

Diplomarbeit

**Annotation medizinischer Begriffe für Anwendungen der künstlichen Intelligenz - Konzepte und Herausforderungen
Literaturübersicht über bekannte Klassifizierungen,
Nomenklatur und Konzepte, die in den Annotationsprozess für Anwendungen des maschinellen Lernens involviert sind.**

Eingereicht von

Martin Michael Hubner

Zur Erlangung des akademischen Grades

Doktor der gesamten Heilkunde

(Dr. med. univ.)

An der

Medizinischen Universität Graz

Ausgeführt an der

Universitätsklinik für Radiologie

ausgeführt an der

**Klinischen Abteilung für Neuroradiologie, vaskuläre und
interventionelle Radiologie**

Unter der Anleitung von

Research Prof. DI Dr. Gernot Reishofer

Univ. FÄ Priv.-Doz. Dr.med.univ. Eva Maria Hassler

Graz, 02.07.2025

Eidesstattliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe, andere als die angegebenen Quellen nicht verwendet habe und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Des Weiteren erkläre ich hiermit, dass, sofern bei der Erstellung dieser Arbeit Künstliche Intelligenz (KI) Werkzeuge zur Generierung und/oder Korrektur bestimmter Textpassagen verwendet wurden, dieser Einsatz unter Einhaltung ethischer Grundsätze, akademischer Integrität und den Vorgaben meiner Universität erfolgte, sowie in Folge dies transparent gemacht und in angemessener Weise gekennzeichnet wurde.

Graz, 02.07.2025

Martin Michael Hubner eh.

Danksagungen

Zuallererst möchte ich meinen Eltern meinen tiefsten Dank aussprechen. Ohne ihre unermüdliche Unterstützung, ihr Vertrauen und ihre Opferbereitschaft wäre diese Arbeit nicht möglich gewesen. Sie haben mir nicht nur den Weg zu dieser akademischen Reise geebnet, sondern mir auch stets den Rückhalt und die Motivation gegeben, meine Ziele zu verfolgen.

Ein herzliches Dankeschön gilt Research Prof. Priv.-Doz. DI Dr. Gernot Reishofer für seine hervorragende Betreuung und Unterstützung in allen technischen Fragen. Seine Expertise und wertvollen Ratschläge haben maßgeblich zur Qualität dieser Arbeit beigetragen.

Ebenso möchte ich Univ. FÄ Priv.-Doz. Dr. med. univ. Eva Maria Hassler für ihre fachliche Betreuung und Unterstützung bei den medizinischen Fragestellungen danken. Ihre Einblicke und ihr Engagement haben mir geholfen, die medizinischen Aspekte dieser Arbeit fundiert auszuarbeiten.

Mein Dank gilt auch allen weiteren Personen, die mich in dieser Zeit begleitet und unterstützt haben – sei es durch fachliche Diskussionen, wertvolles Feedback oder einfach durch aufmunternde Worte in herausfordernden Momenten.

Nicht zuletzt danke ich meinen Freunden und meiner Familie für ihre Geduld, ihr Verständnis und ihre ermutigenden Worte während dieser intensiven Zeit. Ihr Zuspruch war eine unverzichtbare Stütze.

Diese Arbeit ist das Ergebnis vieler Stunden des Denkens, Schreibens und Recherchierens – und ohne die Unterstützung all dieser Menschen wäre sie nicht in dieser Form zustande gekommen.

Graz,

Martin Hubner

02.07.2025

Zusammenfassung

Ob bei der automatisierten Auswertung von medizinischen Texten, der Analyse radiologischer Bilder oder der Unterstützung bei Diagnosen – künstliche Intelligenz (KI) findet zunehmend Anwendung in der Medizin. Damit solche Systeme zuverlässig funktionieren, müssen sie medizinische Inhalte korrekt erkennen und interpretieren. Die Grundlage hierfür bildet die sogenannte Annotation – die gezielte Kennzeichnung und Strukturierung medizinischer Daten.. Diese Arbeit bietet eine umfassende Literaturübersicht über Klassifikationen, Terminologien und Ontologien, die in der medizinischen Annotation verwendet werden, und analysiert deren Einfluss auf maschinelle Lernanwendungen. Besonders hervorgehoben wird die zentrale Rolle qualitativ hochwertiger, konsistenter Annotationen, da diese maßgeblich zur Leistungsfähigkeit und Sicherheit von KI-Modellen in der klinischen Praxis beitragen. Die Arbeit untersucht bestehende medizinische Klassifikationssysteme wie *Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme* (ICD), *Systematisierte Nomenklatur der Medizin – Klinische Termini* (SNOMED-CT), *Medical Subject Headings* (MeSH) und andere und bewertet deren Rolle bei der Strukturierung medizinischer Daten für KI-gestützte Anwendungen. Darüber hinaus werden die Bedeutung von Natural Language Processing (NLP) sowie der MATTER-Annotationszyklus im Annotationsprozess erläutert. Neben praxisbezogenen Anwendungen von NLP und KI in der Medizin – darunter klinische Dokumentation, Bildverarbeitung, Patient*inneneinbindung und diagnostische Unterstützung – werden auch ethische und regulatorische Herausforderungen analysiert. Der Fokus liegt hierbei insbesondere auf der europäischen Gesetzgebung sowie auf spezifischen Regularien in Österreich. Die Ergebnisse unterstreichen die Relevanz standardisierter medizinischer Daten und regulatorischer Vorgaben für eine sichere und effektive Implementierung von KI im Gesundheitswesen.

Abstract

Whether it is the automated evaluation of doctor's letters, the analysis of radiological images or support in diagnoses – artificial intelligence (AI) is increasingly being used in medicine. For such systems to function reliably, they must correctly recognise and interpret medical content. The basis for this is the so-called annotation – the targeted labelling and structuring of medical data. This paper provides a comprehensive literature review of classifications, terminologies and ontologies used in medical annotation and analyses their influence on machine learning applications. It highlights the central role of high-quality, consistent annotations, as these significantly contribute to the performance and safety of AI models in clinical practice. The paper examines existing medical classification systems such as *International Statistical Classification of Diseases and Related Health Problems* (ICD), *Systematized Nomenclature of Medicine* (SNOMED-CT) and *Medical Subject Heading* (MeSH) and assesses their role in structuring medical data for AI-based applications. It also explains the importance of natural language processing (NLP) and the MATTER annotation cycle in the annotation process. In addition to practical applications of NLP and AI in medicine – including clinical documentation, image processing, patient engagement and diagnostic support – the project also analyses ethical and regulatory challenges. The focus here is particularly on European legislation and on specific regulations in Austria. The results underscore the relevance of standardised medical data and regulatory requirements for the safe and effective implementation of AI in healthcare.

Inhaltsverzeichnis

Eidesstattliche Erklärung	II
Danksagungen	III
Zusammenfassung	IV
Abstract	V
Tabellenverzeichnis	X
1 Einleitung	1
1.1 Forschungsfragen	1
1.2 Ziele.....	1
1.3 Verarbeitungsgrundlage.....	2
1.4 Inhaltsübersicht	2
2 Grundlagen	4
2.1 Klassifikationen, Terminologien und Ontologien in der Medizin.....	4
2.1.1 ICD-10 und ICD-11 – Klassifikation von Krankheiten	4
2.1.2 International Classification of Health Interventions (ICHI)	6
2.1.3 MeSH – Thesaurus für medizinische Fachbegriffe	8
2.1.4 SNOMED-CT – Struktur, nationale Umsetzung und Anwendungsgebiete	9
2.1.5 FMA- Ontologie der menschlichen Anatomie	11
2.1.6 UMLS – Integration medizinischer Terminologien	13
2.1.7 LOINC – Standardisierung medizinischer Labordaten	17
2.1.8 RxNorm – Arzneimittelinformation in standardisierter Form.....	18
2.2 Grundlagen der natürlichen Sprachverarbeitung.....	22
2.2.1 Historischer Hintergrund.....	22
2.2.2 Verwendungszweck von NLP	23
2.2.3 Aktueller Stand und zukünftige Ausrichtung.....	25
2.3 Der Zyklus der Annotationsentwicklung: M.A.T.T.E.R.....	27
2.3.1 Erstellung des Modells	27

2.3.2	Der Trainingskorpus wird durch Annotation definiert.	28
2.3.3	Der Algorithmus wird trainiert und getestet.	29
2.3.4	Die Ergebnisse werden ausgewertet.	30
2.3.5	Modell und Algorithmus werden überarbeitet.	30
3	Strategien und Anwendungen von Annotation und NLP/KI in der Medizin	32
3.1	Bedeutung hochwertiger annotierter Daten.....	32
3.2	Die Rolle von Annotationen bei der Erstellung von Trainingsdatensätzen.....	34
3.3	Überblick über unterschiedliche Anwendungsmöglichkeiten.....	39
3.4	Praxisnahe Anwendungen von NLP in der Medizin	49
3.4.1	Your.MD.....	49
3.4.2	CURATE.AI	50
3.4.3	Watson for Oncology.....	51
3.4.4	Woebot.....	53
3.5	Verzerrte Ausgaben und Halluzinationen: Grenzen medizinischer KI-Systeme.....	55
3.6	Potential von KI und NLP in der medizinischen Datenverarbeitung.....	57
3.6.1	Veränderungen im Gesundheitssystem durch Digitalisierung: Einsatzmöglichkeiten für NLP	57
3.6.2	Fallstudien und Projekte.....	58
3.7	Regulierungen	66
3.7.1	Überblick über die bestehenden Rechtsrahmen in der EU und in Österreich ..	66
3.7.2	Ethische Aspekte des Einsatzes von KI in der Medizin.....	67
3.7.3	Analyse von Haftungsfragen und Sicherheitsrisiken	68
3.7.4	Vergleich verschiedener Regulierungsrahmen	69
4	Diskussion	70
4.1	Warum ist es für MedizinerInnen relevant, sich mit KI-Anwendungen auseinanderzusetzen?	71
4.2	Wie können MedizinerInnen im Prozess der KI-Entwicklung teilnehmen?	71

4.3	Welche Auswirkungen hat/wird es auf die Arbeit von MedizinerInnen in der Praxis und Theorie haben?	71
4.4	Vergleich der Strategien: Top-to-Bottom und Bottom-to-Top	72
4.4.1	Top-to-Bottom-Ansatz	72
4.4.2	Bottom-to-Top-Ansatz	72
4.5	Hybridansatz: Top-to-Bottom-to-Top.....	73
4.6	Sicherheits- und Qualitätsaspekte	73
4.7	Ausblick	74
5	Zusammenfassung.....	75
5.1	Medizinischer Fokus: Klinische Anwendungen und Auswirkungen auf die Patient*innenversorgung.....	75
5.1.1	Diagnostik und Frühwarnsysteme.....	75
5.1.2	Personalisierte Medizin und KI-gestützte Therapie	76
5.1.3	Administrative Prozesse und Effizienzsteigerung.....	76
5.1.4	Auswirkungen auf die Beziehung zwischen Ärzt*innen und Patient*innen ...	77
5.2	Zugang zu qualitativ hochwertigen Daten	77
5.3	Regulatorische Anforderungen und deren Integration	78
5.4	Ethische Herausforderungen	79
5.5	Zukunftsaussichten.....	80
5.6	Schlusswort	80
	Quellenverzeichnis	82

Abkürzungen

AI	Artificial Intelligence
ALPAC	Automatic Language Processing Advisory Committee
AUI	Number of concept names
HER	Electronic Health Record
FMA	Foundational Model of Anatomy
GTelG	Gesundheitstelematikgesetz
ICD	International Classification of Diseases
ICD-10	International Classification of Diseases, 10th Revision
ICD-11	International Classification of Diseases, 11th Revision
ICF	International Classification of Functioning, Disability, Health
ICHI	International Classification of Health Interventions
IE	Information Extraction
LOINC	Logical Observation Identifiers Names and Codes
LUI	Number of distinct normalized concept names
MeSH	Medical Subject Headings
MLA	Machine Learning Application
MLDS	Member Licensing and Distribution Service
MIPD	Model-informed Precision Dosing
MT	Machine Translation
NCR	National Release Center
NLA	Natural Language Annotation
NLP	Natural Language Processing
RxNorm	Standardized Nomenclature for Clinical Drugs
SCR	Supplementary Concept Records
SNOMED-CT	Systematized Nomenclature of Medicine – Clinical Terms
SUI	Number of distinct concept names
TDM	Therapeutic Drug Monitoring
UMLS	Unified Medical Language System

Tabellenverzeichnis

TABELLE 1: UNTERSCHIEDE ZWISCHEN ICD-10 UND ICD-11	6
TABELLE 2: GENAUIGKEITSMATRIX	30
TABELLE 3: VERGLEICH UNTERSCHIEDLICHER REGULATORISCHER RAHMENBEDINGUNGEN	69

1 Einleitung

1.1 Forschungsfragen

Die Annotation von Texten ist eine essenzielle Methode zur Anreicherung von Daten mit Metadaten, um die Möglichkeiten der natürlichen Sprachverarbeitung (NLP) zu erweitern. Annotationen dienen als Hinweisgeber (pointers), indem sie relevante Textabschnitte, Begriffe und Phrasen markieren. Dadurch können präzisere Daten aus großen Datenpools extrahiert und für Machine Learning Applications (MLAs) zugänglich gemacht werden (Pustejovsky and Stubbs, 2012).

MLAs gewinnen in der medizinischen Forschung zunehmend an Bedeutung (Miotto et al., 2017). Bereits in den 1970er Jahren wurden medizinische Bilder digitalisiert und erste Algorithmen zur Verarbeitung entwickelt (Litjens et al., 2017). Heute sind MLAs in der Lage, klinische Texte zu analysieren und komplexe Muster zu erkennen. Entscheidend für die Leistungsfähigkeit dieser Systeme ist jedoch der Zugang zu hochwertigen, annotierten Daten. Die Qualität dieser Annotationen hat direkten Einfluss auf die Präzision der Modelle – eine unzureichende Annotation kann zu fehlerhaften Analysen führen, was in der Medizin schwerwiegende Folgen haben kann (Krenzer et al., 2022).

1.2 Ziele

Die vorliegende Arbeit verfolgt das Ziel, einen umfassenden Überblick über die Bedeutung von Annotationen für NLP-gestützte medizinische Anwendungen zu geben. Dabei stehen folgende Aspekte im Fokus:

- Untersuchung bestehender Klassifikationen, Nomenklaturen und Ontologien im medizinischen Bereich und deren Einfluss auf die Annotationseffizienz.
- Analyse des Zusammenhangs zwischen Annotationsqualität und der Leistungsfähigkeit von MLAs in der Medizin.
- Identifikation zentraler Herausforderungen im Annotationsprozess und deren Optimierungsmöglichkeiten durch bestehende medizinische Standards wie ICD, SNOMED-CT, MeSH oder UMLS.
- Betrachtung von NLP-Technologien und ihrer Rolle in der medizinischen Forschung und Praxis.
- Evaluation der regulatorischen und ethischen Rahmenbedingungen für den Einsatz von KI in der Medizin, insbesondere in der EU und Österreich.

Die Arbeit stützt sich auf eine detaillierte Literaturanalyse, um sowohl technische als auch praxisbezogene Herausforderungen zu identifizieren. Dabei wird nicht nur die methodische Umsetzung von NLP in der Medizin analysiert, sondern auch deren praktische Relevanz für den klinischen Alltag diskutiert. Besondere Aufmerksamkeit gilt der Qualität medizinischer Annotationen, da diese die Basis für verlässliche MLAs bildet.

1.3 Verarbeitungsgrundlage

Im Fokus dieser Arbeit steht eine umfassende Literaturrecherche zu etablierten medizinischen Klassifikationen wie ICD, SNOMED CT und MeSH, zu ontologischen Systemen wie FMA oder UMLS sowie zu aktuellen NLP-Methoden, die für die semantische Erfassung medizinischer Sprache verwendet werden. Ziel ist es, deren jeweilige Struktur, Zielsetzung und Eignung für die Annotation medizinischer Inhalte systematisch zu analysieren und ihre Rolle im Entwicklungsprozess KI-gestützter Systeme kritisch zu bewerten.

Ein wesentlicher Teil dieser Arbeit ist die systematische Analyse von Klassifikationssystemen wie zum Beispiel ICD, SNOMED-CT, MeSH und UMLS, um deren Anwendungsmöglichkeiten im Annotationsprozess zu bewerten. Darüber hinaus werden aktuelle NLP-Methoden in der medizinischen Datenverarbeitung analysiert.

Die Forschung basiert auf einer Literaturrecherche, bei der wissenschaftliche Publikationen, regulatorische Dokumente und medizinische Standards systematisch im Hinblick auf ihre Inhalte, strukturellen Eigenschaften und Anwendungsbereiche, sowie der praktischen Erfahrung bei der Annotierung für diverse Forschungsprojekte ausgewertet werden. Ziel ist es, bestehende Herausforderungen bei der Annotation medizinischer Daten herauszuarbeiten und Strategien zu identifizieren, wie diese durch geeignete Klassifikationen, Ontologien und NLP-Verfahren optimiert werden können.

1.4 Inhaltsübersicht

Zuerst werden die theoretischen und methodischen Grundlagen dieser Arbeit behandelt. Im Zentrum steht die Analyse von medizinischen Klassifikationen, Terminologien und Ontologien wie ICD-10, ICD-11, SNOMED CT und MeSH, die eine zentrale Rolle bei der Strukturierung und Standardisierung medizinischer Daten im Annotationsprozess einnehmen. Darüber hinaus wird auf weiterführende Systeme wie FMA, UMLS, LOINC und RxNorm eingegangen, um ein umfassendes Verständnis der semantischen Dimension medizinischer Informationen zu ermöglichen.

Ein weiterer Fokus liegt auf den Technologien des Natural Language Processing (NLP) sowie deren Entwicklung, Anwendungsfeldern und technischen Grundlagen. Das MATTER-Modell wird als strukturierter Entwicklungszyklus für medizinische Annotationen vorgestellt und in seiner praktischen Relevanz erläutert.

