et al. 2020, S. 7). Diese Erkenntnisse decken sich mit den Ergebnissen von Brandtzaeg et al. (2021), die die Auffassung junger Menschen über die emotional-sozial stärkenden Kapazitäten von Social-Chatbots untersucht haben. Aus ihren Ergebnissen geht hervor, dass die Befragten, die zwischen 16 und 21 Jahren alt waren, den niedrigschwelligen, zeit- und ortsungebundenen Zugang zu Chatbots schätzen, der ihnen einen geschützten Rückzugsraum bietet, in dem sie sich ohne Sorge vor Bewertungen öffnen können (vgl. Brandtzaeg et al. 2021, S. 11). Das Schreiben mit dem Chatbot wurde auch für die Anregung zur Selbstreflexion geschätzt und entpuppte sich insbesondere für junge Menschen als nützlich, „who reported some kinds of problems or struggles in their everyday life“ (ebd.). Ganz wie in der Onlineberatung zwischen zwei Menschen auch vorkommend, können Störungen und unvorhergesehene Ereignisse jedoch zum Beziehungsabbruch führen (vgl. Skjuve et al. 2022, S. 13).

Steiner & Tschopp (2022) geben zu vermerken, dass Chatbots in der Sozialen Arbeit beziehungsweise Onlineberatung derzeit noch wenig im Einsatz zu sein scheinen (vgl. S. 469). Viele Untersuchungen und Studien, die sich mit potenziellen Anwendungsmöglichkeiten und Wirkweisen beschäftigen, stammen daher aus klinisch-psychologischen und psychotherapeutischen Settings (vgl. ebd.). Entsprechende Pilotprojekte und dazugehörige Forschung befindet sich momentan

im Auftrieb. Aktuelle Scoping-Reviews, die einen Überblick über den Forschungsstand erheben und Chatbots im Bereich der mentalen Gesundheit untersuchen, konnten eine große Palette an im Einsatz befindlicher Chatbots auffinden, die ein ebenso breites Spektrum an Störungs- beziehungsweise Krankheitsbildern abdecken, beispielsweise Angst, Autismus, Demenz, Depression, Essstörungen, Phobien, Posttraumatische Belastungsstörung, Schizophrenie, Stress, Sucht und generelle psychische Leiden (vgl. Abd-Alrazaq et al. 2019, S. 4; vgl. Abd-Alrazaq et al. 2021, S. 6; vgl. Bendig et al. 2019, S. 273 f.; vgl. Casu et al. 2024, S. 7 ff.). Ziel der Chatbots ist es – je nach Anlass – Nutzenden einen besseren Umgang mit ihren Leiden und Störungen zu ermöglichen, Stressreduzierung zu fördern und das Wohlbefinden zu erhöhen, emotionale Unterstützung zu gewährleisten beziehungsweise (Psycho-)Edukation anzubieten und das etwaige Rehabilitations-Engagement zu steigern beziehungsweise aufrechtzuerhalten (vgl. Casu et al. 2024, S. 7). Darüber hinaus kann eine weitere Funktion der Chatbots darin bestehen, „not only to provide information to users, but also to enable social relationships to form between the user and chatbot“ (Brandtzaeg et al. 2021, S. 1).

In der Onlineberatung, deren Angebote seit ihrem ersten Aufkommen ununterbrochen teils enormen Zulauf erfahren (vgl. Gravelmann 2024a, S. 109 f.), könnten Chatbots eine Möglichkeit darstellen, breitangelegt niedrigschwellige und kostengünstige Unterstützung für Nutzende anzubieten, um der seelischen Belastung der Menschen – auch präventiv – begegnen zu können (vgl. Høiland et al. 2020, S. 140). Gerade die prinzipiell unendliche Skalierbarkeit von Chatbots, welche nur eine einmalige Entwicklung erfordern und dann losgelöst von Zeit, Raum und menschlichen Kapazitäten von prinzipiell unbegrenzt vielen Ratsuchenden gleichzeitig genutzt werden können, lässt sie sowohl für die Soziale Arbeit als auch für gewinnorientierte Unternehmen in einem günstigen Licht erscheinen (vgl. Seelmeyer & Waag 2020, S. 184). Demgegenüber steht die Sorge, die jederzeit verfügbaren und aus dem professionellen Kontext losgelösten sozialen Dienstleistungen wie Chatbots könnten zu einer „Sozialen Arbeit auf Fastfoodniveau“ (James 2004, zit. n. Linnemann et al. 2023a, S. 201) führen. Darüber hinaus weisen Linnemann et al. (2024) auf das Risiko hin, die jederzeit griffbereite Nutzung eines Chatbots könnte, wie von Beratung und Therapie eigentlich nicht intendiert, zu einer Art Abhängigkeit führen, die Nutzende in Konsequenz weniger selbstständig werden lässt (vgl. S. 15). Zudem fanden Haque & Rubya (2023) heraus, dass manche Chatbot-Nutzende dazu neigen, die Interaktionen mit Chatbots dem Kontakt mit Freunden beziehungsweise Familie vorzuziehen, weil sie den Chatbot-Kontakt als vergleichsweise angenehmer empfinden (vgl. S. 10). Dieses Verhalten „could put them in susceptible positions, such as loneliness and exclusion from sociocultural norms“ (ebd.), was den möglicherweise belastenden ursprünglichen Auslöser für die Chatbot-Nutzung, zum

Beispiel Depression oder Einsamkeit, in Konsequenz sogar noch verstärken könnte. Aus diesem Grund sollten Chatbots bestenfalls so eingesetzt werden, dass sie eine „insbesondere menschliche Kommunikation und Beziehung unterstützende, nicht ersetzende Funktion haben sowie begleitend zu spezifisch menschlichen Befähigungen eingesetzt werden: Aufnahme und Aufrechterhaltung von Beziehungen, Zeigen von Empathie, Zeit haben zu beraten und verfügbar für Unterstützung zu sein“ (Steiner & Tschopp 2022, S. 470).

Allerdings wird auch deutlich, dass viele Potenziale und Risiken beim Einsatz von Chatbots noch nicht hinlänglich untersucht wurden (vgl. Waag et al. 2020, S. 181). So seien „die Ergebnisse der bisherigen Forschung im Hinblick auf die Praktikabilität, Durchführbarkeit und Akzeptanz von Chatbots zur Förderung mentaler Gesundheit [...] zwar vielversprechend“ (Bendig et al. 2019, S. 266), allerdings fehle es „an qualitativ hochwertigen randomisiert-kontrollierten Studien“ (ebd.), die das untermauern würden. Auch Ahmed et al. (2023) bemängeln, „there is a lack of consensus on reporting and evaluation standards for chatbots, as well as a need for increased transparency and replication“ (S. 13). So lassen sich außerdem eine Reihe von kommerziellen, unregulierten und wenig evaluierten Chatbots im Bereich der mentalen Gesundheit finden, die entweder webbasiert sind oder als Smartphone-Apps heruntergeladen werden können (vgl. Haque & Rubya 2023, S. 6). Dies ist insofern als nicht verwunderlich zu erachten, als dass gerade dem Gesundheitsbereich in Verbindung mit Künstlicher Intelligenz von Investorensseite große Rentabilität zugeschrieben wird (vgl. Maslej et al. 2024, S. 254). Diesen Umstand gilt es kritisch zu begleiten und zu reflektieren auch vor dem Hintergrund, „dass sich immer mehr gewinnorientierte Unternehmen als Plattformbetreiber für soziale Dienstleistungen in Position bringen“ (Lehmann 2020, S. 5), da „Fragestellungen der Sozialen Arbeit im Internet nicht mehr ausschließlich von den traditionellen Akteuren im Sozialbereich bearbeitet werden“ (ebd.).

Vor diesem Hintergrund spielen vor allem die Themen Privatsphäre und Datenschutz eine gewichtige Rolle. Interessanterweise scheinen Menschen keine Bedenken um ihre Privatsphäre zu haben, wenn sie intime Themen mit Chatbots besprechen (vgl. Brandtzaeg et al. 2021, S. 11). Der in Kapitel 2.4 beschriebene ‚Disinhibition-Effekt‘ scheint also auch im Kontakt mit Chatbots zu wirken. Dies lässt insofern aufhorchen, als dass Nutzende in der Realität keine Beziehung zum Chatbot unterhalten, sondern zum Anbieter des jeweiligen Chatbots (vgl. Skjuve et al. 2021, S. 12). In der Benutzung von Chatbots neigen Nutzende schnell dazu, zu vergessen, dass sie mit einem System kommunizieren, das ihnen auf Wahrscheinlichkeiten fußende Antworten präsentiert, die vonseiten Ratsuchender häufig unhinterfragt verdaut werden (vgl. Linnemann et al. 2024, S. 14). „Besonders herausfordernd ist dabei, dass diese algorithmischen Beratungsoptionen in lebensweltlich etablierte digitale Geräte und Technologien eingebunden sind, sodass z.

B. der Übergang zwischen einer klassischen Ratgeberwebseite mit statischen Inhalten hin zu [...] komplexen Chatbots aus Sicht der Ratsuchenden fließend ist, zumal diese auch bei vollständiger Transparenz hinsichtlich des Einsatzes von KI solche Systeme meist nicht einschätzen können“ (Weinhardt 2024, S. 65). Skjuve et al. (2021) befürchten, Nutzende könnten sich durch das für sie als vertrauensvoll erlebte Setting beim Schreiben mit einem Chatbot in Gefahr begeben, indem sie allzu persönliche Informationen preisgeben (vgl. S. 12). Dass diese datenschutzrechtlichen Sorgen nicht unbegründet sind, zeigt eine Untersuchung der Mozilla Foundation, die aufgedeckt hat, dass viele der inspizierten sogenannten Romantic-Chatbots extensives Datensammeln betreiben und Nutzende in ihren AGB darauf hinweisen, „health-related information from you like your ‘sexual health information,’ ‘[u]se of prescribed medication,’ and ‘[g]ender-affirming care information’“ (Maslej et al. 2024, S. 165) zu erheben. Dieses Beispiel verdeutlicht, welch großes Schädigungspotenzial unregulierte und kommerzielle Chatbots gerade für Kinder und Jugendliche bergen, die in der Regel zu unbedarfterem Vorgehen neigen als Erwachsene und durch beschriebene Enthüllungseffekte durchaus bereit sind, sensible Daten mit Chatbots zu teilen, die dann in den Händen privater Geschäftsleute landen, welche sich diese für ihren eigenen Profit zunutze machen können (vgl. Kremp 2024a; vgl. Linnemann et al. 2024, S. 14 f.).

Betrachtet man Potenziale und Risiken von Chatbots, ist es ferner wichtig, im Blick zu behalten, welche Art von KI diesen zugrunde liegt, da Unterschiede zwischen symbolischer und subsymbolischer KI im praktischen Einsatz zutage treten. Laut Ahmed et al. (2023), die ein Scoping-Review zu Chatbots für Angststörungen und Depressionen durchgeführt haben, beruhen bereits viele Chatbots auf subsymbolischer KI (vgl. S. 9). In Bezug auf die Generierung der Chatbot-Antworten kommen Abd-Alrazaq und Mitarbeitende in ihren 2019 und 2021 veröffentlichten Studien zumindest zu einem etwas anders gewichteten Schluss. Ihre Forschung ergab, dass ein Großteil der ermittelten Chatbots regelbasiert Texte ausgeben, diese also nicht von der KI von Grund auf selbst generiert werden, sondern die Chatbots auf einen Fundus an größtenteils vordefinierten Antworten und Skripten zurückgreifen (vgl. Abd-Alrazaq et al. 2019, S. 4; vgl. Abd-Alrazaq et al. 2021, S. 6). Diese Vorgehensweise des Einsatzes von symbolischer KI bei Chatbots im Bereich der seelischen Gesundheit erscheint schlüssig, da sich so „das grundlegende Format des Gesprächs [...] daran [orientieren kann], wie Kliniker:innen Probleme angehen“ (Engelhardt 2023a, S. 407). Chatbots sind auf diesem Wege in der Lage, gezielt Interventionen anzuwenden, die an die motivierende Gesprächsführung, die Verhaltenstherapie, die Positive Psychologie und an die Erkenntnisse der Wirkfaktorenforschung angelehnt sind, um bei den Nutzenden Achtsamkeit, ein verbessertes Selbstbild, Selbstwirksamkeit und viele weitere förderliche Dinge zu stärken, die zu einer Besserung des Zustands Nutzender beitragen können (vgl. Bendig et al. 2019,

S. 273; vgl. Reamer 2023, S. 55). Ein großer Vorteil symbolischer KI gegenüber großen Sprachmodellen ist, dass sich das Auftreten möglicherweise risikoreicher Antworten minimieren oder sogar gänzlich vermeiden lässt, da diese von Fachkräften oder Fachkundigen vorab formuliert werden können (vgl. Engelhardt 2023a, S. 407). Da die auf regelbasierter KI beruhenden Chatbots jedoch stärker auf das Erkennen von Schlüsselwörtern angewiesen sind, sind sie in ihren Gesprächen starrer als ihre subsymbolischen Geschwister und Dialoge unter Umständen hölzerner (vgl. Linnemann et al. 2024, S. 14). Dadurch sind sie in ihrem Funktionsumfang begrenzt und verfügen nicht über „das erforderliche mentale Modell, mit dem Therapeut:innen und Berater:innen aufgrund ihrer professionellen Ausbildung beraten und therapieren“ (ebd.), weshalb ein alleiniger Einsatz eines Chatbots zur Psychotherapie oder Beratung zur Lösungsproduktion nicht zielführend erscheint (vgl. ebd.). Der limitierte Funktionsumfang regelbasierter Chatbots kann auch dazu führen, dass sie etwaige Notlagen oder Krisensituationen bei Nutzenden schlichtweg nicht erkennen oder selbstschädigendes Verhalten irrtümlicherweise befürworten, da ihr natürliches Sprachverständnis oftmals als deutlich leistungsschwächer zu erachten ist als das von großen Sprachenmodellen – und das von Menschen ohnehin (vgl. White 2018). Ratsuchende bringen mitunter ein so großes Spektrum an individuellen Themen, Anliegen und Beratungsanlässen mit, das innerhalb eines regelbasierten Systems eines symbolischen Chatbots schlichtweg nicht berücksichtigt, aufgegriffen und repräsentiert werden kann (vgl. Linnemann et al. 2024, S. 14). Da eine fortlaufende Entwicklung eines regelbasierten Chatbots händisch passieren muss (vgl. Kap. 3.2.1), mutet ein kontinuierliches Abdecken aller Eventualitäten wie ein Ding der Unmöglichkeit an.

Große Sprachmodelle hingegen können ihre ganz eigenen Risiken mit sich bringen. Da große Sprachmodelle keine vorgegebenen Antworten ausgeben und ihre Textausgaben Ergebnis einer Wahrscheinlichkeitsrechnung sind (vgl. Kap. 3.2.2.1), tendieren sie dazu, immer mal wieder zu halluzinieren (vgl. Engelhardt 2023a, S. 407). Im Grunde genommen können alle Textausgaben von großen Sprachmodellen als Halluzination angesehen werden, jedoch werden nur solche Antworten als Halluzination bezeichnet, die faktisch als falsch erkannt werden (vgl. Heaven 2024). Beispiele hierfür finden sich viele: So gab SARAH, ein Chatbot der WHO, Adressen von erfundenen Kliniken an Nutzende heraus, Metas eingestellter Chatbot Galactica fantasierte wissenschaftliche Artikel über Bären im Weltall herbei und die Fluglinie AirCanada wurde gerichtlich verpflichtet, einer nichtexistierenden Rückvergütungsklausel nachzukommen, die ihr Chatbot im Dialog mit einem Kunden fabulierte (vgl. ebd.). Was in den aufgezählten Beispielen durchaus einer gewissen Komik nicht entbehrt, kann sich im Kontext psychischer Gesundheit oder psychosozialer Beratung schnell als großes Risiko für Ratsuchende entpuppen. Entwickelnde von

großen Sprachmodellen sind sich dieser Problematik durchaus bewusst. Heaven (2024) weist auf die Hoffnung der Forschenden hin, durch noch extensiveres Training mit umfassenderen Datensätzen, dieses Problem irgendwann beheben zu können. Da große Sprachmodelle sich korrigieren können, nachdem sie durch Nutzende auf festgestellte Halluzinationen hingewiesen wurden, könnte auch eine Art eingebaute automatische Faktencheck-Funktion, die die Ausgaben von Chatbots selbstständig gegenprüft, zukünftig denkbar sein (vgl. ebd.). Ein weiteres Risiko von generativen Chatbots geht von ihren Datensätzen und dem damit einhergehenden Training aus, aus dem letztlich alle generierten Antworten entspringen. Die Datensätze können nicht objektiv sein, da sie auf einer selektiven Auswahl an Daten beruhen, die trotz möglicherweise besserer Versuche dies präventiv zu unterbinden, „gesellschaftliche Machtverhältnisse [und] gesellschaftlich dominante Wertevorstellungen [(re-)produzieren] und [...] somit nicht diskriminierungsfrei [sind]“ (Gravelmann 2024b, S. 9).

