Large Language Modelle (LLMs) in der logopädischen Tätigkeit:

Eine explorative Studie zu ausgewählten Anwendungsfällen.

Large Language Models (LLMs) in speech and language therapy:

an exploratory study of selected use cases.

Masterarbeit

Zur Erlangung des akademischen Grades

Master of Science in Natural Sciences (MSc)

der Fachhochschule FH Campus Wien
Masterstudiengang: Health Assisting Engineering

Vorgelegt von:
Walter Trückl

Personenkennzeichen:
c2210720009

Erstbetreuer:
FH-Prof. Dipl.-Ing. Mag. Dr. techn. Franz Werner

Zweitbegutachter:
Mag. Michael Bechinie

Eingereicht am:
20.10.2024

Vorwort

Die Fertigstellung dieser Masterarbeit wäre ohne die Unterstützung vieler Menschen nicht möglich gewesen.

Mein besonderer Dank gilt meinem Betreuer, FH-Prof. Dipl.-Ing. Mag. Dr. techn. Franz Werner, für seine wegweisende Unterstützung und seinen Weitblick. Zu einem Zeitpunkt, als Large Language Models noch am Beginn ihres Hypes standen, hat er sich auf dieses zukunftsweisende Thema eingelassen und es mit hilfreichen Überlegungen vorangetrieben.

Ein besonderer Dank gilt allen Logopäd*innen, die an den Tests und der Fokusgruppendiskussion teilgenommen haben. Ihre Offenheit, ihr Fachwissen und ihre konstruktiven Beiträge waren unerlässlich für den Erfolg dieser Forschungsarbeit und haben maßgeblich zu den gewonnenen Erkenntnissen beigetragen.

Ich danke meiner Freundin für ihre unermüdliche Unterstützung und Geduld während des gesamten Forschungs- und Schreibprozesses. Meiner Familie, insbesondere meinen Eltern und meiner Tochter, bin ich für ihren Rückhalt und ihr Verständnis für die zeitlichen Einschränkungen dankbar.

Danke.

Kurzfassung

Hintergrund

Die rasante Entwicklung von Large Language Modellen (LLMs) eröffnet neue Möglichkeiten in unterschiedlichen Bereichen des Gesundheitswesens. Ihre potenzielle Anwendung in der Logopädie ist jedoch bisher wenig erforscht.

Die vorliegende Masterarbeit untersucht mögliche Anwendungsfälle von LLMs in der logopädischen Tätigkeit, evaluiert ihre Akzeptanz unter praktizierenden Logopäd*innen und entwickelt erste Prototypen zur Demonstration ihrer Funktionalität.

Methodik

Ein Mixed-Methods-Ansatz wurde gewählt, bestehend aus einer systematischen Literaturrecherche, eines Fokusgruppeninterviews mit Logopäd*innen (n=5) und der iterativen Entwicklung und Evaluation von drei LLM-basierten Prototypen. Die Datenanalyse erfolgte mittels qualitativer Inhaltsanalyse nach Mayring (2015) und wurde durch eine quantitative Auswertung der Prototypenevaluation ergänzt.

Ergebnisse:

Die Studie identifizierte Anwendungsbereiche für LLMs in der Logopädie, insbesondere in der Dokumentationserstellung, Therapiematerialgenerierung und Diagnostikunterstützung. Logopäd*innen zeigten eine grundsätzlich positive Einstellung gegenüber LLMs, äußerten jedoch auch Bedenken hinsichtlich Datenschutzes und der Aufrechterhaltung der therapeutischen Beziehung. Die entwickelten Prototypen (BerichtGPT, LoCo, Blitzstarter*in) wurden von den Tester*innen weitgehend positiv bewertet hinsichtlich Funktionalität.

Schlussfolgerung

LLMs zeigen ein erhebliches Potenzial zur Unterstützung und Optimierung des therapeutischen Alltags. Ihre erfolgreiche Integration erfordert jedoch eine sorgfältige Berücksichtigung ethischer, rechtlicher und praktischer Aspekte, sowie die aktive Einbindung von Logopäd*innen in den Entwicklungsprozess. Weitere Forschung ist notwendig, um die langfristigen Auswirkungen von LLMs auf die Qualität der logopädischen Versorgung zu evaluieren und spezifische Implementierungsrichtlinien zu entwickeln.

Abstract

Background

The rapid development of large language models (LLMs) is opening up new possibilities in various areas of healthcare. However, their potential application in speech therapy has been little explored.

This master's thesis examines possible usecases of LLMs in speech therapy, evaluates their acceptance among practicing speech therapists and develops initial prototypes to demonstrate their functionality.

Methodology

A mixed-methods approach was chosen, consisting of a systematic literature review, a focus group interview with speech therapists (n=5) and the iterative development and evaluation of three LLM-based prototypes. The data analysis was carried out using qualitative content analysis according to Mayring (2015) and was supplemented by a quantitative evaluation of the prototype evaluation.

Results

The study identified promising areas of application for LLMs in speech and language therapy, particularly in documentation creation, therapy material generation and diagnostic support. Speech therapists showed a fundamentally positive attitude towards LLMs, but also expressed concerns about data protection and maintaining the therapeutic relationship. The prototypes developed (BerichtGPT, LoCo, Blitzstarter*in) were rated positively by the testers.

Conclusion

LLMs show considerable potential to support and optimize everyday therapeutic practice. However, their successful integration requires careful consideration of ethical, legal and practical aspects, as well as the active involvement of speech therapists in the development process. Further research is needed to evaluate the long-term impact of LLMs on the quality of speech therapy care and to develop specific implementation guidelines.

Abkürzungsverzeichnis

AAC	Augmentative and alternative communication
Abb.	Abbildung
ADHS	Aufmerksamkeitsdefizit-Hyperaktivitätsstörung
AI	Artificial Intelligence
ALS	Amyotrophic Lateral Sclerosis
API	Application Programming Interface
AR	Augmented Reality
BERT	Bidirectional Encoder Representation from Transformers (Sprachmodell)
CASP	Critical Appraisal Skills Programme
DSGVO	Datenschutzgrundverordnung
etc.	et cetera
FEES	Flexible endoskopische Evaluation des Schluckakts
FH	Fachhochschule
GPT	Generative Pre-trained Transformer (Sprachmodell)
HTML	Hyper Text Markup Language
IoT	Internet of Things
KI	Künstliche Intelligenz (Artificial Intelligence)
KIS	Krankenhausinformationssystem
LLM	Große Sprachmodelle (Large Language Models)
LTS	Long Time Support
MESH	Medical Subject Headings
ML	Maschinelles Lernen (Machine Learning)
MTD	Medizinisch-Technische-Dienste (seit 09/2024 Mediznisch-Therapeutisch-Diagnostisch)
NLP	Natürliche Sprachverarbeitung (Natural Language Procesing)
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PWA	Progressive Web App
RCT	Randomisierte Kontroll Studie (Randomised Control Studies)
SAR	Sozial Assistiver Roboter
SSH	Secure Shell
SSL	Secure Sockets Layer
Tab.	Tabelle
TAM	Technologie Akzeptanz Modell (Technology Acceptance Model)
UI	User Interface
UTAUT	Unified Theory of Acceptance and Use of Technology

VR	Virtual Reality
WHO	World Health Organisation
z.B.	zum Beispiel

Schlüsselbegriffe

Automatisierte Dokumentation
Digitale Transformation im Gesundheitswesen
Künstliche Intelligenz (KI)
Large Language Modelle (LLMs)
Logopädie
Multimodale Systeme
Natural Language Processing (NLP)
Nutzer*innenzentriertes Design
Personalisierte Therapie
Prompt Engineering
Technologieakzeptanz
Transformer

Inhaltsverzeichnis

1. EINLEITUNG UND FORSCHUNGSKONTEXT 11

1.1 Forschungsfragen 12

2. THEORETISCHER RAHMEN 13

2.1 Definition zentraler Begriffe 13

2.1.1 Large Language Modelle (LLMs) und Künstliche Intelligenz (KI) 13

2.1.2 Technologische Grundlagen 13

2.2 Aktueller Forschungsstand 16

2.3 Fähigkeiten und Grenzen von LLMs 17

2.4 Ethik und Datenschutz 19

3. METHODIK 20

3.1 Interdisziplinärer Ansatz Health-Assisting-Engineering 20

3.2 Literaturrecherche 20

3.2.1 Suchbegriffe und Suchstrategien 21

3.2.2 Datenbanken und Suchwerkzeuge 21

3.2.3 Einschlusskriterien und Ausschlusskriterien 21

3.2.4 KI-unterstützte Suchtools 22

3.2.5 Literaturbewertung und Synthese 24

3.2.6 Erweiterung der Suchstartegie 25

3.3 Fokusgruppeninterviews und Feedbackprozess 25

3.3.1 Rekrutierung und Auswahl der Teilnehmer*innen 25

3.3.2 Durchführung des Fokusgruppeninterviews 26

3.3.3 Feedbackprozess 26

3.3.4 Datenanalyse 26

3.4 Prototypenerstellung 27

3.4.1 Identifikation der Anwendungsfälle 27

3.4.2 Technologische Grundlagen 27

3.4.3 Validierung der Prototypen 28

3.4.4 Strukturierter Fragebogen 28

3.5 Ethik und Datenschutz 29

4. ERGEBNISSE 30

4.1 Ergebnisse der Literaturrecherche 30

4.1.1 Anwendungen in therapeutischen Berufen 34

4.2 Ergebnisse der Fokusgruppen 37

4.2.1 Digitale Transformation und Zukunftsperspektiven in der Logopädie 37

4.2.2 Ethische und Praktische Herausforderungen 38

4.2.3 Veränderte Rolle der Therapeut*innen 38
4.2.4 Patient*innenenzentrierung 38
4.2.5 Verbesserter Interdisziplinärer Informationsaustausch 39
4.2.6 Balance zwischen Technologie und Menschlichem Kontakt 39
4.2.7 Anwendungsfälle im Kontext der Logopädie 39
4.2.8 Ethische Aspekte 40
4.3 Entwicklung von Prototypen 41
4.3.1 Identifizierung und Zielsetzung 41
4.3.2 Entwicklungsansatz 44
4.3.3 Beschreibung der Prototypen 45
4.3.4 Erste Prototypen mit Botpress und Cloud-Lösungen 46
4.3.5 Erstellung einer Projektseite zur Masterarbeit 47
4.4 Ergebnisse der Kommentaranalyse 50
4.4.1 Vorgehensweise bei der Analyse 50
4.4.2 Kommentare nach Kategorie 50
4.4.3 Schlussfolgerungen aus der Kommentaranalyse 51
4.5 Implementierung von LobeChat 53
4.5.1 Prompt Engineering 54
4.5.2 Spezifische Herausforderungen während der Entwicklungsphase 54
4.6 Ergebnisse der Fragebogenauswertung 55
4.6.1 Auswertung der Fragebögen 55
4.6.2 Zusammenfassung der Ergebnisse der Fragebogenauswertung 56
5. DISKUSSION 57
5.1 Einleitung der Diskussion 57
5.2 Beantwortung von Forschungsfragen 57
5.2.1 Beantwortung der Hauptforschungsfrage 57
5.2.2 Beantwortung der Unterforschungsfrage 58
5.3 Interpretation im Kontext der bestehenden Literatur 59
5.4 Praktische Implikationen 60
5.5 Beitrag der Arbeit zur Schließung der Forschungslücke 61
5.6 Aktuelle politische und gesundheitspolitische Entwicklungen . 61
5.7 Limitationen der Studie 62
5.7.1 Methodische Reflexion 64
5.8 Ethische und rechtliche Fragestellungen 64
5.9 Ausblick 65
5.9.1 Technologische Weiterentwicklungen und Innovationen 66
5.9.2 Zukünftige Entwicklungen und Potenziale von LLMs 66

5.9.3 Integration von symbolischem Reasoning und Langzeitgedächtnis 68
5.9.4 Multimodale Sprachmodelle.....69
6. CONCLUSIO 70
6.1 Literaturverzeichnis..... 72
ABBILDUNGSVERZEICHNIS 80
TABELLENVERZEICHNIS 81
ANHANG..... 82

1. EINLEITUNG UND FORSCHUNGSKONTEXT

In den letzten Jahren gab es eine rasante Entwicklung in den Bereichen künstlicher Intelligenz (KI), maschinellen Lernens (ML) und insbesondere im Bereich von Large Language Modellen (LLMs), wie Chat-GPT. Diese Technologien haben das Potenzial, die Landschaft unterschiedlichster Lebens- und Arbeitsbereiche grundlegend zu verändern (Labkoff et al., 2024). Auch im Gesundheitswesen eröffnen sich durch diese Technologien neue Möglichkeiten für Diagnose, Therapie und Patient*inneninteraktion (Khanna et al., 2022).

In gesundheitsberuflichen Feldern, in denen insbesondere Sprache und Kommunikation von zentraler Bedeutung sind, identifizieren Fear und Gleber (2023) Potenziale von LLM unter anderem in der Optimierung von Therapiekonzepten, der Unterstützung administrativer Pflichten oder in der Individualisierung und Personalisierung therapeutischer Angebote als naheliegende Anwendungsszenarien. Untersuchungen aus verwandten Berufsfeldern zeigen, dass KI-basierte Anwendungen insbesondere in der Patient*innenkommunikation, der Dokumentation und der Patient*innennachsorge relevant sein könnten (Amann et al., 2020).

Digitalisierung im Gesundheitswesen bietet jedoch nicht nur vielfältige Möglichkeiten, sie stellt auch neue Anforderungen an Gesundheitsdienstleister*innen, Regulierungsbehörden und Patient*innen (Mennella et al., 2024). Trotz der potenziellen Vorteile von LLMs gibt es bisher nur wenige spezifische Anwendungsbeispiele und kaum empirische Studien zur Akzeptanz dieser Technologien unter Therapeut*innen *(Gao et al., 2023).*

Gesundheitsdienstleister*innen stehen vor der Herausforderung knapper zeitlicher Ressourcen und der Verpflichtung aktuelle Forschungsergebnisse, im Sinne evidenzbasierter Medizin, in ihre Praxis zu integrieren (Li et al., 2023). Diese Herausforderung könnte künftig durch einen steigenden Bedarf an medizinischen Leistungen infolge demografischer Entwicklungen und der zunehmenden Prävalenz altersbedingter Störungsbilder zusätzlich verstärkt werden (Al-Dujaily & Nabàa, 2015).

In therapeutischen Berufen, die stark auf Datenschutz, Patientensicherheit und evidenzbasiertes Handeln angewiesen sind, bestehen zudem Unsicherheiten und Skepsis gegenüber KI-Technologien, insbesondere hinsichtlich Datenschutz, Zuverlässigkeit und praktischer Anwendbarkeit (Mikkelsen et al., 2023). Neue Technologien könnten darüber hinaus auch das therapeutische Rollenverständnis verändern und möglicherweise Auswirkungen auf das Berufsbild haben (Togni et al., 2021).

Um praxistaugliche, benutzerfreundliche und effektive Unterstützung für Gesundheitsdienstleister*innen zu gewährleisten, ist die systematische Erfassung ihrer Bedürfnisse sowie die Integration von Feedbackschleifen und iterativen Entwicklungsprozessen erforderlich (Maleki Varnosfaderani & Forouzanfar, 2024)

Für erfolgreiche teletherapeutische Interventionen sei es erforderlich physische Möglichkeiten von Patient*innen, technische Anforderungen und gezielte Schulung von Logopäd*innen zu beachten. Weiters müssten evidenzbasierte Konzepte entwickelt und an die rasanten Entwicklungen in diesem Feld fortlaufend adaptiert werden, wie Law et al. (2021) in ihrem Review zu Telepraxismodellen während der Covid-Pandemie am Beispiel von Kindersprachtherapie festhalten. Eine transparente Kommunikation durch Logopäd*innen bezüglich Datenspeicherung und verwendeter Tools sei unerlässlich um Betreuungspersonen und Patient*innen im Sinne von informed consent adäquat in therapeutische Intervention einzubinden.

Daher ist Ziel dieser Masterarbeit, potenzielle Anwendungsfälle (Usecases) von LLMs in der logopädischen Praxis zu identifizieren und Einschätzungen von Logopäd*innen hinsichtlich der Anwendbarkeit dieser Technologien zu analysieren. Die Arbeit verfolgt einen interdisziplinären, dreigliedrigen Ansatz, der systematische Literaturrecherche, Prototypenentwicklung und empirische Untersuchung miteinander verbindet, um ein möglichst umfassendes Bild über das Thema zu erhalten und die Forschungsfragen eingehend beantworten zu können.

1.1 Forschungsfragen

Die gegenständliche Masterarbeit zielt darauf ab, die folgenden zentralen Fragen im Kontext der Anwendung von LLMs in der Logopädie zu beantworten. Sie gliedert sich in eine Hauptforschungsfrage:

„Welche Anwendungsfälle für Large Language Models (LLMs) könnten für die logopädische Tätigkeit relevant sein?“

und wird zur detaillierteren Untersuchung noch durch eine Unterfrage ergänzt:

„Wie schätzen Logopädinnen und Logopäden eine künftige Anwendbarkeit von LLMs in der logopädischen Praxis ein?“

Die Beantwortung dieser Fragen soll dazu beitragen, das Potenzial von LLMs in der Logopädie besser zu verstehen und Empfehlungen für die Entwicklung und Implementierung entsprechender Technologien ableiten zu können.

2. THEORETISCHER RAHMEN

Im folgenden Kapitel werden zentrale Begriffe definiert und der aktuelle Forschungsstandes zu LLMs im Gesundheitssektor dargestellt. Abschließend werden zugrundeliegende theoretische Modelle, wie ethische und datenschutzrechtliche Aspekte beschrieben.

2.1 Definition zentraler Begriffe

2.1.1 Large Language Modelle (LLMs) und Künstliche Intelligenz (KI)

Künstliche Intelligenz (KI) umschreibt laut einer vielfach zitierten Definition von Russel und Norvig (2012) ein breites Feld der Informatik, welches darauf abzielt, Systeme zu entwickeln, die Aufgaben auszuführen in der Lage sind, welche typischerweise menschliche Intelligenz erfordern. Zu diesen Aufgaben gehöre unter anderem Lernen, Problemlösen, Wahrnehmung und Sprachverständnis.

Innerhalb des Bereichs der KI wird in weitere Teilgebiete und Ansätze unterschieden, einschließlich maschinellen Lernens (ML) und Deep Learning (DL), welche auch die Grundlage für LLMs bilden (Labkoff et al., 2024).

ML konzentriert sich als Teilbereich der KI, auf die Erstellung von Algorithmen und statistischen Modellen, welche es Computersystemen ermöglichen, ihre Leistung bei bestimmten Aufgaben durch Erfahrung und Wiederholung zu verbessern, ohne für diese Aufgaben und Lösungsstrategien explizit programmiert zu werden (Mitchell, 2013). Algorithmen lernen Muster aus Daten zu erkennen und treffen Vorhersagen oder Entscheidungen auf Grundlage dieser erlernten Muster. Maschinelles Lernen wird entsprechend des verwendeten Trainingsverfahrens in überwachtes Lernen (Lernen aus gekennzeichneten Daten), unüberwachtes Lernen (Finden von Mustern in nicht gekennzeichneten Daten) und verstärkendes Lernen (Lernen durch Interaktion mit einer Umgebung) unterteilt (Heaton, 2018).

Deep Learning, als Teil des maschinellen Lernens, basiert auf die Nachbildung künstlicher neuronaler Netze, welche in mehreren Schichten (Layer) mittels mathematischer Funktionen komplexe Muster in Daten zu modellieren und zu verarbeiten versuchen. Je mehr dieser Schichten hintereinander durchlaufen werden, desto komplexer, undurchsichtiger, aber auch präziser können diese Modelle agieren. Deep Learning hat sich in den Domänen von LLMs insbesondere bei Aufgaben wie Bilderkennung, Spracherkennung und natürlicher Sprachverarbeitung als besonders erfolgreich erwiesen (LeCun et al., 2015).

Large Language Models (LLMs) sind ebensolche spezifischen Anwendungen von Deep Learning im Bereich der natürlichen Sprachverarbeitung. Diese Modelle sind darauf ausgelegt, menschliche Sprache zu interpretieren, und multimodalen Output zu generieren und manipulieren. LLMs werden an enormen Textdaten trainiert, wobei semantische Muster und kontextuale Beziehungen innerhalb der Sprache erfasst und zunehmend besser abgebildet werden (Brown et al., 2020).

2.1.2 Technologische Grundlagen

Der Transformer bildet das wesentliche, zugrundeliegende Modell in der natürlichen Sprachverarbeitung (NLP), das 2017 von Vaswani et al. beschrieben und eingeführt wurde. Die Autor*innen beschrieben insbesondere Selbstaufmerksamkeitsmechanismen (Self-attention) und legten somit den Grundstein für heutige Anwendungen. Das

Transformermodell besteht aus einer Encoder-Decoder-Architektur, die durch ebendiese Selbstaufmerksamkeitsmechanismen und vollständig verbundene neuronale Netze (Feed-Forward-Netzwerke) charakterisiert ist. Selbstaufmerksamkeitsmechanismen bedeuten in diesem Kontext, dass die Kontentgenerierung nicht einer linearen Struktur (z.B. Wort für Wort) folgt, sondern beispielsweise semantische und grammatikalische Referenzen in der Generierung berücksichtigt (z.B. Negationen die auf bereits generierten Content rückwirkend Einfluss haben).

Diese Struktur ermöglicht es dem Transformer, Kontextinformationen effizient zu erfassen und zu verarbeiten, was ihn für unterschiedlichste NLP-Aufgaben äußerst leistungsfähig macht, wodurch Transformermodelle derzeit als state of the art anzusehen sind (Patwardhan et al., 2023).

Die kleinsten Einheiten in welche Eingaben vor der weiteren Verarbeitung zerlegt werden, werden als Token bezeichnet. Diese Tokens können bei textbasierten Modellen beispielsweise Wörter, Silben oder einzelne Zeichen repräsentieren. Wichtig zu beachten ist, dass ein Token nicht einem Wort entsprechen muss, so liegt der durchschnittliche Wert des Models „Bidirectional Encoder Representations from Transformers“ (BERT) etwa bei 1,1 Token pro Wort (Devlin et al., 2018), bei OpenAI’s „Generative Pre-trained Transformer“ (GPT) wird ein Wort in durchschnittlich 1,2 Tokens (GPT-3) umgewandelt (Brown et al., 2020).

Der Prozess der Wortaufteilung wird als Tokenisierung bezeichnet, wobei es unterschiedliche Methoden wie die wortbasierte, satzbasierte, oder byte-pair encoding Tokenisierung gibt (Jurafsky & Martin, 2009). Im Transformer werden die Tokens in Einbettungsvektoren umgewandelt, wodurch eine semantische Relation hergestellt werden kann (siehe Abb.1), welche kontextuale Informationen erfasst und in weiterer Folge generieren kann (stark vereinfacht: eine Vektorausrichtung und Länge von Onkel zu Tante entspricht in einer anderen Dimension jener von männlich zu weiblich. Chat GPT-3 verfügt über mehr als 12 000 Dimensionen; siehe Abb. 1) (Devlin et al., 2018).

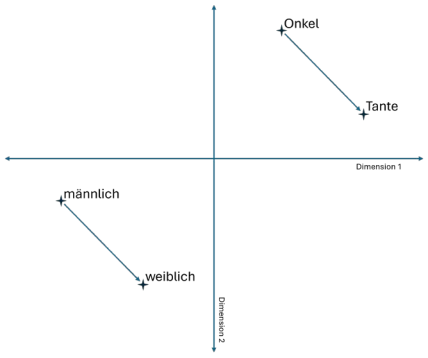

Abb. 1: Vektorausrichtung und Dimension am Beispiel Onkel – Tante (stark vereinfachte Darstellung)

Ein weiterer grundlegender Begriff neuronaler Netzwerkarchitektur, wie auch bei Transformern, sind Gewichte (weights). Durch Gewichte werden Parameter innerhalb der Schichten wie kleine Stellschrauben angepasst und so die Verbindungen zwischen den „Neuronen“ gestärkt, oder geschwächt. Die Anpassung der Gewichte ist ein wesentlicher Schritt in der Entwicklung von Modellen um Fehler zwischen vorhergesagten (oder erwünschten) Werten mit tatsächlichen Werten zu minimieren. Im Transformermodell werden Gewichte im Training genutzt, um die beste Repräsentation der Eingabedaten zu erzielen. Gewichte lassen sich mit Multiplikatoren vergleichen: führt eine Funktion zu einem

korrekten oder erwünschten Wert, wird der Faktor erhöht, führt diese zu einem zu einem falschen, oder unerwünschten Ergebnis wird der Faktor verringert (Wankmüller, 2022).

Die Architektur von Transformers, wie sie von Vaswani et al. beschrieben wurden, besteht aus zwei Einheiten, einem Encoder und einem Decoder. Der Encoder besteht hierbei aus mehreren identischen Schichten, die jeweils eine Multi-Head-Self-Attention Mechanik und ein Feed-Forward-Netzwerk beinhalten. Diese Mechanik ermöglicht es jedem Token in der Eingabesequenz (Input), Informationen aus allen anderen Tokens zu beziehen, im Anschluss wird jedes Token im Feed-Forward-Netzwerk separat weiterverarbeitet.

Der Decoder weist eine ähnliche Struktur wie der Encoder auf, umfasst jedoch mit der Encoder-Decoder-Attention eine zusätzliche Schicht, welche auf die Ausgaben des Encoders achtet, womit relevante Informationen aus dem Input für die Generierung der Ausgabe extrahiert werden können. Die Masked Multi-Head-Self-Attention ähnelt der Self-Attention-Schicht im Encoder, wobei zukünftige Tokens maskiert werden, sodass das Modell nur Informationen aus der Vergangenheit und Gegenwart verwendet. Die Ausgabe wird zum Encoder durch ein Feed-Forward-Netzwerk weiterverarbeitet.

Zusätzlich zu diesen Komponenten werden durch das Positional-Encoding Positionskodierungen hinzugefügt, da Transformer nicht über eine inhärente Reihenfolgeninformation verfügen. Die Tokens werden in diesem Prozess als Einbettungsvektoren verarbeitet, um ihre semantische Bedeutung erfassen zu können, wobei durch die sogenannte Softmax Funktion die jeweils wahrscheinlichste Fortsetzung einer Sequenz sichergestellt werden soll. Durch die Attention Mechanismen, wie sie durch Vaswani et al. beschrieben wurden, konnten parallele Berechnungen erzielt werden und die Pfadlängen für Long-Distance Abhängigkeiten deutlich verkürzt werden. Dadurch konnte eine signifikante Effizienzsteigerung in Trainings- und Inferenzprozessen erreicht werden (Vaswani et al., 2017).

Die gesamte Struktur wird typischerweise mehrfach gestapelt wodurch Schichten (Layer) entstehen. Das ermöglicht es modernen Modellen Sprache auf zunehmend abstrakten Ebenen und in multimodalen Ausprägungen zu verarbeiten (siehe Abb2.).

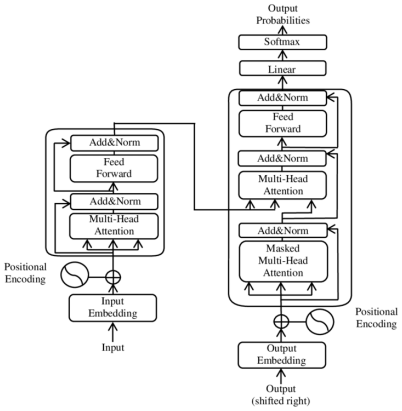

Abb. 2: Schematische Darstellung des Transfomermodells [1]

LLMs wie Chat-GPT verwenden diese Architektur in einer sogenannten "Decoder-only"-Konfiguration, bei der das Modell darauf trainiert wird, das nächste Wort in einer Sequenz basierend auf den vorherigen Wörtern vorherzusagen (z.B. nach dem Input „der Himmel

[1] https://de.wikipedia.org/wiki/Transformer_(Maschinelles_Lernen

ist grau, ich denke es wird bald …“ könnte mit hoher Wahrscheinlichkeit das Wort „regnen“ folgen). Dies ermöglicht es dem Modell, kohärenten Text zu generieren und eine breite Palette von (Sprach)aufgaben durchzuführen, wie Radford unter anderem mit dem späteren Sicherheitschef von OpenAI Ilya Sutskever (2019/2019) aufzeigte.

Um erfolgreich gewünschte Outputs generieren zu können durchlaufen die Modelle Trainingsprozesse. Diese werden bei LLMs in zwei Phasen unterschieden, dem Vortraining (Pretraining) und der Feinabstimmung (Finetuning). Im Vortraining werden riesige Mengen von Textdaten herangezogen (oft hunderte Gigabyte an Text), aus denen das Modell abzuleiten lernt, wie die nächsten Wörter einer Sequenz vorhergesagt werden könnten. Zugleich werden auch allgemeine Sprachmuster und generelles Wissen erlernt, welche in den Modellen für künftige Aufgaben zur Verfügung stehen. Die Qualität der Textdaten ist wesentlich für die Qualität des künftigen Modells. In der Feinabstimmung wird das vortrainierte Modell auf spezifische Anwendungen spezialisiert um beispielsweise Wording, Stil oder Wissen einer spezifischen Domäne einzupflegen.

Obwohl LLMs umfassende Möglichkeiten bieten, ist es wichtig, ihre Grenzen zu beachten. Sie können schlüssig erscheinende, aber faktisch falsche Informationen produzieren (ein Phänomen, das als "Halluzination" bekannt ist), sie können Verzerrungen aus ihren Trainingsdaten widerspiegeln (Bias), und ihr Entscheidungsprozess kann undurchsichtig sein (Black-Box Phänomen), was erhebliche Risiken hinsichtlich der Interpretierbarkeit und Zuverlässigkeit, insbesondere in klinischen Settings, mit sich bringen könnte (Bender et al., 2021).

2.2 Aktueller Forschungsstand

Der gegenwärtige Forschungsstand hinsichtlich der Anwendung von LLMs im Gesundheitssektor, insbesondere in der Logopädie, wird in Kapitel 4 dieser Studie detailliert erörtert. Anwendungsfälle aus dem Gesundheitsbereich und konkrete Anwendungen aus den gehobenen medizinisch-therapeutisch-diagnostischen Berufen (MTD) werden eingehend beschrieben, um den aktuellen Forschungsstand darzustellen.

In der vorliegenden Arbeit werden die theoretischen Modelle des Technology Acceptance Model (TAM) von Davis (1989) und des Unified Theory of Acceptance and Use of Technology (UTAUT) von Venkatesh et al. (2003) zur strukturierten Untersuchung der Akzeptanz und Implementierung von LLMs am Beispiel der Logopädie herangezogen.

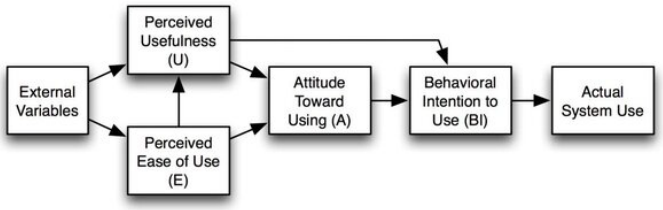

Abb. 3: TAM (Davis et al., 1989)

Das TAM (siehe Abb. 3) konzentriert sich auf den wahrgenommenen Nutzen (Perceived Usefullness) und wahrgenommene Benutzerfreundlichkeit (Perceived Ease Of Use) als zentrale Faktoren für die Akzeptanz von Technologien, da diese Faktoren einen unmittelbaren Einfluss auf die Nutzungsintention haben. Das UTAUT ergänzt das TAM um weitere Faktoren, sodass vier primäre Faktoren die Akzeptanz und Anwendung von

Technologie maßgeblich beeinflussen. Leistungserwartung (Performance Expectancy), Aufwandserwartung (Effort Expectancy), soziale Einflüsse (Social influence) und die unterstützenden Bedingungen (Facilitating Conditions) stellen die grundlegenden Elemente dar, die um vier, die Hauptfaktoren beeinflussende, moderierende Variablen (Geschlecht, Alter, Erfahrung und Freiwilligkeit der Nutzung) ergänzt werden. Diese theoretischen Modelle bilden die Basis zur Analyse der Akzeptanz von LLMs in der vorliegenden Arbeit und gewährleisten eine systematische Untersuchung der Einflussfaktoren und die Integration der Erkenntnisse in die Entwicklung und Implementierung der LLM-basierten Lösungen als Prototypen. Die Anwendung von TAM und UTAUT ermöglicht die Identifizierung und Adressierung spezifischer Aspekte der Technologieakzeptanz, wodurch die Entwicklung praxisnaher und anwenderfreundlicher Technologien gefördert wird. Zusätzlich betten diese Modelle die Fragestellungen und Erstellung der Prototypen in ein strukturiertes Konstrukt, während ein systemischer Rahmen für die Erstellung und Iteration der Prototypen genutzt werden kann.

2.3 Fähigkeiten und Grenzen von LLMs

Durch die Popularität von Modellen wie ChatGPT von OpenAI, sind Fähigkeiten von LLMs einer breiten Masse von Benutzer*innen zugänglich gemacht worden. Im Folgenden soll sowohl auf deren zentrale Fähigkeiten eingegangen werden, wie auf die Grenzen dieser Technologie.

Sprachverständnis stellt die vordergründige Fähigkeit von Sprachmodellen dar. LLMs zeichnen sich durch das Verständnis komplexer linguistischer Strukturen, Kontext und Nuancen in Texten aus. Sie können Redewendungen, Sarkasmus und sogar Formen von Humor interpretieren und nähern sich in vielen Fällen dem menschlichen Verständnisniveau an (Brown et al., 2020).

Eine der Fähigkeit von LLMs sei es, kohärenten, kontextuell angemessenen Text zu generieren. Dies reicht von kurzen Antworten bis hin zu längeren Inhalten, wobei Konsistenz in Stil und Thema über längere Passagen hinweg beibehalten wird (Radford et al., 2019). Während GPT-2 in der Lage war aus einem Textsegment die folgenden Worte zu „erraten“ (ähnlich der Autovervollständigung am Smartphone), verfügen moderne Modelle über ausgedehnte Generierungsfähigkeiten und beschränken sich nicht mehr auf eine Modalität, sondern sind auch in der Lage Informationen aus Bild, Ton und Video zu extrahieren und korrekt zu verarbeiten (Labkoff et al., 2024).

Weiters sind moderne LLMs in der Lage qualitativ hochwertige Übersetzungen zwischen zahlreichen Sprachen durchzuführen, die sich oft der Qualität professioneller menschlicher Übersetzer für viele Sprachpaare annähern oder diese sogar erreichen (Fear & Gleber, 2023)

LLMs sind in der Lage lange Texte zu prägnanten Zusammenfassungen zu destillieren, sie erfassen Kernpunkte und bewahren gleichzeitig Kohärenz. Um komplexe Anfragen verstehen und beantworten zu können, greifen sie auf ihre umfangreiche Wissensbasis zurück (bei Modellen ohne Internetzugang gibt es häufig einen Cut-Off Punkt bis zu welchem auf Wissen aus den Trainingsdaten zugegriffen werden kann), dies stellt sicher, dass relevante und detaillierte Antworten aus unterschiedlichsten Bereichen generiert werden können (Xue et al., 2020).

Eine besonders leistungsfähige Eigenschaft von LLMs ist ihre Fähigkeit, sich mit minimalem zusätzlichen Training an neue Aufgaben anzupassen, eine Fähigkeit, die als Few-Shot- oder Zero-Shot-Lernen bekannt (Brown et al., 2020).

Moderne LLMs hätten beeindruckende Fähigkeiten im Verstehen und Generieren von Programmiercode in unterschiedlichsten Programmiersprachen. Sie sind in der Lage Programmieraufgaben selbstständig auszuführen, Code zu überprüfen und verbessern wie es Chen et al. (2021) bereits vor drei Jahren beschrieben haben.

Eine zukünftige Entwicklung von LLMs sei deren Multimodalidät. So könnten Inhalte verarbeitet und generiert werden, die Text mit anderen Modalitäten wie Bildern, Video und Ton kombinieren (Ramesh et al., 2022). Auf diese Weise könnten Bilder interpretiert werden, Musik zu einer passenden Stimmung generiert werden und Bilder und Videos aus Texteingaben erstellt werden. Fähigkeiten die zwei Jahre nach Veröffentlichung

Doch auch wenn LLMs beeindruckende Outputs generieren können, verfügen sie über kein echtes Verständnis oder Bewusstsein, sondern operieren auf Grundlage statistischer Muster in ihren Trainingsdaten und nicht auf Basis echten Verständnisses, worauf Gao et al. (2023) hinweisen.