Im Anschluss wird die Bedeutung qualitativ hochwertiger Annotationen für den Erfolg KI-gestützter Anwendungen thematisiert. Reale Anwendungsfälle wie die Analyse elektronischer Gesundheitsakten, medizinische Bildverarbeitung und die Entwicklung diagnostischer Unterstützungssysteme verdeutlichen die praktische Relevanz. Ergänzend dazu werden die regulatorischen und ethischen Rahmenbedingungen beleuchtet – insbesondere im Hinblick auf die europäische Gesetzgebung (AI Act, DSGVO) und spezifische Vorgaben in Österreich. Dabei stehen Themen wie Verantwortung, Datenschutz und algorithmische Fairness im Mittelpunkt.

Das Kapitel schließt mit einer kritischen Reflexion der zentralen Erkenntnisse und einem Ausblick auf zukünftige Entwicklungen im Bereich KI und Annotation in der Medizin.

2 Grundlagen

2.1 Klassifikationen, Terminologien und Ontologien in der Medizin

Klassifikationen, Terminologien und Ontologien sind eingehende Begleiter des medizinischen Alltags. Medizinische Klassifikationen dienen der Umwandlung von Beschreibungen medizinischer Diagnosen oder Verfahren in standardisierte Codes in einem Prozess. (World Health, 2004) Diagnoseklassifikationen stellen ein System zur Erfassung von Diagnosecodes dar, welches zur Identifizierung eines umfangreichen Spektrums von Gesundheitszuständen dient. Dazu zählen z.B. chronische Erkrankungen wie Diabetes mellitus und koronare Herzkrankheit sowie Infektionskrankheiten wie Influenza, COVID-19 und Norovirus.. Parallel dazu umfassen Verfahrensklassifikationen Codes, die spezifische medizinische Interventionen erfassen. Diese Diagnose- und Verfahrenscodes dienen als grundlegende Bausteine für eine Vielzahl von Anwendungen in der Medizin, im öffentlichen Gesundheitswesen und in der medizinischen Informatik. Sie werden von Gesundheitsdienstleistern, staatlichen Gesundheitsprogrammen, privaten Krankenversicherungen, Unfallversicherungsträgern, Softwareentwicklern und anderen maßgeblichen Akteuren genutzt, um Daten zu erfassen, Entscheidungen zu treffen und die Qualität der Gesundheitsversorgung zu verbessern. Die medizinische Terminologie wird als Sprache zur präzisen Beschreibung des menschlichen Körpers mit all seinen Bestandteilen, Prozessen, Zuständen und Verfahren verwendet. Diese weist eine recht regelmäßige Morphologie auf, wie die Verwendung von denselben Präfixen und Suffixen, um Wortstämmen eine Bedeutung zu verleihen. (OpenStax, 2025) Wurzeln, Präfixe und Suffixe stammen häufig aus dem Griechischen oder Lateinischen und weichen oft stark vom üblichen Sprachgebrauch ab. (Aapc, 2025) Eine Ontologie ist eine formale Darstellung einer Vielzahl heterogener Konzepte. Klinische Ontologien gehen jedoch über die bloße Abbildung klassifizierter Terminologien hinaus, indem sie auch Klassifikationen und Nomenklaturen einschließen. Sie bieten nicht nur ein kontrolliertes Vokabular, sondern etablieren auch Beziehungen zwischen den Konzepten, die es Computersystemen ermöglichen, Schlussfolgerungen zu ziehen. Dadurch können die verschiedenen Akteure im Gesundheitssektor komplexe Anfragen effizient beantworten. (National Library of, 2025b)

2.1.1 ICD-10 und ICD-11 – Klassifikation von Krankheiten

Die Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (ICD) ist ein Referenzstandard für die Erfassung von Krankheiten und Gesundheitszuständen. Sie dient als Grundlage für die Identifizierung und den internationalen

Vergleich von Gesundheitstrends und Gesundheitsstatistiken (Otero Varela et al., 2021). Bei jedem Kontakt mit dem Gesundheitssystem werden die gesundheitspezifischen Informationen der/des Patient*in in der administrativen Gesundheitsdatenbank mit Hilfe der ICD kodiert. Diese Daten können wiederum für Krankheitsüberwachung, Forschung, Entscheidungsfindung, Ressourcenallokation, Qualität und Sicherheit verwendet werden (Cadarette and Wong, 2015). Für die Forschung bieten die ICD-Codes die Möglichkeit, MLAs spezifisch für bestimmte Krankheiten und auch Bereiche zu bilden.

Die zehnte Version der ICD (ICD-10) ist die weltweit am häufigsten verwendete Klassifikation, deren exakte Ausgestaltung von Land zu Land unterschiedlich ist und für die klinische Modifikationen vorgenommen wurden, um den länderspezifischen Bedürfnissen gerecht zu werden (z. B. Australien ICD-10-AM, USA ICD-10-CM, Deutschland ICD-10-GM, Kanada ICD-10-CA, Thailand ICD-10-TM, Korea ICD-10-CM). Auch nach den Versuchen der WHO, die länderspezifischen Änderungen der ICD-10 zu regeln, sind gravierende Unterschiede zwischen den Versionen aufgetreten. Strukturelle Unterschiede zeigen sich unter anderem in einer größeren Anzahl von Codes im Vergleich zur ursprünglichen Version der WHO. Ein weiterer Unterschied zwischen den internationalen ICD-10 Modifikationen ist die Definition der "main condition" (Hauptkrankheit). In der zweiten Revision der WHO wird die "Hauptkrankheit" als eine Erkrankung beschrieben, die am Ende der Behandlungsepisode diagnostiziert wird und die in erster Linie für die Behandlung oder Untersuchung der/des Patient*in verantwortlich ist. Kanada und Thailand folgen der WHO-Definition. Deutschland, die USA und Australien verwenden stattdessen den Begriff "Hauptdiagnose" und definieren ihn als die Krankheit, die nach der Untersuchung als Hauptursache für den Krankenhausaufenthalt festgestellt wurde (Jetté et al., 2010).

Im Jahr 2022 wurde die elfte Version der ICD von der WHO veröffentlicht. Es wird erwartet, dass alle Mitgliedsstaaten der WHO die aktuelle Version der ICD für die Meldung von Todes- und Krankheitsfällen verwenden (World Health, 1967).

ICD-11 gewährleistet dabei die Interoperabilität von Einzeldaten und Wiederverwendbarkeit gespeicherter Daten für Anwendungen außerhalb der Gesundheitsstatistik. Es umfasst beispielsweise die Entscheidungshilfe, die Ressourcenzuteilung, die Kostenerstattung und Leitlinien. Im Vergleich zur zehnten Version der ICD ergeben sich folgende wesentliche Änderungen, die in der folgenden Tabelle aufgeführt sind:

Eigenschaften	ICD-10	ICD-11
Kodierschema	Kapitelnummerierung in römischen Ziffern	Kapitelnummerierung in Arabisch
	Dreistellige Kategorien, die weiter in bis zu 10 vierstellige Unterkategorien unterteilt werden können	Stammcode (Kategorie) besteht aus vier Zeichen mit zwei Ebenen von Unterkategorien
	Alphanumerischer Code mit einem Buchstaben an erster Stelle und einer Zahl an den folgenden Positionen, wobei der vierte Zeichen einem Dezimalpunkt folgt	Alphanumerischer Code mit einem Buchstaben an zweiter Stelle und einer Zahl an der dritten Zeichenposition, beginnend mit einer erzwungenen Zahl, um unerwünschte Wörter zu verhindern. Erweiterungs-codes beginnen mit 'X'
Kodierstruktur	Der erste Zeichen des Codes ist ein Buchstabe, der nicht mit der Kapitelnummer in Beziehung steht	Der erste Zeichen des Codes bezieht sich auf das Kapitel
	Restkategorie durch .8 gekennzeichnet, nicht näher bezeichnete Kategorie durch .9	Restkategorie durch Endbuchstabe 'Y' für 'anderweitig spezifiziert' und 'Z' für 'nicht näher bezeichnet'
Terminologie	Bereich von Ausdrücken für kausale Beziehungen im Codetitel	Bevorzugter Begriff 'bedingt durch' für kausale Beziehungen
	Bereich von Ausdrücken für das gleichzeitige Auftreten von Bedingungen im Codetitel	Bevorzugter Begriff 'assoziiert mit' für das gleichzeitige Auftreten von Bedingungen im Codetitel

TABELLE 1: UNTERSCHIEDE ZWISCHEN ICD-10 UND ICD-11

Weitere allgemeine Unterschiede zeigen sich in der Beschreibung der Kategorien. Alle Kategorien der ICD-11 haben eine kurze und eine lange Beschreibung. Die Kurzbeschreibung erklärt in maximal 100 Wörtern die Bedeutung der Kategorie und erscheint auch in der gedruckten Form der ICD-11. Die lange Beschreibung hingegen hat keine Längenbeschränkung und enthält detaillierte Informationen, die wiederum Auskunft über Anatomie, Ätiologie und andere Aspekte geben (World Health, 2022).

2.1.2 International Classification of Health Interventions (ICHI)

Die Internationale Klassifikation für Prozeduren (ICHI) ist ein umfassendes System zur systematischen Kategorisierung von medizinischen Verfahrenscodes, das von der Weltgesundheitsorganisation (WHO) konzipiert und seit 2012 entwickelt wird. Diese Initiative entstand als Antwort auf die Notwendigkeit eines aktualisierten und standardisierten Systems,

um die Klassifizierung von diagnostischen, medizinischen und chirurgischen Eingriffen zu verbessern. Frühere Bemühungen, wie die Internationale Klassifikation für Verfahren in der Medizin (ICPM), die seit 1978 in Gebrauch war, konnten nicht mehr den modernen Anforderungen gerecht werden und wurden daher eingestellt. (World Health, 2025)

Die Entwicklung der ICHI als Ersatz für die veraltete ICPM markierte einen bedeutenden Schritt in der internationalen Gesundheitsklassifikation. Infolgedessen begannen Länder, die auf individuellen Systemen wie der OPS-301 basierten, der eine Weiterentwicklung der niederländischen Erweiterung (ICPM-DE) der WHO darstellt. (Federal Institute for and Medical, 2025) Dies verdeutlichte die Notwendigkeit eines einheitlichen und aktuellen Standards auf globaler Ebene.

Die ICHI zielt darauf ab, eine umfassende und einheitliche Grundlage für die Codierung von medizinischen Verfahren bereitzustellen, die es Gesundheitssystemen auf der ganzen Welt ermöglicht, einheitliche und vergleichbare Daten zu erfassen und zu analysieren. Ihre fortlaufende Entwicklung spiegelt das Engagement der WHO wider, die Qualität und Effizienz der Gesundheitsversorgung weltweit zu verbessern und den Bedürfnissen einer sich wandelnden medizinischen Landschaft gerecht zu werden.

Die *Internationale Klassifikation für Gesundheitseingriffe (ICHI)* ist eine statistische Klassifikation, die eine breite Palette von Interventionen in sämtlichen Bereichen der Gesundheitssysteme abdeckt und damit im Einklang mit den beiden anderen WHO-FIC-Referenzklassifikationen, ICD und ICF, steht. Zusätzlich zu Verfahrenscodes für diagnostische, medizinische und chirurgische Eingriffe umfasst die ICHI auch Interventionen im Bereich der Psychiatrie, der primären Gesundheitsversorgung, der Gesundheitsförderung, der Funktionsunterstützung, der Rehabilitation und der Prävention. Sie bietet somit eine umfassende Darstellung der Maßnahmen im Gesundheitswesen. Die in der ICHI enthaltenen Interventionen werden als Stammcodes bezeichnet und umfassen eine Vielzahl von Maßnahmen, die anhand von drei Achsen beschrieben werden:

- Target: Die Entität, an der die Action ausgeführt wird.
- Action: Die Tat, die vom Akteur an der zuvor genannten Entität ausgeführt wird.
- Means: Die Verfahren und Methoden, mit denen die Action ausgeführt wird.

Jede Achse besteht dabei aus einer kodierte Liste von beschriebenen Kategorien. Ein ICHI-Stammcode setzt sich aus einem Titel und einem eindeutigen siebenstelligen Code zusammen, der die Achsenkategorien für die jeweilige Intervention angibt. Zum Beispiel steht der Code "HBA.AD.AF" für "Percutaneous transluminal left ventricular biopsy", wobei "HBA" für den

linken kardialen Ventrikel, "AD" für eine Biopsie und "AF" für einen perkutanen transluminalen Zugang stehen.

Bei Bedarf können Erweiterungs-codes verwendet werden, um zusätzliche Informationen über eine Intervention hinzuzufügen. Diese beinhalten Codes für therapeutische und unterstützende Produkte sowie Informationen zur Dosierung Lateralität und detaillierten Beschreibung der Anatomie. Zusätzlich können Codes von anderen Klassifikationen wie LOINC (International Standard Industrial Classification and the Central Product Classification) als Erweiterungscode genutzt werden. (World Health, 2022)

2.1.3 MeSH – Thesaurus für medizinische Fachbegriffe

Die Medical Subject Headings (MeSH) sind ein polyhierarchischer, konzeptbasierter Thesaurus, der von der National Library of Medicine (NLM) herausgegeben und gepflegt wird. Er dient zur systematischen Erfassung und internationalen Klassifikation von medizinischen Inhalten. MeSH soll dabei folgende Aktivitäten unterstützen (Ivanović and Budimac, 2014):

- Artikel aus den führenden biomedizinischen Zeitschriften werden indiziert und in den Datenbanken MEDLINE/PubMed, NLM Catalog und weiteren NLM-Datenbanken verfügbar gemacht.
- Erstellung einer Datenbank, die als Katalog für die von der Bibliothek erworbenen Bücher, Dokumente und audiovisuellen Medien fungiert.

Im Jahr 2024 umfasst die Medical Subject Headings (MeSH) insgesamt 30.764 Hauptüberschriften, wovon 311 als neu hinzugefügt gelten. Ebenso verzeichnet man 322.838 Supplementary Concept Records (SCRs), auch als SCRs bekannt, von denen 367 als neu registriert wurden. Die jüngsten Erweiterungen des MeSH-Thesaurus sind in verschiedenste Zweige unterteilt, wobei eine signifikante Anzahl in den Bereich der Chemikalien und Drogen fällt. (National Library of, 2024)

Die MeSH-Deskriptoren sind in 16 Hauptkategorien unterteilt:

- a. Anatomische Begriffe (Kategorie A)
- b. Organismen (Kategorie B)
- c. Krankheiten (Kategorie C)
- d. Arzneimittel und Chemikalien (Kategorie D)
- e. Analytische, diagnostische und therapeutische Techniken und Geräte (Kategorie E)
- f. Psychiatrie und Psychologie (Kategorie F)
- g. Phänomene und Vorgänge (Kategorie G)
- h. Disziplinen und Berufe (Kategorie H)

- i. Anthropologie, Pädagogik, Soziologie und soziale Phänomene (Kategorie I)
- j. Technik, Industrie und Landwirtschaft (Kategorie J)
- k. Geisteswissenschaften (Kategorie K)
- l. Informationswissenschaft (Kategorie L)
- m. Benannte Gruppen (Kategorie M)
- n. Gesundheitsfürsorge (Kategorie N)
- 1. Publikationsmerkmale (Kategorie V)
- o. Geographisches (Kategorie Z)

Innerhalb dieser Hauptkategorien sind die Deskriptoren hierarchisch von der allgemeinsten bis zur spezifischsten Ebene organisiert. Aufgrund der verzweigten Struktur dieser Hierarchien werden sie gelegentlich als "Bäume" bezeichnet. Jeder MeSH-Deskriptor erscheint mindestens an einer Stelle in den Bäumen und kann an so vielen weiteren Stellen erscheinen, wie es sachdienlich ist. Bei der Verwendung von MeSH ist es ratsam, den spezifischsten MeSH-Deskriptor zu finden, der dem jeweiligen Konzept entspricht. (National Library of, 2025a)

2.1.4 SNOMED-CT – Struktur, nationale Umsetzung und Anwendungsgebiete

SNOMED-CT ist die Abkürzung für Systematized Nomenclature of Medicine - Clinical Terms. SNOMED CT stellt gegenwärtig die umfangreichste Gesundheitsterminologie weltweit dar, bestehend aus über 350.000 eindeutig identifizierten und logisch definierten Begriffen, die hierarchisch strukturiert sind. Aufgrund ihrer umfassenden Abdeckung nimmt sie eine entscheidende Position bei der Gestaltung einer Plattform ein, die eine präzise Interaktion mit klinischen Daten sowie deren Analyse ermöglicht. Die Verwaltung und Verbreitung dieser Terminologie liegt in der Verantwortung von SNOMED International (ehemals IHTSDO), einer gemeinnützigen internationalen Organisation, die im Besitz ihrer Mitglieder ist und sich engagiert für die Förderung und Entwicklung im Gesundheitsbereich einsetzt. Österreich ist seit dem Jahr 2018 ein Mitglied von SNOMED International. In diesem Zusammenhang stellt das Management von Terminologien, darunter die Entwicklung, Verwaltung, Pflege und Veröffentlichung, eine essenzielle Grundvoraussetzung für die Errichtung und den Betrieb von elektronischen Gesundheitsakten wie beispielsweise ELGA dar.

SNOMED-CT ist eine Klassenhierarchie von Konzepten, die verschiedene hochrangige Kategorien umfasst, darunter klinischer Befund/Störung, Verfahren/Interventionen, beobachtbare Entitäten, Körperstruktur, Organismus, Substanz, pharmazeutisches/biologisches Produkt, Probe, spezielles Konzept, physisches Objekt, physische Kraft, Ereignis, Umwelt-/geografischer Ort, sozialer Kontext und Stadien/Skalen. Diese Kategorien dienen dazu,

medizinische und gesundheitsbezogene Begriffe zu organisieren und zu strukturieren, um eine einheitliche und umfassende Beschreibung von medizinischen Informationen zu ermöglichen. In Österreich kann SNOMED CT innerhalb der nationalen Infrastruktur uneingeschränkt genutzt werden, wodurch ein semantischer Austausch relevanter Informationen zwischen den Gesundheitsanbietern ermöglicht wird. Dabei fungiert das National Release-Center (NRC) als Vermittler und erste Anlaufstelle für Anfragen von Benutzer*innen und anderen Mitgliedsländern. Die Einführung von SNOMED CT in Österreich hat die Interoperabilität der Gesundheitsdaten erweitert.