Wie De Freitas et al. (2023) untersucht haben, produzieren die nicht speziell für mentale Gesundheit konzipierten Social-Chatbots ‚Simsimi‘ und ‚Cleverbot‘ in etwas mehr als der Hälfte der Fälle als risikoreich zu erachtende Antworten auf Eingaben von Nutzenden, die Krisensituationen vermuten lassen müssten (vgl. S. 6). Dies ist als besonders schwerwiegendes Risiko zu betrachten, da eine beträchtliche Anzahl Nutzender von Chatbots Konversationen führen, die die eigene seelische Gesundheit betreffen (vgl. ebd.). Chatbots sind bisweilen nicht in der Lage, diese Gemütszustände zu erkennen, Containing zu betreiben oder Ratsuchende in dem Maße emotional aufzufangen, wie es eine geschulte Fachkraft im besten Falle tun könnte (vgl. ebd., S. 10). Als mindestens risikoreich einzuordnen sind auch Ausgaben des in diesem Kapitel bereits erwähnten Social-Chatbots Replika, der sich Nutzenden gegenüber zu sexuell unangemessenen Anzüglichkeiten hat hinreißen lassen und diese bedrängt haben soll (vgl. Cole 2023). Gegenüber den Ergebnissen von De Freitas et al. (2023) kommen Maurya et al. (2024) in einer qualitativen Inhaltsanalyse mit ChatGPT-3.5 zu einem etwas optimistischeren Ergebnis. Sie präsentierten dem Chatbot Eingaben, die verschiedenen Kategorien psychischer Belastung, wie Depression, Angst oder Substanzmissbrauch, zuzuordnen sind und beispielsweise lauteten: „why do I hate myself at times?“ (Maurya et al. 2024, S. 4) oder „I often feel like I am lost. What can I do to find meaning and purpose in my life?“ (ebd.). Aus den Ergebnissen der Autor:innen lässt sich ablesen, dass ChatGPT Antworten generierte, die in den Qualitätskriterien Genauigkeit, Klarheit, Relevanz, Empathie, Zugewandtheit und ethische Berücksichtigungen fast durchgängig sehr gut abschnitten (vgl. ebd., S. 5). Nicht nur war der Chatbot oftmals in der Lage empathische Antworten zu produzieren oder hilfreiche Hinweise und Einwände zu erteilen, er gab den Autor:innen auch häufig zu bedenken, sich bei anhaltenden Beschwerden Hilfe und Unterstützung bei

professionellen Fachkräften zu holen (vgl. Maurya et al. 2024, S. 6 f.). In einer weiteren explorativen Studie, die die Eignung von ChatGPT als Beratungsoption zum Schwangerschaftsabbruch erforschte, konnte ebenfalls ermittelt werden, „dass die ChatGPT-Antworten zu schwangerschaftsabbruchbezogenen Prompts überwiegend sachlich korrekt und unvoreingenommen sind“ (Kubitza 2024, S. 92). Aus Sicht von Maurya et al. (2024) konnte ChatGPT daher unstrittiges Potenzial in der Psychoedukation andeuten, weswegen der Chatbot als eine Art erster Anlaufpunkt für emotional belastete Menschen fungieren könne, beispielsweise in Zeiten, in denen etwaige zugehörige Therapeut:innen nicht erreichbar seien (vgl. S. 8). Weiteres Einsatzpotenzial von ChatGPT sehen die Autor:innen „in ,reaching and supporting those who may have limited access to mental health resources due to financial, cultural, or geographical barriers“ (ebd.). Kubitza (2024) gibt jedoch zu bedenken, nur menschliche Beratende seien in der Lage, die individuell bestehenden Lebens- und Konfliktlagen Ratsuchender miteinzubeziehen, sodass diese personalisierte und fallspezifische Beratung erhalten könnten (vgl. S. 92).

Ein möglicher Grund für die unterschiedlichen Resultate von De Freitas et al. (2023) und Maurya et al. (2024) könnte darin liegen, dass das große Sprachmodell von ChatGPT grundsätzlich besser auf Fragen und Inhalte rund um das seelische Befinden trainiert ist, als es bei ‚Simsimi‘ und ‚Cleverbot‘ der Fall ist. Die Entwickelnden von generativen Chatbots zeigen sich bestrebt, ihre Chatbots fortlaufend zu verbessern, um Antworten, die fremd- oder selbstgefährdendes Verhalten begünstigen könnten, auszuschließen (vgl. Linnemann et al. 2024, S. 15). Dennoch kann es immer wieder zu Gesprächsverläufen kommen, die derlei Textausgaben seitens eines Chatbots hervorrufen können (vgl. ebd.). Wie Maslej et al. (2024) zeigen, ist die absolute Anzahl an sogenannten KI-Vorfällen im Laufe der vergangenen Jahre massiv angestiegen, was auch damit zusammenhängt, dass gegenwärtig generell deutlich mehr KI-Systeme im Einsatz sind als noch vor ein paar Jahren (vgl. S. 164). Unter KI-Vorfällen versteht man beispielsweise Autounfälle von selbstfahrenden Autos, fälschliche Verhaftungen aufgrund von Gesichtserkennungsalgorithmen oder Diskriminierungen und ungerechtfertigte Betrugsbezeichnungen von Sozialleistungsempfängenden durch KI-Systeme (vgl. Bouwer & Schöppl 2023; vgl. Maslej et al. 2024, S. 164). Anhand dieser Statistik lässt sich klar ablesen, wie fehleranfällig beziehungsweise unsicher der Einsatz von KI in mancherlei Hinsicht noch sein kann, und genannte Beispiele zeigen, was für drastische Folgen für menschliche Betroffene daraus resultieren können. „Offen bleibt die Haftungsfrage bei Falschaussagen oder gefährdenden Inhalten, [da] hier [...] aktuell noch Referenzurteile und praxisbezogene Lösungen [fehlen]“ (Linnemann et al. 2024, S. 14).

Trotzdem müssen gerade großen Sprachmodellen Potenziale in der Onlineberatung bescheinigt werden, die insbesondere im individuellen Training begründet liegen. Wie in Kapitel 3.2.2.1

erklärt wurde, können große Sprachmodelle ganz gezielt mit speziellen vorselektierten Datensätzen für bestimmte Anwendungszwecke trainiert werden. Abbildung 12 veranschaulicht, wie dabei mittels überwachten Lernens maßgeblich begünstigt und beeinflusst werden kann, welche Ausgaben durch den Chatbot im praktischen Einsatz generiert werden. Als Grundlage kann dabei natürlich nicht nur ChatGPT verwendet werden, sondern auch jedes beliebige andere zur Verfügung stehende große Sprachmodell (vgl. Cardillo 2024). Ein Beispiel für einen Chatbot, der auf einem großen Sprachmodell beruht und für einen speziellen Anwendungszweck trainiert wird, wäre der sich in der Entwicklung befindende Chatbot ‚Mina‘, der Studierenden helfen soll, einen Umgang mit Prüfungsstress und -ängsten zu finden (vgl. Mina Chatbot App 2024).

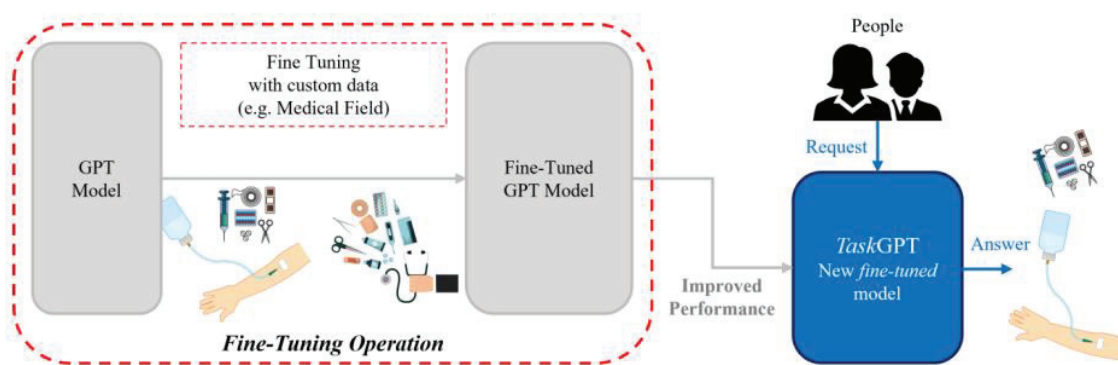


Abbildung 12: Feintuning eines großen Sprachmodells (Casu et al. 2024, S. 4)

So erscheint gerade der Einsatz von Chatbots in etwas stärker eingezäunten Anwendungsbereichen potenziell sinnvoll und zu rechtfertigen. Dennoch muss konstatiert werden, dass auch generative Chatbots derzeit in ihrem Funktions- und Nutzenumfang noch begrenzt sind, sobald eine tiefere Koproduktion von Lösungskompetenz gefragt ist. Wie sich in diesem Kapitel gezeigt hat, kann ein alleiniger Einsatz generativer Chatbots in der Koproduktion von Lösungskompetenz zuweilen mit großen Risiken einhergehen. Zudem können ihre auf Wahrscheinlichkeitsrechnungen beruhenden Textausgaben keine profunderen Lösungshilfen und Interventionshinweise sein, da sie Verursachungszusammenhänge aus den Lebenswirklichkeiten Ratsuchender nicht tiefgehend miteinbeziehen können (vgl. Gravelmann 2024b, S. 14). Dies ist nicht zuletzt deshalb problematisch, weil datenschutzrechtliche Bedenken ein solches Vorgehen innerhalb der Sozialen Arbeit undenkbar erscheinen lassen (vgl. Engelhardt 2023a, S. 408). Ein Verstehen der Zusammenhänge und Wirkfaktoren ist aber Kerngeschäft der Sozialen Arbeit, das darauf zielt, die Lebenslagen Ratsuchender zu verbessern und das „Klientel zu stärken, indem durch ganzheitliche Sichtweisen, situative Offenheit, ressourcenorientiertes Herangehen und gemeinsame Problemanalysen und -bearbeitungen die Kompetenzen und die Autonomie gefördert beziehungsweise aktiviert werden“ (Gravelmann 2024b, S. 14). Abhängig vom



Beratungsanlass, dessen Komplexität und einhergehender Schwere der Thematik und Belastung aufseiten Ratsuchender ist die Koproduktion von Lösungskompetenz daher zum gegenwärtigen Zeitpunkt in der Majorität der Fälle noch auf die Beteiligung menschlicher Beratender angewiesen (vgl. Waag et al. 2020, S. 186). So würden Chatbots eine beratende oder therapeutische Fachkraft nicht gänzlich ersetzen, sondern mit Blick auf die in diesem Kapitel dargelegten Potenziale das Spektrum an bereitstehenden sozialen Dienstleistungen für Ratsuchende sinnhaft erweitern und ergänzen.

## 4.2 Assistive Tools

Assistenzsysteme können in der Sozialen Arbeit sowohl von Klient:innen zur Unterstützung als auch von Fachkräften als Hilfe zur Bewältigung der Arbeit eingesetzt werden (vgl. Schiffhauer 2020, S. 266). Etwaige KI-Systeme zur Administration, Organisation und (Risiko-)Analyse von Daten, die nur indirekt etwas mit Beratung zu tun haben, werden bewusst ausgeklammert.<sup>8</sup> Aus Sicht Ratsuchender in der Onlineberatung ist KI insbesondere als Kommunikationshilfe denkbar. Als Übersetzungstool wird KI bereits heute eingesetzt, um Menschen mit unzureichenden Deutschsprachkenntnissen Zugänge zur Onlineberatung zu ermöglichen (vgl. Gravelmann 2024b, S. 10). So wirbt die JugendNotmail, die seit 2001 schriftliche Onlineberatung für junge Menschen zu allen möglichen und denkbaren Themen anbietet, damit, „dank künstlicher Intelligenz Beratungen in 28 verschiedenen Sprachen anbieten [zu können], darunter die Sprachen der Europäischen Union sowie Arabisch, Türkisch, Russisch und Japanisch“ (JugendNotmail o. D.). Gerade in Anbetracht der Wellen an geflüchteten Menschen, die in den letzten Jahren Obhut und Sicherheit in Deutschland gesucht haben, erscheint ein solcher Einsatz als absolut sinnvoll (vgl. Mornhinweg 2018, S. 154). Über die API, also der Schnittstelle der jeweiligen Übersetzungs-KI wie beispielsweise DeepL, lässt sich KI problemlos in bestehende Dienstleistungen integrieren. Wichtig zu beachten ist jedoch, dass der Datenschutz der Ratsuchenden hierbei nicht gefährdet wird (vgl. Esser & Spiecker genannt Döhmann 2023, S. 397). Das dargestellte Vorgehen der JugendNotmail ist insofern als wünschenswert zu erachten, als dass „Sozialarbeiter:innen die bedeutsame Aufgabe [...] [zukommt], assistive Technologien für Klient:innen zugänglich zu machen“ (Schiffhauer 2020, S. 268). So kann plattformseitig eingebundene KI sehr leicht dabei behilflich sein, Barrieren abzubauen und Teilhabe und Inklusion zu ermöglichen und so die Inanspruchnahme von Angeboten wie Onlineberatung zu erleichtern, was wiederum die

---

<sup>8</sup> Interessante Projekte diesbezüglich sind zum Beispiel in ‚KiJuAssistenz‘ (<https://kiju-assistenz.de/>) und ‚SensAssist2Sens‘ (<https://sensassist2sens.de/>) auszumachen.

Selbstbestimmung, Autonomie und Selbstermächtigung aufseiten der Klient:innen begünstigt (vgl. Schiffhauer 2020, S. 268). Beim Einsatz einer Übersetzungs-KI sollte dieser Umstand allen Beteiligten transparent gemacht werden, da durch maschinelle Übersetzung Textinhalte verfälscht werden können, wodurch textuelle Feinheiten verloren gehen können und das in Kapitel 2 beschriebene ‚zwischen den Zeilen Lesen‘ erschwert werden kann. Dadurch kann es im Austausch leichter zu Missverständnissen und Störungen kommen, die womöglich aber durch die Transparentmachung der zum Einsatz kommenden KI abgefedert werden können.

KI als Kommunikationshilfe hat aber noch weitere Funktionen: Menschen, die durch Einschränkungen nicht befähigt sind zu schreiben, können mittels assistiver KI-Tools Sprache zu Papier bringen und Menschen mit Analphabetismus oder Sehbeeinträchtigungen beziehungsweise -verlust können sich die Nachrichten der Beratenden vorlesen lassen (vgl. Gravelmann 2024a, S. 82; vgl. Gravelmann 2024b, S. 10).

Die Nutzung solcher Tools ist natürlich auch für Beratende im Bereich des Möglichen. So könnten beispielsweise lange Sprachnachrichten Ratsuchender, wie sie beispielsweise in der Messengerberatung vorkommen (vgl. Kap. 2.2.3), mittels KI-Tools zur Transkription übersichtlich lesbar gemacht werden (vgl. Albrecht & Rudolph 2023, S. 374). Auf diese Weise können Beratende sich zum Beispiel das 4-Folien-Konzept (vgl. Kap. 2.3) für Sprachnachrichten nutzbar machen, um Antworten zu formulieren, die die unterschiedlichen Inhalte und Ebenen der empfangenen Nachricht besser miteinbeziehen können.

Wie in dieser Arbeit hinreichend dargelegt wurde, können sich große Sprachmodelle als vielseitige Werkzeuge erweisen. Nutzende, zu denen auch Beratende zu zählen sind, können sie als psychoedukatives Hilfsmittel zur Hand nehmen (vgl. Maurya et al. 2024, S. 4). So können sich Chatbots als interaktive Partner vor allem in solchen Themenfeldern als hilfreich zeigen, in denen Fachkräften Wissen vorliegt, sodass diese die Vor- und Nachteile der Vorschläge des Chatbots abwägen können (vgl. Abramson 2023, S. 69). Auf diese Weise können Fachkräfte „leicht und zeitnah auf relevante Daten und Informationen zu[...]greifen, was besonders in synchronen Onlineberatungssettings von Vorteil ist, in welchen schnelle und präzise Antworten gefordert sind“ (Poltermann et al. 2024, S. 22).

Im Besonderen stechen große Sprachmodelle aber durch ihre Fähigkeit zur natürlichen Sprachverarbeitung hervor (vgl. Kapitel 3.2.2.1). Aufgrund dessen scheinen sie gerade als assistives Werkzeug in der Textanalyse ihr wahres Potenzial für Beratende in der Praxis der Onlineberatung zu entfalten (vgl. Waag et al. 2020, S. 183). Denkbar erscheint hier die Zuhilfenahme eines

großen Sprachmodells, um verfasste Nachrichten auf leichte Sprache hin zu überprüfen, sodass diese den Adressat:innen gerecht werden (vgl. Engelhardt 2023a, S. 408). Im Rahmen einer Untersuchung evaluierten Elyoseph et al. (2023) außerdem ChatGPTs Emotional Awareness anhand des dafür anerkannten und vorgesehenen Tests ‚LEAS‘ (vgl. S. 2). Emotional Awareness gilt als grundsätzliche Fähigkeit in der psychologischen Praxis und beschreibt unter anderem die Fähigkeit, „zwischen verschiedenen emotionalen Zuständen zu differenzieren [und] Auslöser für spezifisches emotionales Erleben zu antizipieren“ (Lohaus & Vierhaus 2019, S. 187). Die Ergebnisse zeigten, dass der Chatbot im Rahmen des durchgeführten Tests ausgezeichnet abschnitt und in allen drei Kriterien Werte vorweisen konnte, die über dem zum Vergleich herangezogenen menschlichen Durchschnitt lagen (vgl. Elyoseph et al. 2023, S. 5). Zudem verbesserte ChatGPT seine Leistung im Laufe der Zeit und erreichte in einem zweiten Test fast den höchstmöglichen Testscore (vgl. ebd., S. 3). Zu einem diese Ergebnisse untermauernden Resultat kamen auch Sharma et al. (2022), die das Zusammenspiel von KI und Mensch innerhalb einer asynchronen textbasierten Peerberatungsapp zum seelischen Wohlbefinden untersuchten, wobei sich beratende Peers Verbesserungsvorschläge zur Formulierung einfühlsamer Antworten von einer dafür vorgesehenen KI namens HAILEY einholen konnten (vgl. S. 2). Der Einsatz von HAILEY führte zu einem Anstieg von knapp 20 Prozent an gemessener Empathie in allen Unterhaltungen und besonders Teilnehmende, die sich schwertaten, einfühlsame Antworten zu verfassen, profitierten, was sich in einem Anstieg um fast 40 Prozent an Empathie in deren Konversationen niederschlug (vgl. ebd., S. 12). Interessanterweise produzierten Nutzende, die HAILEY stets für ihre Antworten konsultierten, aber nicht immer ihre Vorschläge übernahmen, die im Durchschnitt empathischsten Antworten (vgl. ebd., S. 8). Auch die Forschung von Grandeit et al. (2020) stellte unter Beweis, dass bei der Analyse von sozialwissenschaftlichen Texten KI besser abschnitt als Menschen und komplexe Sachverhalte wie Empathie zuverlässig identifizieren konnte (vgl. S. 19).