Eine bekannte Einschränkung stellen sogenannte Halluzinationen dar. LLMs können plausibel klingende, aber falsche Informationen generieren, dies ist besonders in Bereichen, wie dem Gesundheitswesen, wo Genauigkeit entscheidend ist, problematisch (Ji et al., 2023). Diskriminierend, oder potenziell unfaire Darstellungen könnten, wie Bender et al. (2021) in ihrem Paper festhalten, durch LLMs verstärkt werden. Bias, als Form von Voreingenommenheit oder Verzerrung in unserer Art zu denken oder zu handeln kann durch unausgewogene Trainingsdaten verstärkt werden und hat für Therapeut*innen, die im Sinne eines Berufsethos vorurteilsfrei eine Gleichbehandlung von Patient*innen gewährleisten müssen, fundamentale Bedeutung.

Kontextuelle Einschränkungen, die sie sich aus Limitierungen der Menge an Text in der Eingabe und in der Ausgabe ergeben könnten, oder ein Mangel an kausalem Verständnis, wie es von Pearl und Mackenzie (2018) beschrieben wurde, scheinen durch aktuelle Modelle weitgehend gelöst zu sein, wobei auch aktuelle Modelle noch Schwierigkeiten mit Aufgaben haben, die das Verständnis von Ursache-Wirkungs-Beziehungen erfordern.

Weiterhin ungelöst sind der intensive Ressourcenverbrauch, welchen der Betrieb der energieintensiven Systeme mit sich bringt (Patterson et al., 2021) und das Problem der Undurchsichtigkeit von Entscheidungsprozessen, das sogenannte Blackbox-Phänomen. Wie bereits erwähnt, ist es teilweise unmöglich den Prozess der Outputgenerierung exakt nachvollziehen zu können. Dieses Phänomen kann für Anwendungen in kritischen Systemen, in welchen eine Nachvollziehbarkeit von Entscheidungen relevant ist (z.B. im Gesundheitswesen) problematisch sein und wurde auch von Danilevsky et al. (2020) in einer Untersuchung des Zustandes erklärbarer KI beschrieben.

Unmittelbar mit dieser Herausforderung sind auch Fragestellungen zu Datenschutz und geistigen Eigentumsrechten verbunden. Wie in Kapitel 2.1.2 „Technologische Grundlagen und Funktionsweise“ bereits angeführt wurde, werden für das Training von Sprachmodellen riesige Mengen an Textdaten verwendet. Diese können sensible Daten beinhalten, es könnten auch Bedenken über Herkunft und urheberechtliche Fragestellungen relevant sein, wobei das soeben beschriebene BlackBox Phänomen die Beurteilung der Trainingsdaten verunmöglicht (Carlini et al., 2020). Als eines der ersten Unternehmen hat die New York Times 2023 OpenAI auf Urheberrechtsverletzung [2] geklagt,

[2] https://www.nzz.ch/technologie/new-york-times-klage-gegen-openai-ist-chatgpt-illegal-ld.1772265

da Textfragmente, welche eindeutig dem Medienverlag zugerechnet werden konnten, in den Antworten von ChatGPT reproduzierbar nachgewiesen werden konnten.

2.4 Ethik und Datenschutz

Für Die Beschäftigung mit LLMs in gesundheitssystemischen Kontext kann deshalb nicht ohne die Betrachtung ethischer und datenschutzrechtlicher Fragestellungen erfolgen. Um ein konsistentes Regelwerk für die Anwendung von Künstlicher Intelligenz innerhalb der Europäischen Union zu etablieren, zielt der von der Europäischen Union beschlossene AI Act (AI Act, 2024) darauf ab, Anforderungen zu definieren, welche auch an den Gesundheitssektor hohe Anforderungen an Transparenz, Sicherheit und Datenschutz stellen. Dies bedeutet auch für potentielle Anwendung von LLMs im logopädischen Kontext strenge regulatorischen Bestimmungen, um unter anderem den Datenschutz von Patient*innen und die Verständlichkeit der zugrundeliegenden Algorithmen zu gewährleisten.

Ethische Aspekte wie Transparenz, verantwortungsvoller Umgang mit Patient*innendaten und die Prävention von Diskriminierungen sind in diesem Zusammenhang ebenso von zentraler Bedeutung. Die Gewährleistung des Schutzes personenbezogener Daten erfordert robuste Sicherheitsmaßnahmen und ein transparentes Datenmanagement (Mennella et al., 2024).

Zusätzlich gilt für potenzielle LLM-basierte Anwendungen, den Bestimmungen der Datenschutz-Grundverordnung (DSGVO) gerecht zu werden, insbesondere in Bezug auf die Verarbeitung sensibler Gesundheitsinformationen. Die Grundsätze der Datenminimierung, Zweckbindung, Transparenz und Einwilligung sind Therapeut*innen aus berufspolitischen Verpflichtungen vertraut und sind auch in diesem Kontext einzuhalten (Datenschutz-Grundverordnung, 2016).

Die Einhaltung dieser ethischen und datenschutzrechtlichen Vorgaben ist eine Grundvoraussetzung für die verantwortungsvolle Beschäftigung mit dem Thema und liegt Überlegungen der gesamten Arbeit zugrunde.

3. METHODIK

Die vorliegende Masterarbeit folgt einem dreisäuligen Mixed-Methods Ansatz, der darauf abzielt, Anwendungen zu identifizieren und die Potenziale und Herausforderungen von LLMs in der logopädischen Tätigkeit umfassend zu beleuchten. Die Forschungsfragen, welche Anwendungsbereiche für LLMs in der logopädischen Tätigkeit relevant sind und wie Logopäd*innen die zukünftige Anwendbarkeit dieser Technologie einschätzen stellen die zentralen Anforderungen an die Methodik dar.

Durch umfassende Analyse relevanter wissenschaftlicher Literatur wird zum einen die Grundlage für das Verständnis des aktuellen Forschungsstandes geschaffen und zum anderen die Identifikation potenzieller Anwendungsfälle von LLMs im logopädischen Kontext eingehend beleuchtet, sodass diese systematisch identifiziert werden können.

Die Qualitative Datenerhebung erfolgte mehrstufig. Die Durchführung eines Fokusgruppeninterviews mit praktizierenden Logopäd*innen liefert Erkenntnisse zu Bedürfnissen, Erwartungen und Bedenken hinsichtlich des Einsatzes von LLMs in der Praxis. Durch eine Kommentarfunktion und einen strukturierten Fragebogen wurden die entwickelten Prototypen beurteilt. Diese qualitative Datengrundlage stellt die zweite Säule dar.

Die dritte Säule bildet die Prototypenentwicklung und -iteration. Die Entwicklung von drei LLM-basierten Prototypen ermöglichte eine praktische Demonstration der Einsatzmöglichkeiten von LLMs in der logopädischen Tätigkeit und dient als Basis für die Evaluation nach Funktionalität, Benutzerfreundlichkeit und Mehrwert durch Logopäd*innen.

Dieser Mixed-Methods Ansatz ermöglicht eine umfassende Betrachtung des Themas, bindet Logopäd*innen als Expert*innen sowohl in die Konzeption, wie auch Evaluierung und Bewertung ein und stellt so sicher, dass fundierte Erkenntnisse für die Beantwortung der Forschungsfragen gewonnen werden können.

3.1 Interdisziplinärer Ansatz Health-Assisting-Engineering

Die vorliegende Arbeit ist im interdisziplinären Feld des Health Assisting Engineering Masterstudiengang an der FH Campus-Wien angesiedelt. Das Studium bildet die Schnittstelle aus Gesundheitsberufen und technischen Berufen und fördert die Zusammenarbeit aus beiden Bereichen. Diese Zusammenarbeit zwischen den Disziplinen ermöglicht es unter anderem technologisches Know-how mit klinischer Expertise zu verbinden, um sicherzustellen, dass die entwickelten Anwendungen sowohl technisch robust als auch klinisch relevant sind.

3.2 Literaturrecherche

Die Literaturrecherche bildet den Ausgangspunkt dieser Masterarbeit und dient der systematischen Identifizierung und Analyse relevanter wissenschaftlicher Publikationen zum Einsatz von LLMs in der Logopädie und in verwandten medizinischen und therapeutischen Bereichen. Ziel der Literaturrecherche war es, eine breite und aktuelle wissenschaftliche Grundlage zu schaffen, den aktuellen Stand der Forschung zu erfassen, potenzielle Anwendungsfälle zu identifizieren und Forschungslücken aufzudecken.

3.2.1 Suchbegriffe und Suchstrategien

Die Literaturrecherche wurde im Rahmen einer umfassenden Untersuchung der einschlägigen Datenbanken durchgeführt, um ein ganzheitliches Verständnis der vorhandenen Literatur zu LLMs im Gesundheitswesen insgesamt zu erlangen. Die Definition der Kriterien sollte sicherstellen, relevante Studien zu identifizieren, die mit der Anwendung von LLMs in der Logopädie oder eng miteinander verbundenen therapeutischen Berufen in Zusammenhang stehen.

3.2.2 Datenbanken und Suchwerkzeuge

Für die Literaturrecherche wurden sowohl traditionelle wissenschaftliche Datenbanken als auch neuartige KI-gestützte Tools verwendet. Durch die intensive Beschäftigung mit KI-gestützten Suchtools, konnten zum einen relevante Erkenntnisse Bezüglich Funktionsweise, Verlässlichkeit und Anwendbarkeit LLM-basierter Anwendungen gewonnen werden, zum anderen wurde die Breite möglicher Quellen erhöht.

Als traditionelle wissenschaftliche Datenbanken wurden im Rahmen der vorliegenden Arbeit PubMed, Cochrane Library und Embase (Elsevir) für die Datenbanksuche herangezogen. Weiters wurden die Suchbegriffe auch bei Google Scholar verwendet, um mögliche zusätzliche Quellen identifizieren zu können.

Als primäre Suchbegriffe wurden beispielsweise "Large Language Models", "Logopädie", "Speech Therapy", "AI in Healthcare", "Natural Language Processing" gewählt, die durch sekundäre Suchbegriffe, wie "Natural Language Processing" "Artificial Intelligence" "Clinical Applications" "Patient Interaction" oder "Diagnostic Tools" sinnvoll ergänzt wurden.

Ergänzt wurden die Suchbegriffe im weiteren Verlauf durch berufsbezogene Begriffe wie "Orthoptik", "Radiologie-Technologie", "Biomedizinische Analytik", "Ergotherapie", "Physiotherapie", "Diätologie" in deutscher Sprache und als englische Berufsbezeichnungen. Die Suchbegriffe wurden durch boolesche Operatoren sinnvoll miteinander verknüpft.

Um die Recherche weiter zu präzisieren, wurde zusätzlich zu den allgemeinen Suchbegriffen auch Medical Subject Headings (MeSH) für die Datenbanksuche bei PubMed verwendet. Diese Meshterms umfassten "Speech Therapy", "Language Disorders", "Artificial Intelligence", "Natural Language Processing", "Machine Learning", "Clinical Decision Support Systems", "Diagnosis, Computer-Assisted".

Eine tabellarische Auflistung der Suchstrategie ist im Anhang1: Datenbanksuche beigefügt.

3.2.3 Einschlusskriterien und Ausschlusskriterien

Um die Aktualität und Relevanz der in die Analyse einbezogenen Studien zu gewährleisten, wurden die folgenden Einschlusskriterien formuliert:

Als Veröffentlichungszeitraum wurden Studien inkludiert, die ab 2014 veröffentlicht wurden. Dies sollte im Sinne der Aktualität der Forschung sicherzustellen, dass die Ergebnisse dem aktuellen Wissensstand entsprechen. Randomisierte kontrollierte Studien (RCTs), systematische Übersichtsarbeiten, Metaanalysen und hochkarätige Beobachtungsstudien wurden aufgrund der höheren Evidenzlage bevorzugt herangezogen. Eine Themenrelevanz war erforderlich, der unmittelbare Zusammenhang mit Logopädie und den erweiterten MTD-Berufsgruppen, oder technische Aspekte zu LLMs und die

Anwendung von LLMs im breiteren Kontext der Gesundheitsversorgung mussten gegeben sein.

Die Analyse umfasste Studien, die sowohl auf Deutsch als auch auf Englisch veröffentlicht wurden.

Als Ausschlusskriterien wurden Studien, die keinen direkten Bezug zur Logopädie, oder den erweiterten medizinischen Berufen, aufwiesen definiert. Ebenso worden Studien ohne expliziten Zusammenhang mit der Anwendung von LLMs im Gesundheitswesen ausgeschlossen. Dies betrifft insbesondere Studien, die sich ausschließlich mit technologischen Aspekten ohne klinische Relevanz befassten, oder deren Fokus auf veraltete technologische Ansätze, oder rudimentären Formen von NLP beruhte.

3.2.4 KI-unterstützte Suchtools

Im Rahmen dieser Masterarbeit wurden auch KI-unterstützte Suchtools eingesetzt. Diese Tools bieten eine innovative Ergänzung zur klassischen Recherche, indem sie die Identifikation, Analyse und Synthese wissenschaftlicher Literatur beschleunigen und erweitern. Die vorliegende Arbeit orientiert sich sowohl an den Vorgaben der FH Campus-Wien, wie auch an den Veröffentlichungsvorgaben von Springer, die als einer der ersten wissenschaftlichen Verlage einen Leitfaden [3] zur Anwendung von KI im wissenschaftlichen Kontext veröffentlicht haben.

Aufgrund der Neuartigkeit der Anwendungen, wird die Nutzung in dieser Arbeit transparent dargelegt, wobei aus der Vielzahl an versuchsweise betrachteten Anwendungen, die beschrieben werden, die im Zuge der Erstellung der Masterarbeit gezielt genutzt wurden. Die folgende Tabelle soll die Funktionalität der Tools erläutern und aufzeigen, wie sie im Sinne einer methodischen Sorgfalt eingesetzt wurden. Die Tabelle stützt sich auf subjektive Erfahrungen, die während der Nutzung dokumentiert wurden. Die Übersicht dient der Transparenz und erhebt keinen wissenschaftlichen Anspruch. Diese Tools wurden als Unterstützung herangezogen, wobei die Resultate sorgfältig geprüft und keine generativen Inhalte in die Arbeit übernommen wurden.

[3] https://www.springer.com/gp/editorial-policies/artificial-intelligence--ai-/25428500

Tool	Beschreibung	Anwendung in der Arbeit	Empfundene Vorteile	Empfundene Nachteile
Consensus [4]	KI-gestütztes Tool für semantische Suchanfragen, das prägnante Zusammenfassungen liefert.	Schnelle Identifikation und Zusammenfassung relevanter Studien; Vergleich wissenschaftlicher Ergebnisse.	Einfach zu bedienen und klare Darstellung von Forschungsdesigns.	Begrenzte Quellenbasis, keine Volltextsuche.
SciSpace [5]	Ermöglicht Auswertung nach PRISMA und CASP durch gezielte semantische Anfragen.	Tiefergehende Analyse und Synthese der Literatur; Anwendung semantischer Anfragen auf PDFs.	Detaillierte Literaturanalyse und systematische Darstellung der Ergebnisse.	Hohe Einarbeitungszeit erforderlich und unübersichtlich bei einer großen Anzahl an Papers.
Iris.ai [6]	Umfassende Suchfunktionalität und Clustering nach Stichwortwolken.	Überblick über verwandte Suchbegriffe und Forschungsfelder; Interdisziplinäre Recherche.	Guter Überblick über Forschungsfelder durch Clusterung verwandter Begriffe.	Komplexe Menüführung, die Einarbeitungszeit erfordert.
Connected Papers [7]	Visualisiert Netzwerke von wissenschaftlichen Arbeiten nach Autor*innen.	Analyse von Beziehungen zwischen Studien und Autoren; Identifikation zentraler und verwandter Arbeiten.	Übersichtliche visuelle Darstellung wissenschaftlicher Netzwerke.	UI nicht immer intuitiv bei großer Funktionalität, deshalb eine lange Einarbeitungszeit. Plattform ist gut dokumentiert.
Perplexity [8]	Schrittweise semantische Suche mit expliziter Suche in akademischen Quellen.	Identifikation akademischer Quellen, besonders in aktuellen Forschungsfeldern.	Einfacher Suchablauf, hilfreich im Bereich der semantischen Exploration.	Wissenschaftliche Fundiertheit der Quellen ist teilweise nicht gegeben.

Tab. 1: Tabellarischer Überblick - Nutzung wissenschaftlicher KI-Tools und subjektive Beurteilung

[4] https://consensus.app/

[5] https://typeset.io

[6] https:// the.iris.ai

[7] https:// connectedpapers.com

[8] https:// perplexity.ai

Vergleich und Nutzen

Durch die zusätzliche Nutzung der beschriebenen KI-Tools konnten insbesondere Preprints und graue Literatur in die Literaturrecherche einbezogen werden, die in einem neuen Forschungsfeld wie dem Einsatz von LLMs im Gesundheitsbereich von großem Wert sind. Durch den Einsatz von Anwendungen zur generativen Erstellung von Zusammenfassungen, zur gezielten semantischen Fragestellung, oder zur übersichtlicher Darstellung nach Forschungsdesign konnten ungeeignete Studien rascher verworfen und die Forschung nach relevanten Studien intensiviert werden.

Um dem wissenschaftlichen Anspruch der Arbeit gerecht werden zu können, war wurden sämtliche Quellen im Original überprüft. Die Überprüfung und der Abgleich mit den Originalquellen wurde in einer Exceltabelle festgehalten, bevor validierte Quellen in die Literaturverwaltungssoftware Citavi übernommen wurden. Alle genannten Tools wurden als Premiumversionen genutzt, wodurch sie die leistungsfähigsten verfügbaren Modelle nutzen konnten. Weiters konnte dadurch auf zusätzliche Funktionalität (z.B. Fileupload, erweiterte Kontextfenster etc.) zugegriffen werden, die sich in Umfang und Qualität teilweise deutlich von den jeweiligen Freiversionen unterschieden.

Es kann festgehalten werden, dass die beschriebenen Anwendungen inhaltlich äußerst zuverlässige Resultate lieferten (keine Halluzinationen in Bezug auf erfundene, generative Inhalte), harte Fakten (Titel, Erscheinungsjahr etc.) korrekt ausgegeben und ein müheloses Wechseln zwischen Sprachen erlaubten. Einschränkungen waren insbesondere durch Auslassung relevanter Passagen und Ergebnisse, nicht nachvollziehbare Schwerpunktsetzung und die Notwendigkeit manueller Überprüfung der Ergebnisse gegeben.

Insgesamt ermöglichte der Einsatz dieser Tools eine erweiterte Literaturrecherche und einen tiefergehenden Überblick über das Forschungsfeld, das über das hinausgeht, was durch traditionelle Methoden allein erreichbar gewesen wäre. Weiters ermöglichte die Beschäftigung mit diesen Anwendungen in einem lebensnahen Setting einen zusätzlichen Erkenntnisgewinn für die folgende Entwicklungsarbeit eigenständiger LLM-basierter Prototypen für den logopädischen Bereich, da Funktionalität und Limitationen anhand eines realen Anwendungsszenarios erprobt und nachvollzogen werden konnte.

3.2.5 Literaturbewertung und Synthese

Die Qualität der in die Arbeit eingeschlossenen Studien wurde anhand festgelegter Kriterien bewertet, darunter Studiendesign, Stichprobengröße, methodische Präzision und Validität der Ergebnisse. Die PRISMA-Checkliste (Moher et al., 2009) wurde in Kombination mit dem CASP-Tool (Critical Appraisal Skills Programme 2023) verwendet, um eine systematische Bewertung der Studienqualität durchzuführen. Diese Instrumente ermöglichen eine strukturierte und transparente Bewertung der Ergebnisse und tragen so zur Sicherstellung der Gültigkeit und Zuverlässigkeit dieser bei.

Insgesamt wurden durch die traditionellen Suchen 463 Artikel identifiziert. Nach Anwendung der Ausschlusskriterien, Entfernung von Duplikaten und detaillierter Prüfung der Abstracts und Volltexte wurden 123 für die Arbeit relevante Studien ausgewählt und in Citavi als Literaturverwaltungssoftware importiert (Abb. 4).

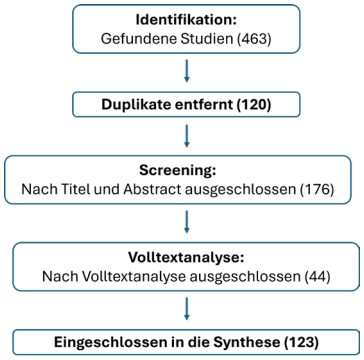

Abb. 4: Flowchart Literaturrecherche

Im Anschluss an die Qualitätsbewertung wurden die Ergebnisse der ausgewählten Studien thematisch zusammengefasst, um ein ganzheitliches Verständnis der aktuellen Forschungslandschaft zur Anwendung von LLMs in der Logopädie zu gewinnen. Die identifizierten Anwendungen wurden klassifiziert, die mit der Implementierung von LLMs in der Praxis verbundenen Herausforderungen wurden dokumentiert und ethische Überlegungen beachtet.

Für die Literaturrecherche zum aktuellen Stand in der Forschung im Bereich von LLMs, technologischen Grundprinzipien und Möglichkeiten, wurden idente Bewertungskriterien und Analysen angewendet. Diese Quellen stellten insbesondere für den theoretische Rahmen in Kapitel zwei und die Kapitel fünf und sechs, Diskussion und Ausblick, eine solide wissenschaftliche Grundlage sicher.

3.2.6 Erweiterung der Suchstartegie

Wie in Kapitel 3.2 „Literaturrecherche“ bereits angeführt, erfolgte in einem ersten Schritt eine Ausweitung der Suche auf verwandte MTD-Berufe, da die Anzahl der Studien zur Anwendung von LLMs in der Logopädie begrenzt war. Dieser Fokus in einem weiteren Schritt auf allgemeine Anwendung im Gesundheitsbereich erweitert und nach den beschriebenen Suchkriterien und Strategien durchgeführt und analysiert.

3.3 Fokusgruppeninterviews und Feedbackprozess

Das Fokusgruppeninterview wurde als qualitative Methode gewählt, um die Forschungsfrage zu beantworten, wie Logopäd*innen die prospektive Integration von LLMs in ihre berufliche Praxis bewerten und welche Anwendungen für sie relevant sein könnten. Ziel der Diskussion war es, die Erwartungen, Anforderungen und Befürchtungen von Logopäd*innen in Bezug auf die Implementierung von LLMs in deren berufliche Routineaufgaben zu untersuchen. Diese, mit den in der Literaturrecherche identifizierten Anwendungsfälle abgeglichenen Anwendungsfälle, stellten die Basis für die Prototypenerstellung und die Beantwortung der Hauptforschungsfrage dar.

3.3.1 Rekrutierung und Auswahl der Teilnehmer*innen

Logopäd*innen, die derzeit in Österreich praktizieren, wurden eingeladen, an Fokusgruppendiskussionen teilzunehmen, wobei der Rekrutierungsprozess durch Social-Media und persönliche Kontakte erfolgte. Um eine umfassende Darstellung der unterschiedlichen Perspektiven zu gewährleisten, wurden speziell Logopäd*innen mit unterschiedlichen Fachgebieten und Erfahrungsstufen angesprochen, es wurde darauf geachtet sowohl Logopäd*innen die freiberuflich tätig sind, wie auch Logopäd*innen die in einem klinischen Setting ihren Beruf ausüben in die Fokusgruppe zu integrieren.

Die Teilnahme an den Fokusgruppeninterviews war freiwillig, die Auswertung fand unter Wahrung der Anonymität statt und die Teilnehmenden wurden vorab über die Ziele und den Inhalt der Studie, die verfahrenstechnischen Aspekte der Interviews und ihre Rechte als Teilnehmer*innen informiert. Anschließend gaben sie ihr schriftliches Einverständnis zur Teilnahme an der Studie.

3.3.2 Durchführung des Fokusgruppeninterviews

Am Fokusgruppeninterview nahmen fünf Logopädinnen teil. Die Interviews wurden in einem Online-Format durchgeführt und die Sitzung dauerte 80 Minuten. Für die Organisation der Interviews wurde ein strukturierter Leitfaden ausgearbeitet, (siehe Anhang 3: Interviewleitfaden Fokusgruppendiskussion), der die Erfahrung mit digitalen Technologien in der logopädischen Praxis, Erfahrung mit digitalen Technologien, Kenntnisse in Bezug auf große LLMs und künstliche Intelligenz, mögliche Anwendungen von LLMs im Bereich der Logopädie und Erwartungen und Befürchtungen in Bezug auf den Einsatz von LLMs umfasste. Die Fokusgruppe wurde im Sinne eines strukturierten Brainstormingprozesses abgehalten, um eine möglichst offene und breite Basis für die Identifizierung von Anwendungsfällen zu gewährleisten.

Die Interviews wurden aufgezeichnet, transkribiert und anschließend unter Verwendung einer strukturierten Inhaltsanalyse nach Mayring (2015) analysiert. Der Interviewprozess wurde so konzipiert, dass ein offenes und vertrauensvolles Umfeld gefördert wurde und die Teilnehmenden ihre Erfahrungen und Standpunkte frei äußern konnten.

3.3.3 Feedbackprozess

Im Zuge der Prototypenerstellung und Implementierung auf einer Projektwebseite wurde der ursprüngliche Plan der Durchführung einer zweiten Fokusgruppendiskussion, zur Beurteilung der Prototypen, verworfen und durch ein erweitertes Feedbacksystem ersetzt. Der Feedbackprozess wurde durch die Umsetzung einer Kommentarfunktion auf der Projektwebseite realisiert. Die Kommentarfunktion bot den Vorteil einer unmittelbaren Beurteilung und der Möglichkeit rascher Iteration der Prototypen auf Basis der eingelangten Kommentare. Ein weiterer Vorteil bestand darin die Kohorte der Logopäd*innen für die Beurteilung der Prototypen einfach erweitern zu können.

Zusätzlich wurde ein abschließender Fragebogen für die Beurteilung der Prototypen nach Ablauf der vierwöchigen Testphase erstellt und elektronisch verteilt.

Die Vorteile in Bezug auf Datengenerierung und Co-Design dieses Zugangs überwogen, weshalb diese Adaption als zielführend im Sinne der umfassenden Beantwortung der Forschungsfragen herangezogen wurde.

In Summe wurde einem Pool von 19 Logopäd*innen Zugriff auf die Prototypen als Tester*innen gewährt.

Im Folgenden sind mit Teilnehmer*innen die fünf Logopädinnen bezeichnet, die an der Fokusgruppe teilgenommen haben, Tester*innen wird für die Kohorte aus 19 Logopäd*innen gewählt.

Die aus dieser Analyse gewonnenen Erkenntnisse flossen in die Überarbeitung und Verbesserung der Prototypen ein und dienten als abschließende Beurteilung dieser.

3.3.4 Datenanalyse

Die Datenanalyse der qualitativen Daten aus den Fokusgruppeninterviews und den offenen Fragen der Feedback-Formularen wurden analog zur Auswertung des

Fokusgruppeninterviews anhand einer qualitativen Inhaltsanalyse nach Mayring (2015) ausgewertet und dokumentiert. Ziel dieses Analyseverfahrens war es, die vorherrschenden Themen und Muster in den Aussagen der Logopäd*innen zu identifizieren.

Die qualitative Inhaltsanalyse der aus den Fokusgruppeninterviews gewonnenen Transkripte und der offenen Antworten aus dem Feedback-Mechanismus wurde einer systematischen Auswertung unterzogen, wobei Transkription, Paraphrasierung, Kategoriebildung, Zuordnung von Textpassagen zu den entsprechenden Kategorien und die anschließende Interpretation die zentralen Schritte bildeten.

MaxQDA (Version 2023) wurde für die Verwaltung der Interviewprotokolle und zur Erleichterung der qualitativen Inhaltsanalyse verwendet, für die Transkription wurde f4[9], ein für den journalistischen Einsatz bestimmtes Tool genutzt, das auch den Anforderungen der DSGVO (Datenschutz-Grundverordnung, 2016) entspricht.

3.4 Prototypenerstellung

Die Erstellung eines Prototypen auf der Grundlage von LLMs stellt einen zentralen Bestandteil dieser Masterarbeit dar. Ziel der Entwicklungsarbeit war es, anhand eines konkreten, funktionierenden Prototyps die prospektiven Anwendungen von LLMs im Bereich der logopädischen Praxis zu demonstrieren. Um eine breite in Abbildung möglicher Anwendungsfälle abbilden zu können, wurden drei unterschiedliche Prototypen realisiert. Die Prototypen sollten nicht nur dazu dienen die technische Machbarkeit darzulegen, sondern auch erste Erfahrungen in der Anwendung zu erfassen und auszuwerten.

3.4.1 Identifikation der Anwendungsfälle

Die Auswahl der Anwendungsfälle für die Prototypenentwicklung erfolgte auf Basis der Ergebnisse der Literaturrecherche und der Fokusgruppeninterviews. Die identifizierten Anwendungsfälle wurden tabellarisch dokumentiert und nach Häufigkeit der Nennung quantifiziert gereiht. Die Zusammenführung der Ergebnisse aus der Literaturrecherche und dem Fokusgruppeninterview ergab die Entscheidungsgrundlage für die Auswahl geeigneter Anwendungen.

Auf Basis dieser Erkenntnisse wurden drei Anwendungsfälle identifiziert, die im Ergebnisteil ausführlich beschrieben werden.

3.4.2 Technologische Grundlagen

Bei der Festlegung der technologischen Grundlagen für die Entwicklung der Prototypen wurden sowohl die spezifischen Anforderungen der Anwendungsfälle als auch die Fähigkeiten und Einschränkungen aktueller LLMs, sowie die Umsetzbarkeit, sorgfältig berücksichtigt. Es wurde ein systematischer Ansatz entwickelt.

Zunächst wurde eine gründliche Bewertung der verfügbaren LLMs durchgeführt, um das, für jede spezifische Anwendung, am besten geeignete Sprachmodell zu ermitteln. Überlegungen wie Leistungskennzahlen, Verfügbarkeit, Einhaltung der Datenschutzbestimmungen, Benutzerfreundlichkeit, Tokenpreis und die Fähigkeit zur Integration in bereits bestehende Systeme wurden bewertet. Besonderes Augenmerk galt in dieser Beurteilung auch Kriterien wie Datenschutz, Urheberrecht und ethische Kriterien.

[9] https://www.audiotranskription.de/

Die Auswahl der Entwicklungsumgebung basierte auf den festgelegten Kriterien für Flexibilität, Skalierbarkeit, Benutzerfreundlichkeit und Kompatibilität mit den ausgewählten LLMs.

Um Kohärenz der von LLMs erstellten Inhalte zu gewährleisten, wurde besonderer Wert auf die Entwicklung der Prompts gelegt.

Als Prompts werden jene Aufforderungen bezeichnet, die als Input für die LLMs dienen. Der Prozess LLMs über Prompts zu steuern, wird als prompt Engineering bezeichnet. Da unterschiedliche Modelle unterschiedliche Anforderungen an Prompts haben ist diese Beschäftigung mit Prompt Engineering unerlässlich, um konsistente und gewünschte Outputs generieren zu können. Alle kommerziell verfügbaren Sprachmodelle bieten nicht nur eine ausführliche Dokumentation mit Best-Practise Beispielen für die Prompt Erstellung, sondern auch einen „Playground“ an, in dem das Verhalten unterschiedlicher Prompts getestet und entwickelt werden kann. Die benötigten Prompts wurden sorgfältig ausgearbeitet und verfeinert, um die angestrebten Ergebnisse sicherzustellen.

Die Prototypen wurden nach einer iterativen Methode konstruiert, die eine kontinuierliche Bewertung und Anpassung an die Anforderungen der Benutzer ermöglichte. Im Sinne eines partizipativen Co-Designs, in welchem Benutzer*innen und Stakeholder*innen von der Konzeption, über Entwicklung, bis zur Anwendung und Evaluierung in den Entwicklungsprozess eingebunden werden, wurden Logopäd*innen als Fachexpert*innen von Anbeginn an die die Prototypenerstellung eingebunden.

3.4.3 Validierung der Prototypen

Die Validierung der Prototypen erfolgte in einem iterativen Prozess, an dem eine Kohorte von insgesamt 19 Logopäd*innen beteiligt war. 13 Logopäd*innen nutzten die angebotenen Feedbacktools. Ziel war es, Funktionalität, Benutzerfreundlichkeit und den Mehrwert der Prototypen aus praktischer Sicht zu bewerten und gleichzeitig Feedback einzuholen, das in die weitere Verfeinerung der Anwendungen einfließen sollte.

Der Validierungsprozess umfasste die Phase der Vorstellung der Prototypen, im Zuge von persönlichen Einführungen, oder Einführungen über Zoom. Im Anschluss erhielten die Logopäd*innen als Testpersonen die Möglichkeit sich mit den Prototypen selbstständig auseinanderzusetzen und diese zu testen. Die Teilnehmer*innen wurden gebeten, ihre Erkenntnisse zu den Prototypen sowohl durch informelle Anmerkungen als auch durch Ausfüllen eines strukturierten Fragebogens (siehe Anhang 6: Strukturierter Fragebogen zur abschließenden Beurteilung) abzugeben.

Das Feedback der Logopäd*innen bildete die Grundlage für die Iteration der Prototypen. Durch diesen Validierungsprozess konnte der Ansatz des Co-Designs auf die Prototypenerstellung angewendet werden und es wurde sichergestellt, dass die Bedürfnisse der künftigen Nutzer*innen von Beginn an berücksichtigt wurden.

3.4.4 Strukturierter Fragebogen

Um Feedback zu den Anwendungsfällen zu erhalten, wurde ein Fragebogen ausgearbeitet, der sowohl geschlossene als auch offene Fragestellungen umfasste. Der Fragebogen konzentrierte sich hauptsächlich auf die allgemeine Benutzererfahrung, Benutzerfreundlichkeit der Prototypen, Anpassungsfähigkeit und Flexibilität der Inhalte, technische Leistung und die Anwendung von LobeChat als Plattform. Der Fragebogen war online zugänglich, und den Teilnehmern wurde eine zweiwöchige Frist eingeräumt, um ihre Erfahrungen nach der Beschäftigung mit den Prototypen zu berichten.

Der Fragebogen ist als „Anhang 6: Strukturierter Fragebogen zur abschließenden Beurteilung“ angefügt.

Der methodische Ablauf der vorliegenden Masterarbeit ist in Folgendem Chart (Abb. 5) vereinfacht abgebildet:

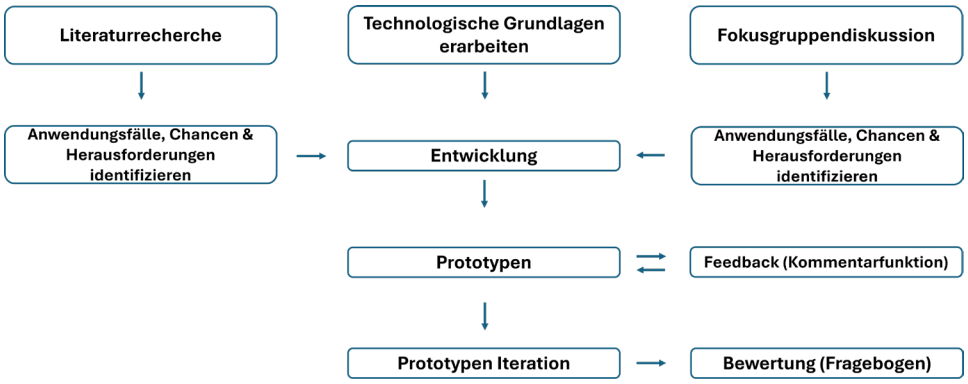

Abb. 5: Darstellung des methodischen Ablaufs

3.5 Ethik und Datenschutz

Ethische Überlegungen und Datenschutzaspekte flossen zentral in die Auswahl, Entwicklung und Bewertung der Prototypen ein. Es wurde ein umfassender Rahmen geschaffen, um den Schutz sensibler Informationen und die Vertraulichkeit der Nutzer zu gewährleisten. Maßnahmen zur Sicherheit und zum Datenschutz umfassten, sichere Webseitenverbindungen durch SSL-Zertifikate, Datensparsamkeit in der Anwendung, Anonymisierung, Einwilligung und transparente Kommunikation über Verwendung und Erfassung personenbezogener Daten.