Die Hauptaufgaben des National Release Center (NRC) umfassen den Kontaktpunkt, Lizenzmanagement, Change Request Management und National Extensions.

Der Kontaktpunkt stellt die erste Anlaufstelle bei Fragen jeglicher Art zu SNOMED-CT dar.

Die Nutzung von SNOMED CT in Österreich erfordert entweder eine Affiliate-Lizenz oder eine Sublizenz, die über den Member Licensing and Distribution Service (MLDS) beantragt werden kann. Affiliate-Lizenzen gestatten die direkte Weitergabe von Sublizenzen an Endbenutzer*innen unter Einhaltung der Bedingungen des Affiliate License Agreements.

SNOMED International (ehemals IHTSDO) bietet das „Content Request Service (CRS)“ für Change-Request-Anträge an. Fehlende oder inkorrekte Konzepte, Beschreibungen oder Beziehungen sollten an das National Release-Center (NRC) gerichtet werden. Genehmigte Änderungen werden nach einer Prüfung durch die ELGA GmbH und SNOMED International in zukünftige internationale Veröffentlichungen integriert.

Falls Konzepte, die länderspezifisch sind, für die internationale Version nicht geeignet sind, besteht die Option, diese in einer sogenannten „National Extension“ zu integrieren. Die Erstellung solcher Erweiterungen obliegt dem entsprechenden NRC und erfolgt bei Bedarf über das Member Licensing and Distribution Service (MLDS). Diese Konzepte, die ihren Ursprung in den National Extensions haben, werden mit einem spezifischen nationalen "SNOMED-CT Identifier" versehen, der es ermöglicht, sie den verschiedenen Ländern zuzuordnen.

Im Rahmen der Mitgliedschaft bei SNOMED International übernehmen die Mitgliedsländer in der Regel bestimmte Aufgaben, die zur effektiven Nutzung und Verwaltung von SNOMED-CT beitragen. Dazu zählen insbesondere:

- Einrichten einer zentralen Anlaufstelle innerhalb eines Mitgliedslandes für SNOMED-CT-Anwender. Diese bietet Unterstützung in Bezug auf Implementierung und Nutzung, sammelt Feedback, bearbeitet Change Requests und fungiert als Vermittler zwischen Affiliates und der SNOMED-CT Community.

- Diese fungiert als Hauptansprechpartner für SNOMED International bezüglich des Terminologiemanagements und fördert die Beteiligung von Experten sowie die Kommunikation mit Affiliates und Fachleuten.
- Entwicklung von Mechanismen, um den Benutzer*innen die Distribution und Lizenzierung von SNOMED-CT zu erleichtern, einschließlich der Verwaltung der jährlichen Lizenzberichte.
- Gewährleistung der Einhaltung der SNOMED International Standards für sämtliche Terminologieprodukte.

In elektronischen Gesundheitsakten findet SNOMED-CT Anwendung zur Eingabe, Aufbewahrung und Anzeige/Abruf von Daten, zur Datenanalyse sowie zur Kommunikation zwischen Fachpersonal. Als Terminologieservice für die Suche nach Begriffen, die Erstellung von Abfragen und Ableitungen sowie deren Integration in Anwendungen. Zur Aggregation, Analyse und Forschung von Daten, zur Erstellung von Statistiken sowie zur Entwicklung von Richtlinien.

Vergleichsweise gebrauchen folgende Länder SNOMED-CT großflächig:

- Neuseeland nutzt SNOMED-CT in Ambulanzdiensten im ganzen Land zur mobilen elektronischen Erfassung von auf SNOMED-basierten klinischen Eindrücken, Beobachtungen und Interventionen mittels Tablets. (International, 2025c)
- Australien verwendet SNOMED-CT als National Clinical Terminology zur Kodierung von in Australien gebräuchlichen Medikamenten mit Integrierung der Australian Medicines Terminology in SNOMED-CT-AU. (International, 2025a)
- In Singapur wurde eine Big Data-Softwareplattform erfolgreich implementiert und ist nun in Betrieb. Diese Plattform ermöglicht das Durchsuchen von Gesundheitsdaten, die Gruppierung von Patient*innen mit ähnlichen Charakteristika in Kohorten und die Identifizierung von Trends sowie Korrelationen. SNOMED CT wird hier verwendet, um klinische Informationen zu speichern und abzurufen, Abfragen durchzuführen und aufgrund seiner Mappings zu bestehenden Terminologien und Klassifikationen (wie ICD-10, ICD-10-CM, ICD-O, ICD-9 und ICD-9-CM) eine umfassende Analyse zu ermöglichen. (International, 2025b)

2.1.5 FMA- Ontologie der menschlichen Anatomie

Die Ontologie Foundational Model of Anatomy (FMA) fungiert als etablierte und umfassende Wissensressource im Bereich der biomedizinischen Informatik. Sie bietet Klassen, Typen und deren Beziehungen an, die entscheidend sind, um die phänotypische Struktur des menschlichen

Körpers symbolisch darzustellen. Diese Darstellung ist sowohl für Menschen als auch für maschinelle Systeme zugänglich, navigierbar, analysierbar und interpretierbar. Die FMA ist eine Domänenontologie, die das explizite deklarative Wissen über die menschliche Anatomie in einem kohärenten Gesamtbild präsentiert. Ihr ontologischer Rahmen kann auf andere Spezies angewendet und erweitert werden.

Die Ontologie des Grundmodells der Anatomie ist eine umfangreiche Wissensquelle mit über 75.000 Klassen und 120.000 Begriffen sowie 2,1 Millionen Beziehungsinstanzen. Sie bietet ein äußerst kohärentes und symbolisches Modell mit einer Vielzahl von Beziehungstypen. Die Anatomie-Taxonomie (AT) ist die umfassendste Komponente und umfasst dominierende Klassen wie die Anatomische Struktur, die eine breite Palette biologischer Entitäten abdeckt. Makroskopische anatomische Strukturen werden ausführlich behandelt, wobei biologische Moleküle nur zur Veranschaulichung des strukturellen Kontinuums verwendet werden. Das Ziel des Foundational Model Explorers/Conducted Tours ist es, ein tiefes Verständnis für die verschiedenen Klassen zu vermitteln und sie in einer kontinuierlichen Taxonomie zusammenzufassen. Die FMA definiert Körpersubstanzen, Räume und andere relevante Konzepte in Bezug auf ihre Beziehung zu anatomischen Strukturen. Durch die Klassifizierung und Definition von Klassen oder Typen anatomischer Einheiten, die in traditionellen Quellen nicht explizit behandelt werden, ergänzt die FMA die medizinische Forschung. Es besteht eine zunehmende Übereinstimmung zwischen der FMA und traditionellen Quellen, was die Kompetenz und Autorität der FMA in der Anatomie unterstreicht. Die FMA enthält im Wesentlichen alle Begriffe der Terminologia Anatomica und zeigt deren Ableitung aus dieser Quelle auf.

Im Gegensatz zu herkömmlichen Quellen wie Atlanten, Lehrbüchern, Wörterbüchern, Thesauri oder Begriffslisten, bietet das FMA eine umfassendere und präzisere Darstellung. Während Atlanten und Lehrbücher sich an eine bestimmte Benutzer*innengruppe richten, ist das FMA für alle zugänglich und bietet eine Möglichkeit anatomische Informationen zu erfassen. So bietet diese umfassenden Informationen für Krankenpflegeschüler, Medizinstudenten, Chirurgen und Radiologen gleichermaßen. Das FMA stellt anatomische Informationen auf eine einzigartige Weise dar, indem er nicht nur beschriftete Bilder verwendet, sondern auch nicht-grafische Symbole wie Begriffe und numerische Identifikatoren einsetzt. In Kombination mit volumetrischen Datensätzen wie dem Visible Human oder Bildern aus Atlanten der makroskopischen und mikroskopischen Anatomie können so noch detailliertere Informationen vermittelt werden. Anatomie-Lehrbücher konzentrieren sich neben den anatomischen Informationen zusätzlich auf nicht-anatomische Informationen wie physiologische Funktionen,

pathologische Läsionen oder klinische Fälle. Die FMA modelliert die strukturelle Organisation des Körpers in einem bewusst strukturellen Kontext und konzentriert sich dabei auf die strukturellen Aspekte. In Anwendungen, die für spezifische Benutzer*innengruppen entwickelt wurden, wie beispielsweise Medizinstudenten, können nicht-anatomische Informationen mit dem anatomischen Inhalt der FMA verknüpft werden. Im Vergleich zu Wörterbüchern, die begriffsorientiert sind, ist das FMA klassen- oder typorientiert und ordnet seine Klassen in einer Vererbungshierarchie oder Taxonomie in einem streng strukturellen Kontext an. Die Klassen oder Typen dieser Taxonomie werden auf der Grundlage der Struktureigenschaften festgelegt, die die Mitglieder einer Klasse gemeinsam haben.

Im Gegensatz dazu ordnen Thesauri ihren Inhalt nach der Bedeutung ihrer Begriffe an. Diese sind zwar nicht explizit definiert, aber jede/jeder Benutzer*in kann sie auf der Grundlage der wahrgenommenen Ähnlichkeiten und Unterschiede zwischen den Begriffen ableiten. Die FMA definiert stattdessen die Klassen ihrer Taxonomie explizit und verknüpft alle diese Klassen über eine Vererbungshierarchie mit einer gemeinsamen Wurzel.

Die Ontologie des FMA umfasst vier miteinander verbundene Komponenten:

1. Anatomische Taxonomie (At): Klassifiziert anatomische Einheiten basierend auf gemeinsamen Merkmalen (Genus) und Unterscheidungsmerkmalen (Differentia). Früher als Anatomie-Ontologie (Ao) bekannt.
2. Anatomische strukturelle Abstraktion (ASA): Spezifiziert Teil-Ganzes- und räumliche Beziehungen zwischen den in At dargestellten Entitäten.
3. Anatomische Transformationsabstraktion (ATA): Beschreibt morphologische Transformationen von Entitäten in At während der pränatalen Entwicklung und des postnatalen Lebenszyklus.
4. Metawissen (Mk): Definiert die Prinzipien, Regeln und Definitionen, die die Darstellung von Klassen und Beziehungen in den anderen drei Komponenten der FMA leiten.

2.1.6 UMLS – Integration medizinischer Terminologien

Das UMLS steht für Unified Medical Language System. Dieses ist eine umfassende Zusammenstellung von Dateien und Software, wobei Letztere speziell entwickelte Tools und Schnittstellen umfasst, die der Integration und Verknüpfung medizinischer Terminologien dienen. Durch eine Zuordnungsstruktur zwischen diesen Vokabularen wird eine Übersetzung zwischen den unterschiedlichen Terminologiesystemen ermöglicht. In einem erweiterten Kontext kann das UMLS daher als eine umfassende Thesaurus- und Ontologie-Sammlung

biomedizinischer Konzepte angesehen werden. Des Weiteren bietet das UMLS eine Fülle an Vokabular, Daten und Relationen, die die Durchführung von Prozessen und die Fortschritte im Bereich der Natural Language Processing (NLP) unterstützen.

Die Statistik der UMLS-Publikation 2023AB enthält eine Vielzahl von Metadaten. Die Version 2023AB liegt im RRF-Format vor und enthält 3.361.615 Konzepte. Diese Konzepte werden durch 15.888.629 Atom Unique Identifiers (AUIs) repräsentiert, wobei 13.147.630 String Unique Identifiers (SUIs) und 11.972.841 Lexical (term) Unique Identifiers (LUIs) unterschieden werden können. Insgesamt tragen 166 Quellenfamilien aus 185 Quellen zur Benennung der Konzepte bei, die in 27 verschiedenen Sprachen vertreten sind.

Die Auszählung der Namen nach Sprachen zeigt, dass Englisch den größten Anteil ausmacht, gefolgt von Spanisch, Portugiesisch, Französisch und Japanisch. Der Source Restriction Level (SRL) zeigt, dass die meisten Konzepte keiner Beschränkung unterliegen. Die meisten Konzepte sind nicht unterdrückbar, gefolgt von teilweise unterdrückbaren Konzepten. Die Verteilung der Quellen nach Sprachen zeigt, dass Englisch die dominierende Sprache ist, gefolgt von Spanisch, Deutsch, Französisch und Niederländisch. Andere Sprachen haben einen geringeren Anteil. Dies ist insofern bedeutsam, als die Sprache eine wesentliche Rolle beim Annotierungsprozess spielt – insbesondere in Bezug auf sogenannte Low-resource languages, bei denen aufgrund begrenzter terminologischer Ressourcen und Vokabularabdeckungen zusätzliche Herausforderungen für die semantische Erfassung und automatisierte Verarbeitung bestehen (National Library of, 2025d)

Das UMLS besteht aus drei verschiedenen Wissensquellen: Metathesaurus, Semantisches Netzwerk und SPECIALIST Lexikon. Die größte Komponente ist dabei der Metathesaurus. Dieser umfasst dabei 187 Einträge, ausgewählte Beispiele für die integrierten kontrollierten Vokabel sind: SNOMED-CT, ICD, MeSH, LOINC, HPO, Gene Ontology. (National Library of, 2025e)

Der Metathesaurus ist ein umfassender biomedizinischer Thesaurus, der Konzepte oder Semantiken systematisch organisiert und synonyme Bezeichnungen für identische Konzepte miteinander verknüpft. Dies wird erreicht, indem jedem Konzept spezifische Attribute zugewiesen werden, die nicht nur seine Bedeutung definieren, sondern auch die Verknüpfung mit verschiedenen Quellvokabularen ermöglichen. Auf diese Weise können zahlreiche Beziehungen zwischen Konzepten dargestellt werden, darunter hierarchische Beziehungen wie "isa" für Unterklassen und "is part of" für Untereinheiten sowie assoziative Beziehungen wie "is caused by" oder "in the literature often occurs close to". (National Library of, 2009a)

Das semantische Netzwerk ist ein Verzeichnis, das eine Vielzahl weitreichender Themenkategorien oder semantischer Typen sowie eine Reihe nützlicher und bedeutender Beziehungen oder semantischer Verknüpfungen zwischen diesen Typen umfasst. Die semantischen Typen bieten eine konsistente Kategorisierung aller Konzepte, die im UMLS-Metathesaurus vertreten sind. Insgesamt umfasst das Netzwerk 127 semantische Typen und 54 Beziehungen. Die semantischen Typen sind die Knoten im Netzwerk und die Beziehungen zwischen ihnen sind die Links/Verbindungen. Ein Metathesaurus enthält hauptsächlich Typen wie Organismen, anatomische Strukturen, biologische Funktionen, Chemikalien, Ereignisse, physische Objekte und Konzepte oder Ideen. Die Hierarchie der Typen wird durch die 'isa'-Verknüpfung festgelegt, welche die primäre und wichtigste Verbindung im Netz darstellt. Diese dienen der Entscheidung über den spezifischsten semantischen Typ zur Zuordnung zu einem Metathesaurus-Konzept. Die Informationen über einen semantischen Typ umfassen eine Bezeichnung, Definition, Beispiele, hierarchische Informationen über den/die zugehörigen semantischen Typ(en) und assoziative Beziehungen. Assoziative Beziehungen innerhalb des Netzwerkes sind schwach. Sie erfassen lediglich die Tatsache, dass eine Instanz des ersten Typs mit einer Instanz des zweiten Typs verbunden sein kann. (National Library of, 2009b)

Das SPECIALIST Lexikon ist ein allgemeines Lexikon, das viele biomedizinische Begriffe enthält. Es wurde entwickelt, um ausreichend lexikalische Informationen für das SPECIALIST NLP bereitzustellen. Es umfasst daher sowohl häufig vorkommende englische Wörter als auch biomedizinisches Vokabular. Jeder Eintrag enthält syntaktische, morphologische und orthografische Informationen.

Wörter reflektieren oft mehrere Formen, die als Instanzen desselben Wortes gelten würden. Damit dieser hohe Grad an Variabilität bei Wörtern und Begriffen der natürlichen Sprache berücksichtigt werden, wurden die "Lexical Tools" entwickelt. Eine Reihe von Programmen, hauptsächlich Java-Programme, nutzen das Lexikon, um die Variationen in biomedizinischen Texten durch die Zuordnung von Wörtern zu ihren Wortarten zu verarbeiten, was bei der Suche im Internet oder in einer elektronischen Patient*innenakte hilfreich sein kann. Für jede Schreibweise oder Schreibvariante in einer bestimmten Wortart enthält das Lexikon einen Eintrag, wodurch eine Reihe von lexikalischen Einträgen entsteht. "Mehrwortbegriffe", die sich aus anderen/mehreren Wörtern zusammensetzen, können lexikalische Einträge sein, wenn der "mult-term" Begriff in allgemein englischen andernfalls in medizinischen Wörterbüchern oder in medizinischen Thesauri wie MeSH als Eintrag vorhanden ist.

Die Datenelemente des Lexikons bilden das Rückgrat des relationalen Formats, das zur strukturierten Darstellung von lexikalischen Einträgen dient. Folgende Liste beschreibt die Zusammenhänge der verschiedenen Elemente:

- Die String-Eigenschaften repräsentieren das Feld, das die verschiedenen Formen eines lexikalischen Eintrags enthält. Dies umfasst sowohl die Flexionsformen als auch die Zitierformen, wobei Groß- und Kleinschreibung sowie Interpunktion und Leerzeichen beachtet werden. Das Feld AGR (Agreement/Inflection Code) codiert Informationen zur Übereinstimmung und Flexion, einschließlich Person und Numerus für Substantive, Verben, Pronomen und Determinanten.
- Das Feld EUI (Unique Identifier Number for Lexical Entries) spielt eine zentrale Rolle bei den Eintrageigenschaften. Es stellt einen eindeutigen Identifikator für einen lexikalischen Eintrag bereit und verfolgt auch dessen Varianten über verschiedene Wortarten hinweg. Die Felder CIT (Citation Form) und BAS (Base Form) spezifizieren die Zitierform und die Grundform eines Eintrags. Die Zitierform fungiert als Singular für Substantive, Infinitiv für Verben und Positiv für Adjektive und Adverbien. Das Feld SCA gibt die Wortart oder syntaktische Kategorie des Eintrags an, während das Feld TYP die Flexionsmuster für verschiedene Wortarten definiert.
- Die Eintragsbeziehungen werden durch das Feld ABR (Acronym or Abbreviation) und SPV (Spelling Variant) repräsentiert, wobei erstere angibt, ob ein Begriff ein Akronym oder eine Abkürzung ist, und letztere die Zitierform einer bestimmten Rechtschreibvariante enthält.
- Die Datenbeschreibung beinhaltet verschiedene Elemente wie WRD (Word), das Zeichenketten in "Wörter" indiziert, sowie DES (Description), welches eine kurze Definition für Dateien oder Felder bereitstellt. FMT (Format) listet die geordneten Felder in einer Datei auf, während RWS (Number of Rows) die Anzahl der Zeilen angibt. FIL (File Name(s)) verweist auf die Datei(en), die die relationalen Tabellen enthalten, während BTS (Size in Bytes) die Größe einer Datei in Bytes darstellt. Die Anzahl der Spalten wird durch CLS (Number of Columns) angegeben, während COL (Three Letter Field Name) eine Kurzbezeichnung für ein Feld bereitstellt. Schließlich bietet REF (Cross Reference to Document) einen Querverweis auf einen Abschnitt des Dokuments für zusätzliche Informationen.