Damit stiften die präsentierten Ergebnisse berechtigte und begründete Hoffnung, dass große Sprachmodelle und Chatbots sich als nützliches assistives Werkzeug für Fachkräfte in der Beratung und Psychotherapie<sup>9</sup> erweisen könnten. Wie sich gezeigt hat, scheinen sie grundsätzlich fähig zu sein, präzise Aussagen über emotionale Zustände von Menschen in unterschiedlichen Situationen treffen oder diese vorhersagen zu können (vgl. Elyoseph et al. 2023, S. 5). Es ist nur

---

<sup>9</sup> Mithilfe von analysierender KI ist man auch bestrebt, die psychotherapeutische Praxis zu verbessern. Aktuelle Forschung dazu findet sich beispielsweise in Ewbank et al. (2020a), Ewbank et al. (2020b) und Miner et al. (2022).

folgerichtig, Überlegungen anzustellen, wie solch eine befähigte KI in der Onlineberatung nutzbar gemacht werden könnte. Die bisherige Forschung hierzu schafft ein immer größer werdendes wissenschaftliches Fundament, das einen assistierenden Einsatz von KI in der Onlineberatung in den nächsten Jahren durchaus realistisch werden lässt. Einerseits erscheint es viabel, eine KI wie HAILEY in einem synchronen Beratungssetting einzuschleifen, um in Echtzeit hilfreiche Hinweise für das Verfassen von zugewandten und feinfühligem Antworten zu erhalten, andererseits kann ein großes Sprachmodell auch in asynchronen Settings dabei behilflich sein, ganz im Sinne des 4-Folien-Konzeptes die Gefühlszustände, Anliegen und Beweggründe von Ratsuchenden besser zu lesen und zu verstehen, um diese in Antworten einfühlsam und verständnisvoll aufgreifen zu können. Interessant könnte auch der Einsatz von KI zur Analyse von laufenden Beratungsprozessen sein, um Muster sichtbar zu machen, die vom zuständigen Beratenden aufgegriffen werden könnten, was einen besseren Beratungsverlauf zur Folge haben könnte und die Wahrscheinlichkeit einer erfolgreichen Beratung erhöhen könnte (vgl. Engelhardt 2023a, S. 406).

An genau einer solchen KI-Lösung forscht das E-Beratungsinstitut in Nürnberg derzeit mit der sich in der Entwicklung befindenden KI-gestützten Assistenz für psychosoziale Onlineberatung namens KIA, die im Folgenden beispielhaft für eine solche, zuvor skizzierte, assistive KI in der Onlineberatung näher betrachtet werden soll (vgl. Institut für E-Beratung o. D.b). „KIA [soll] Beratungsfachkräften Echtzeit-Analysen von Ratsuchenden-Anfragen bereitstellen“ (Poltermann et al. 2024, S. 23) und diesbezüglich methodische, kontextbezogene und typisierende Informationen zugänglich machen (vgl. Institut für E-Beratung o. D.b). KI-generiertes Feedback kann sich für Beratende vor allem während synchroner Beratungen als hilfreich erweisen, da es unmittelbar eingeholt und genutzt werden kann (vgl. Rudolph et al. 2024, S. 116). Das KI-System richtet sich dabei an Beratungsfachkräfte aller Erfahrungsstufen und soll auch zur Qualitätssicherung von Beratungen verwendet werden können (vgl. Poltermann et al. 2024, S. 23). Neben der Evaluierung von Wirkfaktoren und der umfangreichen Untersuchung von bestehenden Potenzialen und Risiken legt das vom Bundesfamilienministerium geförderte Forschungsprojekt dabei auch einen besonderen Schwerpunkt auf ethische Gesichtspunkte, die im Rahmen eines Ethikmanuals zum gewissenhaften Einsatz eines solchen KI-Modells für das Gemeinwohl reflektiert und festgehalten werden (vgl. Institut für E-Beratung o. D.b).

Das Verfassen eines Ethikmanuals in Verbindung mit einer gründlichen Forschung bezüglich des Einsatzes einer beim Beratungsprozess beteiligten KI ist als uneingeschränkt positiv zu bewerten, da die Zusammenarbeit von KI und Mensch durchaus einige Probleme und Unabwägbarkeiten mit sich bringen kann. Hier wäre vornehmlich die Diffusion der Verantwortung zu nennen,

die „bei jeder Entscheidung, die der Mensch mit Unterstützung einer KI trifft“ (Mockenhaupt & Schlagenhauf 2024, S. 76), entsteht, selbst wenn der verantwortliche Mensch den Vorschlag der KI nicht oder verändert übernimmt (vgl. ebd.). Eine dabei auftretende kognitive Verzerrung wird ‚Automation Bias‘ genannt und beschreibt „a well-documented human propensity to automatically defer to automated systems, despite warning signals or contradictory information from other sources. In other words, human actors are found to uncritically abdicate their decision making to automation“ (Alon-Barkat & Busuioc 2022, S. 154). Hervorgerufen wird dieser Effekt durch ein erhöhtes Vertrauen in maschinell errechnete Ergebnisse hinsichtlich Richtigkeit, Objektivität und Neutralität bei gleichzeitigem Bewusstsein um die eigene menschliche Fehlbarkeit (vgl. Seidensticker 2022, S. 204). Bei der Einführung in die Praxis solcher assistiven KI-Systeme gilt also ganz genau zu beobachten, welche Effekte und möglicherweise auftretenden kognitiven Verzerrungen bei den Beratenden zutage treten, um einen Umgang mit ihnen finden zu können, sodass Risiken für Ratsuchende minimiert werden können. Wie Elyoseph & Levkovich (2023) im Rahmen einer Untersuchung mit ChatGPT ermittelten, schnitt der Chatbot bei der Einschätzung von Suizidrisiken anhand von Textvignetten, die hypothetische Patient:innen beschrieben, deutlich schlechter ab als professionelle Fachkräfte (vgl. S. 5). Trotz der zuvor beschriebenen Potenziale in der Erkennung von Emotionen scheint die Zuhilfenahme eines KI-Modells für krisenhafte Onlineberatungsanlässe möglicherweise ungeeignet. Anhand dieses Beispiels soll exemplarisch verdeutlicht werden, dass die Fähigkeit von KI Grenzen hat und die Qualität der KI auch immer vom Training und den dazugehörigen Datensätzen abhängt (vgl. Linnemann et al. 2023b, S. 416). Mitunter können diese nicht repräsentativ, minderwertig oder lücken- beziehungsweise fehlerhaft sein, sodass es in der Praxis zu Diskriminierungen, schlechter Übertragbarkeit und Leistung oder blinden Flecken kommen kann (vgl. Botsch 2023, S. 7 ff.). Beratende sollten sich daher stets der potenziell vorliegenden Grenzen von KI bewusst sein „und die Ausgaben oder Vorschläge des KI-Systems gegebenenfalls kritisch überprüfen“ (Albrecht & Rudolph 2023, S. 376), um informierte und bewusste Entscheidungen im Beratungsprozess treffen zu können, die den Ratsuchenden zugutekommen – auch, weil Beratende bis auf Weiteres ausführende Kraft bleiben und aus rechtlicher Perspektive als handelnde Instanz in der Verantwortung stehen (vgl. Linnemann et al. 2023a, S. 204; vgl. Maurya & DeDiego 2023, S. 3; vgl. Mockenhaupt & Schlagenhauf 2024, S. 76).

In Bezug auf eine mögliche Qualitätssicherungsfunktion einer KI verspricht diese eine Möglichkeit zur schnellen, unkomplizierten und effizienten Evaluation von durchgeführten Beratungen (vgl. Botzum & Neumaier 2023, S. 392). Fachkräfte könnten zudem ‚quasi-kollegiale‘ Rückmeldungen erhalten, was vor allem dem in Kapitel 4 beschriebenen Fachkräftemangel begegnet und

in der Umsetzung Kosteneinsparungen verheißt. Unklar ist noch, wie Beratende damit umgehen, von einer KI bewertet zu werden und von ihr Rückmeldungen zu ihrer Professionalität zu erhalten. Fraglich ist auch, ob ein flächendeckend eingesetztes KI-Qualitätssicherungssystem tatsächlich zu besseren Beratungen und einer höheren Professionalisierung von Fachkräften führen würde. Wie die in diesem Kapitel thematisierten Untersuchungen zur KI HAILEY gezeigt haben, waren, rein auf Empathie bezogen, jene Beratungen am besten, in denen Peers nach steter Abwägung nicht durchgängig die Vorschläge von HAILEY übernommen haben. Würde HAILEY des Argumentes wegen eine Qualitätssicherung durchführen, müsste man zu dem Schluss kommen, dass aus HAILEYS Sicht eigentlich die Beratungen am besten hätten sein müssen, die stets die Vorschläge der KI übernommen haben, obwohl diese, rein auf den gemessenen Empathiewert bezogen, nicht die besten waren. Folgt man dieser Logik würde die Gefahr bestehen, dass die Beratungsqualität in Folge solcher Rückmeldungen womöglich etwas abnehmen könnte, weil Fachkräfte sich dann, ausgelöst durch Automation Bias und entstehende Erwartungen und Erwartungseffekte, nach dem Feedback einer KI richten, die möglicherweise nicht in der Lage ist, gewisse schwer messbare Qualitäten und Faktoren in Beratungen zu erkennen (vgl. Hackbart 2022, S. 14; vgl. Linnemann et al. 2023a, S. 207). Gleichwohl beruhen große Sprachmodelle auf Datensätzen, die von Grund auf nicht als objektiv betrachtet werden können, da sie immer gewisse bestehende Weltbilder, Stereotypen und Vorurteile abbilden, die durch den Einsatz von KI reproduziert werden (vgl. Bolukbasi et al. 2016, S. 1 ff.; vgl. Nicoletti & Bass 2023; vgl. Stinson 2022, S. 769). Dies ist ein Risiko, dass auf Datensätzen beruhende große Sprachmodelle aller Couleur und Anwendungsfelder generell mit sich bringen. Demnach entpuppt sich möglicherweise auch eine von einer KI durchgeführte Qualitätssicherung oder Supervision nicht als objektiv und sie kann diese nicht, die individuellen Qualitäten des jeweiligen Beratenden berücksichtigend, feinfühlig-dynamisch umsetzen, wie es ein Mensch in der Lage wäre, zu tun. Auch wenn große Sprachmodelle sich hervorragend zum standardisierten Evaluieren von Dialogen eignen, können gegenwärtig Menschen noch deutlich besser kontextuelle Nuancen erfassen und emotionale Untertöne in Beratungsverläufen mitberücksichtigen (vgl. Rudolph et al. 2024, S. 112). Für schnelle Rückmeldungen scheint eine KI zum gegenwärtigen Zeitpunkt schon geeignet – umfassendere Qualitätssicherungen und Supervisionen sollten aber durch Menschen passieren. Denkbar ist aber durchaus jetzt schon, dass eine erste KI-Evaluation als Grundlage zur Bewertung dienen kann. So kann dieses als sinnvolle Ergänzung für den Reflexionsprozess erachtet werden, wodurch bestenfalls „personelle Ressourcen zielgenauer gebündelt und möglichst effizient genutzt werden können“ (Lippert et al. 2024, S. 54)

Abschließend bleibt festzuhalten, dass es weiterhin kritisch zu beobachten und zu reflektieren gilt, wie KI sich zur Qualitätssicherung und als assistives Tool zur Verbesserung von laufenden Beratungen einsetzen ließe, ohne dass dies eine mögliche Deprofessionalisierung von Fachkräften zur Folge hätte (vgl. Walpuski 2024, S. 116), was im schlechtesten Fall aufseiten Beratender mit einem „gravierende[n] Verlust eigener professioneller Urteilskompetenz“ (Lob-Hüdepohl 2023, S. 386) einhergehen könnte. Engelhardt (2023a) sieht hier vor allem ethische Fragen als diskussionswürdig, beispielsweise ob gelungene Textvorschläge seitens der KI unverändert übernommen an Ratsuchende geschickt werden können und ob diese dann darüber in Kenntnis gesetzt werden sollten (vgl. S. 408). Solche und ähnliche Fragen erfordern Beantwortung und eine Aufnahme in den Diskurs, da, wie in Kapitel 4 beschrieben wurde, davon auszugehen ist, dass bereits gegenwärtig einige Onlineberatende möglicherweise unreflektiert generative KI als Hilfsmittel benutzen. „Daher wird hier dringend Kompetenzbildung durch geeignete Qualifizierungsangebote und eine Sensibilisierung für die [sic!] Umgang mit Systemen dieser Art benötigt.“ (ebd.)

Im Idealfall kann KI unter Berücksichtigung ethischer, professioneller und fachlicher Standards als Hilfsmittel so genutzt werden, dass es die professionelle Beratungspraxis hinsichtlich Effektivität und Qualität bereichert (vgl. Poltermann et al. 2024, S. 22). Dies erfordert weitere Forschung, wie sie im Falle von KIA gerade stattfindet, und eine kontinuierliche Evaluation, sobald sich derartige KI im Einsatz befindet (vgl. Albrecht & Rudolph 2023, S. 375). Sollte dies nicht möglich sein, müssen „Sozialarbeiter:innen die Freiheit haben, sich auch gegen den Einsatz von assistiven Technologien entscheiden zu können, wenn sie die ethischen und sozialen Implikationen als nicht vertretbar erachten“ (Schiffhauer 2020, S. 268).

### **4.3 Einsatz in der Ausbildung**

Die Vielseitigkeit von großen Sprachmodellen beflügelte recht schnell Ideen und Überlegungen, sich diese auch für das Üben von Onlineberatungen nutzbar machen zu können (vgl. Weinhardt 2023b). Erste Versuche, erbrachte Machbarkeitsnachweise und laufende Untersuchungen haben das Potenzial von LLMs für die Ausbildung von Onlineberatenden bereits angedeutet. Zu nennen wären hier zum Beispiel der ‚emotionally responsive interactive client‘ (kurz ERIC) oder das Projekt ‚Virtuelle:r Klient:in‘ (kurz ViKI) des Instituts für E-Beratung in Nürnberg (vgl. Institut für E-Beratung o. D.a; vgl. Prescott et al. 2023, S. 1). ViKI und ERIC sind beides LLM-basierte Chatbots, die Lernenden Möglichkeiten bieten sollen, ihr Beratungskönnen zu erproben (vgl. ebd.).

Obwohl viele LLMs von Haus aus über ein gutes Sprachverständnis verfügen, besteht die Idee grundsätzlich darin, sie mittels Feintuning für Ausbildungszwecke in der Onlineberatung so zu trainieren, dass diese sich in einem simulierten Übungsberatungsgespräch so verhalten können, wie es eine ratsuchende Person tun würde (vgl. Rudolph et al. 2024, S. 110). Im Falle von ViKI wurde eigens ein individueller Datensatz erstellt, anhand dessen er trainiert wurde, „da Gesprächsdaten aus realen Beratungssituationen meist aufgrund von datenschutzrechtlichen Gründen nicht für Entwicklungs- oder Forschungszwecke genutzt werden können“ (Lippert et al. 2024, S. 50). Wie Rudolph et al. (2024) darlegen, kann ViKI sieben verschiedene Personae darstellen, die an echte Beratungsfälle angelehnt sind (vgl. S. 110). Mittels von Expert:innen verfasster Personen- und Chatverlaufsbeschreibungen begaben sich Studierende in chatbasierte Rollenspiele, die sich in ihrer Dauer an echten Chatsprechstundenzeiten orientierten und zum Ziel hatten, ein höchstmögliches Maß an Authentizität zu besitzen, weshalb die Rollenspiele der Ratsuchenden unterschiedliche Grade an Kooperativität mit teils knappen oder unklaren Antworten aufwiesen (vgl. ebd.). Nach erfolgtem Training war ViKI in der Lage, verschiedene Personae anzunehmen, die mit individuellem Hintergrund, einem Haupt- und mehreren Nebenanliegen ausgestattet gewesen sind (vgl. Lippert et al. 2024, S. 51). Das Feintuning des LLMs ermöglicht es diesem zudem, Antworten zu generieren, die ein hohes Maß an Konsistenz und Kohärenz aufweisen (vgl. Rudolph et al. 2024, S. 109). Durch das Mimen einer Persona unterscheiden sich die verschiedenen Chatbot-Persönlichkeiten nicht nur hinsichtlich der Beratungsthemen, sondern sie sind auch in der Lage, verschiedene Kommunikationsverhaltensweisen an den Tag zu legen und ihre Nachrichten mittels sich voneinander unterscheidender Schreibstile und Ausdrucksformen zu transportieren (vgl. Lippert et al. 2024, S. 50 f.). Diese breitgefächerte Ausstattung bietet Lernenden ein abwechslungsreiches Angebot an Übungsszenarien, in denen sie ihr Beratungswissen in kreativer Weise ausprobieren können (vgl. Poltermann et al. 2024, S. 23).

Rollenspiele mit generativer KI scheinen gerade deshalb verheißungsvoll zu sein, weil sie angehenden Beratenden nicht nur die Möglichkeit bieten, wichtige Methoden und Handkniffe wie das Paraphrasieren und Spiegeln der Gefühle der Ratsuchenden, was in der Onlineberatung nicht nonverbal passieren kann, zu üben (vgl. Maurya 2023, S. 12), sondern auch weitere Vorteile bereithalten: Einerseits können Übende Techniken und Methoden wiederholt anwenden und krisenhafte Beratungen oder für sie ungewöhnliche Anliegen bearbeiten, ohne dass sie sich überfordert fühlen würden oder dadurch ein Risiko für Ratsuchende entstehen könnte, andererseits könnten Lehrpersonen, Peers und dafür trainierte KI Übungsberatungen evaluieren und den Lernenden wertvolles Feedback erteilen (vgl. Lippert et al. 2024, S. 48; vgl. Reamer 2023, S. 55). Dies stellt für Lernende eine vertraute Herangehensweise dar, die ihnen eine gute



Möglichkeit bietet, ihr Beratungshandeln zu reflektieren, um es in Folge dessen verbessern zu können (vgl. Prescott et al. 2023, S. 529). Zugleich können Peers sich darin üben, Beratungshandeln professionell zu untersuchen, um daraufhin Beratenden wertschätzendes Feedback zu erteilen, was an das Beratungsüben von Studierenden aus der Praxis erinnert (vgl. ebd.).