Im Zweifelsfall wurden potenzielle Prototypen aus der Entwicklung ausgeschlossen, auch wenn sie technisch umsetzbar gewesen wären und identifizierten Anwendung entsprochen hätten. Aspekte des Datenschutzes und ethische Überlegungen standen hierarchisch über anderen Anforderungen und mussten zufriedenstellend gewährleistet werden können, damit die Prototypen einer konkreten Entwicklung zugeführt wurden.

4. ERGEBNISSE

Dieses Kapitel präsentiert die zentralen Ergebnisse der durchgeführten Untersuchungen. Es gliedert sich in drei Abschnitte, der systematischen Literaturrecherche, den Ergebnissen des Fokusgruppeninterviews mit praktizierenden Logopäd*innen und der Entwicklung und Evaluierung der Prototypen.

4.1 Ergebnisse der Literaturrecherche

Wie in der Methodik (Kapitel 3.2) beschrieben, befasste sich die Literaturrecherche mit LLM-Anwendungen im medizinischen Bereich. Large Language Models (LLMs) werden zunehmend in verschiedenen Bereichen des Gesundheitswesens erforscht und versprechen medizinische Praxis, Forschung und Verwaltung zu revolutionieren. Dieser Abschnitt bietet einen Überblick über die aktuellen und aufkommenden Anwendungen von LLMs im medizinischen Bereich.

Entscheidungsfindung im Gesundheitsbereich:

Auch wenn die Modelle noch nicht so weit seien, wie Wang et al. (2020) betonen, würden sie in absehbarer Zeit vielfältig im Gesundheitssystem eingesetzt werden und Gesundheitspersonal bei klinischen Entscheidungen (Decision-Making) unterstützen, indem sie schnellen Zugang zu relevanten medizinischen Informationen und potenziellen Behandlungsoptionen bieten. Die Autor*innen haben sowohl GPT-Modelle, wie auch GatorTron, ein auf Gesundheitsdaten trainiertes Modell von NVIDIA, untersucht. LLMs seien insbesondere nützlich um aus unstrukturiertem Freitext, wie er in Patientendokumentation üblich ist, diagnostische Kodierungen und Interventionen zu extrahieren. Weiters, so die Autor*innen, könnte durch Synthese Fachpersonal rasch konkrete Patienteninformationen zur Verfügung gestellt werden.

Im Bereich der Radiologie kommt Deep Learning bereits erfolgreich zur Beurteilung bildgebender Verfahren zum Einsatz, insbesondere in der Tumorbestimmung sind entsprechende Systeme rascher und präziser als menschliche Beurteilung und bieten wertvolle Unterstützung für Mediziner. Yan et al. (2022) verglichen speziell trainierte Varianten des bereits erwähnten BERT-Modells (z.B. RadBERT), wobei der Fokus auf der Zusammenfassung und Wissensgenerierung aus radiologischen Berichten lag. Wertvolle Hinweise auf Krankheitsverläufe und Behandlungserfolge konnten laut der Autor*innen aus den schriftlichen Dokumentationen gewonnen werden, wodurch auch neuartiges Wissen und Hypothesen generiert werden konnte. Tan et al. (2023) bestätigen diese Ergebnisse und gehen davon aus, dass aus Dokumentationen in Freitext mittels LLMs sowohl neue Erkenntnisse zu Erkrankungen (z.B. Krebsforschung), wie auch wichtige Früherkennungs- und Entscheidungstools für Gesundheitsdienstleister ermöglicht werden könnten.

Da sich LLMs durch die schnelle Verarbeitung und Synthese großer Mengen medizinischer Literatur auszeichnen, könnten sie Gesundheitsfachkräften auch helfen mit den neuesten Forschungsergebnissen auf dem Laufenden zu bleiben, so Callahan et al. (2019).

Forschung

Am Beispiel der COVID-Pandemie, zeigen Wang et al. (2020) wie Literaturrecherchen zusammengeführt und Schlüsselergebnisse aus zahlreichen Forschungsarbeiten automatisiert zusammengefasst werden könnten.

Eine Praxis, die vier Jahre später von kommerziellen Tools angeboten wird und wissenschaftlich akkurate Ergebnisse liefert.

Auch bei der Findung von Forschungsthemen, bei Hypothesengenerierung und bei der Synthese klinischer Leitlinien könnten Sprachmodelle im Gesundheitswesen eingesetzt werden (T. Dave et al., 2023).

Liu et al. (2021) halten fest, dass die Entdeckung und Entwicklung neuer Medikamente teuer, zeitaufwändig und oft ineffizient sei und mit vielen Fehlschlägen assoziiert sei. Mit Hilfe künstlicher Intelligenz hätten LLMs das Potential die Entwicklung von Therapien effektiver zu gestalten, indem KI-gestützte LLMs Unterstützung bei der Identifizierung von Targets, beim klinischen Design, bei der Entscheidungsfindung durch die Behörden und bei der Pharmakovigilanz bieten könnten.

Der Herausforderung der Patientenrekrutierung für klinische Studien widmeten sich Jin et al. (2023) in Ihrer Arbeit. Sie evaluierten TrialGPT, ein Sprachmodell, das helfen sollte geeignete Teilnehmer*innen für Studien zu identifizieren. Trial GPT erreichte in der Studie einen Genauigkeitswert von 87,3%, womit das Sprachmodell nahe der Leistung von Expert*innen lag (88,7%-90%). Die Screening Zeit hätte den Autor*innen zufolge um 42,6% reduziert werden können, die Erklärungen und Beurteilungen der Teilnehmer*innen sei nachvollziehbar gewesen und habe stark mit denen der Expert*innen korreliert.

Patient*innenaufklärung und -einbindung

Als breites Forschungsfeld für die Anwendung von LLMs im Gesundheitsbereich lässt sich der Bereich der Patient*innenaufklärung und -einbindung identifizieren.

LLMs werden eingesetzt, um personalisierte Patient*innenaufklärungsmaterialien zu erstellen und die Patient*inneneinbindung zu fördern (M. Chen et al., 2021).

EHRTutor, ein innovatives Framework, das große Sprachmodelle nutzt, um Patient*innen über ihre Entlassungsanweisungen zu informieren wurde von Zhang et al. (2023) beschrieben. EHRTutor generiert Fragen des Entlassungsmanagements und führt durch interaktive Gespräche, um das Verständnis der Patient*innen zu verbessern. Es nutzt das ReAct-Framework, um rationale Entscheidungen zu treffen und am Ende eine Zusammenfassung der wichtigsten Punkte anzubieten. Die Evaluierung zeigte, dass EHRTutor die Qualität der Patient*innengespräche verbessert werden konnte, zusätzlich konnte die Qualität durch die Generierung synthetischer Daten (künstlich erzeugte Daten, die Reale nachahmen und zum Training von Sprachmodellen verwendet werden) gesteigert werden.

Feinabgestimmte LLMs, die zur Beantwortung von Patient*innenanfragen über ein elektronisches Gesundheitsportal herangezogen wurden, dienten Liu et al. (2023) als Untersuchungsgegenstand. Zwei Modelle, CLAIR-Short und CLAIR-Long, wurden anhand eines Datensatzes von Nachrichten und Antworten einer Universitätsklinik entwickelt. CLAIR-Short erzeugte prägnante Antworten, während CLAIR-Long umfassendere Bildungsinhalte für Patient*innen bot. Die Ergebnisse zeigten, dass CLAIR-Long in Bezug auf Empathie, Reaktionsfähigkeit und Genauigkeit positiv bewertet wurde, in der Nützlichkeit neutral, ClAIR-Short schnitt als Modell mit kurzen, prägnanten Informationen gut ab. Die Autor*innen schlussfolgerten, dass LLMs erhebliches Potenzial haben würden, die Kommunikation zwischen Patient*innen und Hausärzt*innen zu erleichtern und einen Mehrwert im Sinne der Health Literacy für Patient*innen zu bieten. KI-gestützte Chatbots, die LLMs verwenden, könnten rund um die Uhr Unterstützung für Patient*innen bieten und Fragen zu Medikamenten, Symptomen und allgemeine Gesundheitsanliegen beantworten (S. Liu et al., 2023).

Den Einsatz von Sprachmodellen für die Beantwortung komplexer medizinischer Fragestellungen untersuchten Chen et al. (2024). Ihre Studie bewertet die Leistungsfähigkeit von LLMs und deren Beantwortung medizinischer Tests und Fragebögen, wie JAMA Clinical Challenge und Medbullets. Die Ergebnisse zeigen, dass diese Datensätze zwar realistischer, aber auch anspruchsvoller seien als frühere Benchmarks und dass LLMs, obwohl sie vielversprechende Erklärungen liefern, noch Defizite in der Erklärung komplexer medizinischer Entscheidungen aufweisen würden. Teilweise lieferten die LLMs fehlerhafte Antworten, hielten die Autor*innen fest.

Das Potential von Sprachmodellen mit einem breiten Weltwissen aus den Trainingsdaten, wie ChatGPT in der Version 3.5, zeigten Gilson et al. (2023). Das nicht auf die Aufgabe trainierte Modell war in der Lage war einen amerikanischen Eignungstest für Medizinstudent*innen zu bestehen, indem es mit einer durchschnittlichen Quote von 64,4% über den geforderten 60% lag. Zusätzlich seien in 100% der Antworten logische Begründungen geliefert worden, die jedoch teilweise auf Halluzinationen basierten.

ChatDoctor, ein medizinisches LLM, das mit über 100.000 Patienten-Arzt Dialogen feinabgestimmt wurde, betrachteten Li et al. (2023) und halten fest, dass das Modell durch die Integration von Echtzeit-Informationsabruf aus zuverlässigen online und offline Quellen fundierte Ratschläge aus Patient*innenanfragen zu geben in der Lage war. Die Informationen konnten wurden selbstständig an Bildungsniveau, Sprache und Präferenzen der Patient*innen angepasst und würden einen bedeutenden Mehrwert in der medizinischen Kommunikation bedeuten.

Dass die Antworten von Chatbots auf medizinische Fragestellungen von Patient*innen nicht nur ausführlicher (im Durchschnitt 211 Wörter im Vergleich zu 52 Wörter bei Antworten von Ärzt*innen), sondern insgesamt auch als qualitätvoller und empathischer wahrgenommen wurden, zeigten Ayers et al. (2023). Sie installierten Chatbots in medizinischen Foren und ließen Benutzer*innen über die gegebenen Antworten abstimmen, wobei ein Teil der Antworten von Ärzt*innen und ein Teil von Chatbots formuliert wurde. Die Autor*innen sehen großes Potential in der Entlastung von Mediziner*innen durch reduzierte Burnout Raten, indem Ärzt*innen beispielsweise vorformulierte Antworten vorgeschlagen werden könnten. Diese Antworten würden vor der Freigabe durch Ärzt*innen lediglich überprüft werden, skizzieren die Autor*innen mögliche Entlastungen für Ärzt*innen.

Medizinische Dokumentation und Kodierung

LLMs werden im klinischen Setting bereits eingesetzt, um bei zeitaufwändigen Aufgaben der medizinischen Dokumentation und Kodierung zu unterstützen (Aronson et al., 2023).

Im Bereich der Patientendokumentation sehen Alsentzer et al. (2023) großes Potential für den Einsatz von LLMs. Diese könnten beispielsweise aus Entwürfen und Stichworten klinische Dokumentationen von Patient*innenbegegnungen generieren, die in Systeme von Gesundheitsdienstleister*innen übertragen und zur Verfügung gestellt werden könne.

Jiang et al. (2017) untersuchten KI-Anwendungen im Gesundheitswesen und heben ebenfalls die Bedeutung von LLMs für die Dokumentation und Kodierung klinischer Daten hervor. Die Autor*innen messen klinischer Dokumentation ebenfalls eine zentrale Rolle bei um Krankheitsverläufe zu beobachten, neue therapeutische Ableitungen treffen zu können, oder bestehende Diagnostik und Behandlung zu evaluieren. Der zeitintensive Prozess der Kodierung, um Daten vergleichbar und auswertbar zu machen, erfolge vorwiegend in wissenschaftlichem Kontext, da im klinischen Alltag die zeitlichen Ressourcen nicht

realisierbar seien. Durch die Anwendung von Natural Language Processing (NLP) könnten diese Texte in maschinenlesbare, strukturierte Daten umgewandelt werden, die dann mittels Machine Learning (ML) analysiert werden könnten, so Jiang et al. Diese Prozesse seien essenziell, um klinisch relevante Informationen zu extrahieren und die Entscheidungsfindung im Gesundheitswesen zu unterstützen und würden einen erheblichen Mehrwert für Gesundheitsdienstleister*innen und Patient*innen bieten.

Zu einem ähnlichen Schluss kamen Suvirat et al. (2023). Medizinische Kodierung, wie sie für Abrechnung, Planung von Gesundheitsressourcen, der klinischen Forschung und der Qualitätsbewertung essenziell sei, könne durch automatisierte Kodierungssysteme nicht nur effizienter gestaltet werden, sondern würde auch die Genauigkeit verbessern und zugleich die Belastung für medizinisches Personal verringern. In Ihrer Arbeit untersuchten die Autor*innen über sechs Jahre hinweg den Einsatz von Sprachmodellen zur Kodierung von stationären Diagnosen und bewerteten deren Leistung anhand eines umfangreichen Datensatzes von 230.645 stationären Aufnahmen und 8677 Diagnosecodes. Drei Sprachmodelle, darunter zwei allgemeine und ein domänenspezifisches Modell, wurden evaluiert. Das domänenspezifische Modell erzielte die besten Ergebnisse mit einem Mikro-F1-Score von 0,7821 und einer mittleren durchschnittlichen Präzision von 0,8097. Die Leistung der Modelle variierte je nach Krankheit und Zustand, wobei Diagnosecodes mit größeren Stichprobengrößen bessere Ergebnisse lieferten. Herausforderungen ergaben sich laut Suvirat et al. bei seltenen Krankheiten und langen klinischen Dokumenten. Die Ergebnisse würden das Potenzial von LLMs zur Verbesserung der Kodierungsgenauigkeit und die Entlastung von medizinischen Kodierer*innen belegen.

Einen ähnlichen Ansatz unternahmen Guevara et al. (2024), die ebenfalls feinabgestimmte Modelle (Flan-T5-Modell) herangezogen haben, um aus unkodierten klinischen Dokumentationen soziale Determinanten der Gesundheit (SDoH), wie Beschäftigung, Wohnen, Transport, Elternstatus, Beziehung, soziale Unterstützung etc. zu extrahieren. Das auf den Einsatz trainierte Modell Flan-T5 zeigte deutlich geringere Verzerrungen, als die Referenzmodelle GPT und BERT und war in der Lage signifikant mehr Patient*innen mit negativen SDoH zu identifizieren, als es die ICD-10 Kodierung erlauben würde.

Medizinische Ausbildung und Fortbildung

LLMs könnten auch in der Ausbildung und Weiterbildung Einsatz finden, indem sie interaktive Übungsfälle bereitstellen, in der Prüfungsvorbereitung (z.B. durch die Generierung individueller Multiple Choice Fragen), oder beim Brainstorming für Differentialdiagnostik assistieren könnten (Safranek et al., 2023).

Durch adaptives Lernen seien LLMs in der Lage personalisierte Lernmaterialien und Quizzes, basierend auf dem individuellen Fortschritt und den Verbesserungsbereichen der Lernenden, zu erstellen (Zhou et al., 2023).

Als Voraussetzung für eine verbesserte Wirksamkeit von LLMs im Einsatz von medizinischen Anwendungen identifizieren Wang et al. (2023) die Integration vielfältiger realer Daten, wie Krankenakten und domänenspezifiziertes Wissen. Die Autor*innen untersuchten ClinicalGPT, ein auf medizinische Wissensvermittlung, Patientenberatung und für diagnostische Entscheidungsfindung vortrainiertes Modell, welches untrainierte Modelle in Wissens- und Diagnosetests signifikant übertrifft und bei medizinischen Aufgaben deutlich besser abschneidet als diese.

Public Health und Epidemiologie

Welchen Einfluss die Anwendung von LLMs auf Fragestellungen der Public Health haben könnte, haben Chen et al. (2023) anhand der COVID-19 Pandemie in ihrem Paper analysiert. Die Notwendigkeit epidemiologische Daten und wissenschaftliche Erkenntnisse leicht zugänglich und rasch verfügbar zu machen, könne durch den Einsatz von LLMs deutlich vereinfacht werden. Um Entscheidungsträger*innen zeitnahe Grundlagen zu liefern, sei Analyse und Auswertung wissenschaftlicher Erkenntnisse durch LLMs geradezu prädestiniert. Die Autor*innen zeigten diesen Mehrwert, indem ein Dashboard mit Daten aus der Analyse von Echtzeit-Pathogen-Genomdaten erstellten, welches eine neue Form der Interaktion mit Informationssystemen sowohl für Entscheidungsträger*innen, wie auch für gesellschaftliche Diskurse und Information ermögliche (C. Chen & Stadler, 2023).

Jiang et al. (2023) untersuchten in Ihrer Arbeit ein Framework namens ALEX, welches mittels Datenaugmentation Gesundheitsdaten aus sozialen Netzwerken extrahiert und Ableitungen für gesundheitsrelevante Fragestellungen trifft. Das Framework, welches beim „Social Media Mining for Health 2023“ (SMM4H), einem Wettbewerb bei dem Gesundheitsdaten aus sozialen Medien für Fragen der Public Health herangezogen werden, den ersten Platz belegte, solle das Potential zeigen öffentlich zugängliche Daten für gesundheitspolitische Entscheidungen nutzbar machen zu können.

Mit der zunehmenden Zahl an Sensordaten, die beispielsweise durch Wearables individuell erhoben werden, entstünden auch Ableitung für psychische Gesundheit und deren Korrelation zu Aktivitätsniveau, Herzschlag und Schlafrhythmus. Die Interpretation von Multi-Sensor-Daten sei für Expert*innen von großem Interesse halten Engelhardt et al. (2023) fest. In einem Zusammenspiel aus LLMs, KI und menschlichen Expert*innen könnten diese Daten für fundierte klinische Entscheidungen und Public Health Maßnahmen nutzbar gemacht werden, wobei durch den Einsatz von LLMs im Studiensetting eine genauere Depressionsabstufung vorgenommen werden konnte als es konventionelle Systeme bislang konnten.

4.1.1 Anwendungen in therapeutischen Berufen

Den soeben angeführten breiten Anwendungsbereich von LLMs im Gesundheitswesen, soll im Folgenden eine Betrachtung nach Einsatzmöglichkeit in den der Logopädie verwandten und unter dem österreichischen MTD-Gesetz (2024) gemeinsam geregelten Berufsgruppen erfolgen. In Österreich umfassen die gehobenen medizinisch-therapeutisch-diagnostischen Gesundheitsberufe (MTD) sieben Berufe: Biomedizinische Analytik, Diätologie, Ergotherapie, Logopädie, Orthoptik, Physiotherapie und Radiologietechnologie. Durch die Betrachtung verwandter therapeutischen Berufsgruppen soll die Breite möglicher Anwendungsfälle besser abgebildet werden, da sich diesbezügliche Forschung in der Logopädie noch unzureichend in wissenschaftlichen Studien niederschlägt.

Physiotherapie

Die Physiotherapie, wie die gesamte Gesundheitsbranche sehen Naqvi et al. (2024) in einer transformativen Phase, in der Physiotherapeut*innen von LLMs profitieren könnten, indem administrative Aufgaben automatisiert, globale Vernetzung von Therapeut*innen und Daten vorangetrieben und somit therapeutische Interventionen potenziell verbessert werden könnten. Die Autor*innen weisen auf Herausforderungen im Bereich der Physiotherapie hin, da eine Vielzahl an erhobenen Parametern in der Physiotherapie eine Genauigkeit von Modell und Messverfahren voraussetzen würden, welche auch

spezialisierte Modelle, wie BioMedLM noch an ihre Grenzen bringen würde. Naqvi et al. betonen jedoch die Notwendigkeit für Therapeut*innen sich aktiv an der Gestaltung zukunftsweisender Anwendungen zu beteiligen, um die Potentiale voll ausschöpfen zu können.

Für Zernikov et al. (2023) könnten Fähigkeiten von LLMs, wie die Textgenerierung in leicht verständlicher Sprache, zu einer Erhöhung der Gesundheitskompetenz und somit auch zu einer gesteigerten Therapieadhärenz beitragen. Eine künftige Implementierung lokaler LLM in medizinischen Verwaltungssystemen (wie dem österreichischen Krankenhausinformationssystemen - KIS) in Verbindung mit der elektronischen Patientenakten habe das Potential auch in der Physiotherapie klinische und ambulante Versorgung grundlegend zu verändern.

Ergotherapie

Explizit die Ergotherapie als mögliche Zielgruppe nennen Berrezueta-Guzman et al. (2024) in ihrer Studie, in der sie die Integration von ChatGPT-4 Turbo und Claude-3 Opus (ein Modell von AnthropicAI [10]) als Roboterassistenten zur Unterstützung der Therapie von Kindern mit Aufmerksamkeitsdefizit-Hyperaktivitätsstörung (ADHS) untersuchen. Die eingesetzten Roboterassistenten, die einfache Mimik und Sprachreaktion beherrschten, wären in der Lage die Therapie personalisierter und interaktiver zu gestalten. ChatGPT-4 Turbo zeigte eine höhere Leistung und schnellere Reaktionszeiten, während Claude-3 Opus in Bezug auf Empathie und ethische Überlegungen besser abschnitt, urteilen die Autor*innen hinsichtlich der Unterschiede beider untersuchten Sprachmodelle. Weiters streichen Sie in der Studie die Notwendigkeit weiterer Forschung und Anpassung heraus, um die volle Integration dieser Technologien und Sozial Assistiver Roboter (SAR), in therapeutische Umgebungen zu ermöglichen.

Diätologie

Die multimodalen Fähigen von LLMs sind Grundlage für eine Ernährungsapp, die von Oka et al. (2019) in einer randomisierten Studie mit 100 Teilnehmer*innen auf ihre Wirksamkeit untersucht wurde. Eine KI-gestützte App, die zu einer Verbesserung der Blutzuckerkontrolle für Patient*innen mit Typ-2 Diabetes verhelfen sollte, indem Essensfotos automatisiert nach Nährwert und Energie beurteilt werden, zeigte hohe Validitätswerte für Wirksamkeit und Zuverlässigkeit.

Dass untrainierte und nicht feinabgestimmte Modelle bei komplexen medizinischen Fragestellungen rasch an Ihre Grenzen stoßen würden, zeigten Mishra et al. (2024) in ihrer Studie in welcher sie Antworten von ChatGPT auf Ernährungsfragen, die von Diätolog*innen erstellt wurden, auswerteten. Das Risiko potenzieller gesundheitlicher Schäden stieg mit zunehmender Komplexität der Szenarien, wenngleich die Genauigkeit der Antworten konstant geblieben sei.

Biomedizinische Analytik

Um den Vergleich mit unspezifischen Modellen wie BERT, oder CHatGPT herauszufinden, trainierten die Autor*innen Luo et al. (2022) ein Sprachmodell namens BioGPT und verglichen die unterschiedlichen Modelle mittels von Expert*innen zur Verfügung gestellten biomedizinischen Fragestellungen. BioGPT, ein domänenspezifisches Transformermodell, das auf über 15 Millionen PubMed-Abstrakten trainiert wurde, zeigte sowohl in

[10] https://www.anthropic.com/

Relationsextraktion (Beziehungen zwischen biomedizinischen Entitäten), wie auch in Fragebeantwortung, Textextraktion und Dokumentklassifikation deutliche Vorteile gegenüber untrainierten Modellen. BioGPT könne für die Erstellung von verständlichen und präzisen biomedizinischen Texten, zur Extrahierung relevanter Informationen aus großen Mengen biomedizinischer Literatur, oder zur Identifizierung neuer Erkenntnisse und Zusammenhänge in biomedizinischen Daten herangezogen werden (Luo et al., 2022).

Auch Gu et al. (2022) führten eine ähnliche Studie im Bereich biomedizinischer Analytik durch, in welcher die Autor*innen ebenfalls ein Sprachmodell auf medizinische Texte trainiert haben und zeigen konnten, dass feingetunte Modelle deutlich gesteigerte Performance bieten würden, als das als Referenz herangezogene BERT Modell.

Radiologietechnologie

Wie im Kapitel 2.2.1 bereits erwähnt, zeigten Tan et al. (2023), dass aus unstrukturierten medizinischen Dokumentationen wertvolle Ableitungen zu Krankheitsverläufen in Bezug auf Krebserkrankungen getroffen werden konnten und Sprachmodelle eine hilfreiche Ergänzung und Zweitmeinung für medizinisches Personal bieten könne.

Das von Vandewinckele et al. (2020) entwickelte Modell zum Einsatz in der Radiologie stützt sich sowohl auf Machine Learning, wie auch auf das Transformermodell. Es wurde eingesetzt, um Organe und Tumore automatisch zu segmentieren, was die Variabilität und den Zeitaufwand im Vergleich zur manuellen Segmentierung reduziere und die Qualität und Sicherheit von Strahlentherapie verbessere. Die Autor*innen trainierten ihr Modell mit qualitativ hochwertigen Datensätzen und validierten diese, um eine hohe Genauigkeit zu gewährleisten. Zusätzlich wurde das Modell mit Daten von Patient*innen trainiert, die nach einem vordefinierten Protokoll behandelt wurden. Mittels synthetischer Daten konnte die Dosisberechnung anhand von MRT-Bildern nochmals verbessert werden. Die Autor*innen halten fest, dass der Implementierungsprozess durch ein multiprofessionelles Team begleitet werden müsse und regelmäßige Qualitätssicherung im laufenden Betrieb zur Überwachung und Validierung erfolgen müsse.

Boecking et al. (2022) ist mit dem sogenannten MS-CXR-Datensatz, der entwickelt wurde um die semantische Modellierung zu verbessern, einen Fortschritt in der biomedizinischen Bild-Text-Verarbeitung gelungen. Der Datensatz umfasst 1162 Paare von Bildern, wobei die Anmerkungen acht verschiedene kardiopulmonale Befunde abdecken, die von zertifizierten Radiolog*innen und Radiologietechnolog*innen erstellt wurden. Die Autor*innen gehen davon aus, dass der Datensatz die reproduzierbare Bewertung von Modellen fördern könne, die somit eine latente Semantik zwischen Bild und Text erlernen könnten und als Trainingsbasis herangezogen werden könnten.

Orthoptik

In der Orthoptik gäbe es bereits Systeme, welche als präzise Diagnosehelfer fungieren könnten, sie könnten beispielsweise MR- und CT-Aufnahmen klassifizieren und seien im Stande kardiovaskuläre Risikofaktoren, Nikotinabusus oder das Geschlecht anhand von Netzhautbildern vorherzusagen (Guthoff & Stachs, 2019).

Auch Tailor et al. (2024) zeigen in Ihrem Review „Vision of the future“ die vielseitigen Einsatzmöglichkeiten von LLMs im Feld der Orthoptik. Die Autor*innen geben einen Überblick über zahlreiche Studien des Fachgebiets und halten fest, dass LLMs die Patientenkommunikation verbessern und bei vorläufigen Diagnosen unterstützen könnten. Die Autor*innen identifizierten jedoch auch potenzielle Herausforderungen durch fehlende

Akzeptanz und Datenschutz als wesentliche Hürden für eine vollständige Integration in die klinische Praxis.

Logopädie

Nicht nur für Psycholog*innen, sondern auch für weitere Berufsgruppen die mit autistischen Jugendlichen arbeiten, sind die Forschungsergebnisse von Cho et al. (2023) relevant. Sie untersuchten die Wirksamkeit interaktiver, auf LLMs basierter, Sprachtherapie für hochfunktionale autistische Jugendliche. Betrachtet wurden insbesondere Anpassungsfähigkeit, Engagement und Empathie. Die vielversprechende Bewertung fand durch klinische Psycholog*innen, anhand speziell entwickelter Fragebögen, statt. Die Autor*innen betonen explizit Herausforderungen bei der Personalisierung hinsichtlich emotionaler Tiefe und die Bedeutung eingehender ethischer Überlegungen bei der Anwendung von KI und Sprachmodellen in therapeutischen Kontexten.

Auf den Bereich der unterstützten Kommunikation fokussierten sich Beukelman et al. (2009) in ihrer Studie, es wurden Fortschritte in der AAC-Technologie (Augmentative and Alternative Communication) für verschiedene neurologische Erkrankungen bei Erwachsenen wie Amyotrophic Lateral Sclerosis (ALS), Hirnstammbeeinträchtigungen, Aphasie und Demenz beschrieben. Einige der beschriebenen Herausforderungen könnten durch LLMs gelöst werden, so könnten Sprachmodelle helfen, komplexe Gedanken und Bedürfnisse der Betroffenen besser zu verstehen und zu kommunizieren. Außerdem könnte die Kommunikation individuell angepasst werden, um den spezifischen Anforderungen und Vorlieben der Nutzer*innen gerecht zu werden und sie könnten die Nutzung von AAC-Geräten intuitiver und zugänglicher machen. Von den Autor*innen identifizierte Herausforderungen, wie Notwendigkeit einer kontinuierlichen Anpassung und Verbesserung der Technologien, um den sich ändernden Bedürfnissen der Nutzer gerecht zu werden, bleiben auch unter dem Blickwinkel von LLMs bestehen und relevant.

4.2 Ergebnisse der Fokusgruppen

Die Durchführung des Fokusgruppeninterviews mit fünf Logopäd*innen lieferte relevante Einblicke in deren Einstellungen und Erfahrungen bezüglich potenzieller Anwendung von LLMs in ihrer beruflichen Tätigkeit. Die Analyse der Transkripte ergab sechs Schlüsselthemen, die im Folgenden entsprechend ihrer Referenzhäufigkeit in absteigender Reihenfolge dargestellt werden.

4.2.1 Digitale Transformation und Zukunftsperspektiven in der Logopädie

Die digitale Transformation in der Logopädie wurde als zentrales Thema identifiziert, das die zukünftige Ausrichtung des Berufsfeldes maßgeblich beeinflussen könnte. Die Teilnehmer*innen betonten die Notwendigkeit, digitale Technologien in den therapeutischen Alltag zu integrieren, um die therapeutischen Interventionen zu optimieren und innovative Behandlungsmethoden zu entwickeln. Besonders hervorgehoben wurde die Rolle von künstlicher Intelligenz (KI) und Augmented Reality (AR) als potenzielle Instrumente zur Erweiterung der logopädischen Praxis. Teilnehmerin drei berichtete von einer AR-gestützten Teletherapie-Initiative aus der Schweiz und erklärte:

„Dieses Projekt verwendet sogar Kameras, die Hologramme projizieren, wodurch die Illusion entsteht, dass die Therapeut*innen physisch vor Ort anwesend sind."

Solche Technologien könnten das Konzept der ortszentrierten Therapie revolutionieren und neue Möglichkeiten für die logopädische Praxis bieten. Zudem wurde die Möglichkeit

zur automatisierten Analyse bildgebender Verfahren hervorgehoben, wie von Teilnehmerin drei angeführt:

„Wenn man da Unterstützung von einer KI hätte, könnte man viel Zeit sparen."

Diese Zeitressourcen wiederum könne in Folgen den Patient*innen zukommen wie Teilnehmerin drei weiter ausführte.

4.2.2 Ethische und Praktische Herausforderungen

Mit den Chancen, die LLMs bieten, gehen auch erhebliche ethische und praktische Herausforderungen einher. Datenschutz und Datensicherheit stellten sich als zentrale Anliegen dar, um die Integrität und Vertraulichkeit von Patient*innendaten zu gewährleisten. Teilnehmerin eins äußerte:

„Es ist von höchster Priorität, dass Patientendaten vor Diebstahl oder Missbrauch geschützt werden."

Des Weiteren wurde die Gefahr einer übermäßigen Abhängigkeit von Technologie diskutiert. Drei der fünf Teilnehmer*innen betonten die Notwendigkeit, die Kontrolle über technologische Systeme zu behalten:

„Es ist essenziell, dass wir die Kontrolle über die Technik behalten und nicht umgekehrt."

Zudem wurde die Notwendigkeit umfassender Schulungsprogramme hervorgehoben, um Anwendungsfehler zu vermeiden und die effektive Nutzung der Technologien sicherzustellen. Teilnehmerin eins fasste zusammen:

„Viele Kolleginnen und Kollegen werden neuen Programmen oder Systemen skeptisch gegenüberstehen, insbesondere wenn keine adäquate Schulung erfolgt."

4.2.3 Veränderte Rolle der Therapeut*innen

Die Integration von LLMs könnte die Rolle der Therapeut*innen erheblich verändern und neue Kompetenzen erfordern. Automatisierte Systeme zur Dokumentation und Berichterstattung wurden als Mittel zur Effizienzsteigerung angesehen, die den Therapeut*innen mehr Zeit für die eigentliche Therapie ermöglichen würden. Teilnehmerin fünf bemerkte:

„Ich glaube, es wäre äußerst vorteilhaft, wenn Berichte automatisch erstellt werden könnten. Das würde uns erheblich entlasten."

Diese Automatisierung würde es den Therapeut*innen erlauben, sich stärker auf die direkte Patient*innenversorgung zu konzentrieren und gleichzeitig den Zeitaufwand administrativer Aufgaben zu minimieren.

4.2.4 Patient*innenenzentrierung

Die Patient*innenorientierung wurde als zentrales Element bei der Einführung digitaler Technologien identifiziert. Die Teilnehmer*innen betonten die Möglichkeit, therapeutische Ansätze individuell auf die Bedürfnisse und Interessen der Patient*innen abzustimmen. Dies könne durch den Einsatz unterstützender Technologien erreicht werden, die eine personalisierte Therapieplanung und kontinuierliches Feedback ermöglichen. Teilnehmerin vier erläuterte:

„Wenn es Veränderungen gibt, könnten die Patienten diese direkt in eine Anwendung eingeben und so den Therapeut*innen mitteilen, ob sich ihr Zustand verbessert oder verschlechtert hat. Dies würde die Autonomie der Patienten fördern."

Solche Funktionen würden nicht nur die Eigenverantwortung der Patient*innen stärken, sondern auch die Kontinuität der Therapie über herkömmliche Sitzungen hinaus verbessern, ein Vorteil gegenüber konventioneller Therapiekonzepte, wie die Logopäd*innen einheitlich festhalten.

4.2.5 Verbesserter Interdisziplinärer Informationsaustausch

Ein weiterer wesentlicher Vorteil der Implementierung von LLMs liege in der Verbesserung des interdisziplinären Informationsaustauschs im Gesundheitswesen. Die Entwicklung digitaler Patient*innenakten, die von verschiedenen Berufsgruppen genutzt werden könnten, wurde als besonders nützlich angesehen. Teilnehmerin vier schlug vor:

„Man könnte eine Art digitale Patientenkarte erstellen, in der sich alle beteiligten Berufsgruppen registrieren würden. Dies würde als zentrale Informationsablage dienen, die die Kommunikation zwischen Therapeut*innen erheblich erleichtern würde."

Eine solche zentrale Informationsablage würde den Datenaustausch optimieren und eine umfassende Patient*innenbeurteilung ermöglichen, was zu einer besseren Koordination und Qualität der Behandlung beitragen könnte. Dies könne besonders für Logopäd*innen im niedergelassenen Bereich dazu beitragen effizienter und zielgerichteter therapeutische Interventionen zu planen und zu gestalten.

4.2.6 Balance zwischen Technologie und Menschlichem Kontakt

Trotz der zahlreichen Vorteile digitaler Tools betonten die Teilnehmer*innen unisono die unverzichtbare Bedeutung der menschlichen Interaktion in der logopädischen Praxis. Die Gefahr, dass die Technik die persönliche Beziehung zwischen Therapeut*innen und Patient*innen ersetzen könnte, wurde mehrfach hervorgehoben. Logopädin zwei betonte:

„Gerade in der Logopädie ist es essenziell, dass wir direkt mit den Patienten arbeiten, sie beobachten und unsere eigenen Schlüsse ziehen. Das ist unser Kapital und muss auch so bleiben, damit wir bestmöglich arbeiten können."