Die relationalen Tabellen des Lexikons bieten einen systematischen Rahmen für die Organisation lexikalischer Daten. Sie umfassen zehn verschiedene Relationen oder Tabellen umfasst, die jeweils in einer eigenen Datei untergebracht sind.

Die Daten werden für jeden lexikalischen Eintrag in zehn separate Relationen oder Tabellen strukturiert, um die Organisation und den Abruf zu erleichtern. Dieses Design erlaubt die Duplizierung von Daten über verschiedene Relationen hinweg und gibt Entwicklern die Flexibilität, den angemessenen Grad an Redundanz für ihre Anwendungen zu bestimmen. Das relationale Format spiegelt die Struktur des relationalen ASCII-Formats des Metathesaurus wider. Jede Tabelle besteht aus einer festen Anzahl von Spalten, wobei die Anzahl der Zeilen durch den Inhalt einer bestimmten Lexikonversion bestimmt wird. Die Spalten repräsentieren einzelne Datenelemente oder logische Strukturelemente. Längere Elemente mit variabler Länge stehen in der Regel rechts von kürzeren oder Elementen mit fester Länge stehen. Die Zeilen enthalten Werte für ein oder mehrere Datenelemente oder logische Subelemente für einen einzelnen Lexikoneintrag oder eine einzelne Zeichenfolge. Diese werden durch vertikale Balken getrennt sind. Optionale Elemente werden durch leere Werte gekennzeichnet. Jede Zeile schließt mit einem vertikalen Balken, einem Wagenrücklauf und einem Zeilenvorschub ab. (National Library of, 2009c)

2.1.7 LOINC – Standardisierung medizinischer Labordaten

LOINC steht für Logical Observation Identifiers Names and Codes und bezeichnet eine Datenbank sowie einen universellen Standard zur Identifizierung medizinischer Laborbeobachtungen. Im Jahr 1999 wurde er von der HL7 Standards Development Organization als bevorzugter Codesatz für Labortestnamen in Transaktionen zwischen Gesundheitseinrichtungen, Labors, Labortestgeräten und Gesundheitsbehörden festgelegt. (Loinc, 2025b) Die Datenbank wurde als Antwort auf die Nachfrage nach einer elektronischen Datenbank für die klinische Versorgung und das Management entwickelt und ist öffentlich und kostenlos zugänglich. Seit ihrer Einführung wurde die Datenbank kontinuierlich erweitert und umfasst nun nicht nur die Namen medizinischer Laborcodes, sondern zusätzlich Pflegediagnosen, Pflegemaßnahmen, Ergebnisklassifizierungen und Datensätze zur Patient*innenversorgung.

Ziel von LOINC ist die Unterstützung des elektronischen Austausches und der Erfassung klinischer Ergebnisse wie Labortests, klinische Beobachtungen, Ergebnismanagement und Forschung. Dabei verwendet die Datenbank universelle Codenamen und Identifikatoren für die medizinische Terminologie im Kontext elektronischer Gesundheitsakten.

Die zwei Hauptteile sind: das Labor-LOINC und klinisches LOINC, wobei das klinische LOINC einen Teilbereich der Dokumenten-Ontologie beinhaltet und die verschiedenen Arten von klinischen Berichten und Dokumenten kategorisiert. (Huser, Huser et al.)

Als ein universelles Codesystem wird es Einrichtungen und Abteilungen in der ganzen Welt ermöglichen, Ergebnisse aus ihren Bereichen zu empfangen und zu senden, um sie zu vergleichen und zu konsultieren, und kann zu einer größeren Initiative des öffentlichen Gesundheitswesens zur Verbesserung der klinischen Ergebnisse und der Qualität der Pflege beitragen.

LOINC umfasst zwei Hauptbereiche, Labor und Klinik:

- Laboratory umfasst den Bereich von Laboruntersuchungen wie Hämatologie, Serologie und Mikrobiologietests, einschließlich Antibiotika-Empfindlichkeitstests und Zellzählungen.
- Clinical beschreibt dabei verschiedene nicht-laborbezogene Konzepte, darunter EKG, Herzecho, urologische Bildgebung mittels Ultraschalls und geburtshilflicher Ultraschall, sowie gastroendoskopische Prozeduren, radiologische Bildgebungen, mit den jeweiligen Unterbereichen:
 - Klinische Dokumente beinhalten Konzepte für verschiedene Arten von klinischen Berichten, beispielsweise Entlassungsberichte oder Vorsorgeuntersuchungsnotizen für Kinder.
 - Umfrageinstrumente bezeichnen Konzepte für standardisierte Umfragen, wie beispielsweise den Glasgow Coma Score oder die PHQ-9-Depressionsskala, als auch CMS.

LOINC umfasst somit sowohl Labor- als auch klinische Beobachtungen und Dokumentationen und ermöglicht eine standardisierte Kodierung und Kommunikation dieser Informationen.

2.1.8 RxNorm – Arzneimittelinformation in standardisierter Form

RxNorm bezeichnet eine standardisierte Nomenklatur für klinische Medikamente, zu denen sowohl Generika als auch Markenarzneimittel zählen. Die Nomenklatur stellt eine wichtige Komponente im Bereich der Informatik des Gesundheitswesens dar und dient der Unterstützung der semantischen Interoperabilität zwischen Arzneimittelterminologien und pharmazeutischen Wissensbanksystemen. Der Zweck von RxNorm ist die Erfassung und Verarbeitung von Arzneimittelinformationen in Krankenhäusern, Apotheken und anderen Organisationen. RxNorm bietet die Möglichkeit die verschiedenen Bezeichnungen für Arzneimittel zu normalisieren und ihnen eindeutige Bezeichnungen für Medikamente und

Drogen zu geben. Somit ist es das Ziel eine effiziente und eindeutige Kommunikation von arzneimittelbezogenen Informationen zu ermöglichen.

Umfang von RxNorm beinhaltet die Namen von verschreibungspflichtigen und vielen rezeptfreien Medikamenten, Generika und Markenmedikamente:

- Klinische Arzneimittel sind pharmazeutische Produkte, die einer/einem Patient*in mit therapeutischer Absicht verabreicht werden.
- Arzneimittelpackungen enthalten mehrere Arzneimittel oder Medikamente, die in einer bestimmten Reihenfolge verabreicht werden.

Nicht-therapeutische Radiopharmaka, Bulk-Pulver, Kontrastmittel, Lebensmittel, Nahrungsergänzungsmittel, Medizinprodukte und Homöopathika fallen nicht in den Geltungsbereich der RxNorm.

Die RxNorm basiert auf anderen Terminologien und wird von diesen abgeleitet. Es reflektiert und bewahrt die Bedeutungen, Medikamentennamen, Attribute und Beziehungen seiner Quellen. Folgende Quellen liefern die Namen für die Medikamente:

- Anatomical Therapeutic Chemical Classification System (ATC)
- Vaccines Administered (CVX)
- DrugBank (DRUGBANK)
- Gold Standard Drug Database (GS)
- Multum MediSource Lexicon (MMSL)
- Micromedex RED BOOK (MMX)
- Medical Subject Headings (MeSH) - Letztes Erscheinen im Oktober 2023
- CMS Formulary Reference File (MTHCMSFRF)
- FDA Structured Product Labels (MTHSPL)
- FDB MedKnowledge (NDDF)
- RxNorm Normalized Names and Codes (RXNORM)
- US Edition of SNOMED CT (SNOMEDCT_US)
- USP Compendial Nomenclature (USP)
- Veterans Health Administration National Drug File (VANDF)

Die Namen basieren auf den Informationen der oben genannten Quellen und folgen einem einfachen Muster:

- Generische Medikamente bauen sich wie folgt auf: Wirkstoff | Stärke | Darreichungsform
- Markenmedikamente: Wirkstoff | Stärke | Darreichungsform [Markenname]

- Generische Medikamentenpackungen: {# (Wirkstoff | Stärke | Darreichungsform) / # (Wirkstoff | Stärke | Darreichungsform)} per Packung
- Markenmedikamentenpackungen: {# (Wirkstoff | Stärke | Darreichungsform) / # (Wirkstoff | Stärke | Darreichungsform)} per Packung [Markenname]

Als Beispiel "Acetaminophen" wird folgendermaßen als String angezeigt beziehungsweise gespeichert: Acetaminophen 500 MG Oral Tablette.

RxNorm verwendet Termtypen, um generische und markenspezifische Medikamentennamen auf verschiedenen Spezifitätsebenen zu kennzeichnen. Beispiele dafür wären:

- IN - Ingredient: Wirkstoff wie Fluoxetin. Ein Bestandteil, der dem Arzneimittel seine charakteristischen klinischen Eigenschaften verleiht.
- PIN - Precise Ingredient: Spezifischer Wirkstoff wie Fluoxetinhydrochlorid. PIN ist eine bestimmte Form des Inhaltsstoffs, die klinisch aktiv oder inaktiv sein kann. Die meisten präzisen Inhaltsstoffe sind Salze oder isoforme Formen.
- MIN - Multiple Ingredients: Mehrere Wirkstoffe wie Fluoxetin/Olanzapin. Bestehen aus zwei oder mehr Wirkstoffen, die in einer einzigen Arzneimittelzubereitung vorkommen.
- DF - Dose Form: Darreichungsform beispielsweise "Orale Lösung".
- SCDC - Semantic Clinical Drug Component: Wirkstoff + Stärke wie Fluoxetin 4 MG/ML.
- SCDF - Semantic Clinical Drug Form: Wirkstoff + Darreichungsform wie Fluoxetin Orale Lösung.
- SCD - Semantic Clinical Drug: Wirkstoff + Stärke + Darreichungsform wie Fluoxetin 4 MG/ML Orale Lösung.
- BN - Brand Name: Markenname.
- SBD - Semantic Brand Drug: Wirkstoff + Stärke + Darreichungsform + Markenname wie Fluoxetin 4 MG/ML Orale Lösung [Prozac].

RxNorm verwendet Beziehungen, um Konzepte zu verknüpfen, die denselben Wirkstoff oder dieselbe Darreichungsform enthalten. Beispiele für Beziehungen, den sogenannten Relations sind:

- Has_tradename: Verknüpft generische Namen mit Markennamen.
- Part_of: Zeigt an, dass ein Wirkstoff Teil einer Kombination ist.
- Has_dose_form: Verknüpft eine bestimmte Dosisform mit einem Medikament.

Zusätzlich zu den normalen Namen und Beziehungen umfasst RxNorm Attribute wie NDCs (National Drug Codes), die in ein standardisiertes Format gebracht werden.

Diese Struktur ermöglicht es, Medikamenteninformationen konsistent und interoperabel zu verwalten, wodurch die Genauigkeit und Effizienz in verschiedenen Gesundheitssystemen verbessert wird. (National Library of, 2025c)

2.2 Grundlagen der natürlichen Sprachverarbeitung

Natural Language Processing (NLP) spielt eine zunehmend bedeutende Rolle im medizinischen Alltag – etwa bei der automatisierten Analyse von medizinischen Texten der Extraktion relevanter Informationen aus elektronischen Gesundheitsakten oder der sprachbasierten Unterstützung klinischer Entscheidungen. Um diese Anwendungen zu ermöglichen, müssen medizinische Texte strukturiert verarbeitet und semantisch erschlossen werden. Genau hier setzt NLP an: als Schnittstelle zwischen menschlicher Sprache und computergestützter Analyse.

2.2.1 Historischer Hintergrund

Natural Language Processing, NLP, zu Deutsch Natürliche Sprachverarbeitung, hat seinen Ursprung in den 1940er Jahren. Alan Turing veröffentlichte bereits 1940 einen Artikel mit dem Titel "Computing Machinery and Intelligence", in dem er den "Turing-Test" zum Nachweis von Intelligenz vorschlug - besser bekannt als "Imitation Game" (Stanford University - Eric, 2004). Die erste öffentliche Demonstration der Machbarkeit maschineller Übersetzung fand 1954 an der Georgetown University in Zusammenarbeit mit IBM statt. Trotz der Verwendung eines sehr begrenzten Vokabulars und einer sehr begrenzten Grammatik waren die Ergebnisse für die damalige Zeit erstaunlich und führten zu einer weiteren Förderung von Machine Translation (MT)-basierten Projekten. Bis in die 1960er Jahre hinein war der Optimismus groß. Obwohl weiterhin viele Prognosen von immanenten Durchbrüchen die Rede war, wurde 1964 die ALPAC ins Leben gerufen. Diese hatte 1966 einen Bericht veröffentlicht, dass die Fortschritte bezüglich MT langsamer voranschreiten als angenommen.

Als Folge dieses Berichts wurden die Forschungsarbeiten in den USA eingestellt, was wiederum große Auswirkungen auf Länder wie Kanada, Frankreich und Deutschland hatte. In den 1980er Jahren entstanden in einer wachsenden Zahl von Ländern zahlreiche MT-Systeme. Gleichzeitig wurde die Forschung nach fortschrittlicheren Methoden und Techniken fortgesetzt. Eine weitverbreitete Strategie war dabei die indirekte Übersetzung, bei der die Ausgangssprache zunächst in eine abstrakte, sprachunabhängige Repräsentation überführt wurde – also eine strukturierte Bedeutungsdarstellung, die unabhängig von grammatikalischen Eigenheiten einzelner Sprachen funktioniert. Dies erfolgte typischerweise durch semantische, morphologische und syntaktische Analysen. Ein einfacher Satz wie „Der Patient leidet an Diabetes“ wurde dabei in Komponenten wie *Subjekt = Patient, Prädikat = leidet an, Objekt = Krankheit: Diabetes* zerlegt, bevor dieser Inhalt in die Zielsprache, etwa Englisch („The patient suffers from diabetes“), generiert wurde. Diese Methode war insofern bedeutsam, als sie die

Grundlage für die Entwicklung regelbasierter und später statistischer NLP-Verfahren legte und somit den Übergang zu komplexeren Übersetzungssystemen vorbereitete.

Zu Beginn der 1990er Jahre wurden zwei Ansätze verfolgt, ein rein statistischer und ein auf Übersetzungsbeispielen basierender. Bei der Analyse von Texten und der Auswahl von Äquivalenten werden keine syntaktischen oder semantischen Regeln verwendet. Darüber hinaus wurden Translation-Memory-Systeme veröffentlicht, die es Übersetzern ermöglichen, auf bereits übersetzte Texte zurückzugreifen.

Dies war auch der Ausgangspunkt für die Erforschung der Sprachübersetzung, die Spracherkennung, Sprachsynthese und Übersetzungsmodule umfasst.

In den 2000er Jahren führte das Wachstum des Internets zu einer großen Menge an Rohdaten. Daher verlagerte sich die Forschung auf unüberwachte oder halbüberwachte Algorithmen. Es wurden nicht annotierte Daten oder eine Kombination aus annotierten und nicht annotierten Daten verwendet. (Hutchins, 2014)

2.2.2 Verwendungszweck von NLP

Die Anwendungsgebiete von Natural Language Processing (NLP) sind äußerst vielseitig. Sie reichen von der maschinellen Übersetzung über die Textkategorisierung und Spam-Filterung bis hin zur Informationsextraktion, automatischen Textzusammenfassung und der Entwicklung von Dialogsystemen. Auch in der Medizin eröffnen sich durch NLP zahlreiche Möglichkeiten. Im Folgenden werden einige zentrale Anwendungsbereiche näher erläutert:

Maschinelle Übersetzung - Machine Translation

Die Zugänglichkeit von Daten in einer global vernetzten Welt ist von entscheidender Bedeutung. Die Sprachbarriere stellt eine große Herausforderung dar, die durch maschinelle Übersetzungstechnologien überwunden werden soll. Diese basieren auf statistischem maschinellem Lernen, das große Datenmengen analysiert, um die Wahrscheinlichkeit von Übersetzungen zwischen verschiedenen Sprachen zu bestimmen. Es wurden auch Methoden zur automatischen Bewertung der Übersetzungsqualität entwickelt, indem Übersetzungen mit Referenztexten verglichen werden.

Textkategorisierung und Spam-Filterung

Kategorisierungssysteme ordnen große Datenströme vordefinierten Kategorien oder Indizes zu. Diese umfassen unter anderem amtliche Dokumente, Berichte über militärische Unfälle, Marktdaten und Nachrichten.

Die Kategorisierung von Texten, wie sie bei E-Mail-Spamfiltern angewendet wird, ist von großer Bedeutung, da sie die erste Verteidigungslinie gegen unerwünschte E-Mails darstellt.

Die Herausforderung besteht darin, falsch negative und falsch positive Ergebnisse zu minimieren. Fortschritte in der NLP-Technologie erleichtern dies. Spam-Filter in E-Mail-Systemen verwenden verschiedene Protokolle, um Spam von legitimen Nachrichten zu unterscheiden. Es gibt verschiedene Arten von Spam-Filtern:

- Inhaltsfilter: Prüfen den Nachrichteninhalt auf Spam.
- Kopfzeilenfilter: Untersuchen die E-Mail-Kopfzeilen auf Fälschungen.
- Allgemeine Blacklist-Filter: Blockieren E-Mails von auf einer schwarzen Liste stehenden Empfängern.
- Regelbasierte Filter: Nutzen individuelle Kriterien wie Absender oder bestimmte Wörter.
- Erlaubnis-Filter: Erfordern vorab die Zustimmung des Empfängers für den Absender.
- Challenge Response-Filter: Verlangen von Absendern, einen Code einzugeben, um E-Mails zu senden.