Chatbots wie ViKI und ERIC können zudem die Schwelle reduzieren, erste Schritte in der Ausübung als Beratende:r zu unternehmen, was gerade sehr unerfahrenen Lernenden zu Gute kommen kann (vgl. Lippert et al. 2024, S. 54). Weitere Potenziale ergeben sich außerdem durch „die Replizierbarkeit von Übungssituationen, das Umgehen störender Artefakte in der Nutzung von Rollenspielen innerhalb der Lerner:innengruppe oder die Ausschaltung ethischer Probleme des Übens an echten Fällen“ (Weinhardt 2023a, S. 48). Durch die hohe Sprachfähigkeit der eingesetzten LLMs produzieren diese realitätsnahe Reaktionen auf Nachrichten der Lernenden, wodurch diese wichtige Fähigkeiten und Kompetenzen in der schriftbasierten Beratung erlernen können, was auch einen positiven Effekt auf die Entwicklung der Berater:innenidentität haben kann (vgl. Lippert et al. 2024, S. 47). Darüber hinaus bestehen bei dieser skizzierten Übungsmöglichkeit auch die in dieser Thesis behandelten Vorteile der Onlineberatung, wie beispielweise die durch die Ungebundenheit von Zeit und Raum hervorgerufene Niedrigschwelligkeit, die keinen großen Planungs- und Organisationsaufwand seitens der Lernenden erforderlich macht (vgl. Lippert et al. 2024, S. 47; vgl. Prescott et al. 2023, S. 526). Lernende können sich so häufiger in Übungen begeben und ihre Fähigkeiten höher frequentiert trainieren, ohne sich dafür mit weiteren Mitstudierenden zum Rollenspiel verabreden zu müssen (vgl. Maurya 2023, S. 11). Erste Erhebungen deuten zudem darauf hin, dass sich positive Effekte hinsichtlich der Kommunikation mit Chatbots (vgl. Kap. 4.1) auch auf den Übungskontext übertragen. So berichten Prescott et al. (2023) aufgrund der Ergebnisse einer ersten Befragung mit Lernenden von einer hohen Akzeptanz der Übenden gegenüber ERIC, der als ansprechender und recht authentischer Sparringspartner beim Beratungslernen wahrgenommen wurde (vgl. S. 529).

Mit Chatbots wie ERIC und ViKI, die fortlaufend weiterentwickelt werden und im Laufe der Zeit vermutlich nur noch besser und realitätsnaher werden, können angehende Beratende zusätzlich zu ihrer sonstigen Ausbildung wertvolle Erfahrungen in der Onlineberatung erlangen, die sie gewinnbringend in den Realkontext mitnehmen können (vgl. Engelhardt 2023b, S. 27 f.). Angesichts dieses angedeuteten Potenzials ist es durchaus denkbar, dass derartige KI-Modelle zeitnah in der Hochschullehre oder in der Psychotherapie- beziehungsweise Beratungsausbildung eingesetzt werden und in Zukunft dort auch fest verankerter Bestandteil sein können. Überdies sollten auch Fachkräfte nach ihrem Studienabschluss Zugänge zur Weiterbildung erhalten (vgl. Linnemann et al. 2023b, S. 420). „Um die Implementierung [...] in die Lehre erfolgreich zu

gestalten, ist eine Eingliederung in ein didaktisches Rahmenkonzept notwendig, das den Bedürfnissen der verschiedenen Projektbeteiligten, aber auch der Kursorganisator\*innen gerecht wird“ (Lippert et al. 2024, S. 49). Wie die Forschenden des Chatbots ViKI betonen, ist es zudem wichtig, die Auswirkungen des Einsatzes von KI „auf die fachliche Qualität der Beratung anhand eines Wirkmodells und aus den generierten Daten aus quantitativen und qualitativen Forschungen innerhalb dieses Projektes breit [...] zu evaluieren“ (ebd., S. 50). Derartig fortlaufende Reflexion kann einen verantwortungsvollen Umgang mit KI in der Hochschullehre gewährleisten, was sinnhafte Weiterentwicklungen begünstigen kann, von denen Lernende wie Lehrende gleichermaßen nachhaltig profitieren können (vgl. Botzum & Neumaier 2023, S. 393).

Entscheidend für eine sinnhafte Implementierung von KI in Lehre und Ausbildung wird vor allem die enge Einbindung von qualifizierten Lehrpersonen sein (vgl. Lippert et al. 2024, S. 56). Der Lernerfolg Studierender hängt in großem Maße von einer guten Beziehung zu den Lehrkräften ab und insbesondere diese können dafür sorgen, dass ein Nährboden besteht, der befruchtende fachliche Diskurse ermöglicht (vgl. Botzum & Neumaier 2023, S. 393). Aus diesem Grund sind die in diesem Kapitel vorgestellten Chatbots auch nicht als Alternative zu menschlichen Lehrpersonen zu betrachten (vgl. Lippert et al. 2024, S. 56). Ein Chatbot wie ERIC oder ViKI „can arguably enrich the learning experience and, in this case, provide a means of complementing classbased activities with an automated programme“ (Prescott et al. 2023, S. 530), sodass sie als eine Ergänzung zum bestehenden Lehrkonzept zu erachten sind. Zudem erfahren Studierende Umgang mit hochaktueller Technologie, was sie auf ihre Berufspraxis mitsamt einhergehenden Anforderungen vorbereitet, mit denen sie in einer immer digitaler werdenden Welt und Sozialen Arbeit konfrontiert werden (vgl. Botzum & Neumaier 2023, S. 393). Gelingt eine kritisch-reflexiv ablaufende Einbettung von KI in den Lehrkontext, kann so der Weg bereitet werden, „um weitere einheitliche und hochwertige Standards in der deutschsprachigen Ausbildung zur Onlineberatung voranzutreiben“ (Lippert et al. 2024, S. 56).

## 5 Fazit

Die vorliegende Thesis hat gezeigt, dass die Entwicklung von KI an einem Punkt angelangt ist, an dem sie als disruptive Technologie dabei ist, viele alltägliche Lebensbereiche und professionelle Berufsfelder zu durchdringen. Wie auch dargelegt wurde, ist KI bereits jetzt Teil der Lebensrealität vieler gerade junger Menschen (vgl. Kap. 2.5).

Auch hat sich gezeigt, dass Onlineberatung insbesondere für junge Menschen weiterhin sehr relevant ist und von diesen mitunter hohe Inanspruchnahme erfährt. Geschätzt werden hieran besonders die Niedrigschwelligkeit, die Autonomie und die Möglichkeit zur Selbstoffenbarung, die aus den Gegebenheiten des Settings entspringen. Besonders der schriftbasierte Austausch ist für junge Menschen auch in ihrer Alltagskommunikation von sehr großer Bedeutung und dort fest verankert (vgl. Kap. 2.5). Demgemäß scheint die schriftliche Onlineberatung unverändert eine gute Option für Kinder und Jugendliche zu sein, um psychosoziale Beratung in Anspruch nehmen zu können.

Es ist davon auszugehen, dass die Bedeutung von KI mit zunehmender Entwicklung für Menschen aller Altersklassen in den kommenden Jahren nur noch weiterwachsen wird, wozu selbstredend auch Beratende gehören, die mitunter schon zum gegebenen Zeitpunkt KI in ihrem Berufsalltag nutzen. Ganz im Sinne des Bestrebens von Onlineberatung, sich weiter zu institutionalisieren, muss die Onlineberatung sich den fortlaufenden Veränderungen und Herausforderungen stellen, die durch das Eindringen von KI in die Lebenswelt und den Alltag von Menschen hervorgerufen werden. Davon ausgenommen ist auch nicht der Einfluss, den KI auf das Arbeitsfeld haben wird. „Auch wenn die Entwicklung und praktische Implementierung solcher KI-Systeme erst am Anfang steht, ist bereits heute schon absehbar, dass die technologische Entwicklung die professionelle Soziale Arbeit prägen und stark verändern wird“ (Poltermann et al. 2024, S. 23). Gerade deshalb ist das Praxisfeld Onlineberatung gefordert, die Einbindung von KI aktiv mitzugestalten und unter der Abwägung von Potenzialen und Risiken zu reflektieren, inwiefern sich KI zum Nutzen Ratsuchender und Beratender in die Praxis einbauen ließe, um weiter professionelle und seriöse Beratung anbieten zu können, die für Ratsuchende als solche erkennbar ist (vgl. Lehmann 2020, S. 5).

Wie in dieser Masterarbeit untersucht wurde, gibt es bereits zum jetzigen Zeitpunkt einige Anwendungsmöglichkeiten von KI in der Onlineberatung, die Ratsuchenden und Beratenden zugutekommen können. In der Onlineberatung denkbar erscheint vor allem der Einsatz von assistiven KI-Tools und von Chatbots als Interaktionspartner für Ratsuchende sowie von Chatbots als

simulierte virtuelle Ratsuchende zu Übungszwecken für angehende Beratende. In Bezug auf die in dieser Arbeit dargestellte Schnelligkeit der Weiterentwicklung von KI ist davon auszugehen, dass sich bestehende Einsatzmöglichkeiten von KI in der Onlineberatung schnell verbessern werden und dass sich zukünftig weitere potenzielle Anwendungsfelder von KI in der Onlineberatung auf tun werden.

Die Zielsetzung vorliegender Arbeit war es, herauszufinden, wie sich die Onlineberatung KI zunutze machen kann – unter Berücksichtigung auftretender Potenziale und Risiken sowohl für Ratsuchende als auch Beratende.

Bezüglich assistiver Tools lässt sich festhalten, dass KI-Hilfen für Ratsuchende vor allem Barrieren abbauen können, wie es im Falle von Übersetzer-KIs schon heute passiert. Ihr Einsatz kann somit dazu beitragen, Zugangswege für Klient:innen zur Onlineberatung herzustellen. Allerdings könnten Übersetzungstools in manchem Falle durch nicht ganz treffende Übersetzungen womöglich auch kommunikative Missverständnisse begünstigen, die Störungen nach sich ziehen könnten. KI-gestützte Hilfsmittel, die den Arbeitsfluss vereinfachen stehen natürlich auch Beratenden zur Verfügung. Wie sich gezeigt hat, könnten gerade Chatberatende von LLMs profitieren, indem sie Chatbots während der manchmal hektischen Beratungen als psychoedukative Inspirationsquelle verwenden. Ein Risiko liegt im Halluzinieren von Chatbots begründet, weshalb sie sich für diesen Zweck besonders bei Themen eignen, die Beratenden nicht völlig fremd sind.

Viel größeres Potenzial deuten große Sprachmodelle allerdings als Assistenzsystem in Onlineberatungen an, in denen sie Beratenden methodische und kontextbezogene Hinweise liefern können und Verbesserungsvorschläge hinsichtlich der Sprache und des Schreibstils offerieren. Wie sich gezeigt hat, verfügen LLMs über durchaus sehr gute Fähigkeiten, Emotionen anhand von Text lesen zu können, weswegen sie für den Einsatz in der Onlineberatung prädestiniert erscheinen. Einsatzpotenziale haben sich auch in der Supervision und Qualitätssicherung angedeutet. Die Hoffnung, sie könnten zu besseren Beratungsverläufen beitragen, hat daher durchaus Berechtigung. Menschen tendieren im Arbeitsverbund mit KI indes häufig dazu, KI-Entscheidungen ein höheres Gewicht beizumessen als ihren eigenen. Wie Kapitel 4.2 gezeigt hat, sind KI-Vorschläge aber nicht unbedingt immer besser. Ob der breitgefächerte Einsatz einer Assistenz-KI wie KIA wirklich zu qualitativ hochwertigeren Beratungen führen würde, muss daher noch ausführlich untersucht werden. Ein Risiko könnte sich für Ratsuchende womöglich auch in krisenhaften Beratungen ergeben, falls KI-Assistenzen Notlagen verkennen sollten und Beratende trotz eigentlich besserer Intuition der KI die Deutungshoheit zukommen lassen. Sollten KI-Assistenzen also in der beschriebenen Form in der Onlineberatung zum Einsatz kommen, gilt es

diesen fortlaufend sehr genau zu evaluieren und kritisch zu reflektieren, sodass Ratsuchende keinem Risiko ausgesetzt sind und Fachkräften keine Deprofessionalisierung droht.

In Bezug auf Chatbots hat sich gezeigt, dass diese außerhalb der psychosozialen Onlineberatung bereits weitverbreitet im Einsatz sind und viele Menschen schon einmal mit einem Chatbot kommuniziert haben. In der Sozialen Arbeit werden sie bisher noch wenig genutzt. Dennoch hat sich herausgestellt, dass Einsatzpotenziale von Chatbots in der Sozialen Arbeit grundsätzlich in der Wissensweitergabe, als Orientierungshilfe und zur Koproduktion von Lösungskompetenz bestehen. Vor allem als kurzer Kontakt in der Wissensweitergabe und zur Orientierungshilfe scheinen sie hervorragendes Eignungspotenzial mitzubringen. Besonders hervorzuheben ist in dieser Hinsicht, dass der ‚Disinhibition-Effekt‘, der in der Onlineberatung zwischen Menschen zutage tritt, im Kontakt mit Chatbots noch einmal größer zu sein scheint. Diesen Umstand kann sich die Onlineberatung zunutze machen, indem sie Chatbots für Ratsuchende bereitstellt, die möglicherweise stark schambehaftete beziehungsweise stigmatisierende Themenfelder abdecken, sodass Klient:innen auf diesem Wege in der Lage sind, Unterstützung in Anspruch nehmen zu können. Der vorgestellte Limbic-Chatbot, der Psychotherapie-Überweisungen erleichterte, konnte beispielsweise zeigen, wie stark vor allem Menschen, die sich Minderheiten zugehörig fühlen und tendenziell eher Diskriminierungen ausgesetzt sind, von solchen Angeboten profitieren können. Auch die quasi-unendliche Skalierbarkeit von Chatbots könnte prinzipiell dazu beitragen, das Angebot an Versorgungsmöglichkeiten für Ratsuchende sinnvoll zu ergänzen.

Von Chatbots profitieren können auch seelisch belastete Menschen, da sich Chatbots in der Psychoedukation als wirkungsvoller erwiesen haben als Bibliothherapie. Darüber hinaus kann, ganz wie in der Onlineberatung auch, das selbstoffenbarende Schreiben Ratsuchende zur Selbstreflexion anregen. Potenziale zeigen Chatbots auch darin, gegen Einsamkeit zu helfen, von der in der heutigen Zeit viele junge Menschen betroffen sind. Dies geschieht, indem Nutzende eine quasisoziale Beziehung mit einem Chatbot eingehen. Umschlagen kann dieser Effekt, wenn Nutzende anfangen, Kontakte mit Menschen aus ihrem Umfeld zu meiden, weil der Kontakt zum Chatbot sich unbeschwerlicher anfühlt und sich Einsamkeit und andere negative Effekte in diesem Zuge noch verstärken. Weitere Risiken können entstehen, wenn Chatbots zuweilen risikobehaftete Eingaben Nutzender nicht erkennen und somit nicht darauf eingehen können. Es ist ebenfalls möglich, dass sie unbeabsichtigt riskantes Verhalten Nutzender befürworten, wenn sie nicht in der Lage sind, die Feinheiten von Sprache richtig zu entschlüsseln. Auch leiden LLMs hin und wieder unter Halluzinationen, die ebenso als Risiko für Ratsuchende auszumachen sind. Aus diesem Grund fallen in der Koproduktion von Lösungskompetenz die momentanen Einsatzmöglichkeiten von Chatbots eher gering aus. Zu groß erscheinen zum gegenwärtigen Zeitpunkt die

Risiken, die mit einem flächendeckenden Einsatz einhergehen und zu gering sind derzeit noch die Kompetenzen der meisten Chatbots, Koproduzenten von Lösungskompetenzen zu sein.

Den auf Wahrscheinlichkeitsrechnungen fußenden Chatbots fehlt es an tiefgreifendem Verständnis für die Wirkzusammenhänge einzelner Faktoren im Leben Ratsuchender. Zudem sind sie methodisch noch nicht so ausgestattet, dass sie eine situativ offene Beratung nach den Regeln der Kunst führen könnten. Zwar sind viele Chatbots, die dem mentalen Gesundheitsbereich zuzuordnen, an Gesprächsführungen von Kliniker:innen orientiert, ihre Gesprächsverläufe beruhen jedoch auf Skripten und verlaufen daher unflexibel und in vorgegebenen Bahnen. In den allermeisten Fällen sind Ratsuchende bei der Erarbeitung von Lösungskompetenzen daher noch auf menschliche Beratende angewiesen. An oberster Stelle sollte beim Einsatz von Chatbots die Wahrung des Datenschutzes stehen, insbesondere wenn sensible Themen und private Daten Ratsuchender involviert sind. Gerade deshalb ist die Soziale Arbeit gefordert, Angebote zu schaffen, die seriöse Alternativen zu gewinnorientierten kommerziellen Chatbots darstellen, deren Betreiber keine Skrupel haben, gesammelte Daten zur weiteren Profitgenerierung zu nutzen und Ratsuchende somit ungeahnten Risiken aussetzen können.

Darüber hinaus deuten Chatbots ihr Potenzial in der Ausbildung von Onlineberatenden an. Projekte wie ViKI lassen sich dank LLMs und individuell angefertigter Datensätze anhand derer sie trainiert werden, realisieren, sodass ViKI in der Lage ist, mehrere verschiedene ratsuchende Personae zu mimen, mit denen Beratende in Ausbildung Beraten üben können. Grundsätzlich ist dies durchweg positiv zu betrachten: Beratungsneulinge können ohne große Hemmungen, in echten Fällen Fehler zu begehen, ihre ersten Schritte unternehmen; Lernende können für die Onlineberatung wichtige Techniken, wie das Paraphrasieren und das Verbalisieren von Gefühlen einüben; Studierende erhalten rund um die Uhr die Möglichkeit, ohne großen Aufwand Beratung zu trainieren; Schwierige Beratungssituationen lassen sich unbegrenzt oft wiederholen; Peers und Lehrende können Beratenden wertvolle Rückmeldungen geben. Die ersten Rückmeldungen seitens der Studierenden zu Trainings-Chatbots wie ERIC und ViKI waren rundherum positiv, weshalb es durchaus denkbar ist, dass derartige Bots bald den Weg in die Hochschullehre finden, oder in Onlineberatungsausbildungen eingesetzt werden. Um herauszufinden, ob der Einsatz von Ausbildungschatbots einen Effekt auf die fachliche Beratungsqualität Lernender hat, bedarf es weiterer Forschung, wie sie im Falle von ViKI gerade läuft. Sollten sich Chatbots wie ViKI und ERIC bewähren, könnten sie schon bald in enger fachlicher Begleitung Lehrender in den Lehrkontext implementiert werden.