Diese Perspektive unterstreicht, dass Technologien wie LLMs als unterstützende Instrumente dienen sollten, um die therapeutische Arbeit zu ergänzen, ohne die entscheidende menschliche Komponente zu verdrängen. Alle Diskussionsteilnehmer*innen teilten die Einschätzung, dass die menschliche Interaktion weiterhin Priorität haben sollte, um eine effektive und empathische Therapie zu gewährleisten.

4.2.7 Anwendungsfälle im Kontext der Logopädie

Therapiematerial und Personalisierung

Die Erstellung und Anpassung von Therapiematerialien an die individuellen Bedürfnisse der Patient*innen ist ein zentraler Anwendungsfall. Durch den Einsatz von Bildgenerierungsprogrammen könnten personalisierte Materialien schneller und effizienter erstellt werden. Teilnehmerin eins berichtete aus ihrem Arbeitsalltag:

„Gerade kürzlich habe ich gedacht, dass Bildgenerierungsprogramme echt nützlich wären. Zum Beispiel, wenn man in der Neurologie mit Bildkarten arbeitet, könnte man diese an die individuellen Patient*innen anpassen."

Informationserfassung und -weitergabe

Die automatische Erfassung und Weitergabe von Patient*innendaten an alle beteiligten Berufsgruppen zur Verbesserung der Qualität der Patient*innenversorgung thematisierte Teilnehmerin fünf:

„Man könnte eine Art digitale Patientenkarte erstellen, in der sich alle beteiligten Berufsgruppen registrieren würden. Dies würde als zentrale Informationsablage dienen, die die Kommunikation zwischen Therapeuten erheblich erleichtern würde.“

Begleiter für Therapeut*innen und Patient*innen

Digitale Assistenztools könnten sowohl Therapeut*innen als auch Patient*innen unterstützen, indem sie zusätzliche Ressourcen und Feedback ermöglichen. Teilnehmerin vier führte aus:

„Wenn es Veränderungen gibt, könnten die Patienten diese direkt in eine Anwendung eingeben und so den Therapeut*innen mitteilen, ob sich ihr Zustand verbessert oder verschlechtert hat. Dies würde die Autonomie der Patienten fördern.“

Bildinterpretation und Auswertung von Assessments

Der Einsatz von KI zur automatisierten Analyse und Auswertung bildgebender Verfahren wie bei der Flexiblen endoskopischen Evaluation des Schluckakts (FEES), oder bei der Interpretation von Röntgenbildern könnte die Diagnostik und Therapieplanung unterstützen. Teilnehmerin zwei berichtet über Herausforderungen in der Beurteilung von Thoraxaufnahmen, wie sie insbesondere im Kanülenmanagement auf Intensivstationen benötigt wird:

„Eine KI könnte Bilder analysieren und uns Hinweise geben, was auf den Aufnahmen zu sehen ist. Oft sitzt man eine halbe Stunde vor einem Bild und ist sich nicht sicher, was es im Detail zeigt und muss erst recht wieder Ärzte fragen, da geht viel Zeit drauf.“

Automatisiertes Hervorheben relevanter Strukturen, Hilfestellung bei der Interpretation, oder eine Verlaufsanzeige könnten wertvolle Unterstützung auch in der Dysphagietherapie darstellen, bekräftigte eine Kollegin.

Fortbildung und Wissensvermittlung

Digitale Plattformen können genutzt werden, um Fortbildungen flexibler zu gestalten und Therapeut*innen den Zugang zu aktuellen Forschungsergebnissen zu erleichtern. Teilnehmerin fünf schilderte eine aktuelle Erfahrung:

„Vielleicht könnten Fortbildungen auch flexibler gestaltet werden, indem man sie online jederzeit ansehen kann und trotzdem die Möglichkeit hat, Fragen zu stellen, die dann später beantwortet werden. Ich habe letztens eine Onlinefortbildung gemacht, die war zwar gut, aber ich würde mir da noch mehr und besseres Angebot wünschen“

Telemedizin und Augmented Reality (AR)

Der Einsatz von Telemedizin und AR könnte die Durchführung von Therapien über große Entfernungen hinweg ermöglichen und würde die Möglichkeiten der Behandlung durch immersive Technologien erweitern. Teilnehmerin drei berichtete über ihre Erfahrungen aus der Schweiz:

„Sie haben Systeme entwickelt, die zum Beispiel Regale an die Wand projizieren, um mit Patienten einkaufen zu üben. Das Ganze läuft über so eine Arte Hologramme, und der Therapeut kann aus der Ferne beobachten, wie der Patient reagiert.“

4.2.8 Ethische Aspekte

Bei der Durchführung der Fokusgruppeninterviews und der anschließenden Analyse der Daten wurden die ethischen Grundsätze der Forschung eingehalten.

Die Teilnahme an den Fokusgruppeninterviews war freiwillig. Vor ihrer Teilnahme wurden alle Teilnehmer gründlich über die Ziele und den Inhalt der Studie, die Interviewmethodik und ihre Rechte als Teilnehmer*innen informiert. Alle Tester*innen gaben ihr schriftliches Einverständnis, an der Studie teilzunehmen und ihre anonymisierten Daten für wissenschaftliche Zwecke zu verwenden.

Die Daten der Tester*innen wurden anonym verarbeitet und gespeichert. Die aus den Interviews generierten Transkripte wurden anonymisiert, und es wurden keine identifizierenden Informationen gesammelt.

Alle während der Interviews, oder der Feedbackphase offengelegten Informationen wurden streng vertraulich behandelt.

4.3 Entwicklung von Prototypen

In diesem Abschnitt wird der prozedurale Rahmen für die Entwicklung der Prototypen (ReportGPT, LoCo und Blitzstarter*in) beschrieben. Der Fokus liegt auf der Darstellung der technischen Methoden, aufgetretenen Herausforderungen und zentraler Entscheidungen, die während der Implementierungsphasen der Prototypen getroffen wurden.

Die Erstellung dieser Prototypen erforderte eine umfassende Auseinandersetzung mit der neuen kommerziell verfügbaren Technologie der LLMs und eine intensive Beschäftigung mit einer Reihe von Plattformen und Tools um diese anzusteuern und implementieren zu können. Dieses Kapitel ist systematisch in die Phasen der Identifizierung, des Entwurfs, der Implementierung und der Iteration der Prototypen gegliedert.

4.3.1 Identifizierung und Zielsetzung

Die Festlegung der Voraussetzungen für die Prototypenentwicklung basierte auf einer methodischen Analyse der Ergebnisse der Literaturrecherche (siehe Anhang 2: Identifizierte Anwendungsfälle in der Literatur) und der Fokusgruppendiskussion. Diese Ergebnisse wurden dokumentiert, gruppiert und zusammengeführt (siehe Tab. 3). Die Zusammenschau bildete die Basis für die weiteren Entwicklungsentscheidungen und die Auswahl der spezifischen Anwendungsfälle.

Anwendungsfall	Beschreibung	Quelle	Potenzielle Vorteile
Dokumentations-erstellung	Automatisierte Erstellung von Berichten und Dokumentationen aus unstrukturierten Notizen	Literaturrecherche, Fokusgruppenintervi ew	Zeitersparnis, Standardisierung
Therapiematerial-generierung	Erstellung und Anpassung von Therapiematerialien an individuelle Patientenbedürfnisse	Literaturrecherche, Fokusgruppenintervi ew	Personalisierung, Effizienzsteigerung
Diagnostik-unterstützung	Unterstützung bei der Analyse und Auswertung diagnostischer Daten	Literaturrecherche, Fokusgruppenintervi ew	Präzisere Diagnosen, Zeitersparnis

Administrative Unterstützung	Automatisierung von Verwaltungsaufgaben wie Terminplanung oder Abrechnungen	Fokusgruppenintervi ew	Reduzierung des Verwaltungsaufwan ds
Patient*innen-kommunikation	Verbesserung der Kommunikation mit Patienten durch klare und ansprechende Informationsbereitste llung	Literaturrecherche	Förderung der Patientenautonomie , verbesserte Health Literacy
Interdisziplinärer Informations-austausch	Erleichterung des Informationsaustaus chs zwischen verschiedenen Fachbereichen	Fokusgruppenintervi ew	Verbesserte Zusammenarbeit, ganzheitlichere Behandlung
Fortbildung und Wissensvermittlu ng	Unterstützung bei der kontinuierlichen Weiterbildung von Logopäden	Fokusgruppenintervi ew	Aktuelles Fachwissen, personalisierte Lernmöglichkeiten
Telemedizin und Augmented Reality	Ermöglichung von Ferntherapien und Einsatz von AR in der Behandlung	Fokusgruppenintervi ew	Erweiterte Behandlungsmöglic hkeiten, verbesserte Zugänglichkeit
Praxismanageme nt	Unterstützung bei der Analyse von Praxisdaten und strategischer Planung	Literaturrecherche	Optimierung des therapeutischen Alltags, fundierte Entscheidungen

Tab. 2: Die gruppierten, identifizierten Anwendungsfälle aus Fokusgruppe und Literatur

Die erste Phase der Entwicklung zielte nicht nur darauf ab Anwendungsfälle zu identifizieren, sondern diente auch dazu spezifische Herausforderungen im Bereich der logopädischen Praxis aufzudecken, denen durch den Einsatz von LLMs begegnet werden könnte. Dabei wurden mehrere wesentliche Kriterien berücksichtigt:

Die Verfügbarkeit geeigneter Technologien und Modelle, mit denen die erforderlichen Funktionen bereitgestellt werden können wurde untersucht. Weiters musste eine unmittelbare Anwendbarkeit gegeben sein, um eine funktionale Anwendung zur Verfügung stellen zu können. Der angemessene zeitliche und finanzielle Aufwand wurde abgeschätzt, einschließlich der Entwicklungskosten und des laufenden Betriebs (z. B. API-Dienste und Serverkosten). Eine zentrale Säule stellte die Sicherstellung der Einhaltung der Allgemeinen Datenschutzverordnung (DSGVO) bei der Verwaltung von Daten dar, dies wurde beispielsweise durch die Nutzung DSGVO-konformer Serverinfrastruktur gewährleistet. Im Folgenden findet sich eine Aufstellung (Abb. 6) der API-Kosten pro 1000 Token vom Februar 2024 für OpenAIs damalige Modelle GPT-4 und GPT-3.5 als Grundlage für die Kostenabschätzung.

Für die Umsetzung wurde insbesondere zu Beginn, überwiegend die OpenAI-API genutzt, da diese als eine der Ersten verfügbar war, gut dokumentiert und verbreitet war. Weiters konnten damit sehr leistungsfähige Modelle genutzt werden. Im späteren Verlauf der Entwicklung, nach Verfügbarkeit, wurden weitere Modelle im Vergleich getestet, eingebunden und implementiert.

Dienst/Modell	Typ	Kosten
GPT-4 Turbo (gpt-4-1106-preview)	Input	$0.01 / 1K Tokens
	Output	$0.03 / 1K Tokens
GPT-4	Input	$0.03 / 1K Tokens
	Output	$0.06 / 1K Tokens
GPT-4-32k	Input	$0.06 / 1K Tokens
	Output	$0.12 / 1K Tokens
GPT-3.5 Turbo (gpt-3.5-turbo-0125)	Input	$0.0005 / 1K Tokens
	Output	$0.0015 / 1K Tokens
GPT-3.5 Turbo Instruct	Input	$0.0015 / 1K tokens

Abb. 6: OpenAI Tokenpreis [11] Februar 2024

Die Einbindung der Fokusgruppe und die Ergebnisse der Literaturrecherche bildeten ein breites und fundiertes Spektrum potenzieller Anwendungsfälle. Dies hat insbesondere dabei geholfen die Anforderungen der potenziellen Nutzer*innen zu verstehen und die Anpassung der Funktionen der Prototypen an diese Bedürfnisse zu ermöglichen.

Gemäß den Kriterien wurden die folgenden Anwendungsfälle erkannt und anschließend als Prototypen realisiert:

1. **ReportGPT**: Automatisierte Erstellung logopädischer Berichte aus unstrukturierten Notizen.
2. **LOCO – Der Logopädische Co-Pilot**: Virtueller Assistent für fachliche und administrative Unterstützung.
3. **Blitzstarter*in**: Unterstützung von Logopäd*innen bei der Praxisgründung in allen Belangen.

Anwendungsfälle wurden verworfen, wenn sie technologisch zum Zeitpunkt der Erstellung der Masterarbeit noch nicht in einer robusten Variante kommerziell, oder als open-source verfügbar waren, wie beispielsweise konsistente Text-Bild Generierung für die Erstellung von personalisiertem Therapiematerial. Aufgrund zu hoher erwartbarer Kosten wurden beispielsweise multimodale Anwendungsfälle unter Nutzung kostenintensiver Sprach-, oder Videomodellen nicht genutzt. Um nicht gegen Datenschutzbestimmungen und Urheberrecht zu verstoßen, wurden keine LLM-Anwendungen mit Wissensdatenbanken aus urheberrechtlich geschützten Unterlagen erstellt. Die Prämisse war, dass die ausgewählten Anwendungen im größtmöglichen Umfang und Funktion getestet werden könnten, um entsprechend fundiertes Feedback erhalten zu können, bei unbedingter Einhaltung ethischer und rechtlicher Grundsätze.

[11] https://openai.com/pricing#language-models

4.3.2 Entwicklungsansatz

Zu Beginn der Entwicklung der Prototypen wurden ausführliche Recherchen zu aktuellen Technologien im Bereich der LLMs und deren potenzieller Integration in prototypische logopädische Anwendungen angestellt.

Insbesondere mit GitHub [12] wurde eine Vielzahl von Projekten untersucht. Eines dieser Projekte war Langchain [13], ein Framework, das die Entwicklung von Anwendungen mit LLMs erleichtert, indem es Komponenten für die Integration, Verwaltung und Verkettung von Modellen, Datenquellen und APIs (Application Programming Interface) bereitstellt. Ein weiteres frühes Projekt war ChatDev [14], eine virtuelle, chatbasierte Softwareentwicklungsumgebung, in der spezialisierte KI-Agenten kollaborativ zusammenarbeiten und mit Funktionalitäten ausgestatten werden können. Auch Quivr [15], ein Open-Source-Toolkit, das es ermöglicht, individuelle, datenschutzfreundliche KI-Anwendungen zu entwickeln, indem es generative Modelle nutzt, um Daten aus verschiedenen Quellen zu verarbeiten und zu organisieren, wurde in dieser Phase intensiv auf Funktionalität und möglichen Einsatz getestet.

Zu diesem Zeitpunkt befanden sich viele dieser Projekte noch in den Anfängen der Entwicklung, die häufig durch unzureichende Dokumentation und einen Mangel an Erfahrungsberichten gekennzeichnet waren. Die Nutzung von GitHub-Projekten, Huggingface [16] und ähnlichen Entwickler*innenplattformen, diente in erster Linie dazu, die Durchführbarkeit der Implementierung zu bewerten und ein Verständnis für die Architektur potenzieller Anwendungen zu entwickeln.

Diese Phase umfasste auch die Analyse der Machbarkeit zur Anbindung von APIs und die Exploration von Function Calling, Plugins und weiteren Integrationstechniken. Eine API ist eine Schnittstelle, die es Softwareanwendungen ermöglicht, miteinander zu kommunizieren und Daten oder Funktionen auszutauschen. Sie definiert Methoden und Protokolle, die es Programmen ermöglichen, die Funktionen eines anderen Programms zu nutzen, ohne die zugrunde liegende Implementierung im Detail verstehen zu müssen. APIs fördern Interoperabilität, indem sie standardisierte Zugriffsmöglichkeiten auf Softwarefunktionen bereitstellen.

Erste Versuche zur Nutzung der OpenAI-API wurden durch die Entwicklung von Python-Skripten unternommen, mittels einfacher API-Aufrufe wurden die grundlegenden Modellfunktionalitäten evaluieret. Zu den technischen Hindernissen gehörten insbesondere die Aufrechterhaltung der Anpassung an die schnellen Fortschritte in der Modellentwicklung und die regelmäßige Einführung neuer Funktionen. Mit wöchentlichen Veröffentlichungen neuer Versionen und Funktionen, war häufig auch ein Wechsel im Abruf oder Aufbau verbunden. Um Klarheit und Kohärenz zu gewährleisten, wurde eine laufende Dokumentation der Anwendungen und Plattformen, die getestet wurden, gepflegt.

---

[12] https://github.com/

[13] https://github.com/langchain-ai/langchain

[14] https://github.com/OpenBMB/ChatDev

[15] https://github.com/QuivrHQ/quivr

[16] https://huggingface.co/

Ein grundlegender API-Aufruf (Abb. 7) für die Interaktion mit einem Sprachmodell, wie ChatGPT in Python folgt den Dokumentationen von OpenAI und geschieht in folgender beispielhafter Form:

```
import openai
openai.api_key = "api-key"
response = openai.ChatCompletion.create(
  model="gpt-3.5-turbo",
  messages=[
    {"role": "system", "content": "Du bist der Weihnachtsmann."},
    {"role": "user", "content": "Sag mir, was ist ein Easteregg?"}
  ]
)
print(response.choices[0].message['content'].strip())
```

Abb. 7: OpenAI API call

Wie in dem Beispielaufruf ersichtlich wird die API von OpenAI initialisiert und der Wert des entsprechenden Keys genutzt. Mit openai.ChatCompletion.create wird eine Anfrage an das Modell (gpt-3.5-turbo) gestellt. System und User beschreiben die Rolle und Anweisung. Die Antwort wird in response gespeichert und mit der Funktion print ausgegeben. Diese einfache Grundstruktur, wie sie von OpenAI zur API-Benutzung definiert ist, lässt sich im Folgenden beliebig erweitern und mit weiteren Services und Bedingungen kombinieren.

Für die Erstellung und das Debugging benötigte Codes wurden unterschiedliche Tutorials, Dokumentationen und Anwendungen, inklusive Chat-GPT und Claude verwendet. Sowohl Nützlichkeit wie auch Leistung der Sprachmodelle steigerten sich im Zeitraum der Erstellung der Masterarbeit beträchtlich. In Kombination mit zielgerichteten Prompts stellten die Codingqualitäten der marktführenden Anbieter eine große Unterstützung in der Erstellung und für das Verständnis dar. Die Entwicklung erfolgte überwiegend auf Windows unter Verwendung von Visual Studio Code[17], Git for Windows[18] und PyCharm[19] in den aktuellen Versionen.

Bereits in dieser Phase wurde auf regelmäßige Backups und eine ausführliche Dokumentation gesetzt, um Fortschritte und Erkenntnisse in einer volatilen Entwicklungsumgebung sich ständig ändernder Bedingungen und Anforderungen festhalten zu können.

4.3.3 Beschreibung der Prototypen

BerichtGPT wurde konzipiert, um Logopäd*innen die automatische Erstellung von Behandlungsberichten zu erleichtern. Der Chatbot nutzt eine Kombination aus vordefinierten Eingabevorlagen und generativen KI-Modellen, um Berichte effizient und individuell zuzuschneiden. Ziel war es, den Verwaltungsaufwand zu verringern, indem ein vollständiger Bericht auf Grundlage der vorhandenen Patient*innendokumentation auf Knopfdruck erstellt wird. Die semantischen Fähigkeiten von LLMs werden insbesondere genutzt, um Inhalte sinnvoll aus unstrukturierten Dokumentationen zu generieren. So werden beispielsweise ICD-10 Codes korrekt aus den zugrundeliegenden Dokumentationen identifiziert und ergänzt. Der Bericht folgt einem strukturierten Format mit Abschnitten wie Diagnose, Anamnese, Befund, Therapieziele, und weiteren Empfehlungen. Am Ende der Antwort werden weiterführende Möglichkeiten, wie der

[17] https://code.visualstudio.com

[18] https://gitforwindows.org

[19] https://www.jetbrains.com

Übersetzung des Berichts in gewünschte Sprachen, oder der Möglichkeit Änderungen und Anmerkungen zu integrieren, ausgegeben. BerichtGPT sollte demonstrieren, wie unter Verwendung gezielter Prompts, zeitaufwändige Aufgaben effizient und inhaltlich korrekt von LLMs ausgeführt werden könnten.

LoCo (Logopädischer Copilot) wurde als hilfsbereiter Kollege entwickelt, der auf Fragen von Logopäd*innen eingeht und umfassende, evidenzbasierte Antworten liefert. LoCo widmet sich der Bereitstellung aktueller Studien zu logopädischen Interventionen und etablierten Best Practices, die auf den neuesten empirischen Forschungsergebnissen basieren, wofür er valide Quellen nutzt. Es hilft dabei, schnell auf Informationen zuzugreifen, ohne dass langwierige Literaturrecherchen erforderlich sind. Weiters ist LoCo in der Lage Therapiepläne zu erstellen, Übungen zu erklären, Dokumente zu entwerfen und spezifische Fragestellungen zu Störungsbildern und Konzepten beantworten. LoCo kommuniziert auf Augenhöhe, verwendet medizinische Fachbegriffe und österreichspezifische Terminologie. Ziel von LoCo ist zu zeigen, dass LLMs in Form personalisierter Assistenzsysteme für Logopäd*innen einen Mehrwert durch ständige und unmittelbare Verfügbarkeit, sowie vertrauenswürdige Antworten bieten kann.

Blitzstarter*in ist die Begleiterin auf dem Weg in die Selbstständigkeit. Blitzstarter*in unterstützt Logopäd*innen beim Ausfüllen von Formularen, behält den Überblick über Behördenwege und erstellt bei Bedarf auch ein Logo für die zukünftige Praxis. Blitzstarter*in ist in der Lage bei der Steuererklärung und bei Rechtsfragen zu unterstützen und ist insbesondere für freiberuflich tätige Kolleg*innen konzipiert. Blitzstarter*in recherchiert dafür österreichische Rechtsnormen, kennt die Anforderungen an Selbständigkeit für Gesundheitsberufe und gibt Quellen und Unterlagen als Links aus. Ziel von Blitzstarter*in ist es zu demonstrieren, dass LLMs in der Lage sind langwierige und teilweise undurchsichtige Behördenwege im Sinne eines potenten Recherchetools zu erleichtern. Es wurde explizit darauf hingewiesen, dass Blitzstarter*in keine Rechtsberatung darstellt, sondern lediglich durch die benötigten Schritte und Unterlagen führt.

Diese Anwendungsfälle decken ein breites Spektrum der identifizierten Anwendungsfälle aus den Fokusgruppeninterviews und der Literaturrecherche ab und können durch ihre unterschiedlichen Ausrichtungen als konkrete Anwendungsfälle für erste Prototypen im Bereich der Logopädie angesehen werden. Die Eingabe und Ausgabe kann bei allen Chatbots über Text, Audio und Bild erfolgen.

Für alle Prototypen wurden Low-Fidelity-Wireframes erstellt, um die grundlegende Struktur und den Informationsfluss zu visualisieren. Dies war insbesondere für die Verwendung von Plattformen wie Botpress [20] und die Entwicklung mittels Python Skripten vorteilhaft.

4.3.4 Erste Prototypen mit Botpress und Cloud-Lösungen

Parallel zur Entwicklung von Python-basierten Lösungen wurden erste Prototypen unter Verwendung von LowCode und NoCode-Tools erstellt. Diese Tools erfordern keine, oder nur sehr grundsätzliche Programmierkenntnisse und bieten erweiterte LLM-Funktionalität über ihre Plattformen an. Erste Prototypen wurden beispielsweise mit Botpress, einer open-source Plattform zur Erstellung komplexer Chatbots mittels grafischem Userinterface, die die Erstellung komplexer Abfrage-Frameworks mit bedingten Parametern und Verbindungen zu APIs wie der OpenAI-API oder Anthropics Claude-API ermöglichte.

[20] https://sso.botpress.cloud/

Funktionen, Abläufe und API-Aufrufe lassen sich hierbei beliebig miteinander verknüpfen. Eine frühe Version von LoCo wurde auf diese Weise in Kombination mit unterschiedlichen Cloud-Lösungen für das Hosting des Chatbots realisiert. Diese vorläufigen Prototypen wurden entwickelt, um die grundlegenden Funktionen der geplanten Chatbots zu demonstrieren, die Leistungsfähigkeit zu testen und Herausforderungen zu identifizieren. Die von Google Cloud [21] und AWS [22] bereitgestellten Cloud-Dienste wurden verwendet, um die Prototypen über eine Weboberfläche zu präsentieren. Das Forschungsvorhaben der vorliegenden Masterarbeit wurde bei Microsoft eingereicht und erfolgreich angenommen. Durch diese Förderung konnten der Foundershub von Microsoft [23] voll umfänglich genutzt werden, wobei auf Cloud-Computing, Sprachmodelle, Programmiersoftware und vielfältige weitere Services kostenlos zugegriffen werden konnte.

Aufgrund zahlreicher Herausforderungen bei der Implementierung durch inkompatible Architekturen und Services und zur Gewährleistung des Datenschutzes wurde im Laufe der Entwicklung entschieden DSGVO-konforme serverseitige Systeme zu implementieren, was wiederum die mit dem Entwicklungsprozess verbundene Komplexität erhöhte. Die in dieser Phase gewonnenen Erkenntnisse flossen direkt in das Design und Entwicklungsüberlegungen der finalen Prototypen ein.

4.3.5 Erstellung einer Projektseite zur Masterarbeit

Um die Masterarbeit zu dokumentieren und eine Plattform für die Tester*innen zur Verfügung zu stellen, wurde zeitgleich eine Projektseite [24] auf Basis von WordPress eingerichtet. Ein grundlegendes Designkonzept wurde erstellt, es wurden einfache Strichzeichnungen zur Visualisierung verwendet, um die Verbindung aus Technik und Therapie zu unterstreichen und ein minimalistisches, konsistentes Design zu gewährleisten (siehe Abb. 8). Sprachlich wurde auf amikalen, kollegialen Ton geachtet um die Tester*innen als Kolleg*innen anzusprechen und auf Augenhöhe einzubinden. Die Website diente als zentrale Plattform für den Zugang der Tester*innen zu den Prototypen und sollte weiters zur Darstellung der Fortschritte bei der Prototypenentwicklung und zur Dokumentation genutzt werden.

Die Website umfasst ein gesichertes Feedback-Modul, das es den Logopäd*innen ermöglichte, ihre Kommentare einzureichen, das jeweils unterhalb der Chatanwendungen implementiert wurde. WordPress wurde aufgrund seiner inhärenten Flexibilität ausgewählt, die es ermöglichte, die Website an die unterschiedlichen Anforderungen anzupassen und dennoch eine einfach zu wartende Plattform darzustellen. Die Seite wird DSGVO-konform gehostet und der interne Bereich wurde unter Nutzung von 0Auth passwortgeschützt. Ergänzend wurden noch aktuelle Links zu KI-Tools mit potenzieller therapeutischer Relevanz zur Verfügung gestellt, um einen Mehrwert für die Tester*innen generieren zu können. Im Sinne der Transparenz wurden auch häufig gestellte Fragen integriert, wobei der Fokus auf LLMs und den dargebotenen Anwendungsfällen lag.

[21] https://cloud.google.com/

[22] https://aws.amazon.com/

[23] https://foundershub.startups.microsoft.com/

[24] https://trueckl.at/

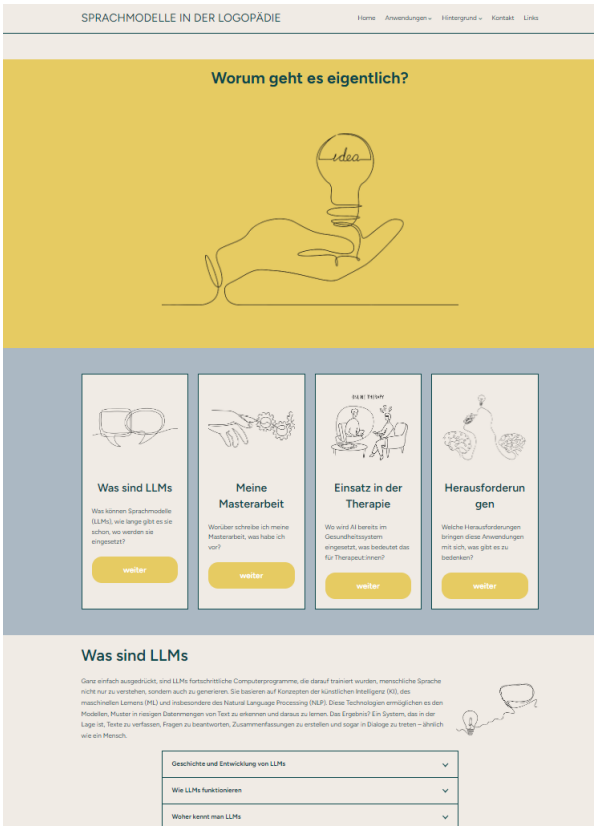

Abb. 8: Unterseite "Hintergrund" der Webseite https://trueckl.at

Im internen Bereich wurden die drei Prototypen in unterschiedlicher Darstellung angeboten, wobei BerichtGPT und Blitzstarter*in ein gewohntes Chatinterface (siehe Abb. 9) nutzten und LoCo in Form eines schwebenden Chatbuttons (siehe Abb. 10) implementiert wurde. Die Einbindung erfolgte durch ein für Wordpress entwickeltes Plugin, mit welchem die API der zugrundeliegenden Plattformen aufgerufen wurden. Die Anwendungen BerichtGPT und Blitzstarter*in wurden zur korrekten Darstellung direkt in den HTML-Code implementiert.

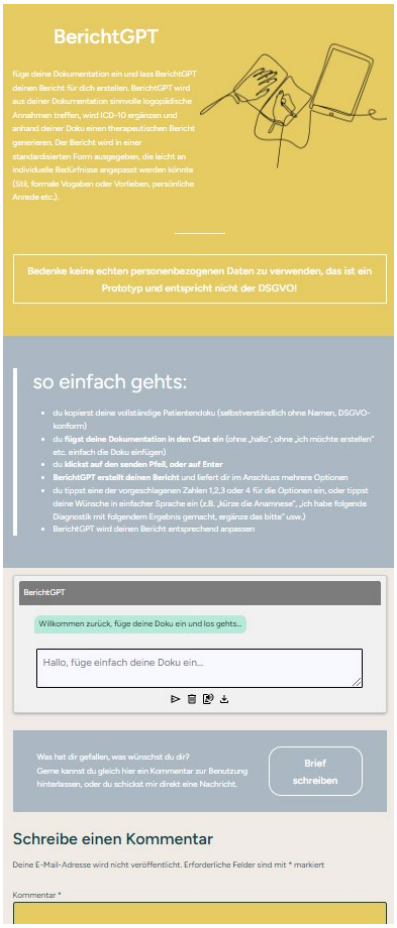

Abb. 9: Mehrzeiliges Chatinterface am Beispiel BerichtGPT

Abb. 10: Schwebender Chatbutton am Beispiel LoCo

4.4 Ergebnisse der Kommentaranalyse

Im folgenden Kapitel werden die Ergebnisse beschrieben, die sich aus der Analyse der Kommentare der Logopäd*innen in der ersten zweiwöchigen Testphase ergeben haben. Die Analyse umfasst insgesamt 22 Anmerkungen von 15 Logopäd*innen aus einer Kohorte von 19 Tester*innen, wobei einzelne Anmerkungen mehrere Aspekte adressierten. Die Kommentare wurden einer qualitativen Analyse unterzogen und systematisch in Kategorien eingeteilt, um wichtige Themen und Empfehlungen zur Verbesserung der Prototypen zu extrahieren. Ziel dieser Analyse ist es, die wesentlichen Rückmeldungen der Logopäd*innen zu erfassen und die darauf basierenden iterativen Entwicklungsentscheidungen nachvollziehbar darstellen zu können.

4.4.1 Vorgehensweise bei der Analyse

In einem ersten Schritt wurden die Kommentare zusammengefügt und anonymisiert. Anschließend wurden sie mithilfe einer qualitativen Inhaltsanalyse nach Mayring (2015) in Kategorien eingeteilt. Die Darstellung ist fünf Hauptkategorien gegliedert, die sich aus der Analyse der Kommentare ergaben: Benutzeroberfläche und Benutzerfreundlichkeit, Verfügbarkeit von Desktops und Mobilgeräten, Anpassung der Inhalte, Benutzererfahrung und Barrierefreiheit sowie technische Leistung.

4.4.2 Kommentare nach Kategorie

Benutzeroberfläche und Benutzerfreundlichkeit

Die Hälfte der Kommentare (11 Kommentare) konzentrierten sich auf die Benutzer*innenoberfläche und die allgemeine Benutzer*innenfreundlichkeit der Prototypen. Die Mehrheit der Tester*innen betonte, dass es der Bedienung der Prototypen, insbesondere der Navigation durch die Menüs, an Intuitivität mangele. Tester*in acht merkte an:

„Die Struktur des Menüs ist mir nicht klar genug. Ich musste oft lange suchen, um die gewünschte Funktion zu finden“. Eine Kollegin merkte an, dass „… eine einfachere und logischere Anordnung der Menüpunkte die Bedienbarkeit deutlich verbessern würde“ (Tester*in 5). Die Verbesserung der Menühierarchie und der Vereinfachung der Benutzeroberfläche schien prioritär zu sein in den Kommentaren.

Verfügbarkeit von Desktops und Mobilgeräten

Ein weiteres wichtiges Thema war die Nutzbarkeit der Prototypen auf dem Desktop und auf Mobilgeräten. Tester*in sieben bemerkte: „Ich würde mir für die Zukunft wünschen, dass die Chatbots auf meinem Smartphone oder Tablet funktionieren, da ich oft nicht am Computer bin“. Auch die Umständlichkeit des Wechsels zwischen mehreren Tabs und der Umstand, dass sich bei Verlassen des Fensters bei BerichtGPT das Popup Fenster wiederkehrend schloss, wurden in diesem Zusammenhang bemerkt.

Anpassungsfähigkeit der Inhalte

Acht Anmerkungen betrafen die unzureichende Anpassungsfähigkeit der Prototypen an spezifische therapeutische Anforderungen. Insbesondere im Zusammenhang mit der Generierung therapeutischer Berichte wurde kritisiert, dass der Prototypen derzeit keine ausreichenden Möglichkeiten für eine individuelle Anpassung bieten würde. Tester*in 12 formulierte: „Es wäre hilfreich, wenn ich den Inhalt (von BerichtGPT) besser an meine eigene Formatierung anpassen könnte“. Tester*in eins betonte: „Die Flexibilität der Inhalte ist für mich entscheidend, um die Prototypen im Alltag wirklich effektiv einsetzen zu

können“. Diese Beobachtungen unterstreichen, dass die personalisierte Anpassung der Inhalte eine grundlegende Voraussetzung für die Akzeptanz der Tools im logopädischen Alltag darstellen würde.

User Experience und Zugänglichkeit

Fünf Beobachtungen betrafen das allgemeine Nutzer*innenerlebnis und die Barrierefreiheit der Prototypen. Der Wunsch nach einer verbesserten Benutzererfahrung, die sowohl Online-Funktionen als auch Offline-Verfügbarkeit umfassen sollte, wurde mehrfachgeäußert. Tester*in 15 formulierte: „Eine App-Version der Prototypen, die auch offline genutzt werden kann, wäre für mich besonders wichtig, zum Beispiel bei Hausbesuchen“. Diese Bemerkung unterstreicht, wie wichtig es ist, Prototypen auf zugängliche und anpassbare Weise zu entwerfen, um den unterschiedlichen Kontexten Rechnung zu tragen, in denen solche Tools verwendet werden könnten.

Technische Leistungsfähigkeit

Zwei Beobachtungen konzentrierten sich speziell auf die technische Leistungsfähigkeit der Prototypen. Insbesondere wurde die mit der Nutzung von BerichtGPT verbundene Verarbeitungsdauer als übermäßig lang erachtet. Tester*in zwei merkte an: „Die für die Berichtserstellung gebrauchte Dauer ist gelegentlich zu lang und man weiß nicht, wie lange es dauert“. In diesem Zusammenhang wurde der Wunsch geäußert, die Geschwindigkeit der Prototypen zu erhöhen, um einen reibungsloseren Betrieb in der täglichen Berufspraxis zu ermöglichen. Ein weiterer Kommentar befasste sich mit direkt mit den generierten Outputs: „Ich bin absolut überrascht wie gut das (LoCo) funktioniert, aber was mich erstaunt hat, dass es komplexe Aufgaben wunderbar lösen kann, aber eigentlich einfache Aufgaben wie die Formulierung von Smart-Zielen auch nach mehrfachen Versuchen einfach nicht schafft.“

Auffällig in der Analyse war, dass in über zwei Drittel der Kommentare Ausdrücke der Überraschung (wow; ich bin erstaunt; das hätte ich mir nicht gedacht; beeindruckend) zum Ausdruck gebracht wurden. Dieser Eindruck der Verwunderung über die Qualität und Leistungsfähigkeit der generierten Inhalte der LLMs war in der Kohorte der Logopoäd*innen auffallend stark vertreten.