Informationsextraktion - Information Extraction

Die Informationsextraktion hat zum Ziel, relevante Phrasen aus Textdaten zu identifizieren. Dabei spielt die Extraktion von Entitäten wie Namen, Orten, Ereignissen, Daten, Uhrzeiten und Preisen eine wichtige Rolle. Die extrahierten Informationen dienen verschiedenen Zwecken, wie beispielsweise der Zusammenfassung von Texten, dem Aufbau von Datenbanken und der Identifizierung von Schlüsselwörtern. Obwohl viele Informationsextraktionssysteme erfolgreich Begriffe aus Dokumenten extrahieren können, ist die Erfassung von Beziehungen zwischen den Begriffen eine Herausforderung. Es ist wichtig, die Komplexität der Aufgabe anzuerkennen und eine ausgewogene Herangehensweise zu wählen, um die bestmöglichen Ergebnisse zu erzielen. Ein effektives Informationsextraktionssystem arbeitet auf verschiedenen Ebenen, von der Worterkennung bis zur Diskursanalyse auf der Ebene des gesamten Dokuments.

Zusammenfassung - Summarisation

Im digitalen Zeitalter ist die Informationsflut eine Realität, die unsere Fähigkeit, Information zu verarbeiten, übersteigt. Es ist wichtig, Daten präzise zusammenzufassen und zu interpretieren, um wichtige Informationen in großen Datenmengen zu identifizieren und tiefere emotionale Bedeutungen zu verstehen. Ein Unternehmen erkennt die Stimmung in sozialen Medien und nutzt diese Informationen, um das Produktangebot zu verbessern.

Dialogsysteme

Dialogsysteme sind weit verbreitet und können von der Unterstützung bis zur Durchführung von Aktionen eingesetzt werden. Dabei erfordern unterstützende Dialogsysteme ein höheres Maß an Kontextbewusstsein als solche, die Aktionen durchführen. Frühere Dialogsysteme

waren hauptsächlich auf kleinere Anwendungen ausgerichtet und nutzten dabei die phonemische und lexikalische Ebene der Sprache.

NLP in der Medizin

NLP in der Medizin umfasst eine Vielzahl von Anwendungsbereichen, darunter insbesondere die klinische Textverarbeitung, die Patient*inneninteraktion und -unterstützung, die Kommunikation zwischen medizinischem Fachpersonal und Patient*innen, die Medikamentenüberwachung und Pharmakovigilanz, die Diagnoseunterstützung sowie die personalisierte Medizin und das Management von Patient*innenprofilen. In bestimmten Fällen kann NLP auch unterstützend bei der medizinischen Bildverarbeitung eingesetzt werden – etwa zur automatisierten Generierung oder Interpretation von radiologischen Befundtexten auf Basis von Bilddaten.

2.2.3 Aktueller Stand und zukünftige Ausrichtung

Als Teilgebiet der Künstlichen Intelligenz beschäftigt sich die NLP mit der Interaktion zwischen Computern und Menschen. Sie ermöglicht die Kommunikation zwischen Menschen und Maschine mittels Techniken, die die Analyse, Erzeugung und Verständigung von menschlicher Sprache umfassen. Bekannte Beispiele, die als integraler Bestandteil vieler Anwendungen dienen, sind unter anderem Sprachübersetzung, Sprachassistent*innen und Chatbots in Echtzeit. Ein bedeutender Fortschritt im Bereich des Natural Language Processing (NLP) war die Entwicklung von Worteinbettungen (word embeddings). Diese stellen Wörter in einem kontinuierlichen Vektorraum dar und erfassen semantische Beziehungen zwischen Phrasen. Ein weiterer wichtiger Fortschritt im NLP ist das Deep Learning, welches zur Entwicklung leistungsstarker neuronaler Netzwerk-Modelle geführt hat, die aus großen Datenmengen lernen können. Ein Beispiel hierfür ist LSTM (Long short-term Memory). In LSTM werden Daten im Kurzzeitgedächtnis wie bei RNN (Recurrent Neural Networks) gespeichert. Wenn der Speicherplatz voll ist, werden die am längsten erhaltenen Informationen gelöscht und durch neue Daten ersetzt. Im Gegensatz dazu werden ausgewählte Daten im Cell State, dem Langzeitgedächtnis, abgelegt. (Aiwakforce, 2025)

Anhand dieser Fortschritte sind folgende NLP-Technologien möglich geworden: (Surusha, 2023)

- Kontextuelles Verständnis und Conversational AI sind dank fortschrittlicher Transformer-Modelle wie GPT-3.5 natürlicher und ansprechender geworden. Dies ist insbesondere in der Kundenbetreuung und personalisierten Interaktionen von Vorteil.

- Multilinguale und übersprachliche Fähigkeiten brechen Sprachbarrieren durch genaue Übersetzungen und die Fähigkeit, Inhalte in mehreren Sprachen zu verstehen und zu generieren. Dadurch wird der internationale Geschäftsverkehr, die Diplomatie und der kulturelle Austausch erleichtert.
- Die Verwendung von emotionaler KI und Sentiment-Analyse ermöglicht es NLP-Modellen, die emotionale Stimmung von Benutzer*innen zu erfassen. Dies bietet wichtige Einblicke für Unternehmen und das Gesundheitswesen.
- Im Gesundheitswesen automatisieren NLP-Technologien die klinische Dokumentation und analysieren große Mengen wissenschaftlicher Literatur. Dadurch wird die Patient*innenversorgung verbessert und die medizinische Forschung beschleunigt.
- Ethik und Bias-Minderung sind essenziell für die Entwicklung vertrauenswürdiger NLP-Modelle. Forschungen in diesem Bereich sind daher von großer Bedeutung.
- Anpassbare und adaptive Sprachmodelle ermöglichen maßgeschneiderte Interaktionen sowohl für Unternehmen als auch für Einzelpersonen. Dadurch wird die Benutzer*innenerfahrung verbessert.
- NLP-gesteuerte Inhaltsgenerierung und -zusammenfassung steigern die Produktivität in der Inhaltsproduktion und erleichtern die Bewältigung großer Datenmengen.
- Echtzeit-Sprachübersetzung in Live-Gesprächen erleichtert die globale Zusammenarbeit und den interkulturellen Austausch.
- Die Integration von Augmented Reality und Virtual Reality eröffnet neue Möglichkeiten für immersive Benutzer*innenerfahrungen und erweitert die Anwendungsbereiche von NLP.
- Fortschritte in der Vorab-Schulung und beim Transferlernen treiben die Vielseitigkeit und Effizienz von NLP-Modellen voran, was die Entwicklung neuer Anwendungen beschleunigt.

2.3 Der Zyklus der Annotationsentwicklung: M.A.T.T.E.R.

Nach Pustejovsky wird zur Entwicklung eines Annotationszyklus die MATTER-Methode verwendet. MATTER stellt dabei ein Akronym dar und lässt sich in folgende Punkte und Bedeutungen gliedern. (Pustejovsky and Stubbs, 2012)

- **Model (Modellieren):** Definition des Phänomens, das annotiert werden soll – z. B. medizinische Begriffe oder Krankheitsverläufe. Dabei wird festgelegt, welche Entitäten und Relationen erfasst werden sollen.
- **Annotate (Annotieren):** Erstellung erster Annotationen anhand eines Leitfadens. Annotator:innen markieren relevante Textstellen entsprechend des entwickelten Modells.
- **Train (Trainieren):** Training eines KI-Modells auf Basis der erstellten Annotationen.
- **Test (Testen):** Überprüfung des trainierten Modells anhand neuer, annotierter Daten.
- **Evaluate (Evaluieren):** Systematische Bewertung der Modelleistung – etwa durch Metriken wie Genauigkeit, Präzision oder F1-Score.
- **Revise (Überarbeiten):** Analyse von Fehlern und Überarbeitung des Modells oder der Annotationen. Danach beginnt der Zyklus erneut.

Um sicherzustellen, dass der zu trainierende Algorithmus das gewünschte Verhalten und Sprachmuster erlernt, müssen die verwendeten Eigenschaften und Merkmale ausreichend umfangreich sein. Hierbei sollten Erklärungen und linguistische Beschreibungen aus umfassenden Sprachtheorien verwendet werden. Diese bilden die Grundlage für das Training und Testen des Identifizierungs- oder Beschriftungsalgorithmus. Anschließend kann das Modell anhand dieser Analyse und Bewertung der Leistung die Regeln der Trainingsgrundlage verbessern und erneut trainieren. (Pustejovsky and Stubbs, 2012)

2.3.1 Erstellung des Modells

Der erste Schritt im MATTER-Zyklus steht für „**Model**“ und beschreibt das **Modellieren des Phänomens**, das annotiert werden soll. In dieser Phase wird genau festgelegt, welche sprachlichen oder inhaltlichen Elemente – etwa medizinische Begriffe, Symptome oder Relationen – durch Annotation erfasst werden sollen. Dabei werden zentrale Parameter und Aufgaben definiert, wie beispielsweise das relevante Vokabular (B), die Beziehungen zwischen diesen Begriffen (R) sowie deren semantische Interpretation (I). Dieses konzeptuelle Modell dient als Grundlage für alle folgenden Annotationsschritte und beeinflusst maßgeblich die Qualität der erzeugten Trainingsdaten.

Das konzeptuelle Modell kann als ein Triplet $\langle B, R, I \rangle$ dargestellt werden, das aus drei Komponenten besteht:

- **B** steht für das Vokabular der zu annotierenden Begriffe,
- **R** für die Relationen zwischen diesen Begriffen,
- **I** für deren semantische Interpretation.

Ein klinisch relevantes Beispiel hilft, dieses Prinzip besser zu verstehen: Angenommen, in einem Korpus medizinischer Texte sollen **Symptome** und deren **zugehörige Diagnosen** annotiert werden, um ein KI-Modell zur Entscheidungshilfe im klinischen Alltag zu trainieren.

- **B** = {Fieber, Husten, Atemnot, COVID-19, Influenza, Pneumonie}
→ Begriffe, die im Text erkannt und annotiert werden sollen
- **R** = {SymptomX → DiagnoseY}
→ Beziehungen, die anzeigen, welches Symptom zu welcher möglichen Diagnose gehört (z. B. „Fieber“ → „COVID-19“)
- **I** = {„Fieber“ = „erhöhte Körpertemperatur“, „COVID-19“ = „Virusinfektion durch SARS-CoV-2“}
→ semantische Bedeutungen der Begriffe zur eindeutigen Interpretation im Kontext

Durch diese strukturierte Modellierung kann ein Annotationsteam klar definieren, welche Elemente markiert werden sollen, wie sie in Beziehung stehen und wie sie medizinisch interpretiert werden. Dies ist essenziell, um später verlässliche Trainingsdaten für KI-gestützte Diagnosesysteme zu erzeugen.

Die Annotation von Dokumenten stellt dabei eine der grundlegendsten Formen der Annotation dar. Bei der Annotation ganzer Texte hingegen wird in der Regel ein komplexeres Modell verwendet, das speziell auf den jeweiligen Anwendungsfall zugeschnitten ist. (Pustejovsky and Stubbs, 2012)

2.3.2 Der Trainingskorpus wird durch Annotation definiert.

A beschreibt den Annotationsprozess mit den zuvor definierten Spezifikationen anhand des Modells. Annotatoren verwenden einen Annotationsleitfaden, um die benötigten Datensätze mit den definierten Tags zu kennzeichnen. Der Prozess erfordert oft mehrere Iterationen von Modellierung und Annotation und wird daher auch als MAMA-Zyklus (Model-Annotate-Model-Annotate) bezeichnet. Beim Hinzufügen von Informationen zu Texten werden zwei Arten von Annotationen unterschieden: Konsumierende und nicht-konsumierende Annotationen. Ein konsumierendes Tag bezieht sich auf ein Metadaten-Tag und ist direkt mit

dem eigentlichen Inhalt der Daten verknüpft. Nicht-konsumierende Tags werden in die Datei eingefügt, sind aber nicht mit einem tatsächlichen Teil des Textes verknüpft und dienen der Organisation. Ein Beispiel aus der Medizin soll dies verdeutlichen: "Liegender Marknägel im rechten Humeruskopf". (Pustejovsky and Stubbs, 2012)

Dieser Satz könnte wie folgt annotiert werden:

- "liegend": konsumierendes Tag, das auf den Zustand der Marknägel hinweist und direkt mit dem realen Inhalt des Textes verknüpft ist.
- "Marknägel": verbrauchender Tag, der auf das Objekt im Text, nämlich die Marknägel, hinweist und direkt mit dem realen Inhalt des Textes verknüpft ist.
- "im rechten Humeruskopf": Konsumierendes Tag, das auf die Position der Marknägel hinweist und direkt mit dem realen Inhalt des Textes verknüpft ist.
- Die Datensatz-ID, z. B. "ID123456", wäre in diesem Fall ein nicht-konsumierendes Tag und dient der Identifizierung des Objekts und der Organisation des Datensatzes.

Für die konsequente Annotation der Datensätze wird ein Standardkorpus benötigt, das im besten Fall von mehreren Personen aus dem Fachgebiet annotiert wird und als Richtlinie für die nachfolgenden Texte, Dokumente, Datensätze, Bilder etc. dient. Das Korpus entsteht durch den Evaluationsprozess und wird als solches für das maschinelle Lernen verwendet. Dabei ist es notwendig, die Annotationen ständig zu vergleichen und die Tags zu korrigieren. (Pustejovsky and Stubbs, 2012)

2.3.3 Der Algorithmus wird trainiert und getestet.

T und T bezeichnet den Prozess des Trainierens und Testens des Algorithmus unter Verwendung des zuvor definierten Korpus. Dieser Korpus wird in zwei Hauptteile aufgeteilt: den Entwicklungskorpus und den Testkorpus. Der Entwicklungskorpus besteht wiederum aus einem Trainingsset und einem Testset. Das Trainingsset wird verwendet, um den Algorithmus für den vorgesehenen Zweck zu trainieren, während das Testset des Entwicklungskorpus zur Fehleranalyse dient. Mithilfe dieses Testsets können nach dem Training Fehler erkannt und fehlerhafte Annotationen identifiziert werden. Diese Fehlerquellen führen dann zu Anpassungen und einem erneuten Training des Algorithmus, gefolgt von einem weiteren Test mit dem Testset. Wenn die Ergebnisse zufriedenstellend sind, wird der Algorithmus abschließend mit dem Testkorpus getestet. Da die Daten des Testkorpus während des Trainingsprozesses nicht verwendet wurden, ermöglicht dieser Test, die Leistung des Algorithmus mit neuen oder aus anderen Quellen stammenden Daten zu bewerten und somit seine Repräsentativität zu überprüfen. (Pustejovsky and Stubbs, 2012)

2.3.4 Die Ergebnisse werden ausgewertet.

Unter Punkt E versteht man die Evaluation der Leistung eines Algorithmus. Sie ist häufig durch die Präzision und den Recall gekennzeichnet, welche die Genauigkeit der Datenkennzeichnung reflektieren. Diese Metriken beruhen auf der Unterscheidung von Datenelementen in zwei Kategorien: als relevant oder irrelevant. Die Präzision (P) misst den Anteil der korrekt gekennzeichneten relevanten Instanzen, während der Recall (R) die Fähigkeit des Algorithmus darstellt, tatsächlich relevante Elemente zu identifizieren. Diese Konzepte werden üblicherweise in einer Tabelle dargestellt, aus der sich die Genauigkeit (Accuracy) ableiten lässt. Die Accuracy ergibt sich aus dem Verhältnis von richtig-positiven und richtig-negativen Ergebnissen zu allen vier möglichen Kategorien von Ergebnissen.

		Voraussichtliche Kennzeichnung	
		Positiv	Negativ
Standardisierte Kennzeichnung	Positiv	Richtig-Positiv	Falsch Negativ
	Negativ	Falsch-Positiv	Richtig-Negativ

TABELLE 2: GENAUIGKEITSMATRIX

Des Weiteren kann die Effektivität eines Algorithmus mittels des F-Scores, auch F-Maß genannt, beurteilt werden, welcher das harmonische Mittel aus Präzision und Recall darstellt. Die Formel für den F-Score lautet:

$$F = \frac{2 * P * R}{P + R}$$

Die Wahl der Evaluationsmaßnahmen kann je nach den spezifischen Anforderungen des Korpus und des zu bewertenden Algorithmus variieren. (Pustejovsky and Stubbs, 2012)

2.3.5 Modell und Algorithmus werden überarbeitet.

Nachdem die Trainings- und Testergebnisse des Algorithmus ausgewertet wurden, erfolgt eine detaillierte Fehleranalyse, um die Leistung des Algorithmus zu verstehen und mögliche Schwachstellen zu identifizieren. Diese Analyse kann mithilfe von Konfusionsmatrizen durchgeführt werden, um die Designaspekte des Modells zu überprüfen und gegebenenfalls zu verbessern. Durch die Identifizierung von Fehlern können neue Tags und Merkmale entwickelt werden, die die Leistung des Lernalgorithmus verbessern und den zuvor definierten Goldstandard besser erfüllen.

Ein konkretes Beispiel aus Unterabschnitt 2.3.2 "Liegender Marknagel im rechten Humeruskopf" illustriert den Prozess der Modellüberarbeitung. In diesem Beispiel wurden die verschiedenen Abschnitte des Textes mit konsumierenden Tags versehen, wodurch es möglich

ist, die darunter liegenden Entitäten zu unterscheiden. Beispielsweise würde "Liegend" als Deskriptor und "Marker" als Gerät klassifiziert werden. "Im rechten Humeruskopf" hingegen beschreibt sowohl eine anatomische Entität als auch eine Lokalisation, und würde dementsprechend in "im rechten" (Lokalisation) und "Humeruskopf" (Anatomische Entität) unterteilt werden.