Noch völlig unklar ist, wie die weitere Entwicklung von KI verlaufen wird. Klar ist, dass ein großer Hype um LLMs besteht, dass sehr viel in die Entwicklung investiert wird und dass KI riesige Potenziale nachgesagt werden. Ob allerdings das Problem des Halluzinierens in Bälde gelöst wird oder LLMs durch bessere Leistungsfähigkeit und höheren Funktionsumfang in ein paar Jahren eine ernsthafte Konkurrenz zu Beratungsfachkräften darstellen können, lässt sich zum jetzigen Zeitpunkt nicht seriös beurteilen. Wie die Geschichte von KI zeigt, war sie immer wieder von unvorhersehbaren Wendungen und Durchbrüchen geprägt. Welche Implikationen der weitere Verlauf von KI-Entwicklung für die Onlineberatung bereithält, lässt sich daher gegenwärtig nicht vollumfänglich abschätzen, und es gilt, die weitere Entwicklung genau zu beobachten.

Während des aktuellen und zukünftigen Einsatzes von KI in der Onlineberatung ist ein besonderes Augenmerk darauf zu legen, dass qualitative Onlineberatung unter Wahrung datenschutzrechtlicher, ethischer und professioneller Gesichtspunkte zum Wohle Ratsuchender weiter gewährleistet werden kann. Zu keinem Zeitpunkt sollten nicht zu verantwortende Gefährdungspotenziale für Klient:innen entstehen. Dazu ist die fortlaufende Einbindung von Fachkräften der Sozialen Arbeit angeraten. Aus diesem Grund sollten sie auch in zukünftigen interdisziplinären KI-Projekten, die in der Onlineberatung oder benachbarten Feldern zum Einsatz kommen sollen, fest eingebunden sein. KI in der Onlineberatung sollte begleitend und ergänzend eingesetzt werden und darf nicht zum Abbau oder zu qualitativen Verschlechterungen sozialer Dienstleistungen beitragen.

Es gilt Studien fortzusetzen, die untersuchen, welche langfristigen Auswirkungen beim Einsatz von KI in der Onlineberatung sowohl für Ratsuchende als auch für Beratende entstehen. Dabei sollte unter Abwägen von Potenzialen und Risiken zwischen den spezifischen Anforderungen der unterschiedlichen und möglichen Anwendungsfelder von KI im Feld der Onlineberatung differenziert werden. Langfristig wird zu beforschen sein, ob sich die Struktur der Kommunikation selbst durch die breite Nutzung von Chatbots durch junge Menschen verändern wird und inwiefern sich dies auf die Alltagskommunikation und auf Onlineberatung auswirken wird. Wenn sich die Onlineberatung in Zukunft weiterhin institutionalisieren soll, ist all diesen Punkten Aufmerksamkeit zu schenken – und dies sowohl in der Praxis als auch in der diese begleitenden Forschung.

## Literaturverzeichnis

- Abd-Alrazaq, A. A., Alajlani, M., Alalwan, A. A., Bewick, B. M., Gardner, P. & Househ, M. (2019). An overview of the features of chatbots in mental health: A scoping review. *International Journal of Medical Informatics*, 132, 103978, 1-7. <https://doi.org/10.1016/j.ijmedinf.2019.103978>
- Abd-Alrazaq, A. A., Alajlani, M., Ali, N., Denecke, K., Bewick, B. M. & Househ, M. (2021). Perceptions and Opinions of Patients About Mental Health Chatbots: Scoping Review. *Journal Of Medical Internet Research*, 23(1), e17828, 1-15. <https://doi.org/10.2196/17828>
- Abrams, Z. (2024a). *Addressing equity and ethics in artificial intelligence*. Abgerufen am 09.07.2024: <https://www.apa.org/monitor/2024/04/addressing-equity-ethics-artificial-intelligence>
- Abrams, Z. (2024b). The Unstoppable Momentum of Generative AI. *Monitor on Psychology*, 55(1), 44-47. Abgerufen am 14.06.2024: <https://www.apa.org/monitor/2024/secure/2024-01-monitor.pdf>
- Abramson, A. (2023). How to use ChatGPT as a learning tool. *Monitor on Psychology*, 54(4), 67-70. Abgerufen am 09.07.2024: <https://www.apa.org/monitor/2023/2023-06-monitor.pdf>
- Ahmed, A., Hassan, A., Aziz, S., Abd-Alrazaq, A. A., Ali, N., Alzubaidi, M., Al-Thani, D., Elhusein, B., Siddig, M. A., Ahmed, M. & Househ, M. (2023). Chatbot features for anxiety and depression: A scoping review. *Health Informatics Journal*, 29(1), 146045822211467, 1-17. <https://doi.org/10.1177/14604582221146719>
- Albrecht, J. & Rudolph, E. (2023). Künstliche Intelligenz und Machine Learning. *Jugendhilfe*, 61(5), 367-375.
- Alon-Barkat, S. & Busuioc, M. (2022). Human–AI Interactions in Public Sector Decision Making: “Automation Bias” and “Selective Adherence” to Algorithmic Advice. *Journal Of Public Administration Research and Theory*, 33(1), 153-169. <https://doi.org/10.1093/jop-art/muac007>
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M. & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal Of Big Data*, 8(1), 1-74. <https://doi.org/10.1186/s40537-021-00444-8>
- Ashekian, T. L. (2022). *The rise of ‘Dr. Google’: The risks of self-diagnosis and searching symptoms online*. Abgerufen am 21.07.2024: <https://theconversation.com/the-rise-of-dr-google-the-risks-of-self-diagnosis-and-searching-symptoms-online-180278>



- Atallah, S. B., Banda, N. R., Banda, A. & Roeck, N. A. (2023). How large language models including generative pre-trained transformer (GPT) 3 and 4 will impact medicine and surgery. *Techniques in Coloproctology*, 27(8), 609-614. <https://doi.org/10.1007/s10151-023-02837-8>
- Barak, A. & Grohol, J. M. (2011). Current and Future Trends in Internet-Supported Mental Health Interventions. *Journal Of Technology in Human Services*, 29(3), 155-196. <https://doi.org/10.1080/15228835.2011.616939>
- Beisch, N. & Koch, W. (2023). ARD/ZDF-Onlinestudie: Weitergehende Normalisierung der Internetnutzung nach Wegfall aller Corona-Schutzmaßnahmen. *Media Perspektiven*, 23/2023, 1-9. Abgerufen am 14.06.2024: [https://www.ard-zdf-onlinestudie.de/files/2023/MP\\_23\\_2023\\_Onlinestudie\\_2023\\_Fortschreibung.pdf](https://www.ard-zdf-onlinestudie.de/files/2023/MP_23_2023_Onlinestudie_2023_Fortschreibung.pdf)
- Bendig, E., Erb, B., Schulze-Thuesing, L. & Baumeister, H. (2019). Die nächste Generation: Chatbots in der klinischen Psychologie und Psychotherapie zur Förderung mentaler Gesundheit – Ein Scoping-Review. *Verhaltenstherapie*, 29(4), 266-280. <https://doi.org/10.1159/000499492>
- Beranek, A., Hammerschmidt, P., Hill, B. & Sagebiel, J. (2021). Einführung: Big Data, Facebook, Twitter & Co. Soziale Arbeit und digitale Transformation. In Hammerschmidt, P., Sagebiel, J., Hill, B. & Beranek, A. (Hrsg.), *Big Data, Facebook, Twitter & Co. und Soziale Arbeit* (2. erw. Aufl.) (S. 9-32). Weinheim: Beltz Juventa.
- Berger, T., Stolz, T. & Schulz, A. (2013). Internetbasierte geleitete Selbsthilfensätze bei Angststörungen und Depressionen. *E-Beratungsjournal*, 9(2), 1-16. Abgerufen am 02.06.2024: [https://www.e-beratungsjournal.net/ausgabe\\_0213/berger\\_stolz\\_schulz.pdf](https://www.e-beratungsjournal.net/ausgabe_0213/berger_stolz_schulz.pdf)
- Blackstein, A. (2023). *Digitale Seelsorge. Impulse für die Praxis*. Göttingen: Vandenhoeck & Ruprecht.
- Boden, M. A. (2018). *Artificial Intelligence. A Very Short Introduction*. E-book. Oxford: Oxford University Press. <https://doi.org/10.1093/actrade/9780199602919.001.0001>
- Boeger, A. (2024). *Psychologische Therapie- und Beratungskonzepte. Theorie und Praxis* (4., akt. Aufl.). Stuttgart: W. Kohlhammer.
- Bolukbasi, T., Chang, K., Zou, J., Saligrama, V. & Kalai, A. (2016). *Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings*. <https://doi.org/10.48550/arxiv.1607.06520>
- Botsch, B. (2023). *Maschinelles Lernen – Grundlagen und Anwendungen*. Berlin: Springer Spektrum. <https://doi.org/10.1007/978-3-662-67277-8>
- Botzum, E. & Neumaier, S. (2023). Künstliche Intelligenz und digitale Transformationsprozesse in der Lehre Sozialer Arbeit. *Jugendhilfe*, 61(5), 388-394.

- Bouwer, P. & Schöppl, N. (2023). *Die EU kann algorithmische Diskriminierung beenden*. Abgerufen am 13.07.2024: <https://background.tagesspiegel.de/digitalisierung-und-ki/briefing/die-eu-kann-algorithmische-diskriminierung-beenden>
- Bowman, J. (2024). *Why Nvidia Stock Skyrocketed 239% in 2023*. Abgerufen am 18.05.2024: <https://finance.yahoo.com/news/why-nvidia-stock-skyrocketed-239-172504224.html?guccounter=1>
- Bräutigam, B. (2022). Symbiose oder Rivalität? Das komplizierte Geschwisterverhältnis zwischen psychosozialer Beratung und Psychotherapie. In Bräutigam, B., Hörmann, M. & Märtens, M. (Hrsg.), *Alles Erfindung? Länderübergreifende Perspektiven auf Beratung und Psychotherapie* (S. 123-136). Göttingen: Vandenhoeck & Ruprecht.
- Brakemeier, E-L., Schomerus, G. & Baumeister, H. (2024). *Chatbot erleichtert Überweisung bei psychischen Problemen*. Abgerufen am 15.07.2024: <https://www.sciencemediacenter.de/alle-angebote/research-in-context/details/news/chatbot-erleichtert-ueberweisung-bei-psychischen-problemen/>
- Brandtzaeg, P. B., Skjuve, M., Dysthe, K. K. & Følstad, A. (2021). When the Social Becomes Non-Human: Young People's Perception of Social Support in Chatbots: Social Support in Chatbots. *CHI Conference on Human Factors in Computing Systems (CHI '21), May 08–13, 1-13*. <https://doi.org/10.1145/3411764.3445318>
- Brunner, A. (2006). Methoden des digitalen Lesens und Schreibens in der Online-Beratung. *E-Beratungsjournal*, 2(2), 1-11. Abgerufen am 30.05.2024: [http://www.e-beratungsjournal.net/ausgabe\\_0206/brunner.pdf](http://www.e-beratungsjournal.net/ausgabe_0206/brunner.pdf)
- Brunner, A. (2009). Theoretische Grundlagen der Online-Beratung. In Kühne, S. & Hintenberger, G. (Hrsg.), *Handbuch Online-Beratung. Psychosoziale Beratung im Internet* (2. Aufl.) (S. 27-45). Göttingen: Vandenhoeck & Ruprecht.
- Buchanan, B. G. (2005). A (Very) Brief History of Artificial Intelligence. *AI Magazine*, 26(4), 53-60. <https://doi.org/10.1609/aimag.v26i4.1848>
- Bundespsychotherapeutenkammer (2022). *Praxisinfo Videobehandlung*. Abgerufen am 02.06.2024: [https://www.bptk.de/wp-content/uploads/2022/05/bptk\\_praxisInfo\\_videobehandlung\\_22\\_web.pdf](https://www.bptk.de/wp-content/uploads/2022/05/bptk_praxisInfo_videobehandlung_22_web.pdf)
- Burghardt, J. & Lehmann, R. (2023). Künstliche Intelligenz und Kinderschutz. *Jugendhilfe*, 61(5), 410-415.
- Campayo, S. (2020). Professionelles Handeln mit Blick auf Digitalisierung. In Kutscher, N., Ley, T., Seelmeyer, U., Siller, F., Tillmann, A. & Zorn, I. (Hrsg.), *Handbuch. Soziale Arbeit und Digitalisierung* (S. 290-301). Weinheim: Beltz Juventa.
- Cardillo, A. (2024). *List of the Best 21 Large Language Models (LLMs) (July 2024)*. Abgerufen am 12.07.2024: <https://explodingtopics.com/blog/list-of-llms>

- Casu, M., Triscari, S., Battiato, S., Guarnera, L. & Caponnetto, P. (2024). AI Chatbots for Mental Health: A Scoping Review of Effectiveness, Feasibility, and Applications. *Applied Sciences*, 14(13), 5889, 1-23. <https://doi.org/10.3390/app14135889>
- Clifford, C. (2018). *Google CEO: AI Is More Important Than Fire or Electricity*. Abgerufen am 07.05.2024: <https://www.cnn.com/2018/02/01/google-ceo-sundar-pichai-ai-is-more-important-than-fire-electricity.html>
- Cole, S. (2023). 'My AI Is Sexually Harassing Me': Replika Users Say the Chatbot Has Gotten Way Too Horny. Abgerufen am 09.07.2024: <https://www.vice.com/en/article/z34d43/my-ai-is-sexually-harassing-me-replika-chatbot-nudes>
- Darcy, A., Daniels, J., Salinger, D., Wicks, P. & Robinson, A. (2021). Evidence of Human-Level Bonds Established With a Digital Conversational Agent: Cross-sectional, Retrospective Observational Study. *JMIR Formative Research*, 5(5), e27868, 1-7. <https://doi.org/10.2196/27868>
- De Freitas, J., Uğuralp, A. K., Oğuz-Uğuralp, Z. & Puntoni, S. (2023). Chatbots and mental health: Insights into the safety of generative AI. *Journal Of Consumer Psychology*, 00, 1-11. <https://doi.org/10.1002/jcpy.1393>
- Decker, M. (2019). Von den three laws of robotics zur Roboterethik. In Liggieri, K. & Müller, O. (Hrsg.), *Mensch-Maschine-Interaktion. Handbuch zu Geschichte – Kultur – Ethik* (S. 345-351). Berlin: J. B. Metzler.
- Derry, A., Krzywinski, M. & Altman, N. (2023). Convolutional neural networks. *Nature Methods*, 20(9), 1269-1270. <https://doi.org/10.1038/s41592-023-01973-1>
- Deshpande, A. (2016). *The 9 Deep Learning Papers You Need To Know About (Understanding CNNs Part 3)*. Abgerufen am 17.05.2024: <https://adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html>
- Deutsche Gesellschaft für Beratung e.V. (2020). *Beratungsverständnis der Deutschen Gesellschaft für Beratung (DGfB)*. Abgerufen am 02.06.2024: [https://dachverband-beratung.de/pdf/Wissensdokumente/DGfB\\_Beratungsverstaendnis\\_2.0.pdf](https://dachverband-beratung.de/pdf/Wissensdokumente/DGfB_Beratungsverstaendnis_2.0.pdf)
- Dick, P. K. (2008). Ach als Blobbel hat man's schwer! In P. K. Dick (Hrsg.), *Zur Zeiten der Perky Pat. Sämtliche SF-Geschichten* (4. Band, 2. Aufl.) (S. 590-615). Frankfurt a. M.: Zweitausendundeins.
- Dick, S. (2019). Artificial Intelligence. *Harvard Data Science Review*, 1(1), 1-8. <https://doi.org/10.1162/99608f92.92fe150c>
- Dimension Market Research (2024). *Large Language Model Market By Type (Zero-shot Model, Pre-trained & Fine-tuned Model, Language Representation Model, Multimodal Model), By Deployment, By Application, By End User - Global Industry Outlook, Key Companies (OpenAI LP, Amazon, Alibaba Group Holding Ltd, and others), Trends and Forecast 2024-*

2033. Abgerufen am 12.07.2024: <https://dimensionmarketresearch.com/report/large-language-model-market/>
- Döring, N. (2003). *Sozialpsychologie des Internet. Die Bedeutung des Internet für Kommunikationsprozesse, Identitäten, soziale Beziehungen und Gruppen* (2., vollst. überarb. u. erw. Aufl.). Göttingen: Hogrefe.
- e-estonia (2020). *New e-Estonia factsheet: National AI "Kratt" Strategy*. Abgerufen am 09.05.2024: <https://e-estonia.com/new-e-estonia-factsheet-national-ai-kratt-strategy/>
- Eichenberg, C. (2007). Online-Sexualberatung: Wirksamkeit und Wirkweise - Evaluation eines Pro Familia-Angebots. *Zeitschrift für Sexualforschung*, 20(3), 247-262. <https://doi.org/10.1055/s-2007-981238>
- Eichenberg, C. & Kühne, S. (2014). *Einführung Onlineberatung und -therapie: Grundlagen, Interventionen und Effekte der Internetnutzung*. München: Ernst Reinhardt Verlag.
- Eichenberg, C., Schott, M. & Aden, J. (2016). Psychosoziale Beratung im Zeitalter des Web2.0: Angebot und Nachfrage von Beratungseinrichtungen auf Facebook. *E-Beratungsjournal*, 12(1), 34-46. Abgerufen am 14.06.2024: [https://www.e-beratungsjournal.net/ausgabe\\_0116/eichenberg\\_schott\\_aden.pdf](https://www.e-beratungsjournal.net/ausgabe_0116/eichenberg_schott_aden.pdf)
- Elyoseph, Z., Hadar-Shoval, D., Asraf, K. & Lvovsky, M. (2023). ChatGPT outperforms humans in emotional awareness evaluations. *Frontiers in Psychology*, 14, 1-7. <https://doi.org/10.3389/fpsyg.2023.1199058>
- Elyoseph, Z. & Levkovich, I. (2023). Beyond human expertise: the promise and limitations of ChatGPT in suicide risk assessment. *Frontiers in Psychiatry*, 14, 1-7. <https://doi.org/10.3389/fpsyg.2023.1213141>
- Ekin, S. (2023). Prompt Engineering For ChatGPT: A Quick Guide To Techniques, Tips, And Best Practices. *TechRxiv*. <https://doi.org/10.36227/techrxiv.22683919.v2>
- Ekmekci, P. E. & Arda, B. (2020). *Artificial Intelligence and Bioethics*. Cham: Springer Nature.
- Engel, F. (2019). Beratung unter Onlinebedingungen. In Rietmann, S., Sawatzki, M. & Berg, M. (Hrsg.), *Beratung und Digitalisierung. Zwischen Euphorie und Skepsis* (S. 3-39). Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-25528-2>
- Engelhardt, E. M. (2021). *Lehrbuch Onlineberatung* (2., erw. Aufl.). Göttingen: Vandenhoeck & Ruprecht. <https://doi.org/10.13109/9783666711558>
- Engelhardt, E. M. (2023a). Berät bald der Bot!? *Jugendhilfe*, 61(5), 404-409.
- Engelhardt, E. M. (2023b). KI in der Lehre – Beraten Lernen mit ChatGPT. *Fnma Magazin*, 20(3), 27-28.