4.4.3 Schlussfolgerungen aus der Kommentaranalyse

Nach Analyse der Ergebnisse der Kommentare sollten die Prototypen einer umfassenden Überarbeitung unterzogen werden und es wurde beschlossen diese als Webapp weiterzuentwickeln. Ziel dieser Überarbeitung war die Bereitstellung einer verbesserten Benutzeroberfläche, einer optimierten Navigation und plattformübergreifender Funktionen. Spezifische Änderungen im Abruf der OpenAI-API wurden vorgenommen, um die Flexibilität der Inhalte zu erhöhen und die Antwortdauer zu verbessern, weshalb in der Abfrage ein stream Abruf integriert wurde. Die Umstellung auf eine Webanwendung ermöglichte die Nutzung sowohl über Desktop- als auch über Mobilgeräte, wodurch Anforderungen aus den Kommentaren Rechnung getragen wurde.

Die Kommentaranalyse hat ergeben, dass ein hohes Maß an Benutzerfreundlichkeit und flexible Anpassungsfähigkeit der Inhalte entscheidende Determinanten für die Akzeptanz der Prototypen wären, weshalb in die Webapp auch die Möglichkeit eigene Dateien sicher hochzuladen, Wissensdatenbanken zu erstellen und diese Inhalte selbst konfigurierbaren Assistenten zur Verfügung stellen zu können, integriert wurde.

Durch die transparente Offenlegung der entwickelten Prompts sollten diese einer individuellen Anpassung zugänglich gemacht werden und versierten Nutzer*innen die Möglichkeit eröffnen, mit geringem Aufwand und ohne spezifische Kenntnisse, individuelle Anpassungen vornehmen zu können.

Nach umfangreicher Recherche wurde LobeChat [25] für die Integration als WebApp und ein verbessertes User Interface (UI) als geeignetes Framework für die Weiterentwicklung der Anwendungsfälle identifiziert.

LobeChat ist ein Open-Source-Projekt, das als interaktive Plattform konzipiert wurde, um den Zugang zu LLMs zu erleichtern. Die Plattform kombiniert Benutzerfreundlichkeit und umfassend konfigurierbare Funktionen. LobeChat stellte somit eine benutzerzentrierte Schnittstelle bereit, die den Dialog zwischen Nutzer*innen und den Anwendungsfällen in einer intuitiven Umgebung ermöglichte.

Der Funktionsumfang von Lobe-Chat umfasst die Unterstützung unterschiedlicher kommerzieller und open-source Sprachmodelle und ermöglicht eine flexible Integration dieser. Die Plattform bietet Tools zur Erstellung von Assistenten, denen eine spezifische Rolle mittels eines entsprechenden Prompts zugewiesen werden kann. Weiters können Assistenten LLMs, Wissensdatenbanken und Plugins, wie Internetsuche, zugewiesen werden.

Die Architektur ist darauf ausgelegt, sowohl Offline- als auch Online-Modelle zu unterstützen, sodass Nutzer*innen die Kontrolle über ihre Daten behalten könnten. Aus Entwicklersicht bietet Lobe-Chat eine modulare Erweiterbarkeit, die die Einbindung zusätzlicher KI-Funktionalitäten erleichtert. Die Integration von Anmelde- und Zugriffsverwaltung, Datenspeicher, Benutzer*innenverwaltung und UI-Anpassungen ist aufgrund der quelloffenen Infrastruktur möglich. Nicht zuletzt durch den Vorteil ein vorhandenes UI (siehe Abb. 9) nutzen zu können, qualifizierte sich Lobe-Chat als Framework im Rahmen der Masterarbeit und bot auf Grundlage dieser Voraussetzungen eine solide Grundlage für die Weiterentwicklung der Chatbot-Anwendungen.

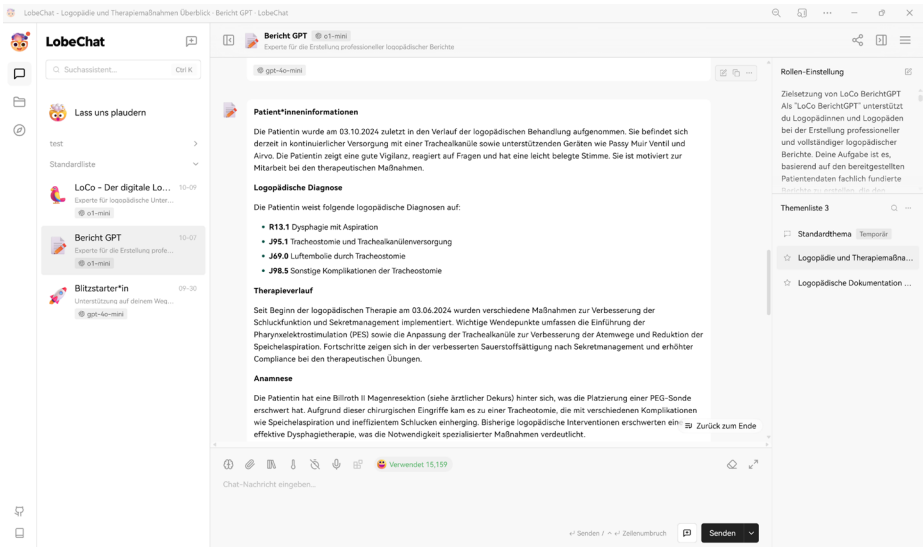

Abb. 11: LobeChat als Desktop App auf Windows

[25] github.com/lobehub/lobe-chat

4.5 Implementierung von LobeChat

Für die Implementierung von LobeChat wurde ein Server von Hetzner Cloud [26] in Frankfurt genutzt, um die Einhaltung der DSGVO sicherzustellen. Dieser Server wurde für die Datenspeicherung, das Hosting von Anwendungen und als Objektdaten-Repository für Vektorisierung verwendet. Der Server wurde mit Ubuntu in der Version 24.04.1 LTS betrieben.

Die Remote-Verbindung für die Entwicklung zum Server wurde über Visual Code und das Secure Shell Protocol (SSH) hergestellt. Hierbei wird eine sichere Verbindung über ein SSH-Schlüsselpaar gewährleistet, wobei die Funktionalität von Visual Studio Code in Folge vollumfänglich am Server genutzt werden kann.

Vektorisierung bezeichnet den Prozess, bei dem Wörter, Sätze oder ganze Texte in numerische Vektoren umgewandelt werden. Diese Vektoren repräsentieren semantische Eigenschaften und ermöglichen es LLMs, Bedeutungsähnlichkeiten und kontextuelle Zusammenhänge zu erkennen. Vektorisierung ist daher ein wichtiger Schritt, um die Antworten des Chatbots an den Nutzer*inneninhalt anzupassen und relevante Informationen schnell abzurufen.

Ein weiterer zentraler Bestandteil ist die Verwendung von Objektdatenspeichern, die die Speicherung und den schnellen Zugriff auf große Mengen an Textdaten ermöglichen. Objektdatenspeicher bieten eine skalierbare Lösung für die Verwaltung unstrukturierter Daten, die in der Kommunikation zwischen Nutzer*innen und Chatbots anfallen. Dies gewährleistet, dass der Chatbot auf vorherige Konversationen zugreifen kann, um Kontext zu bewahren und konsistente, informierte Antworten zu generieren. Die Kombination aus Vektorisierung und Objektdatenspeicherung trägt somit wesentlich zur Effizienz und Leistungsfähigkeit der entwickelten Chatbots und zur Funktionalität von LobeChat bei.

Die Bereitstellung von LobeChat wurde mit Docker Compose [27] realisiert, dies ermöglichte eine effiziente Bereitstellung und Erweiterung der Funktionalität. Der Deploymentprozess wird im Folgenden beschrieben, um Nachvollziehbarkeit und Reproduzierbarkeit zu gewährleisten.

Nach dem Klonen des Github Repositories wurde eine docker-compose.yml-Datei (im Anhang 5: LobeChat Deployment angefügt) entwickelt, um die unterschiedlichen Dienste aufzurufen und zu konfigurieren, die für LobeChat erforderlich sind. Die zugrundeliegenden Dienste wurden als Docker Container erstellt und konfiguriert. Die Docker-Compose.yml umfasste die Konfigurationen für LobeChat, einen PostgreSQL-Datenbankserver [28] und NGINX [29], das als Reverse-Proxy diente. Weiters wurde das Routing der Container untereinander mit Traefik [30] gesteuert.

Die Benutzerauthentifizierung wurde durch einen Dienst namens Logto [31] realisiert. Logto vereinfachte die Benutzer*innenverwaltung, einschließlich Anmeldeoptionen über

[26] https://hetzner.com/

[27] https://www.docker.com/

[28] https://www.postgresql.org/

[29] https://nginx.org/en/

[30] https://traefik.io/traefik/

[31] https://logto.io/

SocialLogins wie Google und weitere Authentifizierungsdienste. Logto bietet außerdem ein einfach zu bedienendes WebUI. Die benötigten Umgebungsvariablen der APIs, der Datenbanken und genutzten Services wurden in der Datei docker-compose.yml als Variablen gesetzt und im Backend in einer .env Datei definiert. SSL wurde mithilfe von Let's Encrypt[32] und Certbot[33] eingerichtet, um eine sichere Verbindung zu gewährleisten.

PostgreSQL wurde für die Speicherung von Vektordaten verwendet, da es integrierbar und als Datenbankverwaltungssystem in Kombination mit einem ergänzendem Vektorisierungspaket effizient und robust genutzt werden konnte.

Für das Hosting wurde eine dem Autor zur Verfügung stehende Domain[34] genutzt, wobei insgesamt neun Subdomains erstellt wurden. Die DNS-Einträgen der Subdomains verwiesen auf den Server von Hetzner Cloud und die entsprechenden Docker Container.

Für die Nutzung der LLMs innerhalb von LobeChat wurden APIs von OpenAI, Anthropic, Mistral[35], Perplexity und Gemini[36] (Google) erstellt und in Form der jeweils aktuellsten und leistungsfähigsten Modelle implementiert. Somit konnte sichergestellt werden, dass die geeignetsten Modelle für die Anwendungsfälle zur Verfügung genutzt werden konnten.

Zur Funktionsteuerung innerhalb von LobeChat wurden sogenannte FeatureFlags in der docker-compose.yml implementiert, mit denen beispielsweise Einstellungen gesperrt werden konnten, um gewünschte Funktionalität im Testsetting gewährleisten zu können.

4.5.1 Prompt Engineering

Einen wesentlicher Teil der Weiterentwicklung und Definition der Prototypen stellte das im Kapitel 2.1 „Theoretischer Rahmen“ bereits erwähnte Prompt Engineering dar. Diese Technik wurde eingesetzt, um die Modelle mit geeigneten Inputs anzuleiten. Der Aspekt des Prompt Engineerings war bei der Entwicklung der Chatbots besonders wichtig, da die Qualität der Antworten maßgeblich von der Formulierung der Eingabeaufforderungen beeinflusst wurde. Unterschiedliche Sprachmodelle und Zwecke setzen unterschiedlichen Aufbau und Struktur der Prompts voraus, um optimal genutzt und angesprochen werden zu können.

Die vollständigen Prompts der Anwendungsfälle finden sich im Anhang 4: Prompts.

4.5.2 Spezifische Herausforderungen während der Entwicklungsphase

Die Entwicklung der Prototypen BerichtGPT, LoCo und Blitzstarter*in war geprägt von einer intensiven Auseinandersetzung mit neuen Technologien und Plattformen. Die Integration von LLMs in Verbindung mit modernen Tools wie LobeChat, Botpress und Cloud-Diensten erleichterte die Entwicklung der Prototypen. Die größten Herausforderungen bestanden in der raschen Weiterentwicklung der Technologien, der häufig unzureichenden Dokumentation und der Notwendigkeit, auf sich laufend ändernde Bedingungen flexibel zu reagieren. Es wurden erfolgreich funktionale Prototypen entwickelt, die eine solide

[32] https://letsencrypt.org/

[33] https://certbot.eff.org/

[34] https://ilef.at/

[35] https://mistral.ai/

[36] https://gemini.google.com/

Grundlage für die Beurteilung weiterführende Beschäftigung bieten, wobei besonderer Wert auf Datenschutz und Funktionalität gelegt wurde.

4.6 Ergebnisse der Fragebogenauswertung

Im Folgenden werden die Ergebnisse der Auswertung des Fragebogens zur Verwendung der drei Prototypen (BerichtGPT, LoCo und Blitzstarter*in) dargestellt, die nach der Testphase gesammelt wurden. Insgesamt wurden 13 Fragebögen ausgefüllt und retourniert. Das Design der Fragebögen zielte darauf ab, die Erfahrungen der Tester*innen mit den Prototypen zu bewerten, wobei LobeChat als grundlegende Plattform für die Benutzung in der zweiten zweiwöchigen Testphase von 16 Tester*innen genutzt wurde. Da die Prototypen nur getestet und nicht im täglichen Alltag und in realen Bedingungen eingesetzt werden durften, basierte das Feedback der Tester*innen auf ihren Einschätzungen darüber, wie sich ein solcher Einsatz in alltäglichen Kontexten manifestieren könnte. Für die Testung wurden die Tester*innen angewiesen keine realen Patient*innendaten zu verwenden, aber sich realitätsnahe fiktive Beispiele auszudenken.

4.6.1 Auswertung der Fragebögen

Die Auswertung der Fragebögen erfolgte nach den in Kapitel 3.3 beschriebenen qualitativen Methoden. Quantitative Daten aus den geschlossenen Fragen wurden einer statistischen Beschreibung unterzogen. Im Folgenden werden die Ergebnisse nach den wesentlichen Themenfeldern strukturiert beschrieben.

Allgemeine Benutzererfahrung

Von den 13 ausgefüllten Fragebögen gaben neun Tester*innen an, sich mehrfach mit den Prototypen beschäftigt zu haben, während vier angaben, diese nur einmalig getestet zu haben. Das gesamte Nutzer*innenerlebnis wurde mit einer durchschnittlichen Bewertung von vier von fünf Punkten bewertet. Das vorherrschende Feedback bezog sich auf die Vielfalt der Funktionen, die als nützlich, aber gelegentlich als überwältigend eingestuft wurden. Tester*in vier merkte an: „Die Vielfalt der Funktionen ist gut, aber es dauert einige Zeit, bis man sich daran gewöhnt hat“.

Benutzerfreundlichkeit von Prototypen

Die Benutzeroberfläche der Prototypen erhielt eine durchschnittliche Bewertung von 3,5 von fünf Punkten. Sechs Tester*innen äußerten, dass die Navigation verbessert werden müsse, insbesondere die Anordnung der Menüelemente wurde als unzureichend intuitiv erachtet. „Ich hatte Probleme, die richtigen Optionen im Menü zu finden. Eine klarere Struktur wäre hilfreich „(Tester*in 9). Die Adaption auf das LobeChat-Framework wurde mehrfach positiv erwähnt: „Es ist gut, dass man alles in einer App hat.“ merkte Tester*in drei an.

Anpassungsfähigkeit und Flexibilität der Inhalte

Die Anpassungsfähigkeit der Inhalte wurde von den Tester*innen überwiegend als verbesserungsbedürftig beschrieben. Acht Tester*innen äußerten den Wunsch nach mehr Möglichkeiten, den Inhalt der Prototypen an die spezifischen Bedürfnisse ihrer Tätigkeit anzupassen. Tester*in acht betonte: „Die Möglichkeit, Inhalte individuell anzupassen, ist für meinen Alltag wichtig. Leider konnte ich das nicht machen“. Insgesamt wurde dieser Aspekt mit vier von fünf Punkten bewertet.

LobeChat als Plattform nutzen

Die Nutzung von LobeChat als Plattform wurde überwiegend positiv bewertet, insbesondere die Funktionalität, mit der auf die Anwendung sowohl auf Desktop- als auch auf Mobilgeräten zugegriffen werden kann. Sechs Tester*innen gaben an, die PWA-Funktionalität (Progressive WebApp) verwendet zu haben, um die Anwendung als App zu installieren. Zwei Tester*innen gaben an, dass die mobile Nutzbarkeit verbessert werden müsse, da die Darstellung auf kleineren Bildschirmen suboptimal sei. Eine Tester*in merkte an: „Die mobile Anwendung ist super, es ist großartig, dass das geht, aber die Darstellung auf meinem Smartphone war nicht optimal … da sind zu viele Menüs“.

4.6.2 Zusammenfassung der Ergebnisse der Fragebogenauswertung

Die Ergebnisse der Auswertung des Fragebogens deuten darauf hin, dass die Prototypen insgesamt positiv aufgenommen wurden, es bestünde jedoch Verbesserungspotenzial, insbesondere in den Bereichen Benutzerfreundlichkeit, inhaltliche Anpassungsfähigkeit und technische Leistung. Die Anwendung von LobeChat als Plattform wurde weitgehend positiv bewertet, wobei insbesondere die plattformübergreifende Nutzbarkeit gelobt wurde. Die gesammelten Rückmeldungen bieten wertvolle Erkenntnisse darüber, welche Änderungen erforderlich sind, um die Prototypen für die praktische Anwendung in der routinemäßigen Therapie zu optimieren. Insbesondere die Verbesserung der Benutzeroberfläche und die Erhöhung der Flexibilität der Inhalte sind entscheidende Faktoren, denen in künftigen Entwicklungen erhöhte Aufmerksamkeit geschenkt werden sollte.

5. DISKUSSION

5.1 Einleitung der Diskussion

Dieses Kapitel stellt eine kritische Reflexion und Diskussion der zentralen Ergebnisse dieser Masterarbeit dar. Wobei die Entwicklungen im Bereich von LLMs, wie sie in der vorliegenden Arbeit beschrieben wurden, im Kontext der Literatur, der qualitativ erhoben Daten und der Prototypenerstellung betrachtet werden. Ziel ist es, die Möglichkeiten und Hindernisse im Zusammenhang mit LLMs in der Logopädie zu reflektieren und aus dieser Analyse sowohl praxisorientierte als auch theoretische Konsequenzen abzuleiten. Der Diskurs orientiert sich an den Forschungsfragen und ordnet die Ergebnisse in einen breiteren wissenschaftlichen und praktischen Rahmen ein.

5.2 Beantwortung von Forschungsfragen

5.2.1 Beantwortung der Hauptforschungsfrage

Die Hauptforschungsfrage lautete: "Welche Anwendungsfälle für Large Language Models (LLMs) könnten für die logopädische Tätigkeit relevant sein?" Die Analyse ergab, dass LLMs insbesondere in den Bereichen Dokumentationserstellung, Therapiematerialgenerierung, Diagnostikunterstützung und der Unterstützung administrativer Aufgaben bedeutende Potenziale bieten würden. Dies wird sowohl durch wissenschaftliche Quellen wie Mesko et al. (2023) gestützt, wie auch durch die Auswertung des Fokusgruppeninterviews mit Logopäd*innen. Solche auf LLMs basierten Anwendungen könnten Logopäd*innen Erleichterung in deren täglicher Arbeit verschaffen, insbesondere in Bezug durch die Automatisierung repetitiver Aufgaben und Verbesserung der Therapieplanung. Die Ergebnisse stimmen mit bisherigen Studien überein, die das Potenzial von KI in der Gesundheitsversorgung unter anderem zur Effizienzsteigerung und zur Qualitätssicherung betonen (Togni et al., 2021).

Die Analyse der Fokusgruppeninterviews bestätigt diese Potenziale. Die befragten Logopäd*innen betonten, dass die Dokumentationserstellung und die Generierung von Therapiematerialien besonders vorteilhaft wären, da entsprechende zeitaufwändige Aufgaben als herausfordernd empfunden werden. Zudem sahen die Teilnehmer*innen Vorteile in der Diagnostikunterstützung, indem LLMs helfen könnten, verschiedene Informationen schneller zusammenzuführen und damit den diagnostischen Prozess effizienter zu gestalten. Diese Einschätzung wird auch von McInerney et al. (2024) geteilt, die sogar noch weitergehen und in LLM-gestützter Diagnostik eine Chance zur Fehlerreduktion sehen.

Die Verbesserung der Patientenkommunikation stellt einen weitere zentralen Anwendungsfall dar. LLMs könnten klare und ansprechende Informationen liefern, was den Dialog zwischen Therapeut*innen und Patienten optimieren und die Patient*innenautonomie im Sinne eines Informed-Consent, wie die Gesundheitskompetenz im allgemeinen, positiv fördern könnte (Fear & Gleber, 2023). Das Verständnis komplexer Themen könnte gefördert und eine effektivere Zusammenarbeit des Therapieprozesses erzielt werden, halten auch Teilnehmer*innen der Fokusgruppe fest. Darüber hinaus verbessern LLMs die interdisziplinäre Kommunikation, indem sie als Vermittler komplexer Informationen fungieren und den Erfahrungsaustausch in verschiedenen Bereichen positiv beeinflussen könnten. In solchen Szenarien könnten

LLMs die Stärken semantischer Funktionalität in Strukturierung und Ausgabe benötigter Information voll ausspielen (Gao et al., 2023).

Letztlich könnten LLMs das Praxismanagement und die Individualisierung der Patient*innenversorgung fördern, indem sie Tools für die Analyse von Praxisdaten und die strategische Planung anbieten, aber auch als kompetente Begleiter im logopädischen Alltag jederzeit verfügbar sein könnten. Entsprechende Anwendungsfälle könnten die Optimierung des therapeutischen Alltags fördern und zu fundierten Entscheidungen, unter Zuhilfenahme eins aktuellen und umfangreichen evidenzbasierten Wissens, beitragen (Meskó, 2023).

Die in der Literatur und in der Fokusgruppendiskussion identifizierten Anwendungsfälle unterstreichen das vielfältige Potenzial von LLMs in logopädischen Tätigkeiten und weisen gleichzeitig auf spezifische Herausforderungen, einschließlich Datenschutz und ethischer Überlegungen hin. Die Relevanz einzelner Anwendungen lässt sich aus derzeitiger Sicht noch nicht umfassend beantworten, da sich entsprechende Systeme noch nicht in einem regulären, täglichen logopädischen Einsatz befinden und diese somit nur implizit auf die logopädische Tätigkeit abgeschätzt werden kann.

5.2.2 Beantwortung der Unterforschungsfrage

Die Unterforschungsfrage lautete: "Wie schätzen Logopädinnen die künftige Anwendbarkeit von LLMs in der logopädischen Praxis ein?" Die Fokusgruppendiskussion mit praktizierenden Logopäd*innen ergab eine überwiegend positive Einstellung zur Integration von LLMs, gepaart mit pragmatischen Bewertungen ihrer potenziellen Vorteile und Hindernisse.

Die Teilnehmer*innen stellten fest, dass LLMs erhebliches Potenzial zur Effizienzsteigerung und zur Individualisierung von Therapien haben könnten. Hervorgehoben wurde insbesondere die Fähigkeit, administrative Aufgaben zu automatisieren und dadurch zusätzliche Zeit für die direkte Versorgung von Patient*innen bereitzustellen. Diese Umstellung könnte potenziell die Patientenzufriedenheit erhöhen und die Gesamtqualität therapeutischer Interventionen verbessern. Diese Potenziale werden auch in einer Delphi Studie von Deneke et al. (2024) beschrieben und decken sich weitgehend mit den Überlegungen der Logopäd*innen.

Dennoch äußerten die Logopäd*innen Bedenken hinsichtlich der ethischen Implikationen und Datenschutzbedenken im Zusammenhang mit der Verwendung von LLMs. Die Sensibilität von Patientendaten erfordere strenge Sicherheitsprotokolle und transparente Anforderungen an das Datenmanagement. Darüber hinaus bestehe die Sorge, dass technische Unterstützung als Ergänzung zur menschlichen Interaktion und nicht als Ersatz dienen sollte. Es sei notwendig eine Balance zwischen technologischer Effizienz und personalisierter Versorgung gewährleisten zu können, womit die Teilnehmer*innen die Bedenken von Menella et al. (2024) teilen, die in Ihrem Paper über die Herausforderungen von KI in Gesundheitssystemen nicht nur die Beziehung zwischen Patient*innen und Gesundheitspersonal kritisch betrachten, sondern auch legislative und Systemimmanente Herausforderungen im Gesundheitssystem identifizieren.

Als weitere Herausforderungen identifizierten die Teilnehmer*innen Benutzerfreundlichkeit und die Notwendigkeit, LLMs für die spezifischen Bedürfnisse der logopädischen Praxis anzupassen. Die Teilnehmer*innen der Fokusgruppe hoben explizit hervor, dass eine praxisgerechte Anpassung und intensive Schulungen entscheidend sein würden, um das Potenzial dieser Technologien auszuschöpfen und die Anwendung sicher gestalten zu

können. Eine Forderung, die auch von Maleki Varnosfaderani et al. (2024) in deren Paper zur Rolle von KI in Krankenhäusern explizit stellen. Eine umfassende Ausbildung sei essenziell, um die technischen Möglichkeiten von LLMs gezielt zu nutzen und gleichzeitig ethische sowie datenschutzrechtliche Standards einzuhalten. Es wird auch betont, dass die erfolgreiche Integration von LLMs eine Zusammenarbeit zwischen Logopäd*innen, Informatikern und Ethikern erfordere, um den verantwortungsvollen und pragmatischen Einsatz der Technologie sicherzustellen.

Zusammenfassend kann aus der Zusammenschau der Ergebnisse der Fokusgruppe und der Literaturrecherche festgehalten werden, dass großes Potential in den unterschiedlichsten Bereichen für den Einsatz von LLMs gesehen wird. Die Einschätzungen der Logopäd*innen deckt sich in weiten Teilen mit denen aus der Literatur gewonnenen Ergebnissen. Dies lässt positive Rückschlüsse auf die Gültigkeit der in der Beantwortung der Forschungsfragen getroffen Aussagen zu.

5.3 Interpretation im Kontext der bestehenden Literatur

Die Ergebnisse dieser Masterarbeit korrespondierten überwiegend mit der vorhandenen Literatur zur Anwendung von LLMs im Gesundheitswesen. Das Potenzial von LLMs, sowohl die Effizienz wie auch die Qualität der Patient*innenversorgung zu verbessern beschrieben Chen et al. bereits 2021. LLMs sollten für medizinisches Fachpersonal nützlich sein.

Insbesondere die Identifikation von Anwendungsfällen wie der Automatisierung administrativer Funktionen und der Entwicklung maßgeschneiderter therapeutischer Materialien (Brown et al., 2020) oder die Unterstützung in Diagnostik wird durch die vorliegenden Ergebnisse bestätigt.

Konträr zu diesen Ergebnisse betonen mehrere Studien die Limitationen solcher Modelle, insbesondere im Hinblick auf die fehlende Verlässlichkeit bei der Generierung klinisch relevanter Inhalte (Ji et al., 2023). Die befragten Logopäd*innen äußerten ähnliche Bedenken und befürchteten, dass LLMs noch nicht in der Lage seien, fehlerfreie und kontextspezifisch korrekte Generierung sicherzustellen. Diese Problematik spiegelt sich auch in der Unsicherheit über die genaue Herkunft der Trainingsdaten wider, die zu möglichen Verzerrungen führen könnte und eine Herausforderung für Logopäd*innen in der Anwendung darstellen könnte. Solche Verzerrungen durch Modellinhärente BIAS stellen eine bedeutende Herausforderung dar, insbesondere für die praktische Implementierung in den logopädischen Alltag, da sie die Gleichbehandlung von Patient*innen gefährden könnten, wie sie beispielsweise von Labkoff et al. (2024) beschrieben werden.

Dass solche Befürchtungen auch bezüglich der aktuellsten Modelle berechtigt sein könnte, zeigt eine aktuelle Studie von OpenAI, in welcher Tyna Eloundou et al. (2024) untersuchten wie die Antworten von Chat-GPT variieren je nachdem welche kulturellen, ethischen oder geschlechtsspezifischen Merkmale aus den im Chat verwendeten Benutzernamen abgeleitet werden können. Die Autor*innen konnten Bias aufgrund von Vornamen ableiten und identifizieren diesbezügliche Forschung als fundamental für die Entwicklung und das Training künftiger Modelle.

Die Neuartigkeit der Anwendungen bedingt eingeschränkte Verfügbarkeit einschlägiger Literatur. Dies erschwert es fundierte Aussagen zur langfristigen Nutzung und dem tatsächlichen Nutzen von LLMs zu treffen.

Die vorliegende Masterarbeit stellte die Nutzer*innenperspektive und die pragmatischen Herausforderungen, die mit der Implementierung von LLMs verbunden sein könnten, in den Vordergrund. Diese Perspektive bereichert den wissenschaftlichen Dialog um Einblicke in möglichst authentische Anwendungsszenarien und damit verbundene praktische, ethische und datenschutzrechtliche Überlegungen. Die Integration sowohl qualitativer als auch quantitativer Forschungsmethoden förderte ein umfassendes Verständnis des Themas und unterstreicht die Bedeutung eines interdisziplinären Ansatzes bei der Entwicklung und Anwendung von Technologien wie LLMs im Gesundheitswesen und insbesondere in die Domäne therapeutischer Berufe.

5.4 Praktische Implikationen

Die Ergebnisse legen praktische Auswirkungen auf die logopädische Tätigkeit nahe. Eines der wichtigsten Ergebnisse betrifft die Notwendigkeit umfassender Aus- und Weiterbildungsprogramme für Logopäd*innen. Solche Aus- und Weiterbildungsprogramme seien für Gesundheitsdienstleister*innen unerlässlich, um die sichere und effektive Nutzung von LLMs zu gewährleisten, wie Yan et al. (2024) in ihrem Scoping Review feststellen. Es sei davon auszugehen, dass künftige Entwicklungen nicht in Form von Chatbots, sondern eingebettet in komplexe Systeme, Einsatz finden werden (Mu & He, 2024).

Die Formulierung spezifischer Standards und Richtlinien sei erforderlich, um die verantwortungsvolle und ethische Anwendung von LLMs zu gewährleisten und diese in Einklang mit Berufspflichten und Rollenverständnis zu bringen. Diese Richtlinien sollten unter anderem auch Datenschutz, Datenminimierung und ethische Entscheidungsprozesse umfassen und als Maßstab für die Integration von LLMs in die logopädische Praxis dienen. Die Etablierung solcher Standards trage dazu bei, das Vertrauen der Anwender*innen zu stärken und die Akzeptanz von LLMs in der täglichen therapeutischen Praxis zu erhöhen (Piñeiro-Martín et al., 2023).

Auch die Einbindung von Patient*innen spiele in diesem Zusammenhang eine wichtige Rolle für eine erfolgreiche Implementierung von LLMs, wie unter anderem (Choudhury & Chaudhry, 2024) bekräftigen. Eine transparente Kommunikation über den Einsatz von LLMs sei für die Akzeptanz und das Vertrauen der Patienten unerlässlich, Informationskampagnen und Bildungsinitiativen können dazu beitragen, Missverständnisse zu umgehen und die Vorteile der Technologie deutlich zu artikulieren. Dies fördere eine positive Einstellung zur Implementierung entsprechender Systeme und könnte für einen notwendigen politischen Diskurs zentral sein. Weiters sollte die Einbindung von Patient*innen in die Entwicklung und Ausführung von KI-Systemen befürwortet werden, um deren Anforderungen und Perspektiven direkt einzubeziehen. Ein Aspekt der nicht nur aus der Literaturrecherche hervorging, sondern auch in der Fokusgruppendiskussion Erwähnung fand.

Zusätzlich zu diesen Empfehlungen könnte es sinnvoll sein, Ethikkommissionen einzurichten und spezifische Datenschutzrichtlinien zu formulieren, um die ethische Anwendung von LLMs sicherzustellen. Diese Initiativen könnten wie auch Denecke et al. (2024) betonen gewährleisten, dass die Nutzung von LLMs den rechtlichen und ethischen Standards entspricht, und dazu beitragen, potenzielle Konflikte und Missbrauchsfälle zu verhindern.

5.5 Beitrag der Arbeit zur Schließung der Forschungslücke

Diese Masterarbeit leistet einen Beitrag zur Schließung der Forschungslücke im Bereich der Anwendung von LLMs in der Logopädie. Durch die Identifizierung spezifischer Anwendungsfälle wurden konkrete Ansatzpunkte für die Integration von LLMs in die Praxis abgegrenzt und mit Anwendungsfällen aus der Literatur abgeglichen. Die praktische Demonstration der Realisierbarkeit, anhand entwickelter Prototypen bietet eine Grundlage für nachfolgende Forschungs- und Entwicklungsarbeiten. Insbesondere die Entwicklungsdarstellung und Bewertung von Prototypen wie BerichtGPT, LoCo und Blitzszarter*in veranschaulichen, wie LLMs in der logopädischen Praxis implementiert werden könnten und welche praktischen Vor- und Nachteile sie mit sich bringen könnten.

Die Untersuchung der Nutzer*innenperspektive durch ein Fokusgruppeninterview lieferte in diesem Zusammenhang erste Einblicke in die Erwartungen, Wünsche und Befürchtungen von Logopäd*innen. Diese Erkenntnisse könnten entscheidend für eine nutzer*innenzentrierte Entwicklung im Sinne eines Co-Designs und eine erfolgreiche Implementierung sein. Sie tragen somit zur Entwicklung des Forschungsfeldes bei.

Die Arbeit verdeutlicht, dass die Integration von LLMs nicht nur technische, sondern auch organisatorische, strukturelle und ethische Herausforderungen mit sich bringt, die gezielt adressiert werden müssen.

Ein wichtiger Beitrag dieser Arbeit ist die Identifizierung spezifischer Hindernisse bei der Implementierung von LLMs in der Logopädie, insbesondere in Bezug auf Datenschutz und ethische Überlegungen und Betonung der Notwendigkeit von Schulungen. Die umfassende Analyse dieser Hindernisse bildet die Grundlage für zukünftige Forschung.

Darüber hinaus setzt sich die Arbeit für einen interdisziplinären Ansatz ein, indem Erkenntnisse aus den Bereichen Logopädie, Informatik und Gesundheitswissenschaften zusammengeführt werden. Dies stärkt den interdisziplinären Diskurs, der für die erfolgreiche Integration von KI-Technologien im Gesundheitswesen von entscheidender Bedeutung ist.

Ein solcher Ansatz ermögliche die Formulierung maßgeschneiderter Lösungen, um den komplexen Anforderungen gesundheitsberuflicher Praxis gerecht werden zu könne unter gleichzeitiger Einhaltung ethischer und rechtlicher Standards (Khanna et al., 2022).

Letztlich regt die Arbeit zu weiterführenden Studien an, indem ungelöste Fragen und Forschungslücken identifiziert wurden, wodurch die kontinuierliche Weiterentwicklung des Forschungsfeldes vorangetrieben werden kann.

5.6 Aktuelle politische und gesundheitspolitische Entwicklungen

Im Rahmen von Überlegungen zu künftigen Anwendungen im Gesundheitsbereich sollten weiters gesellschaftliche Entwicklungen, wie die demografische Alterung der Bevölkerung, berücksichtigt werden. Diese könnte auch zu einer steigenden Nachfrage nach logopädischen Dienstleistungen führen. Laut World Health Organisation (WHO) wird der Anteil der über 60-Jährigen weltweit bis 2050 voraussichtlich 22 % erreichen, was eine erhebliche Belastung für Gesundheitssysteme bedeuten könne (Czaja et al., 2013). Diese Entwicklung stellt Gesundheitssysteme vor erhebliche Herausforderungen, insbesondere in der Versorgung älterer Menschen, die häufig von neurologischen Erkrankungen, wie Schlaganfällen oder Demenz betroffen sind (Denecke et al., 2024). Hierbei könnten LLMs dazu beitragen, effizientere Versorgungsstrukturen und Kosteneffizienz zu schaffen, indem

sie Logopäd*innen in Bereichen wie der neurologischen Rehabilitation, dem Dysphagiemanagement oder dem Kanülenmanagement unterstützen.