Falsche Zuordnungen können zu fehlerhaften Lernergebnissen führen. Wenn beispielsweise "im rechten Humeruskopf" fälschlicherweise als [Anatomical_Entity] markiert würde, könnte das Modell lernen, Marknägel im linken Humeruskopf auch nur als zu identifizieren.

Zusätzlich kann es für Annotatoren herausfordernd sein, komplexe Sätze oder neue Unterkategorien mit den richtigen Tags zu versehen. Daher wird der MATTER-Zyklus wiederholt, bis die Änderungen erfolgreich vom Modell übernommen wurden. (Pustejovsky and Stubbs, 2012)

3 Strategien und Anwendungen von Annotation und NLP/KI in der Medizin

3.1 Bedeutung hochwertiger annotierter Daten

Die Nutzung künstlicher Intelligenz und maschinellen Lernens in der medizinischen Forschung und klinischen Praxis hat in den letzten Jahren erheblich zugenommen. (Alowais et al., 2023) Insbesondere die Verfügung von großen Mengen medizinischer Daten hat die Entwicklung und Anwendung von KI-Modellen ermöglicht, die komplexe medizinische Probleme lösen und klinische Entscheidungsunterstützungssysteme verbessern können. (Alamgir et al., 2021) Eine zentrale Komponente für den Erfolg der Modelle ist die Qualität der verwendeten Daten.

Hochwertig annotierte Daten in der Medizin sind von grundlegender Bedeutung für die Entwicklung und Validierung von KI-Modellen aus mehreren Gründen.

Sie dienen als Grundlage für das Training von ML-Algorithmen, die darauf abzielen, Muster und Zusammenhänge in medizinischen Daten zu erkennen. Ohne präzise Annotationen anhand eines zuvor festgelegten Gold-Standards wäre es für die Algorithmen schwierig, relevante Informationen zu extrahieren und aussagekräftige Modelle zu entwickeln. Insbesondere in der Medizin spielen Genauigkeit und Zuverlässigkeit eine lebenswichtige Rolle, womit hochwertige Annotationen für die Gewährleistung der Qualität der Ergebnisse unerlässlich sind. Darüber hinaus werden hochwertig annotierte Daten für die Validierung der Modelle verwendet. In diesem Schritt wird sichergestellt beziehungsweise überprüft, ob das entwickelte Modell einerseits korrekt funktioniert, andererseits die erwünschten und zuverlässigen Ergebnisse liefert. Hierbei werden die Leistungsfähigkeit des Modells und seine Fähigkeit zur Verallgemeinerung auf neue Daten überprüft. Dabei ist schon während des Trainingsvorgangs eine gewisse Diversität notwendig, um auf die maximalen Varietäten vorbereitet zu sein. Beispielsweise können sich die Quellen, trotz ähnlichen Inhalts, in Semantik und Morphologie unterscheiden. Somit ermöglichen hochwertige Annotationen eine präzise Bewertung der Modellleistung und helfen, potenzielle Fehler oder Schwachstellen aufzudecken, bevor das Modell in der klinischen Praxis eingesetzt wird.

Die Notwendigkeit hochwertig annotierter Daten für die Entwicklung und Validierung von KI-Modellen wird durch zahlreiche Studien und Fallbeispiele unterstützt. Forschungen zeigen, dass mittels KI-Modellen zu verbesserten Ergebnissen führt (Cuocolo et al., 2020, Zhang et al., 2019). Beispielsweise zeigt eine Studie aus dem Jahr 2017 in Bezug auf "Skin Cancer Detection and Tracking", durchgeführt an der Peking University und Stanford University, dass KI-

Modelle und deren Algorithmen dabei unterstützen unter verdächtige Läsionen unter den gutartigen Läsionen zu erkennen und deren zeitliche Veränderung zu beobachten, da diese auf eine Malignität hinweisen könnten. (Li et al., 2016)

Die Relevanz hochwertig annotierter Daten in der Medizin unterstreicht die Notwendigkeit eines sorgfältigen und präzisen Annotationsprozesses, der die Genauigkeit, Konsistenz und Vollständigkeit der Daten gewährleistet. Nur durch hochwertige Annotationen können KI-Modelle entwickelt und validiert werden, die das Potenzial haben, die medizinische Forschung und klinische Praxis zu transformieren und die Patient*innenversorgung zu verbessern.

3.2 Die Rolle von Annotationen bei der Erstellung von Trainingsdatensätzen

Wie bereits dargelegt, bezeichnet die Annotation den Prozess der Ergänzung von Rohdaten mit zusätzlichen Informationen, die für das Training von maschinellen Lernmodellen erforderlich sind. Zur Erstellung von Trainingsdaten werden die Informationen als Labels für die Textklassifikation, Markierungen für benannte Entitäten, Beziehungen zwischen Entitäten oder syntaktische Strukturen kategorisiert. In der medizinischen Domäne müssen NLP-Modelle oft komplexe und spezialisierte Informationen verarbeiten. Dies erfordert eine präzise und umfassende Annotation, die verschiedene Aspekte medizinischer Daten abdeckt. Die Annotation in diesem Bereich umfasst mehrere wichtige Schritte, darunter die Benennung und Klassifikation medizinischer Entitäten, die Markierung von Beziehungen und Interaktionen sowie die Normierung und Kodierung von Begriffen.

Die Benennung und Klassifikation medizinischer Entitäten ist ein zentraler Aspekt der Annotation. Der Prozess der Annotation beinhaltet die Identifizierung und Kennzeichnung verschiedener medizinischer Begriffe und Konzepte innerhalb von Texten. Die wichtigsten Entitäten in diesem Kontext umfassen Krankheiten, Symptome, Medikamente, Laborwerte, anatomische Begriffe, Untersuchungsmethoden sowie medizinische Geräte. Krankheiten und Diagnosen müssen präzise benannt und klassifiziert werden, um es NLP-Modellen zu ermöglichen, spezifische Krankheitsbilder zu erkennen und korrekt zu verarbeiten. Ein Beispiel für die Erkennung und Klassifikation einer Krankheit ist die Nennung von „Diabetes mellitus“ in einem medizinischen Bericht. Symptome hingegen sind oft vielfältig und können je nach Kontext variieren. Die Annotation dieser Symptome erfordert detailliertes Wissen, um sicherzustellen, dass Begriffe wie „Husten“, „Fieber“ oder „Schmerzen“ korrekt identifiziert und klassifiziert werden.

Die Annotation im Bereich der Medikamentierung umfasst die Benennung und Klassifikation von verschreibungspflichtigen und rezeptfreien Medikamenten. Dazu gehören auch die Dosierung, die Form und die Verabreichungsmethode. Diese Information ist von entscheidender Bedeutung, um die korrekte Medikation in den Patient*innendaten zu identifizieren.

Laborwerte sind häufig in numerischer Form vorhanden und müssen in Verbindung mit den entsprechenden Tests und Referenzwerten annotiert werden. Ein Beispiel für einen solchen Text wäre die Angabe des Blutdrucks einer/eines Patient*in, der korrekt als „120/80 mmHg“ erkannt und klassifiziert werden muss. Des Weiteren umfasst die Annotation weitere relevante Begriffe wie medizinische Verfahren, Therapien und Behandlungspläne, die ebenso präzise annotiert

werden müssen, um eine umfassende Analyse medizinischer Texte zu ermöglichen. Des Weiteren werden Deskriptoren und Attribute annotiert, welche medizinische Entitäten genauer beschreiben. Zu den Beispielen zählen unter anderem Angaben zu Konfidenzen bezüglich medizinischer Pathologien, Ortsangaben wie links oder rechts, ventral oder dorsal, Messungen bei unterschiedlichsten Untersuchungen wie der Palpation, CT-Untersuchungen und dermatologischen Lichtmikroskopie. Des Weiteren werden Merkmale und Eigenschaften, Bewertungsskalen, Graduierungen, Typisierungen, Klassifikationen sowie deskriptive Attribute und Zeitangaben erfasst.

Die Markierung von Beziehungen und Interaktionen zwischen Entitäten stellt einen weiteren kritischen Schritt in der Annotation dar. Die Beziehungen können vielfältig und komplex sein, beispielsweise in Bezug auf Wechselwirkungen zwischen Medikamenten, Zusammenhänge zwischen Symptomen und Diagnosen sowie Behandlungsverläufe und kausale Beziehungen. Die gleichzeitige Einnahme von Medikamenten kann zu unterschiedlichen Wirkungen führen. Die Annotation solcher Wechselwirkungen erfordert detaillierte Kenntnisse der Pharmakologie sowie die Fähigkeit, komplexe Beziehungen zu erkennen.

Ein Beispiel für die Markierung von Wechselwirkungen ist die Gegenüberstellung von „Aspirin“ und „Warfarin“. Dadurch lassen sich potenzielle Risiken identifizieren. Symptome können als Indikatoren für bestimmte Diagnosen betrachtet werden. Die Annotation dieser Zusammenhänge ermöglicht die Erkennung von Mustern sowie eine Verbesserung der diagnostischen Genauigkeit. So können beispielsweise Symptome wie Husten und Atemnot in Zusammenhang mit der Diagnose Asthma gebracht werden.

Des Weiteren sind die Beziehungen zwischen verschiedenen Behandlungsschritten und deren Auswirkungen auf den Krankheitsverlauf von Bedeutung. Die Annotation solcher Verläufe kann dazu beitragen, den Behandlungserfolg zu evaluieren und Verbesserungsmöglichkeiten zu identifizieren. Einige medizinische Ereignisse sind kausal miteinander verbunden. So kann beispielsweise die Einnahme eines bestimmten Medikaments zu einem spezifischen Symptom führen. Die Annotation solcher kausalen Beziehungen ist von entscheidender Bedeutung für das Verständnis von Ursache-Wirkung-Zusammenhängen in der Medizin. Zwischen den medizinischen Entitäten und den Deskriptoren und Attributen bestehen Beziehungen, die als Relationen bezeichnet werden. Die Markierungen veranschaulichen die Zugehörigkeit von Deskriptoren und Attributen zu den einzelnen medizinischen Entitäten. Des Weiteren können beispielsweise Wechselwirkungen zwischen Medikamenten oder Zusammenhänge zwischen Symptomen und Diagnosen erfasst werden.

Normierungen und Kodierungen medizinischer Begriffe spielen eine wesentliche Rolle bei der Erstellung konsistenter und interoperabler Datensätze. Begriffe werden zu den standardisierten medizinischen Klassifikationssystemen wie ICD und SNOMED zugeordnet. Eine korrekte Zuordnung von Diagnosen zu den entsprechenden ICD-Codes ermöglicht eine standardisierte Erfassung und Analyse von Gesundheitsdaten. Beispielsweise wird ein "akuter Myokardinfarkt" dem ICD-Code "BA41" zugeordnet.

SNOMED ist ein multihierarchisches Klassifikationssystem, in dem eine detaillierte und präzise Erfassung von medizinischen Informationen ermöglicht wird. Als ein Exempel könnte dem Symptom "Dysurie" dem SNOMED-Code "49650001" zugeordnet werden, des Weiteren finden sich Synonyme und Beschreibungen zu den Einträgen.

Darüber hinaus dient LOINC als Standard zur Identifizierung von Labor- und klinischen Beobachtungen, dabei erleichtern die LOINC-Codes die Konsistenz und Vergleichbarkeit von Labordaten, indem die Laborwerte und Testergebnisse mit Codes versehen werden. Ein Beispiel wäre der LOINC-Code "8480-6" für "Blutdruck systolisch". (Loinc, 2025a)

Die Normierung und Kodierung sind dabei für die Integration und Analyse großer und heterogener Datensätze wichtig. Sie ermöglichen es, Daten aus verschiedenen Quellen zu kombinieren und vergleichbare Analysen durchzuführen.

Die Annotation medizinischer Texte ist ein komplexer und essenzieller Prozess, der die Grundlage für die Entwicklung leistungsfähiger NLP-Modelle in der medizinischen Domäne bildet. Präzise annotierte Trainingsdatensätze unterscheiden sich dabei wesentlich von generalisierten Datensätzen, da sie den Kontext, die Zusammenhänge und die genaue Beschreibung in den Vordergrund stellen. Diese sorgfältige Benennung und Klassifikation von Entitäten, die Markierung von Beziehungen und Interaktionen sowie die Normierung und Kodierung medizinischer Begriffe ermöglichen die Entwicklung von Modellen, die medizinische Texte präzise analysieren und wertvolle Erkenntnisse liefern können.

Ein entscheidender Aspekt dabei ist die Verwendung spezialisierter Datenbanken wie einem Consumer Health Vocabulary (CHV), welches eine Brücke zwischen dem Fachvokabular der Mediziner*innen und einer vereinfachten Sprache schlägt. Dies erlaubt es, den NLP-Modellen nicht nur die medizinischen Fachbegriffe korrekt zu interpretieren, sondern auch deren Bedeutung für Laien verständlich zu machen. Die Annotation berücksichtigt somit nicht nur die exakte medizinische Terminologie, sondern auch die Verständlichkeit und Zugänglichkeit für verschiedene Nutzer*innengruppen.

Die Herausforderungen, die mit der Annotation medizinischer Daten verbunden sind, erfordern eine enge Zusammenarbeit zwischen NLP-Experten und medizinischen Fachkräften.

Fortgeschrittene Technologien zur Unterstützung des Annotationsprozesses sind unerlässlich, um eine hohe Genauigkeit und Verlässlichkeit der Modelle zu gewährleisten. Diese Technologien helfen, den Kontext und die detaillierten Beschreibungen korrekt zu erfassen und in den Trainingsdatensätzen abzubilden.

Die Integration präzise annotierter Daten in Trainingsdatensätze ermöglicht es NLP-Modellen, die Nuancen medizinischer Sprache zu verstehen und relevante Informationen effizient zu extrahieren. Dies führt zu einer verbesserten Diagnostik, personalisierten Behandlungsansätzen und einer allgemein besseren medizinischen Versorgung. Durch die Nutzung von CHV und ähnlichen Ressourcen können die Modelle zudem die Verständlichkeit der medizinischen Informationen für Patient*innen und andere Laien erhöhen, was einen wichtigen Beitrag zur Gesundheitskompetenz der Bevölkerung darstellt.

Daher ist die sorgfältige und umfassende Annotation ein unverzichtbarer Bestandteil der Entwicklung und Implementierung von NLP-Technologien im medizinischen Bereich. Nur durch diese detaillierte und kontextbezogene Annotation können NLP-Modelle die komplexen und oft spezialisierten Informationen in medizinischen Texten korrekt verarbeiten und interpretieren, was letztlich zur Verbesserung der klinischen Entscheidungsfindung und der Patient*innenversorgung beiträgt.

Hochwertige Annotationen sind essenziell für die Entwicklung und Leistungsfähigkeit von Natural Language Processing (NLP)-Modellen im medizinischen Bereich. Sie bilden die Grundlage für eine präzise semantische Analyse medizinischer Texte und sind damit entscheidend für die Extraktion relevanter Informationen, die Generierung strukturierter Daten sowie die Verbesserung von Diagnostik- und Entscheidungsunterstützungssystemen.

Ein zentraler Vorteil hochwertiger Annotationen liegt in der Optimierung der Modellleistung. Durch präzise annotierte Daten lernen NLP-Modelle, medizinische Begriffe, Diagnosen und Therapieansätze mit hoher Genauigkeit zu erkennen und zu interpretieren. Dies verbessert nicht nur die Präzision und Konsistenz der automatisierten Analyse, sondern ermöglicht auch eine effektive Generalisierung auf neue, unbekannte Daten. Gerade in der Medizin, einem sich dynamisch entwickelnden Feld mit ständig neuen Erkenntnissen und sich verändernden Terminologien, sind diese Eigenschaften essenziell, um eine kontinuierliche Anpassung der Modelle zu gewährleisten.

Hochwertige Annotationen tragen zudem maßgeblich zur Erfassung semantischer und kontextueller Relationen bei. Medizinische Texte sind komplex und enthalten oft mehrdeutige oder polysynonyme Begriffe, deren Bedeutung stark vom jeweiligen Kontext abhängt. So erfordert beispielsweise die Identifikation der Beziehungen zwischen Symptomen und

Diagnosen oder zwischen Medikamenten und deren Nebenwirkungen eine tiefgehende semantische Analyse, die nur durch eine sorgfältige Annotation ermöglicht wird. Modelle, die auf diesen annotierten Daten trainiert wurden, liefern belastbare Ergebnisse, die als Entscheidungsgrundlage für medizinisches Fachpersonal dienen können.

Ein weiterer wesentlicher Aspekt ist die Standardisierung durch etablierte medizinische Klassifikations- und Kodierungssysteme wie die International Classification of Diseases (ICD) oder die Systematized Nomenclature of Medicine (SNOMED-CT). Die Integration dieser Standards in den Annotationsprozess gewährleistet nicht nur die Präzision der NLP-Modelle, sondern auch deren Interoperabilität mit bestehenden klinischen Informationssystemen. Zusätzlich profitieren NLP-Modelle von Ontologien wie dem Unified Medical Language System (UMLS), das terminologische Konzepte harmonisiert und eine semantische Vernetzung zwischen verschiedenen medizinischen Begriffen schafft.

Die Relevanz hochwertiger Annotationen zeigt sich auch in ihrer Bedeutung für die evidenzbasierte Medizin. Annotierte medizinische Datenbanken ermöglichen umfassende Analysen, die neue wissenschaftliche Erkenntnisse fördern und die Qualität klinischer Forschung und Praxis steigern. Besonders in Anwendungen wie der personalisierten Medizin, der prädiktiven Analyse oder der automatisierten klinischen Dokumentation sind NLP-Modelle auf robuste, annotierte Daten angewiesen, um präzise und verlässliche Vorhersagen zu generieren.

Zusammenfassend sind hochwertige Annotationen eine unverzichtbare Voraussetzung für den erfolgreichen Einsatz von NLP im medizinischen Bereich. Sie verbessern die Modellgenauigkeit, ermöglichen eine tiefgehende semantische Analyse und gewährleisten eine hohe Generalisierungsfähigkeit. Darüber hinaus sind sie essenziell für die Interoperabilität zwischen verschiedenen medizinischen Systemen sowie für die Unterstützung klinischer Entscheidungen. Ihre zentrale Rolle in der medizinischen Informatik unterstreicht die Notwendigkeit standardisierter, qualitativ hochwertiger Annotationen als Fundament für innovative, KI-gestützte Anwendungen in Forschung und Praxis.