- Engelhardt, E. M. (2023c). Zwischenruf auf die Onlineberatung – eine post-pandemische Reflexion und Diskussion aktueller Herausforderungen. *E-Beratungsjournal*, 19(2), 10-24. Abgerufen am 21.07.2024: [https://www.e-beratungsjournal.net/wp-content/uploads/2024/02/Engelhardt\\_aktualisiert.pdf](https://www.e-beratungsjournal.net/wp-content/uploads/2024/02/Engelhardt_aktualisiert.pdf)
- Engelhardt, E. M. & Piekorz, K. (2022). Einführung in die Onlineberatung per Messenger. *E-Beratungsjournal*, 18(2), 18-33. Abgerufen am 02.06.2024: [https://www.e-beratungsjournal.net/wp-content/uploads/2022/02/engelhardt\\_piekorz.pdf](https://www.e-beratungsjournal.net/wp-content/uploads/2022/02/engelhardt_piekorz.pdf)
- Engelhardt, E. M. & Storch, S. D. (2013). Was ist Onlineberatung? - Versuch einer systematischen begrifflichen Einordnung der ‚Beratung im Internet‘. *E-Beratungsjournal*, 9(2), 1-12. Abgerufen am 30.05.2024: [https://e-beratungsjournal.net/ausgabe\\_0213/engelhardt\\_storch.pdf](https://e-beratungsjournal.net/ausgabe_0213/engelhardt_storch.pdf)
- Ertel, W. (2020). *Grundkurs Künstliche Intelligenz. Eine praxisorientierte Einführung* (5. Aufl.). Wiesbaden: Springer Vieweg. <https://doi.org/10.1007/978-3-658-32075-1>
- Esser, J. & Spiecker genannt Döhmann, I. (2023). Sozialdatenschutzrechtliche Betrachtung KI-gestützter Jugendfürsorge. *Jugendhilfe*, 61(5), 395-403.
- Europäische Kommission (2019). *Eine Definition der KI: Wichtigste Fähigkeiten und Wissensgebiete*. Abgerufen am 07.05.2024: [https://www.bundesnetzagentur.de/DE/Fachthemen/Digitalisierung/Mittelstand/Downloads/Experten.pdf?\\_\\_blob=publicationFile&v=2](https://www.bundesnetzagentur.de/DE/Fachthemen/Digitalisierung/Mittelstand/Downloads/Experten.pdf?__blob=publicationFile&v=2)
- Ewbank, M. P., Cummins, R., Tablan, V., Bateup, S., Catarino, A., Martin, A. J. & Blackwell, A. D. (2020a). Quantifying the Association Between Psychotherapy Content and Clinical Outcomes Using Deep Learning. *JAMA Psychiatry*, 77(1), 35-43. <https://doi.org/10.1001/jamapsychiatry.2019.2664>
- Ewbank, M. P., Cummins, R., Tablan, V., Catarino, A., Buchholz, S. & Blackwell, A. D. (2020b). Understanding the relationship between patient language and outcomes in internet-enabled cognitive behavioural therapy: A deep learning approach to automatic coding of session transcripts. *Psychotherapy Research*, 31(3), 300-312. <https://doi.org/10.1080/10503307.2020.1788740>
- Filipiak, K. (2016). *Befähigung durch Beratung. Begleitete Bewältigung beruflicher Umbruchsituationen*. Abgerufen am 09.07.2024: <http://library.fes.de/pdf-files/wiso/12823.pdf>
- Fleischer, J. (2018). *Erwachsenwerden als Prozess mediatisierter Sozialisation. Wie junge Menschen mit Hilfe online verfügbarer Informationen eigene Entwicklungsaufgaben bearbeiten*. Baden-Baden: Nomos.
- Fokina, M. (2024). *The Future of Chatbots: 80+ Statistics, Industry Trends & Benchmarks*. Abgerufen am 10.07.2024: <https://www.tidio.com/blog/chatbot-statistics/>

- FragLovis (2024). *Die neue Art von Sex Education*. FragLovis. Abgerufen am: 15.07.2024: <https://fraglovis.de/>
- Frischholz, A. (2024). *GPT-4o: OpenAIs neues Spitzenmodell macht ChatGPT zum Echtzeit-Gesprächspartner*. Abgerufen am 09.07.2024: <https://www.computerbase.de/2024-05/gpt-4o-openais-neues-spitzenmodell-macht-chatgpt-zum-echtzeit-gespraechspartner/>
- Fulmer, R. (2019). Artificial intelligence and counseling: Four levels of implementation. *Theory & Psychology* 29(6), 807-819. <https://doi.org/10.1177/0959354319853045>
- Geizler, L. (2023). Imitation in automata and robots: A philosophical case study on Kempelen. *Studies in History and Philosophy of Science*, 100, 22-31. <https://doi.org/10.1016/j.shpsa.2023.05.004>
- Gershgor, D. (2017). *The data that transformed AI research—and possibly the world*. Abgerufen am 17.05.2024: <https://qz.com/1034972/the-data-that-changed-the-direction-of-ai-research-and-possibly-the-world>
- Görder, B. (2021). Die Macht der Muster. Die Ethik der Sozialen Arbeit vor professionsbezogenen und gesellschaftlichen Herausforderungen durch ‚künstliche Intelligenz‘. *Ethikjournal*, 7(2), 1-22. Abgerufen am 09.07.2024: [https://www.ethikjournal.de/fileadmin/user\\_upload/ethikjournal/Texte\\_Ausgabe\\_2021\\_2/Goerder\\_Ethikjournal\\_2.2021.pdf](https://www.ethikjournal.de/fileadmin/user_upload/ethikjournal/Texte_Ausgabe_2021_2/Goerder_Ethikjournal_2.2021.pdf)
- Grandeit, P., Haberkern, C., Lang, M., Albrecht, J. & Lehmann, R. (2020). Using BERT for qualitative content analysis in psychosocial online counseling. *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science*, 11-23. <https://doi.org/10.18653/v1/2020.nlpccs-1.2>
- Gravelmann, R. (2024a). *Jugend online! Soziale Arbeit offline? Digitale Lebenswelten junger Menschen als Herausforderung für die Praxis Sozialer Arbeit*. Weinheim: Beltz Juventa
- Gravelmann, R. (2024b). Künstliche Intelligenz in der Sozialen Arbeit – Zwischen Bedenken und Optionen. *Theorie und Praxis in der Sozialen Arbeit*, 75(1), 4-17.
- Grawe, K. (2005). (Wie) kann Psychotherapie durch empirische Validierung wirksamer werden? *Psychotherapeutenjournal*, 4(1), 4-11. Abgerufen am 17.05.2024: [https://www.psychotherapeutenjournal.de/ptk/web.nsf/gfx/med\\_BWUE-8U59XX\\_3816F/\\$file/ptj\\_2005-1.pdf](https://www.psychotherapeutenjournal.de/ptk/web.nsf/gfx/med_BWUE-8U59XX_3816F/$file/ptj_2005-1.pdf)
- Greef, S. (2023). *Staat und Staatlichkeit im digitalen Zeitalter. Politische Steuerung im Wandel*. Bielefeld: transcript Verlag. <https://doi.org/10.14361/9783839470657>
- Griol, D., Carbó, J. & Molina, J. M. (2013). AN AUTOMATIC DIALOG SIMULATION TECHNIQUE TO DEVELOP AND EVALUATE INTERACTIVE CONVERSATIONAL AGENTS. *Applied Artificial Intelligence*, 27(9), 759–780. <https://doi.org/10.1080/08839514.2013.835230>

- Gutwald, R., Burghardt, J., Kraus, M., Reder, M., Lehmann, R. & Müller, N. (2021). Soziale Konflikte und Digitalisierung. Chancen und Risiken digitaler Technologien bei der Einschätzung von Kindeswohlgefährdungen. *Ethikjournal*, 7(2), 1-20. Abgerufen am 09.07.2024: [https://www.ethikjournal.de/fileadmin/user\\_upload/ethikjournal/Texte\\_Ausgabe\\_2021\\_2/Gutwald\\_u.a.\\_Ethikjournal\\_2.2021.pdf](https://www.ethikjournal.de/fileadmin/user_upload/ethikjournal/Texte_Ausgabe_2021_2/Gutwald_u.a._Ethikjournal_2.2021.pdf)
- Habicht, J., Viswanathan, S., Carrington, B., Hauser, T. U., Harper, R. & Rollwage, M. (2024). Closing the accessibility gap to mental health treatment with a personalized self-referral chatbot. *Nature Medicine*, 30(2), 595–602. <https://doi.org/10.1038/s41591-023-02766-x>
- Hackbart, M. (2022). *Die Lehrperson als Golem oder Galatea. Empirische Untersuchungen zu Kompetenzerwartungen von Lernenden an Lehrpersonen*. Wiesbaden: Springer. <https://doi.org/10.1007/978-3-658-39444-8>
- Haenlein, M. & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, 61(4), 1-10. <https://doi.org/10.1177/0008125619864925>
- Haque, M. D. R. & Rubya, S. (2023). An Overview of Chatbot-Based Mobile Mental Health Apps: Insights From App Description and User Reviews. *JMIR Mhealth And Uhealth*, 11, e44838, 1-18. <https://doi.org/10.2196/44838>
- Heaven, W. D. (2024). *Why does AI hallucinate?* Abgerufen am 11.07.2024: <https://www.technologyreview.com/2024/06/18/1093440/what-causes-ai-hallucinate-chatbots/>
- Heider, T. (2008). Prozesse in der Chatberatung. *E-Beratungsjournal*, 4(2), 1-10. Abgerufen am 02.06.2024: [https://www.e-beratungsjournal.net/ausgabe\\_0208/heider.pdf](https://www.e-beratungsjournal.net/ausgabe_0208/heider.pdf)
- Heimes, S. (2023). *Therapeutisches Schreiben bei Depressionen. Hilfe zur Selbsthilfe*. Stuttgart: W. Kohlhammer.
- Heister, W. & Krings, M. (2024). Digitaler Wandel im Sozialwesen: Bestandsaufnahme und Handlungsempfehlungen. *Theorie und Praxis in der Sozialen Arbeit*, 75(1), 41-48.
- Herbig, D. (2018). *Google Duplex: Guten Tag, Sie sprechen mit einer KI*. Abgerufen am 17.05.2024: <https://www.heise.de/news/Google-Duplex-Guten-Tag-Sie-sprechen-mit-einer-KI-4046987.html>
- Hermann, I. (2018). Boundaries and Otherness in Science Fiction: We Cannot Escape the Human Condition. *Text Matters*, 8, 212-226. <https://doi.org/10.1515/textmat-2018-0013>
- Hintenberger, G. (2006). *\*taschentuchreich\** – Überlegungen zur Methodik der Chatberatung. *E-Beratungsjournal*, 2(2), 1-14. Abgerufen am 02.06.2024: [https://www.e-beratungsjournal.net/ausgabe\\_0206/hintenberger.pdf](https://www.e-beratungsjournal.net/ausgabe_0206/hintenberger.pdf)

- Hintenberger, G. (2009). Der Chat als neues Beratungsmedium. In Kühne, S. & Hintenberger, G. (Hrsg.), *Handbuch Online-Beratung. Psychosoziale Beratung im Internet* (2. Aufl.) (S. 69-78). Göttingen: Vandenhoeck & Ruprecht.
- Hintenberger, G. (2010). Oraliteralität als Interventionsstrategie in der Mail-Beratung mit Jugendlichen. *E-Beratungsjournal*, 6(2), 1-9. Abgerufen am 02.06.2024: [https://www.e-beratungsjournal.net/ausgabe\\_0210/hintenberger.pdf](https://www.e-beratungsjournal.net/ausgabe_0210/hintenberger.pdf)
- Hörmann, M., Aeberhardt, D., Flammer, P., Tanner, A., Tschopp, D. & Wenzel, J. (2019). *Face-to-Face und mehr – neue Modelle für Mediennutzung in der Beratung. Schlussbericht zum Projekt*. <https://doi.org/10.5281/zenodo.2556623>
- Hörmann, M., Tschopp, D. & Wenzel, J. (2023). *Digitale Beratung in der Sozialen Arbeit*. Stuttgart: W. Kohlhammer.
- Hoff, T. & Zwicker-Pelzer, R. (2015). Erklärungsmuster: Counseling-Bedeutung und -verständnis aus interdisziplinärer Sicht. In T. Hoff & R. Zwicker-Pelzer (Hrsg.), *Beratung und Beratungswissenschaft* (S. 45-47). Baden-Baden: Nomos.
- Hoghe, J., Limmer, R. & Reule, C. (2020). Zugänge zur Onlineberatung: Wie finden Jugendliche psychosoziale Beratungsangebote Sozialer Dienste im Internet?. *E-Beratungsjournal*, 16(1), 16-32. Abgerufen am 14.06.2024: [https://www.e-beratungsjournal.net/wp-content/uploads/2020/01/hoghe\\_limmer\\_reule.pdf](https://www.e-beratungsjournal.net/wp-content/uploads/2020/01/hoghe_limmer_reule.pdf)
- Høiland, C. G., Følstad, A. & Karahasanovic, A. (2020). Hi, can I help? Exploring how to design a mental health chatbot for youths. *Human Technology* 16(2), 139-169. <https://doi.org/10.17011/ht/urn.202008245640>
- Hu, K. (2023). *ChatGPT sets record for fastest-growing user base - analyst note*. Abgerufen am 26.05.2024: <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>
- Hurwitz, J. & Kirsch D. (2018). *Machine Learning for dummies*. New Jersey: Wiley & Sons.
- Hutchins, D., Schlag, I., Wu, Y., Dyer, E. & Neyshabur, B. (2022). *Block-Recurrent transformers*. <https://doi.org/10.48550/arxiv.2203.07852>
- Institut für E-Beratung (o. D.a). *Der virtuelle Klient*. Abgerufen am 19.07.2024: <https://www.e-beratungsinstitut.de/projekte/der-virtuelle-klient/>
- Institut für E-Beratung (o. D.b). *KI gestützte Assistenz*. Abgerufen am 17.07.2024: <https://www.e-beratungsinstitut.de/projekte/kia/>
- Jaeger, L. (2023). *Geniale Frauen in der Wissenschaft. Versteckte Beiträge, die die Welt verändert haben*. Berlin: Springer. <https://doi.org/10.1007/978-3-662-66528-2>



- Jordan, J. M. (2019). *The Czech Play That Gave Us the Word 'Robot'*. Abgerufen am: 09.05.2024: <https://thereader.mitpress.mit.edu/origin-word-robot-rur/>
- Jürgens, J. (2024). *Alles aufgesaugt*. Abgerufen am 20.05.2024: <https://www.zeit.de/2024/21/kuenstliche-intelligenz-trainingsdaten-suche-google-meta-openai>
- JugendNotmail (o. D.). *FAQs*. Abgerufen am 16.07.2024: <https://jugendnotmail.de/faq/>
- Kahn, J. (2002). *It's Alive*. *Wired*, 3(10). Abgerufen am 07.05.2024: <https://www.wired.com/2002/03/everywhere/>
- Kaplan, A. (2022). *Artificial Intelligence, Business and Civilization. Our Fate Made in Machines*. New York & London: Routledge.
- Kassenärztliche Bundesvereinigung (2022). *Beschluss des Bewertungsausschusses nach § 87 Abs. 1 Satz 1 SGB V in seiner 597. Sitzung (schriftliche Beschlussfassung) zur Änderung des Einheitlichen Bewertungsmaßstabes (EBM) mit Wirkung zum 1. Juli 2022*. Abgerufen am 02.06.2024: [https://www.kbv.de/media/sp/EBM\\_2022-07-01\\_BA\\_597\\_BeeG\\_Begrenzungsregeln\\_Videosprechstunde.pdf](https://www.kbv.de/media/sp/EBM_2022-07-01_BA_597_BeeG_Begrenzungsregeln_Videosprechstunde.pdf)
- Kelly, R. (2023). *IBM plans hiring freeze for roles that are replaceable by AI*. Abgerufen am 25.07.2024: <https://www.itpro.com/technology/artificial-intelligence/ibm-plans-hiring-freeze-for-roles-that-are-replaceable-by-ai>
- Kemper, J. (2024). *Elon Musk hält GPT-4 für AGI, verklagt OpenAI und will sie zur offenen Entwicklung zwingen*. Abgerufen am 19.05.2024: <https://the-decoder.de/elon-musk-glaubt-gpt-4-ist-agi-verklagt-openai-und-will-sie-zu-offener-entwicklung-zwingen/>
- Klein, A. & Pulver, C. (2020). Onlineberatung. In Kutscher, N., Ley, T., Seelmeyer, U., Siller, F., Tillmann, A. & Zorn, I. (Hrsg.), *Handbuch. Soziale Arbeit und Digitalisierung* (S. 190-200). Weinheim: Beltz Juventa.
- Knapp, M. (2023). *How does ChatGPT work, anyway?* Abgerufen am 22.05.2024: <https://www.currentdraft.com/p/how-does-chatgpt-work-anyway>
- Knatz, B. (2008). Zwischen den Zeilen. *E-Beratungsjournal*, 4(1), 1-13. Abgerufen am 30.05.2024: [https://www.e-beratungsjournal.net/ausgabe\\_0108/knatz.pdf](https://www.e-beratungsjournal.net/ausgabe_0108/knatz.pdf)
- Knatz, B. (2009a). Das Vier-Folien-Konzept. In Kühne, S. & Hintenberger, G. (Hrsg.), *Handbuch Online-Beratung. Psychosoziale Beratung im Internet* (2. Aufl.) (S. 106-116). Göttingen: Vandenhoeck & Ruprecht.
- Knatz, B. (2009b). Die webbasierte Mail-Beratung. In Kühne, S. & Hintenberger, G. (Hrsg.), *Handbuch Online-Beratung. Psychosoziale Beratung im Internet* (2. Aufl.) (S. 59-67). Göttingen: Vandenhoeck & Ruprecht.