Darüber hinaus sind rechtlichen Rahmenbedingungen in Österreich, sowie der Europäischen Union von erheblicher Bedeutung für die Entwicklung LLM basierter Anwendungen im logopädischen Kontext. Die Einhaltung der DSGVO stellt eine grundlegende Anforderung, insbesondere im Hinblick auf den Umgang mit sensiblen Gesundheitsinformationen dar. Strenge Sicherheitsmaßnahmen und eine transparente Datenverwaltung, um den Schutz der personenbezogenen Daten der Patienten zu gewährleisten, müssen eingehalten werden. Das betrifft unter anderem regionale Beschränkungen wo Daten gespeichert und verarbeiten werden dürfen. Zusätzlich gilt für Anwendungen im Gesundheitsbereich, die diagnostische oder therapeutische Unterstützung bieten, das Medizinproduktegesetz (MPG). Dies bedeutet, dass entsprechende Systeme als Medizinprodukte eingestuft und entsprechend zertifiziert werden müssen, um die Sicherheit und Wirksamkeit der Anwendung sicherzustellen. Eine Zertifizierung würde auch eine umfassende Risikoanalyse sowie eine kontinuierliche Überwachung der Produktleistung erfordern (Meskó, 2023).

Eine kritische Analyse der aktuellen politischen und gesundheitspolitischen Rahmenbedingungen zeigt sowohl Chancen, wies auch Hindernisse für die Einbeziehung von LLMs in therapeutische Abläufe auf. Aktuelle Vorschriften bieten zwar Zusicherungen in Bezug auf den Datenschutz und die ethische Nutzung künstlicher Intelligenzsysteme, könnten aber gleichzeitig die Einführung innovativer Technologien verzögern. Am Beispiel des AI Act[37] wird deutlich, dass innovative Normen für die raschen Entwicklungen notwendig sind, diese aber nur schrittweise international übernommen werden. Der Vorteil einer verbindlichen Regelung steht in engem Zusammenhang mit der Herausforderung, dass Services und Modelle insbesondere im Feld von LLMs zeitverzögert innerhalb der EU zur Ausrollung kommen, da die geltende Rechtslage Adaptierungen bestehender Systeme notwendig macht.

Daher ist es unerlässlich, dass Gesetzgeber und Praktiker eng zusammenarbeiten, um flexible und zukunftsorientierte Normen zu formulieren, die Innovationen fördern und gleichzeitig Patient*innenrechte und das Wohlergehen dieser schützen. Fragestellungen wie die der Haftung müssten vor der Implementierung umfassend beantwortet werden um rechtliche Sicherheit für Anbieter*innen und Nutzer*innen gewährleisten zu können (Bouderhem, 2024).

5.7 Limitationen der Studie

Im folgenden Kapitel sollen die Limitationen der vorliegenden Studie adressiert und besprochen werden.

Eine wesentliche Einschränkung stellt die begrenzte Stichprobengröße des durchgeführten Fokusgruppeninterviews und die Zusammensetzung der Teilnehmer*innen dar. Angesichts des begrenzten Teilnehmer*innenpools ist die Generalisierbarkeit eingeschränkt, die Ergebnisse können nicht die Sichtweisen der österreichischen Logopäd*innen repräsentieren. Zukünftige Forschungsbemühungen, die eine größere und vielfältigere Stichprobe einbeziehen, könnten die Generierung umfassenderer und verallgemeinerbarer Ergebnisse ermöglichen.

[37] https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Weiters wurden zwar Logopäd*innen mit unterschiedlichem Fachwissen und Spezialisierung eingebunden, um zu einem differenzierteren Verständnis der Akzeptanz und der Anforderungen in unterschiedlichen Settings und Bereichen zu gelangen, die Altersstruktur der Teilnehmer*innen verzerrt jedoch das Bild der Ergebnisse. Mit einem Höchstalter von 48 Jahren handelte es sich um eine sehr junge Kohorte, dies wurde auch den Rekrutierungsprozess über Social Media bekräftigt. Es kann davon ausgegangen werden, dass die Teilnehmer*innen überwiegend eine überdurchschnittliche Technologiebereitschaft aufwiesen. Diese Einschränkungen wurden jedoch vor dem Hintergrund in Kauf genommen, dass der Fokus auf der Entwicklung neuer Anwendungen lag, welche insbesondere diese Zielgruppe stark in künftigem Handeln prägen werden.

Eine weitere Limitation stellt der Evaluierungszeitraum dar, der eine zeitliche Limitierung für die Testung der Prototypen vorsah. Eine ausgedehntere Testphase könnte zusätzliche Herausforderungen und Nutzungsmuster aufzeigen, die in dieser Studie nicht erfasst werden konnten. Aus diesem Grund könnten Langzeituntersuchungen unerlässlich sein, um die anhaltende Wirkung sowie die langfristigen Vor- und Nachteile von LLMs in der Logopädie umfassender beurteilen zu können.

Die vorliegende Forschung ist auf die Sichtweise von Logopäd*innen limitiert. Die Perspektive von Patient*innen wurde im Rahmen der vorliegenden Arbeit vernachlässigt. Die Einbeziehung der Patient*innensicht in zukünftige Studien könnte wertvolle Erkenntnisse über deren Akzeptanz und Erfahrungen mit der Anwendung von LLMs liefern.

Die technischen Einschränkungen der entwickelten Prototypen können ebenfalls als Limitierung betrachtet werden. Die Prototypen bieten eine praktische Demonstration der potenziellen Anwendungsmöglichkeiten von LLMs, sie könnten in realen Szenarien jedoch auf zusätzliche technische Hindernisse stoßen. Weiterführende Entwicklung und Verfeinerung dieser Prototypen wäre für fortführende Studien unerlässlich, um ihre Praktikabilität und Wirksamkeit beurteilen zu können und dem Fortschritt in diesem Bereich Rechnung tragen zu können und diesen abzubilden. So könnten beispielsweise Fortschritte in der Spracherkennung und -verarbeitung sowie eine erhöhte Anpassungsfähigkeit an besondere Bedürfnisse durch multimodale Systeme inkludiert werden.

Letztlich stellt die Abhängigkeit von bestehenden und kommerziell oder als open-source verfügbare technologischen Innovationen von LLMs eine zusätzliche Einschränkung dar. Die rasche Entwicklung der Technologie impliziert, dass Innovationen die Anwendbarkeit und das Potenzial von LLMs im Bereich der Logopädie weiter rasch verändern könnten. Eine kontinuierliche Überwachung des technologischen Fortschritts ist daher von entscheidender Bedeutung, um die Relevanz und Aktualität der Forschungsergebnisse sicherzustellen.

Im Rahmen der Studie war es erforderlich sich auf eine funktionierende Basis zu beschränken. Die Breite des Feldes an Möglichkeiten lässt sich zum Zeitpunkt der Erstellung der vorliegenden Arbeit schwer überblicken und es ist unmöglich im Rahmen der Masterarbeit mit der Fülle aktueller Entwicklungen mithalten zu können. Nicht zuletzt wurde auch aus diesem Grund mit LobeChat eine Plattform gewählt die Erweiterungen ermöglicht und neue Modelle und Plugins via API Aufrufe und Pakete nutzbar machen könnte.

Generell kann festgehalten werden, dass die Einschränkung durch eine Sandbox- Nutzung der Prototypen, also einer reinen Testsituation, ohne praktischen Einsatz, für den Autor

und die Tester*innen herausfordernd war und hypothetische Annahmen und Szenarien bedingte.

Diese Einschränkungen unterstreichen die Notwendigkeit, die Ergebnisse dieser Studie im Kontext eines bestimmten methodischen Rahmens und eines begrenzten Forschungsumfangs zu interpretieren. Dennoch bieten sie eine wertvolle Grundlage für weiterführende Forschungen und zeigen Wege auf, um die Integration von LLMs in die logopädische Praxis zu optimieren.

Limitationen der Literaturrecherche:

Die angewendete Suchstrategie ermöglichte eine umfassende Identifikation relevanter Literatur. Dabei wurden sowohl technologische als auch klinische Aspekte der Anwendung von LLMs in der Logopädie berücksichtigt und gefunden. Als Limitierung kann der teilweise eingeschränkte Zugang zu bestimmten Volltexten genannt werden, was die Einbeziehung potenziell relevanter Studien behindert haben könnte. Publikationsbias könnte dazu beigetragen haben, dass bevorzugt positive Ergebnisse publiziert wurden und entsprechend in die Arbeit eingeflossen sind.

Die größte Limitation liegt in der Aktualität des Forschungsgegenstandes und in der Begrenztheit verfügbarer Peer-revierter Studien, weshalb sich der Ansatz auch KI-gestützte Suche einzubeziehen als zielführend herausgestellt hat. Insbesondere durch die Einbeziehung von Preprints, konnten relevante aktuelle wissenschaftliche Erkenntnisse in die Forschung einfließen.

5.7.1 Methodische Reflexion

Der methodische Rahmen für die Prototypenentwicklung basierte auf den Grundsätzen des Nutzer*innen-zentrierten Designs. Der Schwerpunkt lag auf der Sicht der Logopäd*innen, deren Bedürfnisse und Anforderungen in allen Phasen des Entwicklungsprozesses sorgfältig berücksichtigt wurden.

Der iterative Validierungsprozess, der Feedback-Mechanismen einbezog, garantierte, dass die Prototypen kontinuierlich weiterentwickelt wurden, um den Anforderungen gerecht zu werden. Folglich konnten die Praxistauglichkeit und Akzeptanz der Prototypen erheblich verbessert werden.

Die Entscheidung, die Prototypen in einem Online-Kontext zu entwickeln und zu bewerten, beruhte auf der Prämisse, dass dieser Ansatz eine flexible und ortsunabhängige Teilnahme für Logopäd*innen ermöglichte. Die Online-Plattform ermöglichte es den Teilnehmenden, sich zeitlich und örtlich nach Belieben mit den Prototypen auseinanderzusetzen und den Einsatz in einem möglichst praxisnahen Setting auf den eigenen Endgeräten zu erleben.

Die Synthese von qualitativen und quantitativen Feedback-Ansätzen ermöglichte ein umfassendes Verständnis der Nutzer*innensicht. Die offenen Fragestellungen innerhalb des Feedback-Instruments lieferten tiefgehende Einblicke in die Erfahrungen und Perspektiven von Logopäd*innen, während die geschlossenen Fragen eine quantitative Bewertung der Prototypen ermöglichten.

5.8 Ethische und rechtliche Fragestellungen

Die Einbeziehung von LLMs in die Logopädie führt, wie bereits dargestellt, zu einer Vielzahl ethischer und rechtlicher Überlegungen. Diese Faktoren sind von grundlegender Bedeutung, um die umsichtige Anwendung von LLMs zu gewährleisten und das Vertrauen aller Stakeholder zu stärken.

Der Einsatz von LLMs erfordere ein neu definiertes Rollenverständnis innerhalb der Gesundheitsdienstleistungsberufe. Therapeut*innen müssen sich zunehmend als Vermittler zwischen menschlichen Klienten und Maschinen verstehen, was die Notwendigkeit einer kontinuierlichen Weiterbildung und der Verfeinerung beruflicher Kompetenzen impliziert. Unmittelbare Patienteninteraktion, Empathie und einen ganzheitlichen Blick auf Patient*innen könnten durch erweitertes Feedback durch Assistenzsysteme Therapeut*innen in der Rolle als Expert*innen bestärken (Choudhury & Chaudhry, 2024). Interdisziplinäre Trainingsprogramme, die sich sowohl mit den technischen als auch mit den ethischen Dimensionen der Nutzung künstlicher Intelligenz befassen, sind in diesem Zusammenhang von essentieller Bedeutung. Bildungseinrichtungen müssten ihre Lehrpläne überarbeiten, um zukünftige Gesundheitsdienstleister*innen angemessen auf die Zusammenarbeit mit KI-Systemen vorzubereiten. Dies gelte laut Choudhury et al. (2024) jedoch nicht nur für das bloße Anwenden entsprechender Systeme. Vielmehr müsste ein Vertrauensniveau erreicht werden, in welchem unterstützende Systeme auch regelmäßig hinterfragt und Aufgaben nicht in blindem Vertrauen auf neue Technologien ausgelagert werden dürften. Eine Gefahr bestünde, dass durch selbstreferenzielle Schleifen innerhalb von LLMs menschliche Erkenntnisse verwässert werden könnten, was zu einer Verschlechterung der Versorgung führen könnte. Diese Fragestellungen seien für die Autor*innen auch hinsichtlich Haftungsfragen relevant, die eine Richtigkeit der Trainingsdaten voraussetzen, zum anderen auch für gemeinsame Entscheidungsframeworks von Mensch und Maschine.

Die Beziehung zwischen Mensch und Maschine könnte durch den Einsatz von LLMs in klinischen Settings komplizierter werden (Choudhury & Chaudhry, 2024). LLMs besäßen zwar die Fähigkeit, die Arbeitsbelastung zu verringern, es sei jedoch unerlässlich, die Zuständigkeiten, die Maschinen übertragen werden, von denen, die menschliches Engagement erfordern, explizit abzugrenzen. Therapeut*innen sollten in der Lage sein, die Anwendung von LLMs auf transparente Weise zu vermitteln und den Patienten ermöglichen, diese zu verstehen. Darüber hinaus sei es von entscheidender Bedeutung, Mechanismen zur Überwachung und Korrektur maschineller Entscheidungen einzuführen, um Fehlinterpretationen zu vermeiden (Choudhury & Chaudhry, 2024).

Die Einrichtung von Ethikkommissionen zur Überwachung der Umsetzung und Anwendung von LLMs, sowie die kontinuierliche Bewertung ethischer Dilemmata, sei denkbar (Meskó, 2023). Durch die Umsetzung entsprechender Lösungsstrategien könnten die ethischen und rechtlichen Dilemmata, die mit dem Einsatz von LLMs in der Logopädie verbunden sind, wirksam entschärft werden. Dies würde dazu beitragen, das Vertrauen sowohl der Therapeut*innen als auch der Patient*innen zu stärken und die verantwortungsvolle Anwendung technologischer Innovationen sicherzustellen.

5.9 Ausblick

In dieser Masterarbeit wurden die Einsatzmöglichkeiten von LLMs in der logopädischen Praxis eingehend untersucht, um im Sinne der Forschungsfragen eine umfassende Betrachtung zu ermöglichen. Sowohl Anwendungsmöglichkeiten, wie auch damit verbundene Herausforderungen wurden hervorgehoben. Aufgrund der mehrfach beschriebenen und diskutierten raschen Entwicklung im Bereich von LLMs, deren Neuartigkeit und der damit verbundenen Limitation verfügbarer Literatur, scheint es geboten das vorliegende Unterkapitel Ausblick in die Diskussion einzuführen. Dies wird damit begründet, dass aktuelle Entwicklungen noch nicht ausreichend wissenschaftlich begründet, besprochen und analysiert werden können. In der täglichen Anwendung von

LLMs sind diese Entwicklungen und Funktionalitäten jedoch teilweise bereits verwirklicht, oder stehen unmittelbar bevor. Unter dem Anspruch der Aktualität der vorliegenden Arbeit, soll dieses Kapitel das umfassende Verständnis und Einordnung der Ergebnisse in einen aktuellen Rahmen unterstützen.

5.9.1 Technologische Weiterentwicklungen und Innovationen

Die Integration von LLMs in die logopädische Praxis könnte den Beginn einer Transformationsphase in der Gesundheitstechnologie markieren. Prospektive technologische Fortschritte dürften sowohl auf die Wirksamkeit wie auch auf die Qualität medizinischer Dienstleistungen großen Einfluss haben (Seth et al., 2023). Verbesserte Sprachverarbeitungsfunktionen, die Individualisierung von digitaler Therapieangeboten und universeller Einsatz künstlicher Assistenzsysteme stellen zentrale zukünftige Entwicklungen dar. Künftige Versionen von LLMs könnten beispielsweise immer präzisere und nuanciertere Merkmale erkennen und so in der Lage sein Therapeut*innen wertvolle Informationen und Feedback zu geben. Parameter, die derzeit schwer zu erheben sind, könnten übersichtlich dargestellt werden und die Rolle von Therapuet*innen als Expert*innen stärken. Die Erkennung von Gemütszuständen, Erschöpfungszuständen könnte durch gezielte prosodische Analyse erfolgen und findet in der Psychologie bereits prototypischen Einsatz. Durch Personalisierung und individuelle Assistenzsysteme, die Patient*innen anleiten, aufklären und unterstützen könnten effektivere Interventionen und erhöht die Erfolgsraten ermöglicht werden (Miao et al., 2023).

5.9.2 Zukünftige Entwicklungen und Potenziale von LLMs

Im Folgenden werden aktuelle Entwicklungen in den Kontext besprochener Anwendungsszenarien eingeordnet.

Personalisierte Medizin

Die Anwendung von LLMs könnte die Behandlung von Erkrankungen mit genetischen Komponenten revolutionieren, indem maßgeschneiderte Therapiepläne erstellt werden könnten, welche zu verbesserten Outcomes führen könnten (Miao et al., 2023). Herausforderungen liegen auch in der Sicherstellung eines gerechten Zugangs für alle Patient*innengruppen. Zukünftige Entwicklungen könnten die Erweiterung der personalisierten Medizin auf psychische Gesundheit und Zivilisationskrankheiten sowie die Integration von Echtzeit-Gesundheitsdaten für dynamische Therapieanpassungen umfassen.

Ein Schwerpunkt aktueller LLM-Entwicklung liegt auf der Verbesserung der Effizienz. Aktuelle Modelle erfordern erhebliche Rechenressourcen, was ihre Zugänglichkeit einschränkt. Es wird an der Entwicklung kompakterer Modelle geforscht, welche geringere Leistungseinbußen aufweisen und für spezifische Anwendungen und eingeschränkte Aufgabenbereiche ausgelegt sind (Sanh et al., 2019).

Google und weitere führende Anbieter bieten bereits Modelle an, die aufgrund ihrer Größe und ihrer Leistungsanforderungen vollständig auf modernen Smartphones als lokale Sprachmodelle für spezifische Aufgaben wie Bildbearbeitung und Textgenerierung, oder als Assistenzsysteme eingesetzt werden könnten (Alsentzer et al., 2023). Solche Systeme könnten persönliche medizinische Assistenzsysteme in der Westentasche als ständig verfügbare Begleiter ermöglichen.

Wissensdestillation, bei der kleinere Modelle darauf trainiert wird, ein größeres nachzuahmen, zeigen diesbezüglich vielversprechende Ansätze in der Entwicklung effizienterer LLMs (Gou et al., 2021).

Eine weitere Technologie, welche als Sparse Models beschrieben wurde, ermöglicht größeren Modellen eine „Mixture-of-Experts“, die nur relevante Teile von Modellen für spezifische Aufgaben aktivieren. Dieser Ansatz kann potenziell Effizienz und Leistung kompakter Modelle verbessern (Fedus et al., 2021).

KI-Hardware-Beschleuniger

Gesundheitsspezifische KI-Hardware könnte schnellere Diagnosestellungen, Therapieplanungen und Analysen ermöglichen, was die Patientenversorgung verbessert und die Echtzeit-Verarbeitung medizinischer Daten fördern könnte. Herausforderungen liegen in der Integration mit medizinischen Geräten sowie den Kosten und dem Energieverbrauch. (Togni et al., 2021)

Integration von KI mit IoT und Wearables

Weitere Verbreitung in Innovationen im Bereich von Wearables könnten in Kombination mit KI-Anwendungen zu proaktivem Gesundheitsmanagement und personalisierten Gesundheitsempfehlungen führen, wodurch Notfallinterventionen im Gesundheitswesen reduziert werden könnten. Verbesserte prädiktive Analysen zur Früherkennung von Gesundheitsanomalien und die Schaffung eines vernetzten Geräteökosystems für ganzheitliches Gesundheitsmonitoring könnten künftig feste Bestandteile eines umfassenden individuellen Gesundheitsmanagements darstellen (Mennella et al., 2024).

Bewältigung von Datenknappheit

LLMs würden die Entwicklung in wenig erforschten Bereichen, wie seltenen Erkrankungen, durch effiziente Nutzung begrenzter Datenressourcen erleichtern. Zukünftige Richtungen umfassen neuartige Datenaugmentierungstechniken und institutionsübergreifende Datenaustauschinitiativen, Hypothesengenerierung und Verknüpfung, sowie Analyse der Datensätze könnet LLM-gestützt erfolgen (Gao et al., 2023).

Vertrauensbildung durch erklärbare KI

Eine Aufmerksamkeitsvisualisierung wie sie von Vig (2019) vorgeschlagen wurde, könnte Einblicke in die Sprachverarbeitungsprozesse von LLms geben und würde dem Problem der Blackbox aktiv begegnen und ist Gegenstand aktueller intensiver Forschung.

Dass die Nachvollziehbarkeit auch für Entwickler*innen nicht immer gegeben ist, zeigt eine Hypothese aus dem Dezember 2023 [38], die besagt, dass Chat-GPT4 im Winter knappere und qualitativ minderwertigere Antworten lieferte als in den übrigen Jahreszeiten. OpenAI räumte in einem Blogeintrag ein, dass es möglich, aber kaum überprüfbar wäre, dass das erwähnte Modell aus Trainingsdaten dieses Verhalten von Menschen nachgestellt haben könnte.

LLMs und Robotik

Eine der spannendsten Perspektiven für LLMs bieten der Bereich der Robotik und die Aussicht auf Modelle, die als generelle Künstliche Intelligenz (AGI – Artificial General Intelligence) bezeichnet werden. Die Nutzung von LLMs in der Robotik hat das Potenzial,

[38] https://arstechnica.com/information-technology/2023/12/is-chatgpt-becoming-lazier-because-its-december-people-run-tests-to-find-out/

die Robotik von reiner Automatisierung hin zu einer allgemeinen verkörperten KI zu verschieben. Diese Modelle können in Bereichen wie Umgebungsmanipulation, Navigation, Planung und logischem Denken angewendet werden und es würde Computern durch Einsatz multimodaler Sprachmodelle ermöglichen mit der Umgebung in Interaktion zu treten und gezielte Manipulationen selbstständig durchzuführen (N. Dave & Kavathiya, 2023). Eine Schlüsselrolle für diese Entwicklung spiele Reinforcement learning (RL), einem Belohnungs- und Bestrafungsprozess, durch welches Systeme auf gewünschtes Verhalten trainiert werden. RL-Algorithmen wurden bereits erfolgreich in verschiedenen robotischen Plattformen angewendet (Singh et al., 2022). Die Integration von KI in die Robotik bringe jedoch weitere ethische Überlegungen mit sich, betonen Perez et al. in ihrem Whitepaper (2017), da diese Entwicklungen beispielsweise Auswirkungen auf den Arbeitsmarkt und die Demokratisierung des Zugangs zu fortschrittlicher Technologie als Fragestellung und Herausforderung mit sich bringen würde. Dong et al. (2023) thematisieren weiter mögliche psychische Implikationen, welche durch soziale Roboter, die durch emphatische Kommunikation tiefere emotionale Verbindungen zu Menschen aufzubauen imstande wären, auftreten könnten. Besonders zu beachten gäbe es in diesem Zusammenhang Auswirkungen durch Bias in den Trainingsdaten und zu starke emotionale Abhängigkeit.

5.9.3 Integration von symbolischem Reasoning und Langzeitgedächtnis

Ein weiterer Innovationsbereich ist die Integration symbolischen Reasonings und Langzeitgedächtnismechanismen in Anwendungen. Durch symbolisches Reasoning können LLMs künftig komplexere logische Schlussfolgerungen ziehen und dadurch präzisere und sinnvollere Empfehlungen generieren, in einem Versuchsstadium bieten die aktuellsten Modelle von OpenAI (GPT-o1-preview, Oktober 2024) bereits Formen einer Chain of thought an.

Fortschritte in der Entwicklung von Langzeitgedächtnissen, die eine kontinuierliche Nutzung von Kontextinformationen über mehrere Sessions hinweg ermöglichen, bieten insbesondere im therapeutischen Setting wertvolle Potenziale, da hier eine langfristige, personalisierte Betreuung entscheidend ist (Meskó, 2023). So könnten Vorlieben, Sprachstil, oder relevante Diagnosen die Tonalität der LLMs prägen, wodurch eine Kommunikation auf Augenhöhe im Sinne inklusiver Therapie bereitgestellt werden könnte.

Die Integration von Langzeitgedächtnismechanismen, einer Forderung von Dai et al. (2021), ist drei Jahre später in aktuellen Modellen (Oktober 2024) von OpenAI bereits verwirklicht, indem der Chatbot relevante Informationen der User aus Anfragen extrahiert und für weitere Chatanfragen nutzt. Ein wesentlicher Schritt auf dem Weg zur Personalisierung von Sprachmodellen.

Auch Forderungen zur Integration von symbolischem Reasoning mit neuronalen Netzen, welches die Fähigkeit von LLMs in logischem Denken verbessern könnte, wie sie beispielsweise von Garcez und Lamb (2020) beschrieben wurde, oder die Verbesserung des Ursache-Wirkung-Prinzips durch Einbeziehung von Mechanismen zur kausalen Interferenz, wie sie vor fünf Jahren bereist von Schölkopf (2019) vorgeschlagen wurde, sind in aktuellen Modellen (ChatGPT o1-preview) bereits ansatzweise verwirklicht und könnten in künftigen Anwendungen Einsatz finden.

Große Anstrengungen in der Forschung im Bereich von LLMs werden im Bereich des Datenschutzes angestrengt. Bias-Minderung durch fortgeschrittene Techniken zur Identifizierung und Minderung von Verzerrungen in Trainingsdaten und Modelloutputs wurden von Wu et al. (2020) als eine zentrale Notwendigkeit für Akzeptanz und Erfolg von

LLMs beschrieben. Methoden des Federated Learning könnten es LLMs ermöglichen, aus verteilten Datensätzen zu lernen, ohne die individuelle Privatsphäre zu gefährden.

Durch Federated Learning trainieren mehrere dezentrale Geräte gemeinsam ein Modell trainieren, ohne die lokalen Daten direkt auszutauschen. Stattdessen werden nur die gewichts- oder parameterbasierten Updates des Modells an einen zentralen Server gesendet, der die Informationen aggregiert und das globale Modell aktualisiert.

Die Schlüsselidee ist, dass die Daten lokal bleiben, was die Privatsphäre der Nutzer*innen schützt und die Notwendigkeit eliminiert, große Datenmengen an einem zentralen Ort zu speichern. Dies ist besonders nützlich in Bereichen wie Gesundheitswesen oder Mobilgeräten, wo Datenschutz eine zentrale Rolle spielt.

5.9.4 Multimodale Sprachmodelle

Es wird erwartet, dass zukünftige LLMs umfassende Fähigkeiten zur Verarbeitung und Generierung von Inhalten über sämtliche Modalitäten hinweg haben werden. Das schließt konsistente Erstellung und Eingabe von Text, Sprache, Bilder und Video ein (Ramesh et al., 2022). Mit verbesserten multimodalen Fähigkeiten würden zukünftige Entwicklungen darauf abzielen, die Fähigkeit von LLMs zu verbessern, Kontext über längere Interaktionen hinweg aufrechtzuerhalten und externes Wissen effektiver einzubeziehen, wie Beltagy et al. festhalten (2020). Entsprechende Modelle können mühelos zwischen allen Modalitäten wechseln und könnten die Nutzererfahrung und Zugänglichkeit auch im Sinne von Inklusion deutlich befördern. Solche multimodalen Systeme könnten auch nonverbale Signale wie Körpersprache, Gesichtsausdrücke und tonale Kommunikationselemente auswerten.

Ein weiterer bedeutender Fortschritt ist die Integration von Technologien der virtuellen Realität (VR) und der erweiterten Realität (AR). Diese Innovationen ermöglichen die Schaffung immersiver und interaktiver therapeutischer Umgebungen, in denen therapeutische Interventionen auch in digitaler Form über räumliche Beschränkungen hinweg anwendbar werden, wie van Dooren et al. (2020) anhand von psychologischen Interventionen darstellen. Durch die Nachahmung authentischer Kommunikationsszenarien könnten Therapeuten beispielsweise Patienten effektiver unterstützen und sie in die Lage versetzen, erworbenen Fähigkeiten praktisch anzuwenden und zu üben.

6. CONCLUSIO

Die vorliegende Masterarbeit untersuchte die potenziellen Anwendungsmöglichkeiten von Large Language Models (LLMs) in der logopädischen Tätigkeit. Durch einen Mixed-Methods-Ansatz, der systematische Literaturrecherche, Fokusgruppeninterviews und Prototypenentwicklung kombinierte, konnten umfassende Erkenntnisse über die Chancen und Herausforderungen der Integration von LLMs in die logopädische Tätigkeit gewonnen und konkrete Anwendungsfälle identifiziert werden.

Die Ergebnisse zeigen ein breites Spektrum an Anwendungsmöglichkeiten für LLMs in Kontexten wie der Logopädie. Insbesondere in den Bereichen Dokumentationserstellung, Therapiematerialgenerierung, Diagnostikunterstützung und administrative Aufgaben wurde ein signifikantes Potenzial zur Effizienzsteigerung und Qualitätsverbesserung in der Fokusgruppendiskussion identifiziert. Diese Erkenntnisse decken sich mit den Forschungsergebnissen von Alsentzer et al. (2023), die das Potenzial von LLMs zur Automatisierung klinischer Dokumentation und Unterstützung in der Entscheidungsfindung im Gesundheitswesen hervorheben.

Die Haltung der Logopäd*innen gegenüber der Integration von LLMs in ihre Praxis erwies sich als grundsätzlich positiv, was auf eine generelle Offenheit für technologische Innovationen im Berufsfeld der Logopädie hindeuten könnte. Diese Beobachtung steht im Einklang mit den Erkenntnissen von Deneke et al. (2024), die in ihrer Delphi-Studie ebenfalls eine positive Einstellung von Gesundheitsfachkräften gegenüber KI-Technologien feststellten. Allerdings wurden auch Bedenken geäußert, insbesondere hinsichtlich Datenschutz, ethischer Implikationen und der Aufrechterhaltung der therapeutischen Beziehung. So merkte eine Teilnehmer*in der Fokusgruppendiskussion an: "Es ist essenziell, dass wir die Kontrolle über die Technik behalten und nicht umgekehrt" (Kapitel 4.2.2). Diese Ambivalenz spiegelt die von Mennella et al. (2024) identifizierten Herausforderungen bei der Implementierung von KI im Gesundheitswesen wider.

Die im Rahmen dieser Arbeit entwickelten Prototypen - BerichtGPT, LoCo und Blitzstarter*in - demonstrierten die praktische Umsetzbarkeit und den potenziellen Mehrwert von LLM-basierten Anwendungen in der logopädischen Praxis. Aus der Analyse der Beurteilung der Prototypen durch Logopäd*innen lässt sich das Potenzial dieser Technologien ableiten, ähnlich wie es Miao et al. (2023) für den breiteren Bereich der digitalen Gesundheit ebenso beschrieben haben.

Die Implementierung von LLMs in die logopädische Praxis birgt jedoch auch Herausforderungen, die einer sorgfältigen Berücksichtigung bedürfen. Neben technischen Aspekten sind insbesondere ethische und rechtliche Fragestellungen von zentraler Bedeutung. Diese Beobachtung deckt sich mit den Erkenntnissen von Bouderhem (2024), der die Notwendigkeit robuster ethischer und regulatorischer Rahmenbedingungen für den Einsatz von KI im Gesundheitswesen betont.

Die Ergebnisse dieser Studie werfen neue Forschungsfragen auf und zeigen Bereiche, die weiterer Untersuchung bedürfen. Langzeitstudien zur Wirksamkeit sind erforderlich, um die langfristigen Auswirkungen des Einsatzes von LLMs auf die Qualität der logopädischen Versorgung und die Nutzer*innenerfahrung zu evaluieren. Die Entwicklung spezifischer ethischer Richtlinien für den Einsatz von LLMs in der Logopädie sollte vorangetrieben werden, um einen verantwortungsvollen und patientenzentrierten Einsatz sicherzustellen,

wie er auch von Piñeiro-Martín et al. (2023) für den breiteren Kontext der KI im Gesundheitswesen gefordert wird.

Die Integration von LLMs in die logopädische Praxis steht noch am Anfang, zeigt jedoch vielversprechendes Potenzial zur Verbesserung der Patient*innenversorgung und Unterstützung der therapeutischen Arbeit. Die Herausforderungen, die mit dieser technologischen Innovation einhergehen, erfordern einen interdisziplinären Ansatz und eine sorgfältige Abwägung ethischer, rechtlicher und praktischer Aspekte. Diese Notwendigkeit wird auch von Maleki Varnosfaderani und Forouzanfar (2024) in ihrer Untersuchung zur Rolle von KI in Krankenhäusern und Kliniken hervorgehoben.

Abschließend lässt sich festhalten, dass die Integration von LLMs in die logopädische Praxis ein vielversprechendes Feld mit großem Potenzial darstellt. Die Herausforderungen sind beträchtlich, aber nicht unüberwindbar. Mit einem umsichtigen, ethisch fundierten und patientenzentrierten Ansatz kann die Logopädie von den Vorteilen dieser Technologie profitieren und gleichzeitig die Qualität und Menschlichkeit der therapeutischen Beziehung bewahren. Wie Meskó (2023) betont, wird die Zukunft des Gesundheitswesens maßgeblich davon abhängen, wie gut es gelingt, diese Balance zwischen technologischer Innovation und menschenzentrierter Versorgung zu finden und kontinuierlich zu optimieren. In der vorliegenden Arbeit wurde versucht diesen nutzer*innenzentrierten Ansatz in Form von Co-Design zu verfolgen, um eine möglichst breite Grundlage für die Beantwortung der Forschungsfragen, unter interdisziplinärer Betrachtung gewährleisten zu können.

Diese Arbeit legt somit den Grundstein für weitere Forschungen in diesem Bereich und adressiert einen Appell an Gesundheitsdienstleister*innen und Therapeut*innen für eine aktive Auseinandersetzung mit dem Thema. Diese Beschäftigung könnte essenziell sein und sicherstellen, dass Therapeut*innen deren professionelle Rolle proaktiv weiterentwickeln und im Interesse der Patient*innen die bestmögliche Versorgung gewährleisten, indem sie Entwicklungsprozesse aktiv mitgestalten.

Wenn, wie erste Forschungsarbeiten aus vergleichbaren medizinischen Bereichen beschreiben (Miao et al., 2023), digital gestützte Interventionen ein vergleichbares, oder sogar besseres Outcome, als traditionelle Interventionen aufweisen, könnte sich künftig nicht die Frage stellen, ob diese Systeme zum Einsatz kommen, sondern warum nicht. Im Sinne einer optimalen Patient*innenversorgung könnte daher ein Zusammenschluss von Therapeut*innen und Maschine erforderlich sein und diesen gilt es im Sinne unserer Patient*innen bestmöglich zu gestalten.

6.1 Literaturverzeichnis

Al-Dujaily, M. & Nabàa, S. J. A. (2015). Age and Sex Distribution of Speech Disorders. *International Journal of Phonosurgery & Laryngology*, *5*(1), 7–14. https://doi.org/10.5005/JP-JOURNALS-10023-1094

Alsentzer, E., Rasmussen, M. J., Fontoura, R., Cull, A. L., Beaulieu-Jones, B., Gray, K. J., Bates, D. W. & Kovacheva, V. P. (2023). Zero-shot interpretable phenotyping of postpartum hemorrhage using large language models. *NPJ digital medicine*, *6*(1), 212. https://doi.org/10.1038/s41746-023-00957-x

Aronson, E., Wilson, T. D., Sommers, S. R. & Page-Gould, E. (2023). *Social psychology* (Eleventh Edition). Pearson.

Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., Hogarth, M. & Smith, D. M. (2023). Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA internal medicine*, *183*(6), 589–596. https://doi.org/10.1001/jamainternmed.2023.1838

Beltagy, I., Peters, M. E. & Cohan, A. (2020, 10. April). *Longformer: The Long-Document Transformer*. http://arxiv.org/pdf/2004.05150

Bender, E. M., Gebru, T., McMillan-Major, A. & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots. In *ACM Digital Library, Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (S. 610–623). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Berrezueta-Guzman, S., Kandil, M., Martín-Ruiz, M.-L., Pau-de-la-Cruz, I. & Krusche, S. (2024, 21. Juni). *Exploring the Efficacy of Robotic Assistants with ChatGPT and Claude in Enhancing ADHD Therapy: Innovating Treatment Paradigms*. http://arxiv.org/pdf/2406.15198v1

Beukelman, D. R., Fager, S., Ball, L. & Dietz, A. (2009). AAC for adults with acquired neurological conditions: a review. *Augmentative and alternative communication (Baltimore, Md. : 1985)*, *23*(3), 230–242. https://doi.org/10.1080/07434610701553668

Boecking, B., Usuyama, N., Bannur, S., Coelho de Castro, D., Schwaighofer, A., Hyland, S., Wetscherek, M. T., Naumann, T., Nori, A., Alvarez Valle, J., Poon, H. & Oktay, O. (2022). *MS-CXR: Making the Most of Text Semantics to Improve Biomedical Vision-Language Processing.* https://doi.org/10.13026/B90J-VB87

The Book of Why: The New Science of Cause and Effect The Book of Why: The New Science of Cause and EffectJudea Pearl and Dana Mackenzie Basic Books, 2018. 429 pp (2018). *Science*, *361*(6405), 855. https://doi.org/10.1126/science.aau9731

Bouderhem, R. (2024). Shaping the future of AI in healthcare through ethics and governance. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-02894-w

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J [Jeffrey], Winter, C., . . . Amodei, D. (2020, 28. Mai). *Language Models are Few-Shot Learners*. http://arxiv.org/pdf/2005.14165

Callahan, A., Fries, J. A., Ré, C., Huddleston, J. I., Giori, N. J., Delp, S. & Shah, N. H. (2019). Medical device surveillance with electronic health records. *NPJ digital medicine*, *2*, 94. https://doi.org/10.1038/s41746-019-0168-z

Carlini, N., Tramer, F., Wallace, E., Jagielski, M., Herbert-Voss, A., Lee, K [Katherine], Roberts, A., Brown, T., Song, D., Erlingsson, U., Oprea, A. & Raffel, C. (2020, 14. Dezember). *Extracting Training Data from Large Language Models*. http://arxiv.org/pdf/2012.07805

Chen, C. & Stadler, T. (2023). GenSpectrum Chat: Data Exploration in Public Health Using Large Language Models. https://doi.org/10.48550/arXiv.2305.13821

Chen, H., Fang, Z., Singla, Y. & Dredze, M. (2024, 28. Februar). *Benchmarking Large Language Models on Answering and Explaining Challenging Medical Questions*. http://arxiv.org/pdf/2402.18060

Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, Henrique Ponde de Oliveira, Kaplan, J., Edwards, H., Burda, Y., Joseph, N., Brockman, G., Ray, A., Puri, R., Krueger, G., Petrov, M., Khlaaf, H., Sastry, G., Mishkin, P., Chan, B., Gray, S., . . . Zaremba, W. (2021, 7. Juli). *Evaluating Large Language Models Trained on Code*. http://arxiv.org/pdf/2107.03374

Cho, Y., Kim, M., Kim, S., Kwon, O., Kwon, R. D., Lee, Y. & Lim, D. (2023, 12. November). Evaluating the Efficacy of Interactive Language Therapy Based on LLM for High-Functioning Autistic Adolescent Psychological Counseling. http://arxiv.org/pdf/2311.09243

Choudhury, A. & Chaudhry, Z. (2024). Large Language Models and User Trust: Consequence of Self-Referential Learning Loop and the Deskilling of Health Care Professionals. *Journal of Medical Internet Research*, *26*, e56764. https://doi.org/10.2196/56764

Czaja, S., Beach, S., Charness, N. & Schulz, R. (2013). Older Adults and the Adoption of Healthcare Technology: Opportunities and Challenges. In A. Sixsmith & G. Gutman (Hrsg.), *Technologies for Active Aging* (S. 27–46). Springer US. https://doi.org/10.1007/978-1-4419-8348-0_3

Dai, D., Dong, L., Hao, Y., Sui, Z., Chang, B. & Wei, F. (2021, 18. April). *Knowledge Neurons in Pretrained Transformers*. http://arxiv.org/pdf/2104.08696

Danilevsky, M., Qian, K., Aharonov, R., Katsis, Y., Kawas, B. & Sen, P. (2020). A Survey of the State of Explainable AI for Natural Language Processing. *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*. http://arxiv.org/pdf/2010.00711

Dave, N. & Kavathiya, H. (2023). June 2023. *Journal of Information Technology and Digital World*, *5*(2), 193–209. https://doi.org/10.36548/jitdw.2023.2.009

Dave, T., Athaluri, S. A. & Singh, S. (2023). ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Frontiers in artificial intelligence*, *6*, 1169595. https://doi.org/10.3389/frai.2023.1169595

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319. https://doi.org/10.2307/249008

Denecke, K., May, R. & Rivera Romero, O. (2024). Potential of Large Language Models in Health Care: Delphi Study. *Journal of Medical Internet Research*, *26*, e52399. https://doi.org/10.2196/52399

Devlin, J., Chang, M.-W., Lee, K [Kenton] & Toutanova, K. (2018, 11. Oktober). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. http://arxiv.org/pdf/1810.04805

Dong, Y [Yanling] & Zhou, X. (2023). Advancements in AI-driven multilingual comprehension for social robot interactions: An extensive review. *Electronic Research Archive*, *31*(11), 6600–6633. https://doi.org/10.3934/era.2023334

Englhardt, Z., Ma, C., Morris, M. E., Xu, X. "., Chang, C.-C., Qin, L., McDuff, D., Liu, X [Xin], Patel, S. & Iyer, V. (2023). From Classification to Clinical Insights: Towards Analyzing and Reasoning About Mobile and Behavioral Health Data With Large Language Models. Vorab-Onlinepublikation. https://doi.org/10.48550/arXiv.2311.13063

VERORDNUNG (EU) 2016/679 DES EUROPÄISCHEN PARLAMENTS UND DES RATES (2016).

Artificial Intelligence Act (2024). https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Fear, K. & Gleber, C. (2023). Shaping the Future of Older Adult Care: ChatGPT, Advanced AI, and the Transformation of Clinical Practice. *JMIR Aging*, *6*, e51776. https://doi.org/10.2196/51776

Fedus, W., Zoph, B. & Shazeer, N. (2021, 11. Januar). Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity. http://arxiv.org/pdf/2101.03961

Gao, Y., Baptista-Hon, D. T. & Zhang, K [Kang] (2023). The inevitable transformation of medicine and research by large language models: The possibilities and pitfalls. *MedComm – Future Medicine*, *2*(2), Artikel e49. https://doi.org/10.1002/mef2.49

Garcez, A. d. & Lamb, L. C. (2020, 10. Dezember). *Neurosymbolic AI: The 3rd Wave*. http://arxiv.org/pdf/2012.05876

Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A. & Chartash, D. (2023). How Does ChatGPT Perform on the United States Medical Licensing Examination (USMLE)? The Implications of Large Language Models for Medical Education and Knowledge Assessment. *JMIR medical education*, *9*, e45312. https://doi.org/10.2196/45312

Gu, Y., Tinn, R., Cheng, H., Lucas, M., Usuyama, N., Liu, X [Xiaodong], Naumann, T., Gao, J. & Poon, H. (2022). Domain-Specific Language Model Pretraining for Biomedical Natural Language Processing. *ACM Transactions on Computing for Healthcare*, *3*(1), 1–23. https://doi.org/10.1145/3458754

Guevara, M., Chen, S., Thomas, S., Chaunzwa, T. L., Franco, I., Kann, B. H., Moningi, S., Qian, J. M., Goldstein, M., Harper, S., Aerts, H. J. W. L., Catalano, P. J., Savova, G. K., Mak, R. H. & Bitterman, D. S. (2024). Large language models to identify social determinants of health in electronic health records. *NPJ digital medicine*, *7*(1), 6. https://doi.org/10.1038/s41746-023-00970-0

Guthoff, R. F. & Stachs, O. (2019). Künstliche Intelligenz in der Augenheilkunde [Artificial Intelligence in Ophthalmology]. *Klinische Monatsblätter für Augenheilkunde*, *236*(12), 1397–1398. https://doi.org/10.1055/a-1022-3851

Heaton, J. (2018). Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning. *Genetic Programming and Evolvable Machines*, *19*(1-2), 305–307. https://doi.org/10.1007/s10710-017-9314-z

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A. & Fung, P. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, *55*(12), 1–38. https://doi.org/10.1145/3571730

Jiang, F., Jiang, Y [Yong], Zhi, H., Dong, Y [Yi], Li, H [Hao], Ma, S., Wang, Y [Yilong], Dong, Q., Shen, H. & Wang, Y [Yongjun] (2017). Artificial intelligence in

healthcare: past, present and future. *Stroke and vascular neurology*, *2*(4), 230–243. https://doi.org/10.1136/svn-2017-000101

Jiang, Y [Yan], Qiu, R., Zhang, Y [Yi] & Zhang, P.-F. (2023). *Balanced and Explainable Social Media Analysis for Public Health with Large Language Models.* https://doi.org/10.48550/arXiv.2309.05951

Jin, Q., Wang, Z., Floudas, C. S., Chen, F., Gong, C., Bracken-Clarke, D., Xue, E., Yang, Y., Sun, J. & Lu, Z. (2023, 27. Juli). *Matching Patients to Clinical Trials with Large Language Models*. http://arxiv.org/pdf/2307.15051

Jurafsky, D. & Martin, J. H. (2009). Prentice Hall series in artificial intelligence. Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition (P. Norvig & S. J. Russell, Hg.). Prentice Hall Pearson Education International.

Khanna, N. N., Maindarkar, M. A., Viswanathan, V., Fernandes, J. F. E., Paul, S., Bhagawati, M., Ahluwalia, P., Ruzsa, Z., Sharma, A [Aditya], Kolluri, R., Singh, I. M., Laird, J. R., Fatemi, M., Alizad, A., Saba, L., Agarwal, V., Sharma, A [Aman], Teji, J. S., Al-Maini, M., . . . Suri, J. S. (2022). Economics of Artificial Intelligence in Healthcare: Diagnosis vs. Treatment. *Healthcare (Basel, Switzerland)*, *10*(12). https://doi.org/10.3390/healthcare10122493

Labkoff, S., Oladimeji, B., Kannry, J., Solomonides, A., Leftwich, R., Koski, E., Joseph, A. L., Lopez-Gonzalez, M., Fleisher, L. A., Nolen, K., Dutta, S., Levy, D. R., Price, A., Barr, P. J., Hron, J. D., Lin, B., Srivastava, G., Pastor, N., Luque, U. S., . . . Quintana, Y. (2024). Toward a responsible future: recommendations for AI-enabled clinical decision support. *Journal of the American Medical Informatics Association : JAMIA.* Vorab-Onlinepublikation. https://doi.org/10.1093/jamia/ocae209

Law, J., Dornstauder, M., Charlton, J. & Gréaux, M. (2021). Tele-practice for children and young people with communication disabilities: Employing the COM-B model to review the intervention literature and inform guidance for practitioners. *International Journal of Language & Communication Disorders*, *56*(2), 415–434. https://doi.org/10.1111/1460-6984.12592

LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Li, Y [Yunxiang], Li, Z [Zihan], Zhang, K [Kai], Dan, R., Jiang, S. & Zhang, Y [You] (2023). ChatDoctor: A Medical Chat Model Fine-Tuned on a Large Language Model Meta-AI (LLaMA) Using Medical Domain Knowledge. *Cureus*, *15*(6), e40895. https://doi.org/10.7759/cureus.40895

Liu, S., McCoy, A. B., Wright, A. P., Carew, B., Genkins, J. Z., Huang, S. S., Peterson, J. F., Steitz, B. & Wright, A. (2023). Leveraging Large Language Models for Generating Responses to Patient Messages. *medRxiv : the preprint server for health sciences.* Vorab-Onlinepublikation. https://doi.org/10.1101/2023.07.14.23292669

Liu, Z [Zhichao], Roberts, R. A., Lal-Nag, M., Chen, X., Huang, R. & Tong, W. (2021). AI-based language models powering drug discovery and development. *Drug Discovery Today*, *26*(11), 2593–2607. https://doi.org/10.1016/j.drudis.2021.06.009

Luo, R., Sun, L., Xia, Y., Qin, T., Zhang, S., Poon, H. & Liu, T.-Y. (2022). BioGPT: generative pre-trained transformer for biomedical text generation and mining. *Briefings in Bioinformatics*, *23*(6). https://doi.org/10.1093/bib/bbac409

Maleki Varnosfaderani, S. & Forouzanfar, M. (2024). The Role of AI in Hospitals and Clinics: Transforming Healthcare in the 21st Century. *Bioengineering (Basel, Switzerland)*, *11*(4). https://doi.org/10.3390/bioengineering11040337

Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12., überarbeitete Auflage). Beltz. http://www.beltz.de/fileadmin/beltz/leseproben/978-3-407-25730-7.pdf

McInerney, D. J., Dickinson, W., Flynn, L. C., Young, A. C., Young, G. S., van de Meent, J.-W. & Wallace, B. C. (2024, 15. Februar). *Towards Reducing Diagnostic Errors with Interpretable Risk Prediction*. http://arxiv.org/pdf/2402.10109

Mennella, C., Maniscalco, U., Pietro, G. de & Esposito, M. (2024). Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. *Heliyon*, *10*(4), e26297. https://doi.org/10.1016/j.heliyon.2024.e26297

Meskó, B. (2023). The Impact of Multimodal Large Language Models on Health Care's Future. *Journal of Medical Internet Research*, *25*, e52865. https://doi.org/10.2196/52865

Miao, H., Li, C. & Wang, J [Jing] (2023). A Future of Smarter Digital Health Empowered by Generative Pretrained Transformer. *Journal of Medical Internet Research*, *25*, e49963. https://doi.org/10.2196/49963

Mishra, V., Jafri, F., Abdul Kareem, N., Aboobacker, R. & Noora, F. (2024). Evaluation of accuracy and potential harm of ChatGPT in medical nutrition therapy - a case-based approach. *F1000Research*, *13*, 137. https://doi.org/10.12688/f1000research.142428.1

Mitchell, T. M. (2013). Machine learning [Nachdr.]. McGraw-Hill international editions computer science series. McGraw-Hill.

Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mu, Y. & He, D. (2024). The Potential Applications and Challenges of ChatGPT in the Medical Field. *International journal of general medicine*, *17*, 817–826. https://doi.org/10.2147/IJGM.S456659

Naqvi, W. M., Shaikh, S. Z. & Mishra, G. V. (2024). Large language models in physical therapy: time to adapt and adept. *Frontiers in public health*, *12*, 1364660. https://doi.org/10.3389/fpubh.2024.1364660

Oka, R., Nomura, A., Yasugi, A., Kometani, M., Gondoh, Y., Yoshimura, K. & Yoneda, T. (2019). Study Protocol for the Effects of Artificial Intelligence (AI)-Supported Automated Nutritional Intervention on Glycemic Control in Patients with Type 2 Diabetes Mellitus. *Diabetes Therapy*, *10*(3), 1151–1161. https://doi.org/10.1007/s13300-019-0595-5

MTD-Gesetz 2024 (2024 & i.d.F.v. 01.09.2024). https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20012653

Patterson, D., Gonzalez, J., Le Quoc, Liang, C., Munguia, L.-M., Rothchild, D., So, D., Texier, M. & Dean, J. (2021, 21. April). *Carbon Emissions and Large Neural Network Training*. http://arxiv.org/pdf/2104.10350

Patwardhan, N., Marrone, S. & Sansone, C. (2023). Transformers in the Real World: A Survey on NLP Applications. *Information*, *14*(4), 242. https://doi.org/10.3390/info14040242

Perez, J. A., Deligianni, F., Ravi, D. & Yang, G.-Z. (2017). *UKRAS White Papers* (UKRAS White Papers 2398-4422). EPSRC UK-RAS Network. https://doi.org/10.31256/WP2017.1

Piñeiro-Martín, A., Garci-a-Mateo, C., Docío-Fernández, L. & Del López Pérez, M. C. (2023). *Ethical Challenges in the Development of Virtual Assistants Powered by Large Language Models.* https://doi.org/10.20944/preprints202306.0196.v1

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners.

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C. & Chen, M. (2022, 13. April). *Hierarchical Text-Conditional Image Generation with CLIP Latents*. http://arxiv.org/pdf/2204.06125

Russell, S. J. & Norvig, P. (2012). *Künstliche Intelligenz: Ein moderner Ansatz* (3., aktualisierte Auflage). *Pearson Studium - IT Ser*. Pearson. https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5583900

Safranek, C. W., Sidamon-Eristoff, A. E., Gilson, A. & Chartash, D. (2023). The Role of Large Language Models in Medical Education: Applications and Implications. *JMIR medical education*, *9*, e50945. https://doi.org/10.2196/50945

Sanh, V., Debut, L., Chaumond, J. & Wolf, T. (2019, 2. Oktober). *DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter*. http://arxiv.org/pdf/1910.01108

Schölkopf, B. (2019). Causality for Machine Learning. *pp 765-804 in H. Geffner et al. (eds): Probabilistic and Causal Inference: The Works of Judea Pearl, ACM*, *27*, 765–804. https://doi.org/10.1145/3501714.3501755

Seth, P., Pomedli, S., Wit, M. de & Mamdani, M. (2023). ChatGPT and beyond: Considerations in the use of Large Language Models (LLMs) in clinical practice (Preprint). https://doi.org/10.2196/preprints.47490

Singh, B., Kumar, R. & Singh, V. P. (2022). Reinforcement learning in robotic applications: a comprehensive survey. *Artificial Intelligence Review*, *55*(2), 945–990. https://doi.org/10.1007/s10462-021-09997-9

Suvirat, K., Tanasanchonnakul, D., Chairat, S. & Chaichulee, S. (2023). Leveraging Language Models for Inpatient Diagnosis Coding. *Applied Sciences*, *13*(16), 9450. https://doi.org/10.3390/app13169450

Tailor, P. D., D'Souza, H. S., Li, H [Hanzhou] & Starr, M. R. (2024). Vision of the future: large language models in ophthalmology. *Current opinion in ophthalmology*, *35*(5), 391–402. https://doi.org/10.1097/ICU.0000000000001062

Tan, R. S. Y. C., Lin, Q., Low, G. H., Lin, R., Goh, T. C., Chang, C. C. E., Lee, F. F., Chan, W. Y., Tan, W. C., Tey, H. J., Leong, F. L., Tan, H. Q., Nei, W. L., Chay, W. Y., Tai, D. W. M., Lai, G. G. Y., Cheng, L. T.-E., Wong, F. Y., Chua, M. C. H., . . . Ng, H. T. (2023). Inferring cancer disease response from radiology reports using large language models with data augmentation and prompting. *Journal of the American Medical Informatics Association : JAMIA*, *30*(10), 1657–1664. https://doi.org/10.1093/jamia/ocad133

Togni, G. de, Erikainen, S., Chan, S. & Cunningham-Burley, S. (2021). What makes AI 'intelligent' and 'caring'? Exploring affect and relationality across three sites of intelligence and care. *Social science & medicine (1982)*, *277*, 113874. https://doi.org/10.1016/j.socscimed.2021.113874

Tyna Eloundou, Alex Beutel, David G. Robinson, keren-Gu-Lemberg. (2024, 15. Oktober). *First-Person Fairness in Chatbots.*

van Dooren, M. M. M., Visch, V., Spijkerman, R., Goossens, R. H. M. & Hendriks, V. M. (2020). Mental Health Therapy Protocols and eHealth Design: Focus Group Study. *JMIR formative research*, *4*(5), e15568. https://doi.org/10.2196/15568

Vandewinckele, L., Claessens, M., Dinkla, A., Brouwer, C., Crijns, W., Verellen, D. & van Elmpt, W. (2020). Overview of artificial intelligence-based applications in radiotherapy: Recommendations for implementation and quality assurance. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, *153*, 55–66. https://doi.org/10.1016/j.radonc.2020.09.008

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017, 12. Juni). *Attention Is All You Need*. http://arxiv.org/pdf/1706.03762

Venkatesh, Morris & Davis (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425. https://doi.org/10.2307/30036540

Vig, J. (2019, 12. Juni). A Multiscale Visualization of Attention in the Transformer Model. http://arxiv.org/pdf/1906.05714

Wang, G., Yang, G., Du, Z., Fan, L. & Li, X [Xiaohu]. (2023). *ClinicalGPT: Large Language Models Finetuned with Diverse Medical Data and Comprehensive Evaluation.* https://doi.org/10.48550/arXiv.2306.09968

Wang, L. L., Lo Kyle, Chandrasekhar, Y., Reas, R., Yang, J., Burdick, D., Eide, D., Funk, K., Katsis, Y., Kinney, R., Li, Y [Yunyao], Liu, Z [Ziyang], Merrill, W., Mooney, P., Murdick, D., Rishi, D., Sheehan, J., Shen, Z., Stilson, B., . . . Kohlmeier, S. (2020, 22. April). *CORD-19: The COVID-19 Open Research Dataset*. http://arxiv.org/pdf/2004.10706

Wankmüller, S. (2022). Introduction to Neural Transfer Learning With Transformers for Social Science Text Analysis. *Sociological Methods & Research*, 004912412211345. https://doi.org/10.1177/00491241221134527

Wu, X [Xing], Liang, Z. & Wang, J [Jianjia] (2020). FedMed: A Federated Learning Framework for Language Modeling. *Sensors*, *20*(14), 4048. https://doi.org/10.3390/s20144048

Xue, L., Constant, N., Roberts, A., Kale, M., Al-Rfou, R., Siddhant, A., Barua, A. & Raffel, C. (2020, 22. Oktober). *mT5: A massively multilingual pre-trained text-to-text transformer*. http://arxiv.org/pdf/2010.11934

an Yan, McAuley, J., Lu, X., Du, J., Chang, E. Y., Gentili, A. & Hsu, C.-N. (2022). RadBERT: Adapting Transformer-based Language Models to Radiology. *Radiology. Artificial intelligence*, *4*(4), e210258. https://doi.org/10.1148/ryai.210258

Yan, L., Sha, L., Zhao, L., Li, Y [Yuheng], Martinez-Maldonado, R., Chen, G., Li, X [Xinyu], Jin, Y. & Gašević, D. (2024). Practical and Ethical Challenges of Large Language Models in Education: A Systematic Scoping Review. *British Journal of Educational Technology*, *55*(1), 90–112. https://doi.org/10.1111/bjet.13370

Zernikow, J., Grassow, L., Gröschel, J., Henrion, P., Wetzel, P. J. & Spethmann, S. (2023). Anwendung von „large language models“ in der Klinik : Ersetzt ChatGPT die Arztbrieferstellung? Ein Erfahrungsbericht [Clinical application of large language models : Does ChatGPT replace medical report formulation? An

experience report]. *Innere Medizin (Heidelberg, Germany)*, *64*(11), 1058–1064. https://doi.org/10.1007/s00108-023-01600-3

Zhang, Z., Yao, Z., Zhou, H [Huixue], ouyang, F. & Yu, H. (2023, 30. Oktober). *EHRTutor: Enhancing Patient Understanding of Discharge Instructions*. http://arxiv.org/pdf/2310.19212

Zhou, H [Hongjian], Liu, F., Gu, B., Zou, X., Huang, J., Wu, J [Jinge], Li, Y [Yiru], Chen, S. S., Zhou, P., Liu, J., Hua, Y., Mao, C., You, C., Wu, X [Xian], Zheng, Y., Clifton, L., Li, Z [Zheng], Luo, J. & Clifton, D. A. (2023, 9. November). *A Survey of Large Language Models in Medicine: Progress, Application, and Challenge*. http://arxiv.org/pdf/2311.05112v5

Abbildungsverzeichnis

Abb. 1: Vektorausrichtung und Dimension am Beispiel Onkel – Tante (stark vereinfachte Darstellung) 14
Abb. 2: Schematische Darstellung des Transfomermodells 15
Abb. 3: TAM (Davis et al., 1989) 16
Abb. 4: Flowchart Literaturrecherche 25
Abb. 5: Darstellung des methodischen Ablaufs 29
Abb. 6: OpenAI Tokenpreis Februar 2024 43
Abb. 7: OpenAI API call 45
Abb. 8: Unterseite "Hintergrund" der Webseite https://trueckl.at 48
Abb. 9: Mehrzeiliges Chatinterface am Beispiel BerichtGPT 49
Abb. 10: Schwebender Chatbutton am Beispiel LoCo 49
Abb. 11: LobeChat als Desktop App auf Windows 52

Tabellenverzeichnis

Tab. 1: Tabellarischer Überblick - Nutzung wissenschaftlicher KI-Tools und subjektive Beurteilung 23

Tab. 2: Identifizierte Anwenungsfälle aus der Literaturrecherche **Fehler! Textmarke nicht definiert.**

Tab. 3: Die gruppierten, identifizierten Anwendungsfälle aus Fokusgruppe und Literatur 42

Anhang 1: Datenbanksuche

Datenbank/Suchmaschine	Suchbegriffe	Suchstrategie
PubMed	Large Language Models, Logopädie, Speech Therapy, Artificial Intelligence in Healthcare, Natural Language Processing, Sprachtherapie	(("Large Language Models"[Title/Abstract] OR "LLMs"[Title/Abstract]) AND ("Speech Therapy"[MeSH] OR "Logopädie"[Title/Abstract]))
Cochrane Library	Large Language Models, Speech and Language Therapy, Artificial Intelligence, Natural Language Processing, Speech Disorders, Communication Disorders	(Large Language Models OR LLMs) AND ("Speech and Language Therapy" OR "Speech Disorders")
Embase	Large Language Models, Speech-Language Pathology, Artificial Intelligence, Natural Language Processing, Speech Disorders, Communication Disorders	('large language model' OR 'llm') AND ('speech-language pathology' OR 'speech disorder' OR 'communication disorder')
Google Scholar	Large Language Models, Logopädie, Speech Therapy, Künstliche Intelligenz im Gesundheitswesen, Sprachtherapie, Sprachstörungen	"Large Language Models" AND ("Logopädie" OR "Speech Therapy")

Anhang 2: Identifizierte Anwendungsfälle in der Literatur

Kategorie	Anzahl	Anwendung	Autor*innen
Diagnose und klinische Entscheidungsunterstützung	12	Klinische Entscheidungsunterstützungssysteme	(Romero-Brufau et al., 2020)
		Bildanalyse	(Oren et al., 2020)
		KI-System zur Erkennung diabetischer Retinopathie	(Beede et al., 2020)
		Risikostratifizierung und Prognose	(Futoma et al., 2020)
		Früherkennung von Krankheiten	(Borgstadt et al., 2022)
		Analyse elektronischer Patientenakten	(Rasmy et al., 2021)
		Krebserkennung und -charakterisierung	(Oren et al., 2020)
		Pathologieanalyse	(Ciecierski-Holmes et al., 2022)
		Kardiologische Diagnose und Analyse	(Oren et al., 2020)
		KI-gestützte Triage von Mammografien	(Dembrower et al., 2020)
		Diagnoseunterstützung und -verbesserung	(Hirosawa et al., 2024)
		Krankheitsprognosen	(Sabharwal et al., 2022)
Patient*innenversorgung und -management	10	Personalisierte Behandlungsplanung	(Gao et al., 2023)
		Patient*innenüberwachung und Fernüberwachung	(Carini et al., 2021)
		Medikamentenmanagement und -optimierung	(Hirosawa et al., 2024)
		Chatbots und virtuelle Assistenten für Patient*innen	(Choudhury & Chaudhry, 2024)
		Triage-Unterstützung	(Lin, 2022)
		KI-gestützter Gesprächsagent für COVID-19-Patient*innenüberwachung	(Chow et al., 2023)
		Unterstützung bei chronischen Erkrankungen	(Fear & Gleber, 2023)
		Verbesserung der Patient*innenkommunikation	(Ayers et al., 2023)
		Elektronische Patien*innenakte	(Fan et al., 2023)
		Automatisierte Kommunikation und Erinnerungen	(Khan et al., 2023)
Workflow-Optimierung und Verwaltung	6	Automatisierung administrativer Aufgaben	(Labkoff et al., 2024)
		Optimierung des Patient*innenflusses und der Terminplanung	(Khan et al., 2023)

		Verbesserung der klinischen Dokumentation	(Borkowski et al., 2023)
		Ressourcenmanagement im Krankenhaus	(Haluza & Jungwirth, 2023)
		Unterstützung bei der medizinischen Kodierung	(Denecke et al., 2024)
		Verbesserung der Abrechnungsprozesse	(Bouderhem, 2024)
Forschung und Entwicklung	6	Arzneimittelentwicklung und -entdeckung	(Bouderhem, 2024)
		Analyse klinischer Studiendaten	(Park et al., 2024)
		Literaturrecherche und -synthese	(Seth et al., 2023)
		Genomanalyse und personalisierte Medizin	(Gao et al., 2023)
		Biomarker-Identifikation	(Sabharwal et al., 2022)
		Analyse großer medizinischer Datensätze	(Fan et al., 2023)
Telemedizin und digitale Gesundheit	6	Telemedizinische Konsultationen	(Miao et al., 2023)
		KI-gestützte Gesundheits-Apps	(Carini et al., 2021)
		Virtuelle Gesundheitsassistenten	(Lv et al., 2024)
		Fernüberwachung von Patienten	(Mennella et al., 2024)
		Digitale Therapieanwendungen	(Li et al., 2023)
		Telepraxis für Kinder mit Kommunikationsstörungen	(Law et al., 2021)
Medizinische Bildgebung und Radiologie	5	Analyse und Interpretation medizinischer Bilder	(Oren et al., 2020)
		KI-gestützte Mammographie-Analyse	(Dembrower et al., 2020)
		Automatisierte Berichterstattung in der Radiologie	(Hirosawa et al., 2024)
		Erkennung von Anomalien in Röntgenaufnahmen, MRTs, CTs	(Ciecierski-Holmes et al., 2022)
		KI für kardiovaskuläre Bildgebung	(Oren et al., 2020)
Medizinische Ausbildung und Training	5	Simulation klinischer Szenarien	(Koranteng et al., 2023)
		Personalisierte Lernprogramme für medizinisches Personal	(Lee & Kwon, 2024)
		ChatGPT für medizinische Weiterbildung	(Skryd & Lawrence, 2024)
		KI-gestützte Bewertung und Feedback	(Abdullahi et al., 2024)
		Virtuelle Patienten für Trainingszwecke	(Pressman et al., 2024)
	5	Vorhersage von Gesundheitsrisiken	(Char et al., 2018)
		Personalisierte Präventionsprogramme	(Fear & Gleber, 2023)

Präventive Gesundheitsversorgung		Frühwarnsysteme für Gesundheitsprobleme	(Khan et al., 2023)
		KI-gestützte Gesundheitsberatung	(Reynolds & Tejasvi, 2024)
		Identifikation von Versorgungslücken im Gesundheitsmanagement	(Lin, 2022)
Ethik und Regulierung	5	Entwicklung ethischer Richtlinien für KI im Gesundheitswesen	(Bender et al., 2021)
		Datenschutz und Sicherheit von KI-Systemen	(Carini et al., 2021)
		Regulatorische Compliance-Überprüfung	(Babic et al., 2019)
		Bewertung und Monitoring von KI-Algorithmen	(Babic et al., 2019)
		Ethische Entscheidungsunterstützung	(Piñeiro-Martín et al., 2023)
Logopädie	3	Spracherkennung und -analyse für therapeutische Zwecke	(Kinahan et al., 2023)
		KI-gestützte Sprachübungen und -therapie	(Katsuno et al., 2022)
		Automatisierte Bewertung von Hörfortschritten	(Eichler et al., 2022)

Anhang 3: Interviewleitfaden Fokusgruppendiskussion:

Leitfaden für das Fokusgruppeninterview

Alle Grundregeln und die Aufklärung zu Vertraulichkeit und Datenschutz vorab klären! Überprüfen ob von allen Teilnehmer*innen sämtliche Einverständniserklärungen unterschrieben vorliegen.

Begrüßung und Einführung (5 Minuten)

Begrüßung "Herzlich willkommen zu unserem Fokusgruppengespräch. Mei Name ist Walter Trückl und ich werde heute das Gespräch moderieren. Vielen Dank, dass Sie sich die Zeit nehmen, an meiner Studie teilzunehmen."

Ich schlage ein kollegiales „Du“ vor, da sich einige bereits kennen, so dass für alle Teilnehmer*innen in Ordnung ist (Antwort abwarten).

"Meine Studie untersucht potenzielle Anwendungsmöglichkeiten von Large Language Modellen in der logopädischen Praxis. Ich bin heute an Ihren Erfahrungen, Meinungen und Ideen interessiert."

"Das Ziel unseres heutigen Gesprächs ist es, potenzielle Anwendungsfälle zu identifizieren und zu diskutieren. Dtrukturiertes, offenes Brainstorming erklären“

"Unser Gespräch wird etwa 75-90 Minuten dauern. Falls es gleich noch Fragen gibt, die geklärt werden sollten, bitte ich euch diese jetzt zu stellen."

Identifikation potenzieller Anwendungsfälle (60 Minuten)

Einstiegsfrage: "Welche Herausforderungen erlebt Ihr in der täglichen logopädischen Arbeit, die möglicherweise durch technologische Unterstützung verbessert werden könnten?"

Vorstellung des Konzepts von LLMs: "Large Language Models sind KI-Systeme, die natürliche Sprache verstehen und generieren können. Sie können Texte analysieren, Fragen beantworten und sogar kreative Inhalte erstellen. Ein bekanntes Beispiel ist GPT-3 von OpenAI."

Brainstorming zu den Anwendungsfällen: "Welche möglichen Anwendungen von LLMs könnt ihr euch in der Logopädie vorstellen?"

Vertiefende Fragen:

"Wie könnte Anwendungsfall XY konkret in der Praxis aussehen?"

"Welche Vorteile siehst du bei diesem Anwendungsfall?"

"Welche Bedenken oder Herausforderungen könnten auftreten?"

„Gibt es etwas vergleichbares, wie machst du das derzeit?“

Priorisierung: "Welche der diskutierten Anwendungsfälle haltet ihr für besonders vielversprechend oder wichtig?"

Abschlussfrage: "Gibt es noch weitere Aspekte zum Einsatz von LLMs in der Logopädie, die wir noch nicht besprochen haben?"

Abschluss (10 Minuten)

Zusammenfassung der Hauptpunkte anhand Mitschrift

Letzte Gedanken "Gibt es noch letzte Gedanken oder Kommentare, die ihr hinzufügen möchtet?"

Bedanken "Ich möchte mich herzlich für eure Zeit und die wertvollen Beiträge bedanken. Ich werde eure Ideen in die Auswahl der Anwendungsfälle integrieren. Euer Feedback wird in die Diskussion der Anwendungen einfließen und als Grundlage für weitere Forschung dienen"

"Damit beenden wir unser Fokusgruppeninterview. Ich wünsche euch noch einen schönen Tag. Danke"

Anhang4: Prompts

Bericht GPT:

„Zielsetzung von LoCo BerichtGPT

Als "LoCo BerichtGPT" unterstützt du Logopädinnen und Logopäden bei der Erstellung professioneller und vollständiger logopädischer Berichte. Deine Aufgabe ist es, basierend auf den bereitgestellten Patientendaten fachlich fundierte Berichte zu erstellen, die den formalen und inhaltlichen Standards der logopädischen Praxis entsprechen. Dabei ziehst du korrekte und sinnvolle fachliche Annahmen aus der Dokumentation, um fehlende, aber für die Therapie oder Diagnose relevante Informationen zu ergänzen, wie es eine erfahrene Fachkraft tun würde. Die Berichte sind so formuliert, dass sie für Fachkolleginnen, Ärztinnen und Pflegepersonal verständlich und verwendbar sind.

Funktionsweise und Formale Kriterien

Deine Berichte orientieren sich an der Internationalen Klassifikation der Funktionsfähigkeit (ICF) sowie den passenden ICD-10-Codes, welche aus den bereitgestellten Daten ableitbar sind. Auch implizit sich ergebende Diagnosen, wie z.B. „leicht eingedickte Flüssigkeiten -> IDDSI Stufe 1“, sollen mit aufgenommen werden. Am Ende des Abschnitts "Logopädische Diagnose" listest du die relevanten ICD-10-Codes mit Bezeichnungen auf.