3.3 Überblick über unterschiedliche Anwendungsmöglichkeiten

Natural Language Processing (NLP) und Künstliche Intelligenz (KI) haben sich in der Medizin zu Schlüsseltechnologien entwickelt, die zahlreiche Einsatzgebiete erschließen. Die automatisierte Analyse von elektronischen Gesundheitsakten, die strukturierte Dokumentation medizinischer Informationen oder die semantische Erschließung von Diagnosen und Behandlungsverläufen zählen zu den wichtigsten Anwendungsfeldern. NLP ermöglicht es dabei, unstrukturierte Daten – wie Freitext in medizinischen Entlassungsbriefen – maschinenlesbar zu machen und klinisch relevante Informationen effizient zu extrahieren. In Kombination mit KI-basierten Modellen lassen sich auf dieser Grundlage diagnostische Entscheidungen unterstützen, Patient*innenverläufe analysieren und präzise Therapieoptionen identifizieren.

Eine elektronische Gesundheitsakte (EHR, Electronic Health Record) ist eine digitale Version von Patient*inneninformationen. Diese enthält umfassende und aktuelle Informationen über die Krankengeschichte, Diagnosen, Medikationen, Allergien, Testergebnisse wie Labordaten und weitere relevante Gesundheitsinformationen. Digitale Patient*innenakten speichern ihre Daten oft in unstrukturierten freien Textfeldern, die nur eine geringe bis keine Standardisierung aufweisen. Natural Language Processing (NLP) wird eingesetzt, um die freien Textfelder der EHRs zu analysieren und dadurch strukturierte Daten aus unstrukturierten Informationen zu extrahieren. Dies ermöglicht eine effizientere Datenverarbeitung und verbessert die Nutzung der EHRs in der klinischen Praxis und Forschung. NLPs sind zusätzlich in der Lage nach Erfassen der unstrukturierten Daten die grammatische Struktur zu analysieren und die Bedeutung der Information zu bestimmen.

Medizinische Bildverarbeitung

Ein weiteres zentrales Anwendungsfeld ergibt sich im Bereich der medizinischen Bildverarbeitung. In der medizinischen Bildgebung haben sich Deep-Learning-Algorithmen bei der Erkennung und Diagnose verschiedener Krankheiten als wirksam erwiesen, da sie Röntgenbilder, MRTs und CT-Scans analysieren können und so vollständige und zeitnahe Informationen über Krankheiten und Verletzungen liefern. Diese Algorithmen können Radiologen bei der Erkennung von Tumoren, Frakturen und anderen komplexen Krankheiten, die mit herkömmlichen Methoden nur schwer zu erklären sind, unterstützen. Darüber hinaus sind sie in der Lage Krankheitsherde zu identifizieren und die am besten geeignete Behandlung für den Zustand einer Person zu bestimmen.

Eine Eigenschaft des Deep Learning besteht darin, dass es die in der Architektur der tiefen neuronalen Netze gefundenen Merkmale erlernt und diese automatisch aus den Daten extrahiert.

Dies erfolgt ohne die Notwendigkeit menschlicher Fähigkeiten, da sich das Deep Learning selbst programmiert. Darüber hinaus zeichnet sich Deep Learning dadurch aus, dass es eine große Datenmenge mit besserer Erfahrung und hoher Leistung verarbeitet, da das Netz die Aufgabe hat, die Daten zu klassifizieren und dann automatisch aus den Datenpraktiken lernt. So gilt Deep Learning als eine der Techniken, die erfolgreich Krankheiten von Hautkrankheiten bis hin zu Herzkrankheiten erklären können, da sie eine genaue Untersuchung durchführen können, um nach bösartigen Tumoren im Körper zu suchen oder Elektrokardiogramme (EKGs) zu verwenden, um unregelmäßige Herzschläge zu erkennen und die Möglichkeit eines Herzstillstands zu identifizieren.

Zu den bekanntesten Methoden des maschinellen Lernens, die häufig bei der Interpretation medizinischer Bilder eingesetzt werden, gehören konvolute neuronale Netze (CNNs – *convolutional neural networks*) und rekurrente neuronale Netze (RNNs – *recurrent neural networks*). Mit ihrer Hilfe können relevante Merkmale aus komplexen Bilddaten extrahiert und analysiert werden, um präzise diagnostische Aussagen zu ermöglichen.

CNNs werden beispielsweise eingesetzt, um Lungenrundherde auf Röntgenbildern zu identifizieren, Hautläsionen in der Dermatologie automatisiert als maligne oder benigne zu klassifizieren oder Tumorgrenzen in MRT-Aufnahmen exakt zu segmentieren.

RNNs kommen insbesondere bei der Auswertung von Bildserien oder in Kombination mit Textdaten zum Einsatz, etwa bei der zeitlichen Analyse von Bildbefunden in der Verlaufskontrolle oder bei der automatisierten Befundgenerierung.

(Maad et al., 2023) (Dhiman et al., 2022)

Während KI eine zentrale Rolle in der medizinischen Bildverarbeitung spielt, kommt Natural Language Processing (NLP) insbesondere in der direkten Interaktion mit Patient*innen zum Einsatz. Bei der Patient*inneneinbindung sind NLP-gestützte KI-Chatbots von wachsender Bedeutung, da sie rund um die Uhr verfügbar sind und unmittelbare Unterstützung bieten können. Diese Systeme lassen sich als Webanwendungen, mobile Apps oder innerhalb von Messenger-Plattformen integrieren. Sie ermöglichen es Patient*innen, in Echtzeit Termine zu vereinbaren, Erinnerungen zur Medikamenteneinnahme einzurichten oder personalisierte Gesundheitsinformationen abzurufen.

Unterstützung der Therapietreue: Eine der größten Herausforderungen im Gesundheitswesen ist die Einhaltung von Behandlungsplänen. KI-gesteuerte Chatbots senden Erinnerungen, Warnungen und Benachrichtigungen an Patient*innen, um sicherzustellen, dass sie ihre Behandlungspläne einhalten, was das Risiko verpasster Einnahmezeiten verringert.

Symptomüberwachung und -management: Patient*innen können ihre Symptome eingeben und Chatbots bieten in Echtzeit Hinweise oder empfehlen Maßnahmen. Diese Unterstützung ermöglicht es den Patient*innen, ihre Symptome besser zu verstehen und bei Bedarf den richtigen Ärzt*in aufzusuchen, wodurch die Gesundheitsversorgung verbessert wird.

Lebensstilberatung und Verhaltensmodifikation: KI-gestützte Chatbots geben personalisierte Empfehlungen zu Bewegung, Ernährung und Stressbewältigung. Kontinuierliche Interaktionen motivieren Patient*innen zu positiven Lebensstiländerungen, was die allgemeine Gesundheit und das Wohlbefinden verbessern.

Fortschrittsverfolgung und Feedback: Chatbots überwachen den Fortschritt der Patient*innen und geben personalisiertes Feedback zu Gesundheitszielen. Diese Feedbackschleife fördert Verantwortlichkeit und Motivation und ermöglicht es Gesundheitsdienstleistern, Behandlungspläne bei Bedarf anzupassen.

Emotionale Unterstützung und Einfühlungsvermögen: Chatbots bieten einfühlsame Gespräche und beruhigende Unterstützung, was die Patient*innenzufriedenheit erhöht und Ängste reduziert. Sie können als unvoreingenommene Zuhörer fungieren und helfen, emotionale Belastungen zu managen.

Befähigung der Patient*innen: Durch zugängliche, personalisierte Informationen ermöglichen Chatbots den Patient*innen, aktive Entscheidungen über ihre Gesundheit zu treffen, was zu höherer Zufriedenheit und Vertrauen führt.

Neben der Optimierung der Terminplanung können die virtuellen Assistent*innen auch Abrechnungsanfragen erleichtern und Verwaltungsaufgaben erledigen, sodass sich das medizinische Personal direkt auf die Patient*innenbetreuung konzentrieren kann.

Zusammengefasst tragen KI-gestützte Chatbots zu verbesserten Patient*innenergebnissen, Zufriedenheit und allgemeinem Wohlbefinden bei und revolutionieren die Patient*inneninteraktion mit Gesundheitssystemen. (Haritha et al., 2024, Manoharan et al., 2024)

Ein damit eng verknüpfter Bereich betrifft die Kommunikation zwischen medizinischem Fachpersonal und Patient*innen. Die Auswirkungen durch die Entwicklung des Internets hat sowohl negative als auch positive auf die Beziehung zwischen Ärzt*innen und Patient*innen. Patient*innen sind beispielsweise in der Lage aus dem Internet gewonnene Meinungen zum ärztlichen Termin mitzubringen, diese können zu einer sachkundigen Diskussion und besseren Entscheidungsfindungen beitragen. Andererseits kann sich auf Grundlage von Internet-Ratschlägen eine starre Vorstellung über die optimale medizinische Behandlung entwickeln, die

wiederum nicht mit der professionellen Meinung des Arztes/ der Ärztin übereinstimmen und die therapeutische Allianz belasten kann.

KI-generierte medizinische Meinungen werden stetig zugänglicher und zuverlässiger. Patient*innen können sich so an die KI-Software wenden und Ratschläge zu Gesundheitsversorgung erhalten. Das Aufkommen von KI-generierten medizinischen Gutachten führt wahrscheinlich in Folge zu einer weiteren Verschiebung der Dynamik zwischen Ärzt*innen und Patient*innen oder Familienmitgliedern. Ein Beispiel für eine auf Medizin spezialisierte KI ist Glass AI, eine auf GPT-4 basierte Technologie, bei der die Nutzer*innen ein klinisches Szenario vorstellen und anschließend eine Differenzialdiagnose oder einen klinischen Behandlungsplan erstellt. Vorteil der KI liegt darin, dass sie große Datenmengen zu verarbeiten und Muster zu erkennen, die für menschliche Kliniker*innen nicht ohne weiteres erkennbar sind. (Esteva et al., 2017, Rajpurkar et al., 2017)

Diese Technologien wurden bereits auf eine Reihe von klinischen Szenarien angewandt und können unter bestimmten Umständen biologische Signaturen in Patient*innendaten erkennen, die sich der menschlichen Interpretation entziehen. Als Beispiel konnte so ein automatisiertes Deep-Learning Modell von Netzhautfundusfotografien aus einer britischen Datenbank das von einem Patient*innen angegebene Geschlecht zuverlässig vorhersagen. (Korot et al., 2021)

Der wissenschaftliche Ansatz betont die Notwendigkeit für Ärzt*innen, den Wandel im Gesundheitswesen aktiv anzunehmen, insbesondere im Kontext der Integration von KI-generierter medizinischer Expertise. Es ist von entscheidender Bedeutung, dass die Angehörigen der Gesundheitsberufe diesen Wandel nicht als Bedrohung, sondern als Chance zur Stärkung der Beziehung zu ihren Patient*innen wahrnehmen. Ein kooperativer Ansatz, der Partnerschaft und gemeinsame Entscheidungsfindung fördert, ist hier von zentraler Bedeutung. Ärzt*innen sollten offen sein für Diskussionen über KI-generierte Meinungen und diese als wertvolle Ergänzung betrachten, während sie gleichzeitig ihre unverzichtbare Rolle als menschliche Experten behalten. Die aktive Einbeziehung von Patient*innen und Angehörigen in den Entscheidungsprozess kann das Vertrauen in und die Akzeptanz von neuen Technologien stärken. (Leslie, 2019)

Die Einbeziehung der Patient*innen in den Entscheidungsprozess ist eine wesentliche Voraussetzung für die verantwortungsvolle Integration von KI in die Gesundheitsversorgung. Die Berücksichtigung von Erkenntnissen, die mittels KI generiert wurden, neben menschlichem Fachwissen und den Präferenzen der Patient*innen, führt zu einer umfassenden Entscheidungsgrundlage. (Aminololama-Shakeri and López, 2019) Die Förderung der Zusammenarbeit zwischen Patient*innen, Familien und Gesundheitsdienstleister*innen stellt

ein wesentliches Element patient*innenzentrierter Pflegemodelle dar und leistet somit einen Beitrag zur Erreichung des genannten Ziels. (Stacey et al., 2017) Es obliegt den Gesundheitsdienstleistern, über die Möglichkeiten und Grenzen der KI informiert zu sein, um ihren Patient*innen die KI-Gutachten adäquat erläutern zu können. (Topol, 2019)

Es sind gezielte Schulungsprogramme und Initiativen zur Patient*innenaufklärung erforderlich, um eine ethische Annahme von KI und informierte Entscheidungen zu fördern. Eine informierte Entscheidung seitens der Patient*innen setzt voraus, dass sie über die Fähigkeiten und Grenzen der KI im Bilde sind. Allerdings können Diskrepanzen zwischen KI-Gutachten und ärztlichen Empfehlungen zu Akzeptanzproblemen führen. Offenheit und Transparenz sind essenziell, um Vertrauen zu fördern und eine kooperative Entscheidungsfindung im Zeitalter der KI sicherzustellen.

Ein weiterer zentraler Bereich, in dem NLP und KI in der Medizin Anwendung finden, ist die Arzneimittelüberwachung und Pharmakovigilanz. Die therapeutische Arzneimittelüberwachung (Therapeutic Drug Monitoring, TDM) hat sich seit den 1960er Jahren als Eckpfeiler der Präzisionsmedizin etabliert. Zu den frühen Innovatoren, die den heutigen Bereich des TDM abdecken, sind Entwicklungen wie Flüssigchromatographie-Massenspektrometrie-Assays, sowie Definitionen von mathematischen Modellen zur Beschreibung der Pharmakokinetik und Pharmakodynamik von Arzneimitteln. (Holford and Sheiner, Jelliffe et al., Sheiner, Sheiner and Beal) Die mathematischen Gleichungen dienten als Grundlage für die Populations-PK/PD-Modellierungsansätze unter der Verwendung der Modellierung nichtlinearer gemischter Effekte, die auch heute noch während des gesamten Arzneimittelentwicklungsprozesses verwendet werden. (Darwich et al., Mould and Upton)^{102]} Beispiele für Künstliche Intelligenz und Machine Learning Prozesse im Bereich TDM und Model-informed Precision Dosing (MIPD) sind Vorhersagen über die Konzentration und Exposition, Dosis-Vorhersage, ML-Anwendungen zur Unterstützung von MIPD-Systemen, ML-Ansätze für die quantitative Systempharmakologie und Anwendungen zur Unterstützung von EHR-Systemen und Datenerfassung.

Ein Ansatz für die Integration von künstlicher Intelligenz in TDM kann durch maschinelles Lernen erfolgen, um Konzentrations- und Expositionsvorhersagen zu erstellen. Woillard et al. (Woillard et al., 2021) verwendeten extreme gradient boosting Modelle zur Vorhersage der Tacrolimus- und Mycophenolsäure-Exposition und fanden, dass diese ML-Modelle den Bayes'schen Schätzungsansatz übertrafen (Marquet et al., 2021). Dies zeigte sich besonders bei Tacrolimus und Mycophenolsäure, wobei große Trainingsdatensätze notwendig waren. Eine weitere Studie zeigte vielversprechende Ergebnisse bei einem künstlichen neuronalen Netzwerk

die Konzentrations-Zeit-Profile vorherzusagen, obwohl es bei der Extrapolation noch Schwächen zeigte. (Bräm et al., 2022) Ein weiteres neuronales Netzwerk mit Langzeitgedächtnis zur Vorhersage von Valproinsäurekonzentrationen bei älteren Erwachsenen übertraf das Populations-PK-Modell, wobei jedoch weitere Arbeiten zur Verbesserung der Vorhersageleistung erforderlich sind. (Soeorg et al., 2022)

Ein weiterer Ansatz sind ML für Vorhersagen im Zusammenhang mit der Dosisoptimierung. In einer Studie wurde ein Klassifizierungs- und Regressionsbaum entwickelt, um die anfängliche Vancomycin-Dosierung bei erwachsenen Patient*innen zu bestimmen. (Imai et al., 2020) Das Modell prognostizierte die anfängliche Vancomycin-Dosis unter Verwendung pharmakokinetisch relevanter Kovariaten (z. B. Alter, BMI und geschätzte glomeruläre Filtrationsrate) mit moderater Genauigkeit und zeigte eine bessere Performance beim Erreichen therapeutischer Bereiche von 10-5 mg/L, 10-20 mg/L und ≥ 20 mg/L als andere MIPD-Methoden. Es sei darauf hingewiesen, dass Klassifizierungs- und Regressionsbäume häufig zu einer Überanpassung neigen. Daher ist eine externe Validierung erforderlich, um die Verallgemeinerbarkeit dieses Ansatzes zu gewährleisten. Ähnliche Anwendungen für andere Medikamente, die eine Dosisoptimierung erfordern, wurden entwickelt. In einer Studie von Zhu et al. (Zhu et al., 2021) wurde ein Extra-Bäume-Regressor-Algorithmus mit anderen evaluierten ML-Modellen hinsichtlich der Vorhersage des Verhältnisses von Lamotrigin-Konzentration zu Dosis bei erwachsenen Patient*innen verglichen. Dabei zeigte sich, dass der Extra-Bäume-Regressor-Algorithmus die beste Performance aufwies. Das Modell zeigte in den höheren Konzentrations-/Dosisverhältnissen eine bessere Performance als in den niedrigeren Bereichen. Des Weiteren wurde von den Autoren ein Vorschlag für eine webbasierte Anwendung unterbreitet, welche mit einem EHR-System verknüpft ist. Diese Anwendung soll eine personalisierte Dosisanpassung auf Basis der vom behandelnden Person eingegebenen demografischen und klinischen Patient*innendaten (z. B. Alter, Geschlecht, Körpergewicht, aktuelle Tagesdosis und gleichzeitig verabreichte Wirkstoffe) sowie der gewünschten Lamotrigin-Konzentration ermöglichen.