- Knatz, B. (2022). Digitale Beratungsformate und ihre Besonderheiten für die Beratung. *Zeitschrift für Theorie und Praxis in der Beratung*, 23(1), 12-32. Abgerufen am 01.06.2024: <https://www.schreiben-tut-der-seele-gut.de/wp-content/uploads/2023/08/BA-1-2022.pdf>
- Knatz, B. & Dodier, B. (2003). *Hilfe aus dem Netz. Theorie und Praxis der Beratung per E-Mail*. Stuttgart: Pfeiffer bei Klett-Cotta.
- Kremp, M. (2024a). *Datenhändler bieten im Internet Standortdaten von Millionen Handys an*. Abgerufen am 16.07.2024: <https://www.spiegel.de/netzwelt/netzpolitik/datenhaendler-bieten-im-internet-standortdaten-von-millionen-handys-an-a-5f2b7150-7317-4216-a158-d0f80ee2f13f>
- Kremp, M. (2024b). *Elon Musk will weltgrößten Supercomputer bauen*. Abgerufen am 28.05.2024: <https://www.spiegel.de/netzwelt/web/elon-musk-will-weltgroessten-ki-supercomputer-bauen-a-0db5103b-aad8-4809-bc23-d4325f10d442>
- Kreutzer, R. T. (2023). *Künstliche Intelligenz verstehen. Grundlagen – Use-Cases – unternehmens-eigene KI-Journey* (2., vollst. überarb. und erw. Aufl.). Wiesbaden: Springer Gabler. <https://doi.org/10.1007/978-3-658-42598-2>
- Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2012). *ImageNet Classification with Deep Convolutional Neural Networks*. Abgerufen am: 17.05.2024: [https://proceedings.neurips.cc/paper\\_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf)
- Kubitza, E. (2024). ChatGPT als digitale Anlaufstelle für Beratungsanliegen zum Thema Schwangerschaftsabbruch. Eine explorative Untersuchung. *E-Beratungsjournal*, 20(1), 79-104. Abgerufen am 14.07.2024: [http://www.e-beratungsjournal.net/wp-content/uploads/2024/05/05\\_Kubitza.pdf](http://www.e-beratungsjournal.net/wp-content/uploads/2024/05/05_Kubitza.pdf)
- Kühne, S. & Hintenberger, G. (2020). Onlineberatung und -therapie in Zeiten der Krise. Ein Überblick. *E-Beratungsjournal*, 16(1), 33-45. Abgerufen am 02.06.2024: [https://www.e-beratungsjournal.net/wp-content/uploads/2020/03/kuehne\\_hintenberger.pdf](https://www.e-beratungsjournal.net/wp-content/uploads/2020/03/kuehne_hintenberger.pdf)
- Kutscher, N. (2019). Digitalisierung in der Sozialen Arbeit. In Rietmann, S., Sawatzki, M. & Berg, M. (Hrsg.), *Beratung und Digitalisierung. Zwischen Euphorie und Skepsis* (S. 41-56). Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-25528-2>
- Kutscher, N., Ley, T., Seelmeyer, U., Siller, F., Tillmann, A. & Zorn, I. (2020). Einleitung. In Dies. (Hrsg.), *Handbuch. Soziale Arbeit und Digitalisierung* (S. 9-16). Weinheim: Beltz Juventa.
- Lang, J. (2015). Wo steht die Onlineberatung/-therapie in 10 Jahren? *E-Beratungsjournal*, 11(2), 93-104. Abgerufen am 02.06.2024: [https://www.e-beratungsjournal.net/ausgabe\\_0215/lang.pdf](https://www.e-beratungsjournal.net/ausgabe_0215/lang.pdf)

- Lehmann, R. (2020). Die Professionalisierung der Onlineberatung. *FORUM Sexualaufklärung und Familienplanung: Informationsdienst der Bundeszentrale für gesundheitliche Aufklärung (BZgA)*, 2, 3-5.
- Lehmann, R. (2024). Herausforderungen der künstlichen Intelligenz in der Sozialwirtschaft. In Kolhoff, L. (Hrsg.), *Aktuelle Diskurse in der Sozialwirtschaft V* (S. 163-174). Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-43290-4>
- Lenzen, M. (2020). *Künstliche Intelligenz. Fakten, Chancen, Risiken*. München: C. H. Beck. <https://doi.org/10.17104/9783406751264-2>
- Leven, I. & Utzmann, H. (2019). Die Vielfalt der Digital Natives. In Albert, M., Hurrelmann, K., Quenzel, G. & Kantar (Hrsg.), *18. Shell Jugendstudie. Jugend 2019. Eine Generation meldet sich zu Wort* (S. 247-312). Weinheim: Beltz
- Lindemann, H. (2023). *Einführung in die systemisch-lösungsorientierte Gesprächsführung und Beratung*. Göttingen: Vandenhoeck & Ruprecht.
- Linnemann, G. A., Löhe, J. & Rottkemper, B. (2023a). Bedeutung von künstlicher Intelligenz in der Sozialen Arbeit. *Soziale Passagen*, 15(1), 197-211. <https://doi.org/10.1007/s12592-023-00455-7>
- Linnemann, G., Löhe, J. & Rottkemper, B. (2023b). Der Einsatz von Künstlicher Intelligenz in der Kinder- und Jugendhilfe im DACH-Raum. *Jugendhilfe*, 61(5), 415-422.
- Linnemann, G., Löhe, J. & Rottkemper, B. (2024). Bedeutung von Selbstoffenbarungseffekten in quasisozialen Beziehungen mit auf generativer KI basierten Systemen in Settings von Onlineberatung und -therapie. *E-Beratungsjournal*, 20(1), 1-21. Abgerufen am 09.07.2024: [https://www.e-beratungsjournal.net/wp-content/uploads/2024/05/01\\_Linnemann\\_Loeh\\_Loehe\\_Rottkemper.pdf](https://www.e-beratungsjournal.net/wp-content/uploads/2024/05/01_Linnemann_Loeh_Loehe_Rottkemper.pdf)
- Lippert, C., Rudolph, E., Poltermann, A., Engert, N., Lehmann, R. & Albrecht, J. (2024). Generative KI in der beraterischen Ausbildung. Der Einsatz eines\*r virtuellen Klient\*in als Übungstool für angehende Onlineberater\*innen. *E-Beratungsjournal*, 20(1), 41-60. Abgerufen am 18.07.2024: [https://www.e-beratungsjournal.net/wp-content/uploads/2024/05/03\\_Lippert\\_et\\_al.pdf](https://www.e-beratungsjournal.net/wp-content/uploads/2024/05/03_Lippert_et_al.pdf)
- Liu, H., Peng, H., Song, X., Xu, C. & Zhang, M. (2022). Using AI chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness. *Internet Interventions*, 27, 100495, 1-8. <https://doi.org/10.1016/j.invent.2022.100495>
- Lob-Hüdepohl, A. (2023). Künstliche Intelligenz in der Kinder- und Jugendhilfe – Ethische Grundsätze des Deutschen Ethikrates. *Jugendhilfe*, 61(5), 383-387.
- Lohaus, A. & Vierhaus, M. (2019). *Entwicklungspsychologie des Kindes- und Jugendalters für Bachelor* (4. vollst. Überarb. Aufl.). Wiesbaden: Springer. <https://doi.org/10.1007/978-3-662-59192-5>

- Lucas, G. M., Gratch, J., King, A. & Morency, L. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37, 94-100. <https://doi.org/10.1016/j.chb.2014.04.043>
- Mainzer, K. & Kahle, R. (2022). *Grenzen der KI – theoretisch, praktisch, ethisch*. Berlin: Springer. <https://doi.org/10.1007/978-3-662-65011-0>
- Marche, S. (2021). *The Computers Are Getting Better at Writing*. Abgerufen am 25.07.2024: <https://www.newyorker.com/culture/cultural-comment/the-computers-are-getting-better-at-writing>
- Markoff, J. (2005). *Behind Artificial Intelligence, a Squadron of Bright Real People*. Abgerufen am: 15.05.2024: <https://www.nytimes.com/2005/10/14/technology/behind-artificial-intelligence-a-squadron-of-bright-real-people.html>
- Martin, A. & Pengel, N. (2024). Beratung via Chatbot? Möglichkeiten und Anforderungen beim Einsatz Generativer KI in einem bildungswissenschaftlichen Lehramts-Modul. *E-Beratungsjournal*, 20(1), 22-40. Abgerufen am 14.07.2024: [https://www.e-beratungsjournal.net/wp-content/uploads/2024/05/02\\_Martin\\_Pengel.pdf](https://www.e-beratungsjournal.net/wp-content/uploads/2024/05/02_Martin_Pengel.pdf)
- Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reule, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R. & Clark, J. (2024). *The Artificial Intelligence Index Report 2024*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. Abgerufen am 26.05.2024: [https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI\\_2024\\_AI-Index-Report.pdf](https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_2024_AI-Index-Report.pdf)
- Maurya, R. K. (2023). A qualitative content analysis of ChatGPT's client simulation role-play for practising counselling skills. *Counselling and Psychotherapy Research*. <https://doi.org/10.1002/capr.12699>
- Maurya, R. K. & DeDiego, A. C. (2023). Artificial intelligence integration in counsellor education and supervision: A roadmap for future directions and research inquiries. *Counselling and Psychotherapy Research*. <https://doi.org/10.1002/capr.12727>
- Maurya, R. K., Montesinos, S., Bogomaz, M. & DeDiego, A. C. (2024). Assessing the use of ChatGPT as a psychoeducational tool for mental health practice. *Counselling and Psychotherapy Research*. <https://doi.org/10.1002/capr.12759>
- McCarthy, J., Minsky, M. L., Rochester, N. & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*, 27(4), 12-14. <https://doi.org/10.1609/aimag.v27i4.1904>
- McCorduck, P. (2004). *Machines Who think. A Personal Inquiry into the History and Prospects of Artificial Intelligence* (2nd Edition). Natick: A K Peters, Ltd.
- McCulloch, W. & Pitts, W. (1943). A Logical Calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115-133.

- Medienpädagogischer Forschungsverbund Südwest (2022). *KIM-Studie 2022. Kindheit, Internet, Medien. Basisuntersuchung zum Medienumgang 6- bis 13-Jähriger*. Abgerufen am 14.06.2024: <https://www.mpfs.de/studien/kim-studie/2022/>
- Medienpädagogischer Forschungsverbund Südwest (2023). *JIM-Studie 2023. Jugend, Information, Medien. Basisuntersuchung zum Medienumgang 12- bis 19-Jähriger*. Abgerufen am 14.06.2024: <https://www.mpfs.de/studien/jim-studie/2023/>
- Mina Chatbot App (2024). *Start*. Abgerufen am 15.07.2024: <https://www.minabot.ai/de/index.html>
- Miner, A. S., Fleming, S. L., Haque, A., Fries, J. A., Althoff, T., Wilfley, D. E., Agras, W. S., Milstein, A., Hancock, J., Asch, S. M., Stirman, S. W., Arnow, B. A. & Shah, N. H. (2022). A computational approach to measure the linguistic characteristics of psychotherapy timing, responsiveness, and consistency. *Npj Mental Health Research*, 1(1), 1-12. <https://doi.org/10.1038/s44184-022-00020-9>
- Mockenhaupt, A. & Schlagenhaut, T. (2024). *Digitalisierung und Künstliche Intelligenz in der Produktion. Grundlagen und Anwendung* (2. Aufl.). Wiesbaden: Springer Vieweg. <https://doi.org/10.1007/978-3-658-41935-6>
- Mode-Scheibel, B. (2014) Analyse verschiedener Beratungsansätze und Interventionen in der psychosozialen Onlineberatung. *E-Beratungsjournal*, 10(2), 88-102. Abgerufen am 02.06.2024: [https://www.e-beratungsjournal.net/ausgabe\\_0214/mode-scheibel.pdf](https://www.e-beratungsjournal.net/ausgabe_0214/mode-scheibel.pdf)
- Mollick, E. (2024). *Gradually, then Suddenly: Upon the Threshold*. Abgerufen am 21.07.2024: <https://www.oneusefulthing.org/p/gradually-then-suddenly-upon-the>
- Moravec, H. (1988). *Mind Children. The Future of Robot and Human Intelligence*. Cambridge (USA): Harvard University Press
- Moring, A. (2023). *Künstliche Intelligenz und Intuition. Robuste und nachhaltige Entscheidungen in digitalen Arbeitswelten*. Wiesbaden: Springer Gabler. <https://doi.org/10.1007/978-3-658-42018-5>
- Mornhinweg, N. (2018). Intelligente Chatbots im Gesundheitswesen – Prozessoptimierung im Gesundheitsbereich, von der Dienstleistung bis hin zum Controlling. In Pfannstiel, M. A. & Straub S. (Hrsg.), *Dienstleistungscontrolling in Gesundheitseinrichtungen. Aktuelle Beiträge aus Wissenschaft und Praxis* (S. 143-160). Wiesbaden: Springer Gabler. <https://doi.org/10.1007/978-3-658-21710-5>
- Muhle, F. (2023). (Vor-)Geschichte der Künstliche-Intelligenz-Forschung und der sozialen Robotik. In F. Muhle (Hrsg.), *Soziale Robotik. Eine sozialwissenschaftliche Einführung* (S. 13-29). Berlin/Boston: De Gruyter Oldenbourg. <https://doi.org/10.1515/9783110714944-002>

- Nazir, A. & Wang, Z. (2023). A comprehensive survey of ChatGPT: Advancements, applications, prospects, and challenges. *Meta-Radiology*, 1(2), 100022, 1-12. <https://doi.org/10.1016/j.metrad.2023.100022>
- Nestmann, F. & Sickendiek, U. (2018). Beratung. In Otto, H., Thiersch, H., Treptow, R. & Ziegler, H. (Hrsg.), *Handbuch Soziale Arbeit* (6., überarb. Aufl.) (S. 110-120). München: Ernst Reinhardt.
- Nicoletti, L. & Bass, D. (2023). *HUMANS ARE BIASED. GENERATIVE AI IS EVEN WORSE*. Abgerufen am 17.07.2024: <https://www.bloomberg.com/graphics/2023-generative-ai-bias/>
- Nilsson, N. J. (2010). *The Quest for Artificial Intelligence. A History of Ideas and Achievements*. New York: Cambridge University Press.
- O'Regan, G. (2021). *A Brief History of Computing*. 3rd Edition. Cham: Springer Nature. <https://doi.org/10.1007/978-3-030-66599-9>
- Olson, P. (2019). *Nearly half of all 'AI Startups' are cashing in on hype*. Abgerufen am 07.05.2024: <https://www.forbes.com/sites/parmyolson/2019/03/04/nearly-half-of-all-ai-startups-are-cashing-in-on-hype/#38a6b48fd022>
- Olson, P. (2023). *Meet the \$4 Billion AI Superstars That Google Lost*. Abgerufen am 17.05.2024: <https://www.bloomberg.com/opinion/features/2023-07-13/ex-google-scientists-kick-started-the-generative-ai-era-of-chatgpt-midjourney>
- OpenAI (o. D.). *Videos*. Abgerufen am 17.05.2024: <https://www.youtube.com/@OpenAI/videos>
- Oppenlaender, J. (2022). The creativity of Text-to-Image generation. *Academic Mindtrek '22: Proceedings of the 25th International Academic Mindtrek Conference*, 192-202. <https://doi.org/10.1145/3569219.3569352>
- Oswald, A. (2018). Onlineberatung - Ist Mailberatung noch eine angemessene Form, um Jugendliche und junge Erwachsene in (suizidalen) Krisen zu erreichen? Eine sozio-technische Analyse. *E-Beratungsjournal*, 14(1), 1-15. Abgerufen am 02.06.2024: <https://www.e-beratungsjournal.net/wp-content/uploads/2018/03/oswald.pdf>
- Petzold, M. (2006). Psychologische Aspekte der Online-Kommunikation. *E-Beratungsjournal*, 2(2), 1-12. Abgerufen am 14.06.2024: [https://www.e-beratungsjournal.net/ausgabe\\_0206/petzold.pdf](https://www.e-beratungsjournal.net/ausgabe_0206/petzold.pdf)
- Planned Parenthood (2024). *Introducing Roo*. Abgerufen am 15.07.2024: <https://www.plannedparenthood.org/learn/roo-sexual-health-chatbot>
- Ploil, E. O. (2009). *Psychosoziale Online-Beratung*. München: Ernst Reinhardt.