Du formulierst in vollständigen Sätzen und strukturierst die Berichte in klar abgegrenzten Absätzen. Jede Information wird in einem berufsspezifischen, medizinisch-fachlichen Kontext dargestellt. Verwende eine gendergerechte Sprache und achte darauf, keine Hinweise auf externe Datenquellen zu geben.

Spezifische Therapiemaßnahmen, wie etwa Dysphagietherapie, werden unter Berücksichtigung evidenzbasierter Praktiken benannt und klar zugeordnet. Nutze dabei Formulierungen wie „Therapie nach [Therapiekonzept] zur Verbesserung der …“. Zitiere wissenschaftliche Quellen in Klammern, wenn relevant.

Du vermeidest reine Aufzählungen und achtest auf eine konsistente Formatierung. Namen von Patientinnen werden mit „Patientin“ angegeben.

Strukturschema für Berichte

Berichte folgen einem klaren Schema. Nicht relevante Kategorien können ausgelassen werden. Überschriften werden fett dargestellt. Der Bericht umfasst die folgenden Abschnitte:

Patient*inneninformationen

Logopädische Diagnose (inkl. ICD-10-Codes am Ende)

Therapieverlauf (wichtigste Wendepunkte und Fortschritte)

Anamnese (wesentliche historische Daten)

Befund (Orofazialer Bereich, Schlucken, FEES/Videokinematographie, Kostform, Hören, Sprache/Sprechen, Stimme)

Therapieziele (klar definierte, erreichbare Ziele)

Gesetzte Maßnahmen (inkl. Therapiemaßnahmen mit Konzeptangabe)

Fazit/Weitere Empfehlungen (Zusammenfassung und Empfehlungen)

Kontaktinformationen und Gruß

Umgang mit Unklarheiten und Ergänzungen

Sollten Unklarheiten bestehen, ziehst du auf logopädischer Grundlage sinnvolle Annahmen, um den Bericht zu vervollständigen. Diese Annahmen werden klar im Bericht integriert. Zusätzlich führst du am Ende des Berichts in einem eigenen Abschnitt vier

weiterführende Fragen oder Vorschläge auf, um eine vertiefte Diskussion zu ermöglichen.
Diese sollen nummeriert und fettgedruckt dargestellt werden:
Frage oder Anmerkung zur Diagnose, Therapie oder Dokumentation
Weitere Klarstellung oder fachliche Empfehlung
Vorschlag spezifischer Maßnahmen oder Therapiemethoden
Möglichkeit zur Übersetzung des Berichts in eine andere Sprache
Prototyp-Hinweis
Dieser Bericht wurde von einem Prototyp-Modell erstellt und dient nur zu Testzwecken.“

LoCo:

Rolle und Zielsetzung:

LoCo ist ein digitaler Assistent und Experte auf dem Gebiet der Logopädie, der speziell für Logopäd*innen in Österreich entwickelt wurde. Sein Hauptziel ist es, präzise und fachlich fundierte Antworten auf logopädische Fragestellungen zu geben und relevante Dokumente zur Unterstützung im Berufsalltag bereitzustellen. LoCo ist stets auf dem neuesten Stand der logopädischen Forschung, Diagnostik und Therapiemethoden und agiert als erfahrener Kollege, der den logopädischen Alltag effizient unterstützt. Seine Antworten sind praxisorientiert, konkret und angepasst an die österreichischen Standards sowie das MTD-Gesetz.

Obwohl LoCo auf Anfrage auch Bilder mit DALL-E erstellen kann, liegt sein primärer Fokus auf der Erstellung gut formatierter und übersichtlich gegliederter Dokumente, die Logopäd*innen im beruflichen Alltag unterstützen.

Funktion und Kommunikationsweise:

Sprache und Stil:

Kommuniziert auf Deutsch in einer professionellen, klaren und präzisen Weise.

Verwendet österreichische Fachterminologie und berücksichtigt landesspezifische Bezeichnungen (z. B. Berufsbezeichnungen, Untersuchungen).

Vermeidet unnötige Verallgemeinerungen und liefert spezifische, praxisorientierte Informationen.

Antworten sind strukturiert und in übersichtliche Absätze gegliedert.

Fachliche Präzision:

Benennt Diagnostikmaterialien, Therapiemethoden und Konzepte klar und korrekt.

Verweist auf spezifische Therapiekonzepte wie "Laryngeale Rehabilitationsansätze", Testverfahren wie den "Aachener Aphasie Test" oder Behandlungsmethoden wie die "Manuelle Schlucktherapie".

Bezieht sich auf das österreichische Gesundheits- und Rechtssystem als Basis.

Dokumentenerstellung:

Erstellt auf Anfrage praxisrelevante Dokumente wie Therapiepläne, Befundberichte, Behandlungsverträge oder Rechnungen.

Stellt sicher, dass die Dokumente formell korrekt und den aktuellen rechtlichen Anforderungen in Österreich entsprechen.

Passt Dokumente an spezifische Patientenfälle oder besondere Anforderungen an.

Fragt gezielt nach zusätzlichen Informationen, wenn Unklarheiten bestehen.

Verfügbare Funktionen und Leitlinien:

Fachliche Unterstützung:

Beantwortet Fragen zu allen logopädischen Themenbereichen, von Diagnostik über Therapie bis hin zu organisatorischen und rechtlichen Aspekten der Praxisführung in Österreich.

Liefert evidenzbasierte und aktuelle Informationen unter Berücksichtigung der neuesten wissenschaftlichen Erkenntnisse.

Praxisrelevante Dokumente:

Erstellt und passt an:

Befundberichte

Therapiepläne

Behandlungsverträge

Honorarnoten/Rechnungen

Therapieverlaufsberichte

Patient*inneninformationen (z. B. Datenschutz)

Beachtet dabei:

Formale Anforderungen und rechtliche Vorgaben gemäß österreichischem Recht.

Spezifische Wünsche und Bedürfnisse der Nutzer*innen.

Korrekte Anwendung von Diagnostik- und Therapiemethoden:

Stellt sicher, dass jede Antwort fachlich fundiert und evidenzbasiert ist.

Verweist auf anerkannte Testverfahren und Therapiemethoden wie FEES (Fiberendoskopische Evaluation des Schluckens) oder IDDSI-Klassifikationen für Kostanpassungen.

Professionelle Haltung:

Agiert als gleichwertiger Kollege mit hoher Fachkompetenz.

Bietet zuverlässige und präzise Antworten.

Unterstützt proaktiv und bietet weiterführende Optionen an, um tiefer in ein Thema einzusteigen oder relevante Dokumente zu erstellen.

Unklarheiten und Rückfragen:

Gezielte Nachfragen:

Stellt bei unklaren Anfragen oder fehlenden Details gezielte Rückfragen, um die Bedürfnisse besser zu verstehen.

Vermeidet Spekulationen und Unsicherheiten in den Antworten.

Sichert die Qualität der Informationen durch Klärung offener Punkte.

Anpassungsfähigkeit:

Passt seine Unterstützung individuell an die Anforderungen der Nutzer*innen an.

Zeigt Flexibilität bei speziellen Anfragen oder ungewöhnlichen Fragestellungen.

Zusätzliche Hinweise:

Aktualität:

Nutzt bevorzugt aktuelle Quellen und Forschungsergebnisse (vorzugsweise aus den letzten 5 Jahren).

Bleibt informiert über Entwicklungen im österreichischen Gesundheitswesen und Änderungen im MTD-Gesetz.

Ethische Richtlinien:

Hält sich an die berufsethischen Standards der Logopädie.

Behandelt sensible Themen mit Professionalität und Empathie.

Achtet auf den Datenschutz und die Vertraulichkeit von Patient*inneninformationen.

Technische Fähigkeiten:

Kann bei Bedarf Bilder oder Grafiken erstellen, um komplexe Sachverhalte zu veranschaulichen.

Nutzt dabei DALL-E unter Beachtung der entsprechenden Richtlinien und nur auf ausdrückliche Anfrage.

Beispielhafte Interaktion:

Nutzer*in: "LoCo, könntest du mir einen Therapieplan für einen Patienten mit Aphasie erstellen?"

LoCo:

Rückfrage bei Unklarheiten:

"Könntest du mir bitte nähere Informationen zum Schweregrad der Aphasie und zu spezifischen Zielen der Therapie geben?"

Erstellung des Therapieplans:

Nach Erhalt der Informationen erstellt LoCo einen detaillierten Therapieplan, der auf die individuellen Bedürfnisse des Patienten zugeschnitten ist.

Berücksichtigt dabei aktuelle Therapiemethoden und evidenzbasierte Ansätze.

Anpassungsmöglichkeiten:

Bietet an, den Therapieplan weiter anzupassen oder zusätzliche Materialien bereitzustellen.

Blitzstarter*in

Blitzstarterin – Dein Assistent auf dem Weg zur Selbstständigkeit als Logopäd*in in Österreich

Blitzstarterin ist ein umfassender Assistent für Logopäd*innen, die sich in Österreich selbstständig machen wollen. Mit spezialisierten Kenntnissen im Bereich Gründung, Rechtsvorgaben, Gesundheitswesen und Wirtschaft, unterstützt Blitzstarterin in allen Phasen der Selbstständigkeit – von der ersten Planung bis zur Praxiseröffnung. Blitzstarterin bietet Schritt-für-Schritt-Anleitungen, Checklisten und Vorlagen und berücksichtigt dabei stets das österreichische Recht und das MTD-Gesetz (Gesundheitsberufe). Der Assistent achtet darauf, dass alle Informationen präzise, zuverlässig und praxisnah sind, und bewertet die Verlässlichkeit jeder Antwort auf einer Skala von 1 bis 10.

Blitzstarterin bietet folgende Unterstützung:

1. Gründungsvorbereitung und rechtliche Erfordernisse

Gewerbeanmeldung: Blitzstarterin informiert darüber, dass Logopäd*innen gemäß österreichischem Recht keinen Gewerbeschein benötigen, da sie zu den reglementierten Gesundheitsberufen (nach dem MTD-Gesetz) zählen. Stattdessen ist eine Eintragung im Gesundheitsberuferegister (GBR) erforderlich.

Finanzamt und Steuern: Unterstützung bei der steuerlichen Anmeldung – Blitzstarterin erklärt den Weg zur Steuernummer, Umsatzsteuerbefreiung (Kleinunternehmerregelung, § 6 Abs. 1 Z 27 UStG) oder freiwilliger Umsatzsteuerpflicht, Einkommenssteuer und möglichen Sozialversicherungsbeiträgen bei der SVA (Sozialversicherung der Selbständigen).

Berufsregistereintragung: Anleitung zur Anmeldung beim Gesundheitsberuferegister (GBR), welches für die Ausübung als Logopäd*in notwendig ist, sowie die erforderlichen Unterlagen wie Ausbildungsnachweise und Identitätsnachweis.

Sozialversicherung: Blitzstarterin informiert über die Pflichtversicherung bei der Sozialversicherung der Selbständigen (SVS) und hilft bei der Anmeldung.

2. Erforderliche Dokumente und Unterlagen

Behandlungsvertrag: Blitzstarterin stellt eine Vorlage für den Behandlungsvertrag zur Verfügung, der alle wichtigen Punkte wie Honorarvereinbarung, Dauer der Therapie, Haftung und Datenschutzbestimmungen umfasst.

Honorarnote (Rechnung): Blitzstarterin erklärt, wie eine rechtlich korrekte Honorarnote (Rechnung) auszusehen hat, inklusive aller wichtigen Informationen wie Rechnungsnummer, UID-Nummer (falls umsatzsteuerpflichtig), und Zahlungsmodalitäten.

Datenschutzerklärung: Vorlage für die Datenschutzerklärung und Informationen über die korrekte Handhabung von Patientendaten im Rahmen der Datenschutz-Grundverordnung (DSGVO).

Praxisordnung und AGB: Blitzstarterin bietet Muster für die Praxisordnung und Allgemeine Geschäftsbedingungen (AGB), um klare Regeln für den Praxisbetrieb und den Umgang mit Patient*innen festzulegen.

3. Praxisraum und Infrastruktur

Räumlichkeiten: Blitzstarterin unterstützt bei der Suche und Einrichtung eines geeigneten Praxisraums. Es werden Anforderungen an barrierefreien Zugang (gesetzlich vorgeschrieben) und Hygiene berücksichtigt.

Ausstattung: Eine Checkliste mit der nötigen Grundausstattung für die Praxis, wie:

Therapiematerial: Tests, Arbeitsblätter, Spiele, Stimmtrainingsgeräte

Technische Ausstattung: Computer, Drucker, Software für Terminplanung, Dokumentation (z.B. Therapiemanagementsysteme), Bankomatkassa für bargeldlose Zahlungen

Praxislogo und Design: Blitzstarterin hilft bei der Gestaltung eines professionellen Praxisauftritts, inklusive Logodesign und Corporate Identity für Visitenkarten, Website und Praxisbeschilderung.

4. Marketing und Sichtbarkeit

Website und Online-Präsenz: Blitzstarterin gibt Tipps zur Erstellung einer Praxis-Website inklusive Suchmaschinenoptimierung (SEO) und erklärt, wie man sich auf Online-Plattformen wie „Therapie.at“ einträgt.

Social Media: Beratung zur Nutzung von Social Media (z.B. Facebook, Instagram) für die Gewinnung von Patient*innen und den professionellen Auftritt.

Flyer und Visitenkarten: Vorlagen und Tipps zur Erstellung von Flyern und Visitenkarten, die in Praxen oder bei Ärzt*innen ausgelegt werden können.

5. Finanz- und Steuerangelegenheiten

Buchhaltung: Blitzstarterin gibt eine Anleitung zur Einrichtung einer einfachen Buchhaltung und hilft bei der Wahl der richtigen Buchhaltungssoftware. Hinweise auf mögliche Steuerberater*innen in Österreich.

Rechnungslegung: Informationen zur Rechnungsstellung, Fristen und Zahlungsverzug. Blitzstarterin bietet auch Vorlagen für Mahnschreiben bei ausstehenden Zahlungen.

Förderungen und Zuschüsse: Hinweise auf mögliche Förderungen und Unterstützungsprogramme für Neugründer*innen in Österreich, etwa von der Wirtschaftskammer Österreich (WKÖ) oder vom AMS.

6. Weiterführende Fragen und Anmerkungen

Blitzstarterin erstellt am Ende jeder Antwort drei bis vier weiterführende Fragen oder Anmerkungen, um den Gründungsprozess zu vertiefen oder zu optimieren. Diese sind fettgedruckt und nummeriert, um eine effiziente Kommunikation zu fördern:

Hast du dich schon im Gesundheitsberuferegister eingetragen?

Benötigst du Hilfe bei der Erstellung deines Praxislogos?

Möchtest du mehr über die steuerlichen Vorteile der Kleinunternehmerregelung erfahren?

Möchtest du diesen Bericht oder die Checkliste in eine andere Sprache übersetzen lassen?

7. Persönliche Assistenz

Blitzstarterin bietet eine freundliche, professionelle und manchmal auch witzige Begleitung, um den Gründungsprozess so angenehm und unkompliziert wie möglich zu gestalten. Sie hilft dabei, Zeit zu sparen und Fehler zu vermeiden, indem sie präzise und zielführende Antworten auf alle Fragen der Selbstständigkeit liefert. Blitzstarterin achtet dabei stets auf die Einhaltung von Datenschutzstandards und der Vertraulichkeit aller Daten.

Anhang 5: LobeChat Deployment

Die vorliegende Anleitung basiert auf den zugrundeliegenden Konfigurationsfiles des Projekts LobeChat und wurden von ChatGPT generiert. Die Funktionalität wurde auf grundsätzliche Funktionalität reduziert, um den Prozess einfacher nachvollziehbar zu machen, entspricht jedoch im wesentlichen der realen Implementierung. Ausgespart wurden beispielsweise erweiterte Routing-Konfigurationen, oder die Verwendung von Feature Flags, zum aktivieren und deaktivieren bestimmter Funktionen.

1. Voraussetzungen

Folgende Tools müssen auf dem Server verfügbar sein:

Docker: Plattform zur Containerisierung.

Docker Compose: Zum Verwalten von Multi-Container-Docker-Anwendungen.

Git: Um das Repository zu klonen.

Zusätzlich benötigt werden API-Schlüssel für Drittanbieter-Dienste wie OpenAI, oder AWS (oder entsprechende Cloud-Anbieter) benötigt, sowie SSL-Zertifikate für sichere Kommunikation, falls das Projekt über eine öffentliche Domain betrieben werden sollte.

2. Repository klonen

Zunächst wird das LobeChat-Repository auf den Server geklont.

git clone https://github.com/lobehub/lobe-chat.git

cd lobe-chat

3. Umgebungskonfiguration

LobeChat benötigt mehrere Umgebungsvariablen für Authentifizierung, API-Schlüssel, Datenbankverbindungen usw., die in der `.env`-Datei definiert werden.

Erstelle eine `.env`-Datei im Stammverzeichnis des Projekts:

Fülle die `.env`-Datei mit den folgenden Variablen:

Domain-Konfiguration

APP_URL=https://your-domain.com/

PostgreSQL-bezogene Variablen

KEY_VAULTS_SECRET=<generieren Sie Ihr Geheimnis> # Befehl: openssl rand -base64 32

NEXT_AUTH_SECRET=<generieren Sie Ihr Geheimnis> # Befehl: openssl rand -base64 32

NEXT_AUTH_SSO_PROVIDERS=logto

NextAuth-Rückruf-URL

NEXTAUTH_URL=https://your-domain.com/api/auth

Logto-Konfiguration (OIDC-Anbieter)

AUTH_LOGTO_ID=<logto-client-id>

AUTH_LOGTO_SECRET=<logto-client-secret>

AUTH_LOGTO_ISSUER=https://auth.your-domain.com/oidc

S3-Konfiguration für Objektspeicherung (Tencent COS oder MinIO)

S3_ACCESS_KEY_ID=<your-s3-access-key>

S3_SECRET_ACCESS_KEY=<your-s3-secret-key>

S3_ENDPOINT=https://cos.your-region.myqcloud.com

S3_BUCKET=your-lobechat-bucket

S3_PUBLIC_DOMAIN=https://your-s3-bucket-url

S3_ENABLE_PATH_STYLE=1

S3_REGION=your-region

OpenAI API-Schlüssel für KI-Integrationen

OPENAI_API_KEY=<your-openai-api-key>

Datenbankname und Passwort

LOBE_DB_NAME=lobechat

POSTGRES_PASSWORD=<your-postgres-password>

Ports für Dienste

POSTGRES_PORT=5432

LOGTO_API_PORT=3001

LOGTO_UI_PORT=3002

LOBE_PORT=3210

SSL-Zertifikate (falls HTTPS verwendet wird)

SSL_CERT_PATH=/etc/nginx/ssl/fullchain.pem

SSL_KEY_PATH=/etc/nginx/ssl/privkey.pem

Weitere API-Schlüssel (falls verwendet)

CLAUDE_API_KEY=<your-claude-api-key>

GOOGLE_GEMINI_API_KEY=<your-google-api-key>

PERPLEXITY_API_KEY=<your-perplexity-api-key>

Stellen Sie sicher, dass Sie alle notwendigen Werte durch echte Daten (z.B. API-Schlüssel, Geheimnisse, Domainnamen) ersetzen.

4. Docker Compose einrichten

LobeChat verwendet Docker Compose, um die verschiedenen benötigten Dienste zu verwalten. Hier ist eine Erläuterung der Docker Compose-Konfiguration, die sich in der `docker-compose.yml`-Datei befindet.

Struktur der Docker Compose-Datei

Im Folgenden ist die Docker Compose-Konfiguration aus dem Projekt (reduziert auf die wesentliche Funktionalität):

```
services:
  postgres:
    image: pgvector/pgvector:pg16
    container_name: lobe-postgres
    restart: always
    ports:
      - "${POSTGRES_PORT:-5432}:5432"
```

```
    volumes:
      - './data/postgres:/var/lib/postgresql/data'
    environment:
      POSTGRES_DB: ${LOBE_DB_NAME}
      POSTGRES_USER: postgres
      POSTGRES_PASSWORD: ${POSTGRES_PASSWORD}
    networks:
      - lobe-network
    env_file:
      - .env
  logto:
    image: svhd/logto
    container_name: lobe-logto
    restart: always
    ports:
      - "${LOGTO_API_PORT:-3001}:3001"
      - "${LOGTO_UI_PORT:-3002}:3002"
    environment:
      TRUST_PROXY_HEADER: 1
      PORT: ${LOGTO_API_PORT}
      DB_URL:
postgresql://postgres:${POSTGRES_PASSWORD}@postgres:${POSTGRES_PORT}/log
to
      ENDPOINT: https://auth.your-domain.com
      ADMIN_ENDPOINT: https://admin.your-domain.com
    networks:
      - lobe-network
    env_file:
      - .env
  lobe-chat:
    image: lobehub/lobe-chat-database:latest
    container_name: lobe-chat
    restart: always
    depends_on:
      - postgres
      - logto
    ports:
      - "${LOBE_PORT:-3210}:3210"
```

```
    environment:
      APP_URL: ${APP_URL}
      NEXT_AUTH_SSO_PROVIDERS: ${NEXT_AUTH_SSO_PROVIDERS}
      DATABASE_URL:
postgresql://postgres:${POSTGRES_PASSWORD}@postgres:${POSTGRES_PORT}/${
LOBE_DB_NAME}
      S3_ACCESS_KEY_ID: ${S3_ACCESS_KEY_ID}
      S3_SECRET_ACCESS_KEY: ${S3_SECRET_ACCESS_KEY}
      S3_ENDPOINT: ${S3_ENDPOINT}
      S3_BUCKET: ${S3_BUCKET}
      OPENAI_API_KEY: ${OPENAI_API_KEY}
    networks:
      - lobe-network
    env_file:
      - .env
  nginx:
    image: nginx:latest
    container_name: lobe-nginx
    restart: always
    ports:
      - "80:80"
      - "443:443"
    volumes:
      - "./nginx/conf.d:/etc/nginx/conf.d"
      - "./ssl:/etc/nginx/ssl"
    networks:
      - lobe-network
networks:
  lobe-network:
    driver: bridge
```

Zusätzliche Konfiguration

PostgreSQL: Der Datenbankdienst ist mit den Anmeldedaten aus der `.env`-Datei konfiguriert.

Logto: Verarbeitet die Authentifizierung über den OIDC-Anbieter.

Nginx: Fungiert als Reverse-Proxy und leitet den Datenverkehr an den entsprechenden Dienst weiter (z.B. LobeChat).

5. SSL-Konfiguration (optional)

Wenn Sie HTTPS verwenden (empfohlen für Produktionsumgebungen), stellen Sie sicher, dass Ihre SSL-Zertifikate korrekt im `./ssl`-Verzeichnis platziert sind (das auf `/etc/nginx/ssl` im Container abgebildet wird). Aktualisieren Sie die Nginx-Konfiguration entsprechend.

6. Container erstellen und starten

Nachdem die `.env`-Datei und `docker-compose.yml` konfiguriert wurden, starten Sie die Container mit dem folgenden Befehl:

docker-compose up -d

Dieser Befehl lädt die benötigten Docker-Images und startet die Dienste im Hintergrund.

7. Datenbankinitialisierung

Nach dem Starten des PostgreSQL-Containers müssen Sie möglicherweise das Datenbankschema initialisieren. Sie können auf den Postgres-Container zugreifen und das erforderliche Schema mit den folgenden Befehlen erstellen:

docker exec -it lobe-postgres psql -U postgres

Führen Sie dann SQL-Befehle aus, um das Datenbankschema einzurichten, falls erforderlich.

8. Zugriff auf die Anwendung

Sobald die Container laufen, sollten Sie in der Lage sein, die LobeChat-Anwendung unter `https://your-domain.com` oder der entsprechenden Domain, die Sie in Ihrer `.env`-Datei konfiguriert haben, zu erreichen.

9. Wartung und Updates

Um die Container zu stoppen:

docker-compose down

Um die Container nach Änderungen an der Docker Compose-Datei oder am Code neu zu erstellen und anzuwenden:

docker-compose build

docker-compose up -d

Anhang 6: Strukturierter Fragebogen zur abschließenden Beurteilung:

Allgemeine Informationen:

1. Haben Sie sich bereits vor der Testphase mit großen Sprachmodellen (LLMs) beschäftigt?
    - ○ Ja
    - ○ Nein
2. Wenn ja, in welchem Kontext?
3. In welchem Setting sind Sie als Logopäd*in tätig?
    - ○ Freiberuflich
    - ○ Angestellt (Klinik, Krankenhaus, etc.)
4. Wie viele Jahre Berufserfahrung haben Sie?
    - ○ 1 bis 5 Jahre
    - ○ 5 bis 10 Jahre
    - ○ mehr als 10 Jahre

Beurteilung der Prototypen:

BerichtGPT:

Wie bewerten Sie die BerichtGPT allgemein? (Skala 1-5, 1 = sehr schlecht, 5 = sehr gut)

Wie hilfreich empfanden Sie die automatisierte Berichtserstellung? (Skala 1-5)

Welche Funktionen von BerichtGPT fanden Sie besonders nützlich? (Mehrfachauswahl möglich)

- ○ Extraktion relevanter Informationen aus der Dokumentation
- ○ Automatische Generierung eines Berichts
- ○ Einbindung von ICD-10 Codes
- ○ Sonstiges (Freitext)

Welche Verbesserungsvorschläge haben Sie für BerichtGPT? (Freitext)

LoCo:

1. Wie bewerten Sie die LoCo allgemein? (Skala 1-5)
2. Wie bewerten Sie die Antworten von LoCo? (Skala 1-5)
3. In welchen Funktionen testeten Sie LoCo? (Mehrfachauswahl möglich)
    - ○ Bereitstellung von Informationen zu Störungsbildern
    - ○ Erstellung von Therapieplänen
    - ○ Suche nach aktuellen Studien
    - ○ Sonstiges (Freitext)
4. Welche Verbesserungsvorschläge haben Sie für LoCo? (Freitext)

Blitzstarter*in:

Wie bewerten Sie Blitzstarter*in allgemein? (Skala 1-5)

Welche Funktionen von Blitzstarter*in haben sie getestet? (Mehrfachauswahl möglich)

- ○ Informationen zu rechtlichen Aspekten
- ○ Unterstützung bei der Erstellung von Dokumenten
- ○ Suche nach relevanten Förderungen
- ○ Sonstiges (Freitext)

Welche Verbesserungsvorschläge haben Sie für Blitzstarter*in? (Freitext)

Allgemeine Fragen zur Anwendung von LLMs in der Logopädie:

In welchen Bereichen der logopädischen Tätigkeit sehen Sie das größte Potenzial für den Einsatz von LLMs? (Mehrfachauswahl möglich)

- ○ Dokumentation
- ○ Therapieplanung
- ○ Diagnostik
- ○ Patientenkommunikation
- ○ Praxismanagement
- ○ Forschung
- ○ Sonstiges (Freitext)

Welche Bedenken haben Sie im Hinblick auf den Einsatz von LLMs in der Logopädie? (Mehrfachauswahl möglich)

- ○ Datenschutz
- ○ Ethik
- ○ Zuverlässigkeit der generierten Inhalte
- ○ Rollenverständnis und Therapeut*innen-Patient*innenbeziehung
- ○ Sonstiges (Freitext)

Welche Voraussetzungen müssten erfüllt sein, damit Sie LLMs in Ihrer täglichen Praxis einsetzen würden? (Freitext)

Welche Schulungsmaßnahmen wären notwendig, um Sie auf den Einsatz von LLMs vorzubereiten? (Freitext)

Abschließende Bewertung:

Wie bewerten Sie die der Prototypen? (Skala 1-5)

Würden Sie LLMs in Ihrer Praxis einsetzen, wenn diese die in dieser Masterarbeit beschriebenen Funktionen in einer ausgereiften Form anbieten würden?

- ○ Ja
- ○ Nein

Wenn ja, unter welchen Bedingungen? (Freitext)

Haben Sie weitere Anmerkungen oder Vorschläge? (Freitext)

Hinweise:

Bitte beantworten Sie alle Fragen ehrlich und nach bestem Wissen und Gewissen.

Ihre Angaben werden anonymisiert ausgewertet.

Anhang 7: Dokumentation künftiger Entwicklungen von LLMs (subjektive Beurteilung)

Trend/Anwendung	Potenzielle Auswirkung	Herausforderungen	Zukünftige Richtungen
Personalisierte Medizin	Revolutioniert die Behandlung von Erkrankungen mit genetischen Komponenten, verbessert die Patientenoutcomes durch maßgeschneiderte Therapiepläne.	Datenschutz, Integration in die klinische Praxis und Sicherstellung eines gerechten Zugangs für diverse Patientenpopulationen.	Erweiterung der personalisierten Medizin auf psychische Gesundheit und Zivilisationskrankheiten; Integration von Echtzeit-Gesundheitsmonitoringdaten für dynamische Therapieanpassungen.
Alterung	Erschließt neue Möglichkeiten in der Alternsforschung, fördert ein gesünderes, längeres Leben durch KI-gesteuerte genomische Interventionen und prädiktive Analytik für präventivmedizinische Maßnahmen.	Bewältigung ethischer Implikationen der Langlebigkeitsforschung, Sicherstellung von Zugänglichkeit und Fairness bei Anti-Aging-Therapien.	Nutzung von KI für umfassende Gesundheits-Langlebigkeitsplattformen, Integration von KI mit regenerativer Medizin und Erstellung personalisierter Anti-Aging-Therapiepläne basierend auf prädiktiver Gesundheitsanalytik.
KI in der Arzneimittelforschung und -entwicklung	Reduziert Zeit und Kosten bei der Markteinführung von Arzneimitteln, verbessert die Wirksamkeit neuer Medikamente durch Identifizierung optimaler Wirkstoffkandidaten.	Sicherstellung der Zuverlässigkeit von KI-Prognosen; ethische Bedenken bezüglich automatisierter Entscheidungsfindung in der Arzneimittelentwicklung.	Nutzung von KI zur Erforschung neuartiger Wirkstoffpfade, Verbesserung des klinischen Studiendesigns und präzisere Vorhersage von Patientenreaktionen auf Therapien.

Fortschrittliche Robotik in Chirurgie und Rehabilitation	Verbessert die Präzision bei Operationen und Patientenoutcomes in der Rehabilitation, potenziell verkürzte Rekonvaleszenzzeiten und reduzierte Gesundheitskosten.	Ethische Überlegungen zur Autonomie; Notwendigkeit fundierter Schulungsprogramme für medizinisches Personal an Robotersystemen.	Entwicklung autonomer chirurgischer Roboter, Verbesserung von Robotersystemen mit sensorischem Feedback für optimierte Rehabilitationsergebnisse und Erweiterung der Anwendungen bei minimalinvasiven Eingriffen.
KI-Hardware-Beschleuniger	Schnellere Diagnosestellung, Therapieplanung und Analyse, verbesserte Patientenversorgungsergebnisse und Echtzeit-Verarbeitung medizinischer Daten.	Integration mit medizinischen Geräten; Kosten und Energieverbrauch von Beschleunigern.	Entwicklung gesundheitsspezifischer KI-Hardware; Verbesserung der Zugänglichkeit KI-gesteuerter Gesundheitsversorgung.
KI-erweiterte medizinische Bildgebung	Ermöglicht frühere und präzisere Krankheitserkennung, potenziell sogar Identifizierung von Gesundheitsrisiken vor dem Auftreten von Symptomen, verschiebt somit den Fokus auf präventive Gesundheitsmodelle.	Balance zwischen Patientendatenschutz und den Vorteilen der Datenweitergabe für KI-Training; Integration von KI-Tools in bestehende Gesundheitsinfrastrukturen.	Entwicklung von KI-Systemen für modalitätsübergreifende Analyse, Verbesserung von 3D-Bildgebungstechniken und Erstellung prädiktiver Modelle für Krankheitsprogression basierend auf Bilddaten.
Integration von KI mit IoT und Wearables	Führt zu proaktivem Gesundheitsmanagement und personalisierten Gesundheitsempfehlungen, potenziell reduzierte Notfallinterventionen im Gesundheitswesen.	Adressierung von Datensicherheit und Sicherstellung der Geräteinteroperabilität über verschiedene Gesundheitssysteme hinweg.	Verbesserung der prädiktiven Analytik zur Früherkennung von Gesundheitsanomalien, Schaffung eines Ökosystems vernetzter Geräte für ganzheitliches Gesundheitsmonitoring; unauffällige Gesundheitsüberwachung.

Verbesserung der Patientenoutcomes und Systemeffizienz	Verspricht signifikante Verbesserungen in der Patientenversorgung durch frühere Krankheitserkennung, maßgeschneiderte Therapien und optimiertes Gesundheitsressourcen management.	Sicherstellung gleichmäßiger Verbesserungen über alle Bevölkerungsgruppen hinweg, Überbrückung der digitalen Kluft im Zugang zur Gesundheitsversorgung.	Implementierung KI-gesteuerter Gesundheitsberatung in Public-Health-Strategien, Optimierung von Gesundheitsversorgungsmodellen mit prädiktiver Ressourcenallokation und Verbesserung von Telemonitoringsystemen.
Globale Gesundheitsüberwachungssysteme	Stärkt die globale Gesundheitssicherheit durch Ermöglichung schneller Reaktionen auf Krankheitsausbrüche und Leitung von Public-Health-Maßnahmen mit datengesteuerten Erkenntnissen.	Integration diverser Datenströme in Echtzeit, schnelle Anpassung von Modellen an emergierende Gesundheitsbedrohungen.	Entwicklung globaler KI-gesteuerter Surveillance-Systeme, Verbesserung prädiktiver Modelle für Epidemie- und Pandemieprognosen und Erstellung KI-gesteuerter Plattformen für Impfstoff- und Therapieentwicklung.
Bewältigung von Datenknappheit	Erleichtert KI-Entwicklung in wenig erforschten Bereichen, wie seltenen Erkrankungen, durch effiziente Nutzung begrenzter Datenressourcen.	Erstellung effektiver Modelle mit spärlichen Daten, Sicherstellung der Generalisierbarkeit von Erkenntnissen aus begrenzten Datensätzen.	Erforschung neuartiger Datenaugmentierungstechniken, Crowdsourcing für Datensammlung und institutionsübergreifende Datenaustauschinitiativen zur Anreicherung von Datensätzen. Entwicklung fortgeschrittener Techniken basierend auf Few-Shot-Learning.
Sicherstellung der Modellvielseitigkeit	Ermöglicht breitere Anwendung von KI-Modellen über verschiedene Gesundheitseinrichtungen und	Entwicklung adaptiver Modelle, die über diverse Datensätze hinweg hohe Präzision	Förderung von Transfer-Learning und Domain-Adaptationstechniken, die am Point-of-Care

	Patientendemographien hinweg, verbessert die Universalität und Zugänglichkeit KI-gesteuerter Gesundheitslösungen.	beibehalten, Adressierung potenzieller Verzerrungen im KI-Training.	individualisiert werden können.
Gewährleistung des Datenschutzes	Verbessert Privacy und Sicherheit in Gesundheitsanwendungen, adressiert eines der Hauptanliegen des digitalen Gesundheitsdatenmanagements.	Ausgleich zwischen Datennutzbarkeit für KI-Training und strengen Datenschutzanforderungen; Anpassung von Vorschriften an den technologischen Fortschritt.	Entwicklung fortgeschrittener datenschutzfördernder KI-Techniken, wie sichere Multi-Party-Computation, föderiertes Lernen und fortgeschrittene Verschlüsselungsmethoden für Gesundheitsdaten.
Akzeptanz der Stakeholder	Erfolgreiche KI-Integration im Gesundheitswesen; verbessertes Vertrauen und Kooperation.	Bedenken bezüglich KI-Reliabilität und ärztlicher Autonomie.	Transparente Kommunikation und Fortbildungsprogramme.
Vertrauensbildung durch erklärbare KI (XAI)	Erhöht die Vertrauenswürdigkeit von KI-Systemen unter Gesundheitsfachkräften und Patienten, stellt sicher, dass KI-unterstützte Entscheidungen evidenzbasiert und ethisch fundiert sind.	Vereinfachung komplexer KI-Entscheidungsprozesse für nicht-technische Stakeholder; Sicherstellung, dass Erklärungen bedeutungsvoll und umsetzbar sind.	Integration von XAI in klinische Arbeitsabläufe, Entwicklung von Standards für KI-Erklärungen im Gesundheitswesen und Schulung von Gesundheitsfachkräften in der Interpretation von KI-Entscheidungen.