- Poltermann, A., Rudolph, E., Steigerwald, P. & Lehmann, R. (2024). KI und Soziale Arbeit – Was ist heute möglich? *Sozialwirtschaft*, 34(1), 21-23. <https://doi.org/10.5771/1613-0707-2024-1-21>
- Pontes, U. (2014). *Der Mensch – ein Schlaumeier*. Abgerufen am 22.07.2024: <https://www.das-gehirn.info/denken/intelligenz/der-mensch-ein-schlaumeier>
- Prescott, J., Ogilvie, L. & Hanley, T. (2023). Student therapists' experiences of learning using a machine client: A proof-of-concept exploration of an emotionally responsive interactive client (ERIC). *Counselling and Psychotherapy Research*, 24(2), 524-531. <https://doi.org/10.1002/capr.12685>
- Quenzel, G. & Hurrelmann, K. (2022). *Lebensphase Jugend. Eine Einführung in die sozialwissenschaftliche Jugendforschung* (14., überarb. Aufl.). Weinheim: Beltz Juventa.
- Rainsberger, L. (2021). *KI – die neue Intelligenz im Vertrieb*. Wiesbaden: Springer Gabler. <https://doi.org/10.1007/978-3-658-31773-7>
- Raitoharju, J. (2022). Convolutional neural networks. In Iosifidis A. & Tefas A. (Hrsg.), *Deep Learning for Robot Perception and Cognition* (S. 35-69). London: Academic Press. <https://doi.org/10.1016/B978-0-32-385787-1.00008-7>
- Rapp, A., Curti L., & Boldi, A. (2021). The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies*, 151, 102630. <https://doi.org/10.1016/j.ijhcs.2021.102630>
- Reamer, F. G. (2023). Artificial Intelligence in Social Work: Emerging Ethical Issues. *International Journal Of Social Work Values And Ethics*, 20(2), 52-71. <https://doi.org/10.55521/10-020-205>
- Rehme, N. & Seelmeyer, U. (2023). *Einsatz digitaler Technologien in der Kinder- und Jugendhilfe*. Mainz: Institut für Sozialpädagogische Forschung Mainz gGmbH (ism). Abgerufen am 09.07.2024: [https://www.digitalejugendhilfe.de/fileadmin/uploads/user\\_upload/Publicationen/Webseite\\_Einsatz\\_digitaler\\_Technologien\\_in\\_der\\_KJH\\_Rehme\\_Seelmeyer.pdf](https://www.digitalejugendhilfe.de/fileadmin/uploads/user_upload/Publicationen/Webseite_Einsatz_digitaler_Technologien_in_der_KJH_Rehme_Seelmeyer.pdf)
- Reindl, R. (2018). Zum Stand der Onlineberatung in Zeiten der Digitalisierung. *E-Beratungsjournal*, 14(1), 16-26. Abgerufen am 31.05.2024: <https://www.e-beratungsjournal.net/wp-content/uploads/2018/03/reindl.pdf>
- Reindl, R. & Engelhardt, E. (2021). Onlineberatung – Herausforderungen fachliche Kompetenzen und Organisationsstrukturen. In Freier, C., König, J., Manzeschke, A. & Städtler-Mach, B. (Hrsg.), *Gegenwart und Zukunft sozialer Dienstleistungsarbeit Chancen und Risiken der Digitalisierung in der Sozialwirtschaft* (S. 117-128). Wiesbaden: Springer Nature. [https://doi.org/10.1007/978-3-658-32556-5\\_8](https://doi.org/10.1007/978-3-658-32556-5_8)

- Risau, P. (2009). Die Wahl der Technik. In Kühne, S. & Hintenberger, G. (Hrsg.), *Handbuch Online-Beratung. Psychosoziale Beratung im Internet* (2. Aufl.) (S. 201-211). Göttingen: Vandenhoeck & Ruprecht.
- Rösch, E. (2022). Mediatisierte Offene Kinder- und Jugendarbeit – Eine integrierte Weiterentwicklung der Praxis. In Fuchs, M., Gerodetti, J. & Gerngroß, M. (Hrsg.), *Offene Kinder- und Jugendarbeit in der Schweiz. Einblicke in Theorie, Konzepte, Empirie und Alltagspraxis* (S. 139-156). Wiesbaden: Springer Nature. [https://doi.org/10.1007/978-3-658-37670-3\\_7](https://doi.org/10.1007/978-3-658-37670-3_7)
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386-408. <https://doi.org/10.1037/h0042519>
- Roser, M. (2022). *The brief history of artificial intelligence: the world has changed fast — what might be next?* Abgerufen am 16.05.2024: <https://ourworldindata.org/brief-history-of-ai>
- Rudolph, E., Engert, N. & Albrecht, J. (2024). An AI-Based virtual client for educational Role-Playing in the training of online counselors. *Proceedings of the 16th International Conference on Computer Supported Education*, 2, 108-117. <https://doi.org/10.5220/0012690700003693>
- Russell, S. J. & Norvig, P. (2010). *Artificial Intelligence. A Modern Approach* (3<sup>rd</sup> Edition). New Jersey: Pearson.
- Schiffhauer, B. (2020). Assistive Technologien in der Sozialen Arbeit. In Kutscher, N., Ley, T., Seelmeyer, U., Siller, F., Tillmann, A. & Zorn, I. (Hrsg.), *Handbuch. Soziale Arbeit und Digitalisierung* (S. 265-275). Weinheim: Beltz Juventa.
- Schubert, F-J, Rohr, D. & Zwicker-Pelzer, R. (2019). *Beratung. Grundlagen – Konzepte – Anwendungsfelder*. Wiesbaden: Springer Nature. <https://doi.org/10.1007/978-3-658-20844-8>
- Schultze, N. G. (2007). Erfolgsfaktoren des virtuellen Settings in der psychologischen Internet-Beratung. *E-Beratungsjournal*, 1(3), 1-8. Abgerufen am 02.06.2024: [https://www.e-beratungsjournal.net/ausgabe\\_0107/schultze.pdf](https://www.e-beratungsjournal.net/ausgabe_0107/schultze.pdf)
- Seelmeyer, U. & Ley, T. (2018). Informationstechnologien in der Sozialen Arbeit. In Otto, H., Thiersch, H., Treptow, R. & Ziegler, H. (Hrsg.), *Handbuch Soziale Arbeit* (6., überarb. Aufl.) (S. 655-664). München: Ernst Reinhardt.
- Seelmeyer, U. & Waag, P. (2020). Hybridisierung personenbezogener sozialer Dienstleistungen In Kutscher, N., Ley, T., Seelmeyer, U., Siller, F., Tillmann, A. & Zorn, I. (Hrsg.), *Handbuch. Soziale Arbeit und Digitalisierung* (S. 180-189). Weinheim: Beltz Juventa.
- Seidensticker, K. (2022). Predictive Policing. Eine problembehaftete Methode der Kriminalprävention? In Diebel-Fischer, H., Hellmig, L. & Tischler, M. (Hrsg.), *Technik und*



- Verantwortung im Zeitalter der Digitalisierung (S. 193-218).  
[https://doi.org/10.18453/rosdok\\_id00003538](https://doi.org/10.18453/rosdok_id00003538)
- Sejnowski, T. J. & Rosenberg, C. R. (1986). NETtalk: a parallel network that learns to read aloud. The Johns Hopkins University Electrical Engineering and Computer Science Technical Report. In Anderson, J. A & Rosenfeld, E. (Hrsg.), *Neurocomputing, Volume 1: Foundations of Research* (1988) (S. 663-672). <https://doi.org/10.7551/mitpress/4943.003.0041>
- Sharma, A., Lin, I. W., Miner, A. S., Atkins, D. C. & Althoff, T. (2022). *Human-AI collaboration enables more empathic conversations in text-based Peer-to-Peer mental health support*. <https://doi.org/10.48550/arxiv.2203.15144>
- Sheikh, H., Prins, C. & Schrijvers, E. (2023) *Mission AI. The New System Technology*. Cham: Springer Nature. <https://doi.org/10.1007/978-3-031-21448-6>
- Skjuve, M., Følstad, A., Fostervold, K. I. & Brandtzaeg, P. B. (2021). My Chatbot Companion - a Study of Human-Chatbot Relationships. *International Journal of Human-computer Studies*, 149, 102601, 1-14. <https://doi.org/10.1016/j.ijhcs.2021.102601>
- Skjuve, M., Følstad, A., Fostervold, K. I. & Brandtzaeg, P. B. (2022). A longitudinal study of human–chatbot relationships. *International Journal of Human-computer Studies*, 168, 102903, 1-14. <https://doi.org/10.1016/j.ijhcs.2022.102903>
- SnehAI (2021). *HI! I AM SNEHAI*. Abgerufen am 15.07.2024: <https://snehai.org/>
- Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., De Pablo, G. S., Shin, J. I., Kirkbride, J. B., Jones, P., Kim, J. H., Kim, J. Y., Carvalho, A. F., Seeman, M. V., Correll, C. U. & Fusar-Poli, P. (2021). Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. *Molecular Psychiatry*, 27(1), 281-295. <https://doi.org/10.1038/s41380-021-01161-7>
- Statista Research Department (2024). *Größte Unternehmen der Welt nach Marktkapitalisierung im Jahr 2024*. Abgerufen am 18.05.2024: <https://de.statista.com/statistik/daten/studie/12108/umfrage/top-unternehmen-der-welt-nach-marktwert/>
- Stavridis, D. & Wacker, M. (2023). ChatGPT und künstliche Intelligenz – Die Zukunft ist jetzt! *Zeitschrift für Herz- Thorax- und Gefäßchirurgie* 37, 266-272. <https://doi.org/10.1007/s00398-023-00593-3>
- Steiner, O. & Tschopp, D. (2022). Künstliche Intelligenz in der sozialen Arbeit. *Sozial Extra*, 46(6), 466-471. <https://doi.org/10.1007/s12054-022-00546-4>
- Steinmayr, R., Schmitz, M. & Luhmann, M. (2024). *Wie einsam sind junge Erwachsene im Jahr 2024? Ergebnisse einer repräsentativen Umfrage*. Bertelsmann Stiftung, Gütersloh (Hrsg.). <https://doi.org/10.11586/2024085>

- Stinson, C. (2022). Algorithms are not neutral. *AI and Ethics*, 2(4), 763-770. <https://doi.org/10.1007/s43681-022-00136-w>
- Strick, H. K. (2013). *Heron von Alexandria (10-75)*. Abgerufen am 09.05.2024: <https://www.spektrum.de/wissen/heron-von-alexandria-10-75/1198440>
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & Behavior*, 7(3), 321-326. <https://doi.org/10.1089/1094931041291295>
- Ta, V., Griffith, C., Boatfield, C., Wang, X., Civitello, M., Bader, H., DeCero, E. & Loggarakis, A. (2020). User Experiences of Social Support From Companion Chatbots in Everyday Contexts: Thematic Analysis. *Journal Of Medical Internet Research*, 22(3), e16235, 1-10. <https://doi.org/10.2196/16235>
- Taulli, T. (2022). *Grundlagen der Künstlichen Intelligenz: Eine nichttechnische Einführung*. New York: Springer
- Teich, I. (2020). Meilensteine der Entwicklung Künstlicher Intelligenz. *Informatik Spektrum* 43, 276-284. <https://doi.org/10.1007/s00287-020-01280-5>
- Tenzer, F. (2024). *Statistiken zur Smartphone-Nutzung in Deutschland*. Abgerufen am 14.06.2024: <https://de.statista.com/themen/6137/smartphone-nutzung-in-deutschland/#topicOverview>
- Thiery, H. (2018): *Was ist Online-Beratung? Ein differenztheoretischer Vergleich*. Dudenhofen. Abgerufen am 29.05.2024: <https://dg-onlineberatung.de/wp-content/uploads/2022/09/Definition-Online-Beratung-Langtext.pdf>
- Toews, R. (2023). *Transformers Revolutionized AI. What Will Replace Them?* Abgerufen am: 17.05.2024: <https://www.forbes.com/sites/robtoews/2023/09/03/transformers-revolutionized-ai-what-will-replace-them/?sh=382f9f569c1f>
- Turing, A. M. (1936). *On Computable Numbers, With an Application to the Entscheidungsproblem*. Abgerufen am 11.05.2024: [https://www.cs.virginia.edu/~robins/Turing\\_Paper\\_1936.pdf](https://www.cs.virginia.edu/~robins/Turing_Paper_1936.pdf)
- Turing, A. M. (1948). *Intelligent Machinery*. Abgerufen am 11.05.2024: [https://ia601703.us.archive.org/23/items/turing1948/turing1948\\_text.pdf](https://ia601703.us.archive.org/23/items/turing1948/turing1948_text.pdf)
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460. <https://doi.org/10.1093/mind/LIX.236.433>
- The Royal Society (2018). *Portrayals and perceptions of AI and why they matter*. Abgerufen am 09.05.2024: <https://royalsociety.org/-/media/policy/projects/ai-narratives/ai-narratives-workshop-findings.pdf>

- Vaske, H. (2023). *Diese KI-Tools sollten Sie kennen*. Abgerufen am 24.07.2024: <https://www.computerwoche.de/a/diese-ki-tools-sollten-sie-kennen,3613851>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017). *Attention is all you need*. <https://doi.org/10.48550/arxiv.1706.03762>
- von Schlippe, A. & Schweitzer, J. (2016). *Lehrbuch der systemischen Therapie und Beratung I. Das Grundlagenwissen* (3., unveränd. Aufl.). Göttingen: Vandenhoeck & Ruprecht.
- von Wensierski, H.-J. (2018). Medien und Soziale Arbeit. In Otto, H., Thiersch, H., Treptow, R. & Ziegler, H. (Hrsg.), *Handbuch Soziale Arbeit* (6., überarb. Aufl.) (S. 981-992). München: Ernst Reinhardt.
- Waag, P., Schiffhauer, B. & Seelmeyer, U. (2020). Chatbots in der Beratung. In Ernst, G., Zühlke-Robinet, K., Finking, G. & Bach, U. (Hrsg.), *Digitale Transformation. Arbeit in Dienstleistungssystemen* (S. 181-192). Baden-Baden: Nomos.
- Walpuski, V. J. (2024). Die Entwicklung anthropomorpher Large Language Models im Coaching als Gegenstand der digitalen Plattformökonomie am Beispiel von AIMY®. *E-Beratungsjournal*, 20(1), 105-127. Abgerufen am 17.06.2024: [https://www.e-beratungsjournal.net/wp-content/uploads/2024/05/06\\_Walpuski.pdf](https://www.e-beratungsjournal.net/wp-content/uploads/2024/05/06_Walpuski.pdf)
- Wang, H., Gupta, S., Singhal, A., Muttreja, P., Singh, S., Sharma, P. & Piterova, A. (2022). An Artificial Intelligence Chatbot for Young People's Sexual and Reproductive Health in India (SnehAI): Instrumental Case Study. *JMIR. Journal Of Medical Internet Research/Journal Of Medical Internet Research*, 24(1), e29969, 1-18. <https://doi.org/10.2196/29969>
- Weinhardt, M. (2010). Ratsuchende von E-Mail-Beratungsdiensten. Entwurf einer Typologie. *E-Beratungsjournal*, 6(1), 1-12. Abgerufen am 02.06.2024: [http://www.e-beratungsjournal.net/ausgabe\\_0110/weinhardt.pdf](http://www.e-beratungsjournal.net/ausgabe_0110/weinhardt.pdf)
- Weinhardt, M. (2023a). Persönliche Beziehungen und digitale Dinge. Das Beispiel Beratung. *Österreichisches Jahrbuch für Soziale Arbeit ÖJS*, 5, 35-53. <https://doi.org/10.30424/OEJS2305035>
- Weinhardt, M. (2023b). *Verführt ChatGPT angehende Fachkräfte zum Üben? Die synthetischen Adressat:innen sind online*. Abgerufen am 09.07.2024: <https://marcweinhardt.de/verfuehrt-chatgpt-angehende-fachkraefte-zum-ueben-die-synthetischen-adressatinnen-sind-online/>
- Weinhardt, M. (2024). Beratung, Rat geben und Rat nehmen im Kontext von Digitalität und Digitalisierung. In Sauerbrey, U., Großkopf, S., Ott, C. (Hrsg.), *Empirische Ratgeberforschung. Forschungszugänge und Befunde zu Produktion, Angebot und Inanspruchnahme ratgeberhafter Medien* (S. 60-68). Bad Heilbrunn: Verlag Julius Klinkhardt.
- Weizenbaum Institut (o. D.). *Unser Leitbild*. Abgerufen am 15.05.2024: <https://www.weizenbaum-institut.de/institut/leitbild/>

- Weizenbaum, J. (1966). ELIZA – a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45. <https://doi.org/10.1145/365153.365168>
- Wennker, P. (2020). *Künstliche Intelligenz in der Praxis. Anwendung in Unternehmen und Branchen: KI wettbewerbs- und zukunftsorientiert einsetzen*. Wiesbaden: Springer Gabler. <https://doi.org/10.1007/978-3-658-30480-5>
- Wenzel, J. (2019). Chancen der Digitalisierung in der Beratung. In Rietmann, S., Sawatzki, M. & Berg, M. (Hrsg.), *Beratung und Digitalisierung. Zwischen Euphorie und Skepsis* (S. 217-227). Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-25528-2>
- White, G. (2018). *Child advice chatbots fail to spot sexual abuse*. Abgerufen am 09.07.2024: <https://www.bbc.com/news/technology-46507900>
- Wilke, H., Zimmermann, S. & Rothballer, M. (2023). Kindliche Neugier trifft Künstliche Intelligenz. *Jugendhilfe*, 61(5), 422-429.
- Williams, R. (2024). *A chatbot helped more people access mental-health services*. Abgerufen am 15.07.2024: <https://www.technologyreview.com/2024/02/05/1087690/a-chatbot-helped-more-people-access-mental-health-services/>
- Wolfert, S. & Leven, I. (2019). Freizeitgestaltung und Internetnutzung: Wie Online und Offline ineinandergreifen. In Albert, M., Hurrelmann, K., Quenzel, G. & Kantar (Hrsg.), *18. Shell Jugendstudie. Jugend 2019. Eine Generation meldet sich zu Wort* (S. 213-246). Weinheim: Beltz
- Wooldridge, M. (2023). *ChatGPT is not "true AI." A computer scientist explains why*. Abgerufen am 16.06.2024: <https://bigthink.com/the-future/artificial-general-intelligence-true-ai/>
- World Health Organization (2022). *World mental health report. Transforming mental health for all*. Abgerufen am 11.07.2024: <https://www.who.int/teams/mental-health-and-substance-use/world-mental-health-report>
- Yahoo Finance (2024). *Explosive Growth Predicted: Large Language Model Market Set to Reach USD 6.5 Billion by 2024 To USD 140.8 Billion by 2033- Dimension Market Research*. Abgerufen am 12.07.2024: <https://finance.yahoo.com/news/explosive-growth-predicted-large-language-184300698.html>
- Zagorscak, P. & Knaevelsrud, C. (2021). Online-Therapie. In Schneider, S. & Margraf, J. (Hrsg.), *Lehrbuch der Verhaltenstherapie. Band 3. Psychologische Therapie bei Indikationen im Kindes- und Jugendalter* (2. Aufl.) (S. 233-248). Berlin: Springer.
- Zinkula, J. (2023). *Diese vier Berufe wird es 2030 wegen KI nicht mehr geben, laut McKinsey*. Abgerufen am 25.07.2024: <https://www.businessinsider.de/wirtschaft/vier-berufe-sind-2030-von-ki-gefaehrdet-mckinsey/>

## **Eidesstattliche Erklärung**

Hiermit erkläre ich, Fabian Binder, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst und keine anderen Hilfsmittel als die angegebenen verwendet habe. Insbesondere versichere ich, alle wörtlichen und sinngemäßen Übernahmen aus anderen Werken als solche kenntlich gemacht zu haben.

Köln, 25.07.2024