Im Bereich der Informatikwerkzeuge, -modelle und -techniken zur Unterstützung von MIPD-Systemen wurden mehrere Entwicklungen vorgenommen. Hughes und Keizer entwickelten einen kombinierten ML- und PK-Ansatz, um vorherzusagen, wann für die Bayes'sche Schätzung abgeflachte Priore verwendet werden sollten. (Hughes and Keizer, 2021) Dadurch wird die Flexibilität bei der Schätzung der Arzneimittelkonzentrationen erhöht. Im Rahmen der Untersuchung wurde ein XGBoost-Modell als das beste Modell identifiziert und in die MIPD-Software InsightRx Nova integriert. Der Ansatz zeigte eine verbesserte Leistung im Vergleich

zur alleinigen Maximum-a-posteriori-Methode für drei Vancomycin-Populations-PK-Modelle, was den klinischen Nutzen und die Verallgemeinerbarkeit unterstreichen.

Ein weiterer Ansatz besteht in der Anwendung von Machine Learning (ML) zur Entwicklung von Populations-PK-Modellen. Sibieude et al. (Sibieude et al., 2021) testeten einen ML-Rahmen, der beim Screening von Kovariaten helfen soll, indem mehrere ML-Modelle wie Support-Vector-Maschinen, Random Forest und neuronale Netze verwendet wurden. Die Studie zeigte, dass der ML-Ansatz vergleichbare Ergebnisse wie traditionelle Verfahren der Kovariatenauswahl lieferte, jedoch deutlich schneller war. Der Ansatz wurde weiter validiert, indem die Cetuximab-Clearance und das Verteilungsvolumen anhand von 30 Kovariaten vorhergesagt wurden. Eine nachfolgende Studie erweiterte dies mit Shapley Additive Explanations (SHAP), um zu veranschaulichen, wie sich Kovariaten auf PK-Parameter für Faktor-VIII-Konzentrat bei Hämophilie-A-Patient*innen auswirken. Im Rahmen der durchgeführten Studie wurde ein Random-Forest-Modell als bestes Modell identifiziert, welches zur Vorhersage der Parameter für Clearance und Verteilungsvolumen verwendet wurde. Dabei wurden die Kovariateneffekte berücksichtigt, die in früheren PK-Studien berichtet wurden.

Zur Unterstützung bei der Modellauswahl haben Lee et al. (Lee et al., 2022) eine ML-Methode zur Bestimmung des besten Vancomycin-PK-Modells anhand virtuell generierter Patient*innen evaluiert. Im Rahmen der Evaluierung wurden drei ML-Algorithmen – Entscheidungsbaum, Random Forest und XGBoost – bewertet. Dabei konnte XGBoost als bestes Modell identifiziert werden. Die Leistungsfähigkeit des Modells variierte in Abhängigkeit von den Szenarien der Konzentrationsentnahme und Dosismodelle, lieferte jedoch konsistentere Ergebnisse im Vergleich zu individuellen Modellen ohne Klassifikator. Die Leistungsfähigkeit des Modells zeigte eine Verbesserung mit zunehmender Anzahl der Konzentrationen, was auf eine potenzielle Anwendung zur Unterstützung des Modellauswahlprozesses hindeutet. Allerdings sind weitere Studien mit umfangreichen realen Datensätzen erforderlich, um die Ergebnisse zu validieren.

Die quantitative Systempharmakologie (QSP) ist ein vielversprechender Forschungsbereich, der von maschinellen Lernansätzen profitieren könnte. Diese sind in der Lage, große Datenmengen effizient zu verarbeiten und genaue Vorhersagen über Krankheits- und Arzneimittelmechanismen zu treffen. Die quantitative Systempharmakologie (QSP) verbindet Systembiologie und Pharmakometrie, um mathematische Modelle zu erstellen, welche die dynamischen Interaktionen zwischen Arzneimitteln, Physiologie und Krankheitssystemen beschreiben. Diese Modelle finden zunehmend Anwendung in der Arzneimittelentwicklung,

um optimale therapeutische Strategien zu ermitteln oder neue therapeutische Ziele zu identifizieren. (Azer and Barrett, Vicini and van der Graaf)

Aghamiri et al. (Aghamiri et al., 2022) haben Anwendungen, die ML und QSP integrieren, umfassend untersucht und die Fähigkeit von ML hervorgehoben, komplexe QSP-Plattformen zu unterstützen. Ein Whitepaper der International Society of Pharmacometrics identifiziert vier Kategorien aktueller Forschungsanwendungen für die Integration von QSP und ML: Parameterschätzung und -extraktion, Modellstruktur, Dimensionsreduktion sowie Stochastizität und virtuelle Populationen. (Zhang et al., 2022)

Ein weiterer Bereich, in dem der Einsatz von KI stetig an Bedeutung gewinnt, ist die diagnostische Unterstützung. Bei der Diagnosestellung können NLP unstrukturierte Daten aus den elektronischen Gesundheitsakten - den EHR wie ärztliche Befundbriefe, Patient*innenanamnese und Laborereignisse, analysieren und relevante Informationen extrahieren. So können KI-Modelle Symptome automatisch erkennen und mit bekannten Krankheitsbildern abzugleichen, um mögliche Diagnosen zu identifizieren. Dabei werden Muster aus großen Datenmengen erkannt, die wiederum auf gewisse Krankheiten hinweisen und so die Diagnose von komplexen oder seltenen Krankheiten erleichtert beziehungsweise bei der Diagnose nützlich ist.

Zudem können KI-Systeme auf umfangreiche medizinische Datenbanken zugreifen und ähnliche Fälle analysieren, um alternative Diagnosen vorzuschlagen, die vom medizinischen Personal möglicherweise übersehen wurden. Die Vorschlagstellung von Differentialdiagnosen dienen als Expertensysteme und nutzen dabei regelbasierte Ansätze und medizinisches Fachwissen, um Differentialdiagnosen zu generieren, wobei sie Symptome, Laborereignisse und andere relevante klinische Daten berücksichtigen.

In der Präzisionsmedizin analysieren KI genetische, klinische und Lebensstil-Daten um bei der Erstellung von personalisierten Behandlungsplänen die medizinischen Entscheidungen zu unterstützen. Dabei ist es möglich Risiken für Krankheitsverläufe und -komplikationen prognostizieren und so präventive Maßnahmen zu unterstützen und Behandlungsstrategien anzupassen. (Malik et al., 2024) Klinische Entscheidungsunterstützungssysteme bieten evidenzbasierte Empfehlungen und aktuelle Forschungsergebnisse in Echtzeit, um Ärzt*innen bei der Entscheidungsfindung zu unterstützen. Alarm- und Benachrichtigungssysteme informieren Ärzt*innen über abnormale Werte oder kritische Veränderungen im Zustand eines Patient*innen, die sofortiges Handeln erfordern. So untersuchte Bhaker et al. Neuere, auf computergestützter Intelligenz basierende Ansätze zur Ermittlung des Schweregrades von Krankheiten. In der Studie wurde sich auf die Parkinson-Krankheit und diabetische

Retinopathie konzentriert, wobei auch auf weitere Erkrankungen wie COVID-19, Osteoarthritis des Knie, autonome Neuropathie, Tuberkulose und Sepsis eingegangen wurde. Über die Parkinson-Krankheit wurde festgestellt, dass die Genauigkeit durch die Verwendung nicht-motorischer Symptome erheblich verbessert werden kann. Bei der diabetischen Retinopathie kann eine frühzeitige Erkennung den Patient*innen helfen, sich schneller zu erholen. (Bhakar et al., 2023)

In diesem Zusammenhang präsentieren Gil-Rios et al. (Gil-Rios et al., 2024) eine neue Methode zur automatischen Klassifizierung von Koronarstenosen. Die von ihm vorgeschlagene Methode erreichte mit vier Merkmalen eine Unterscheidungsrate von 99%. Das lässt darauf schließen, dass einem klinischen Umfeld diese Methode eine nützliche Ergänzung in Entscheidungsfällen sein könnte.

Die personalisierte Medizin profitiert in erheblichem Maße von der Integration von Künstlicher Intelligenz (KI), insbesondere durch die Nutzung detaillierter Patient*innenprofile. Diese Profile basieren auf der genauen Analyse individueller genetischer, biochemischer und verhaltensbezogener Daten, die in personalisierten Behandlungskonzepten entscheidend sind. KI und NLP ermöglichen es, diese komplexen und oft unstrukturierten Daten effizient zu verarbeiten, um gezielte Behandlungsstrategien für jeden Patient*innen zu entwickeln.

Eine zentrale Anwendung von NLP in der personalisierten Medizin liegt in der Analyse von Patient*innenakten und anderen textbasierten Gesundheitsdaten, die essentielle Informationen zu Symptomen, Krankheitsverläufen und genetischen Dispositionen enthalten. NLP-Algorithmen können aus diesen Datenströmen präzise und relevante Informationen extrahieren und in strukturierte Formate umwandeln, die für die Diagnose und Behandlung nutzbar sind. Ein Beispiel ist die Analyse von Genexpressionsdaten oder von spezifischen Mutationen in der DNA, die mittels NLP erfasst und systematisch ausgewertet werden können, um gezielte medikamentöse Therapien anzupassen oder die Reaktion des Körpers auf bestimmte Medikamente zu prognostizieren.

In der personalisierten Medizin wird außerdem die Möglichkeit geschaffen, Interventionen durch detaillierte Patient*innenprofile zu optimieren. Beispielsweise können genetische oder molekulare Analysen durch NLP-Prozesse beschleunigt werden, sodass Ärzt*innen schnell die am besten geeigneten Behandlungsmethoden für die jeweilige genetische Konstellation eines Patient*innen auswählen können. (Crowell et al., 2018) Dieser Ansatz führt zu einer individualisierten Therapie, die auf die spezifischen Bedürfnisse und Risiken jedes Einzelnen zugeschnitten ist und damit die Effektivität der Behandlung signifikant steigert.

Darüber hinaus spielt NLP eine Rolle bei der Echtzeitanpassung von Behandlungsplänen. Da Patient*innenprofile kontinuierlich aktualisiert werden, können NLP-gestützte Systeme die gewonnenen Daten analysieren und den Behandlungsfortschritt in Echtzeit überwachen. Auf diese Weise lässt sich nicht nur der Krankheitsverlauf besser vorhersagen, sondern auch frühzeitig auf mögliche Nebenwirkungen oder Veränderungen im Gesundheitszustand reagieren, was den therapeutischen Erfolg weiter verbessert.

Die Nutzung von KI und NLP zur Erstellung und Verwaltung umfassender Patient*innenprofile revolutioniert somit die personalisierte Medizin, indem sie eine präzisere und schnellere Diagnose sowie eine optimierte Therapie ermöglicht. Diese Technologien machen es möglich, das gesamte Spektrum der Patient*inneninformationen zu erfassen und zu analysieren, was die Anpassung von Gesundheitsinterventionen an individuelle Patient*innenbedürfnisse erheblich verbessert und die Grundlage für eine zukunftsweisende, datengestützte Gesundheitsversorgung schafft.(Schork, 2019)

3.4 Praxisnahe Anwendungen von NLP in der Medizin

Um die Vielfalt aktueller Entwicklungen zu verdeutlichen, stellt dieses Kapitel vier exemplarische Projekte vor, die unterschiedliche Anwendungsbereiche und methodische Ansätze abdecken.

Die Auswahl umfasst Anwendungen zur Selbstdiagnose und Patient*innenberatung (Your.MD), zur personalisierten Medikationsanpassung (CURATE.AI), zur ärztlichen Entscheidungsunterstützung in der Onkologie (Watson for Oncology) sowie zur psychischen Gesundheitsförderung mittels KI-Chatbot (Woebot).

Gemeinsam ist diesen Projekten, dass sie den Einsatz von KI in klinisch relevanten Kontexten demonstrieren, dabei unterschiedliche Technologien (NLP, Entscheidungsbäume, Optimierungsalgorithmen) verwenden und zugleich zentrale Chancen und Herausforderungen der KI-Implementierung in der Praxis aufzeigen.

3.4.1 Your.MD

Your.MD, auch bekannt unter dem Namen Healthily, ist ein digitales Gesundheitsunternehmen, das Künstliche Intelligenz (KI) nutzt, um Nutzer*innen über einen Chatbot personalisierte Gesundheitsinformationen zur Verfügung zu stellen. Die Plattform, die 2013 von Henrik Pettersen in Oslo gegründet wurde und ihren Hauptsitz heute in London hat, verspricht Nutzer*innen einfachen Zugang zu selbstdiagnostischen Informationen und gesundheitlicher Beratung. Das Unternehmen verzeichnete 2020 eine rasante Nutzer*innenzunahme, die vor allem auf das COVID-19 Symptom-Mapper-Tool zurückzuführen ist, das während der Pandemie viel Aufmerksamkeit erhielt und die App zu einem populären Mittel zur Eigenanalyse machte. (The Economist, 2025)

Das Prinzip von Your.MD basiert auf einem interaktiven Chatbot, der durch gezielte Fragen und mithilfe von KI-Algorithmen Symptome analysiert und individuelle Gesundheitshinweise gibt. Um die Genauigkeit der medizinischen Daten zu erhöhen, wurde 2018 eine Partnerschaft mit BMJ Best Practice eingegangen, um die medizinischen Inhalte der App zu validieren. Gleichzeitig hebt sich die Plattform durch ihre umfassende Datensicherheitszertifizierung hervor, da sie das ePrivacy-Siegel erhielt und damit die europäischen Datenschutzgesetze vollständig erfüllt.

In einer Zwei-Jahres-Folgeuntersuchung zur Leistung von selbstdiagnostischen Apps bewertete Aleksandar Ćirković Your.MD zusammen mit weiteren führenden Apps wie Ada und Babylon. Die Untersuchung, die auf der Diagnose von Augenkrankheiten wie Glaukomanfall und Netzhautablösung basierte, ergab, dass Your.MD im Vergleich zu den anderen Apps eine solide

Diagnoseleistung aufwies. Besonders 2018 schnitt Your.MD in Bezug auf die korrekte Diagnose und die Einschätzung der Dringlichkeit von Behandlungsmaßnahmen am besten ab. Im Jahr 2020 zeigte sich jedoch eine leichte Verschlechterung bei der Präzision bestimmter Diagnosen. Beispielsweise wurde für das nicht-akute Syndrom der trockenen Augen anstelle einer Selbstbehandlung eine Notfallbehandlung empfohlen – ein Ergebnis, das auf die Variabilität von Algorithmen und mögliche Unschärfen in der Langzeitoptimierung der KI hinweist (Ćirković, 2020). (Ćirković, 2020)

3.4.2 CURATE.AI

Die Studie von Blasiak et al. (2019) untersucht den Einsatz von CURATE.AI, einer künstlichen Intelligenz (KI)-Plattform, die personalisierte Medizin durch optimierte Dosierung und Medikamentenkombinationen unterstützt. Hier sind die zentralen Ergebnisse und Mechanismen der CURATE.AI-Plattform:

CURATE.AI setzt auf den Ansatz der phenotypischen personalisierten Medizin, der es ermöglicht, Behandlungsstrategien auf die individuellen Bedürfnisse eines Patient*innen anzupassen. Dabei basiert der Algorithmus auf minimalen Eingabedaten (drei Datenpunkte), um die optimale Dosis eines Medikaments zu bestimmen. Dieser Ansatz minimiert den Bedarf an großen populationsbasierten Datensätzen und ermöglicht dynamische Anpassungen während der Behandlung.

Das Kernstück von CURATE.AI ist die Quadratische Optimierungsplattform (QPOP), die die Beziehung zwischen Medikamentendosierung und patient*innenspezifischer Reaktion in einer mathematischen Parabel darstellt. Diese Parabel wird kontinuierlich mit neuen Datenpunkten aktualisiert, um Dosisanpassungen in Echtzeit zu ermöglichen. Dabei berücksichtigt der Algorithmus sowohl die Wirkung einzelner Medikamente als auch potenzielle Wechselwirkungen zwischen verschiedenen Medikamenten. (Blasiak et al., 2020)

In der klinischen Studie von Mukhopadhyay et al. (2022) wurde CURATE.AI zur Behandlung von Patient*innen mit Hypertonie und Typ-II-Diabetes getestet. Bei der Hypertonie-Behandlung erreichten 70 % der Patient*innen innerhalb von vier Monaten eine effektive Kontrolle ihres Blutdrucks, definiert als Werte unter 135/85 mmHg bei häuslichen Messungen oder unter 140/90 mmHg in der Praxis. Besonders hervorzuheben ist, dass die durchschnittliche Zeit bis zur Erreichung der Blutdruckkontrolle weniger als 30 Tage nach der initialen Kalibrierung des CURATE.AI-Systems betrug. Dies verdeutlicht die Effektivität der dynamischen Dosisanpassungen, die durch die Quadratische Optimierungsplattform unterstützt werden. Patient*innen, deren Dosierung vollständig auf CURATE.AI basierte, erzielten zudem

bessere Ergebnisse als jene, bei denen Ärzt*innen die vorgeschlagenen Anpassungen teilweise überstimmten, was die präzisen und patient*innenspezifischen Empfehlungen des Systems unterstreicht.

In der Behandlung von Typ-II-Diabetes zeigte CURATE.AI ebenfalls vielversprechende Ergebnisse. Bei 65 % der Teilnehmenden verbesserten sich die HbA1c-Werte und näherten sich innerhalb von vier Monaten dem Zielwert von weniger als 7 % an. Durch die kontinuierliche Überwachung und Anpassung der Dosierung konnte zudem die Zeit im normoglykämischen Bereich (über 70 % im Bereich von 4,0–10,0 mmol/L) signifikant erhöht werden. Bemerkenswert ist die hohe Anpassungsfähigkeit von CURATE.AI, selbst bei Patient*innen mit Komorbiditäten, was die Vielseitigkeit und das Potenzial des Systems zur Optimierung der Therapie bei komplexen Krankheitsbildern unterstreicht. (Mukhopadhyay et al., 2022)

Die Anwendung von CURATE.AI steht vor mehreren Herausforderungen. Viele Ärzt*innen vertrauen den algorithmischen Empfehlungen nicht vollständig, was zu suboptimalen Ergebnissen führen kann. Zudem fehlt die Integration in elektronische Gesundheitsakten, wodurch der Workflow erschwert wird. Ohne Kontrollgruppe können die Studienergebnisse nicht direkt mit traditionellen Methoden verglichen werden. Auch regulatorische Anforderungen und begrenzte Akzeptanz erschweren eine breite Einführung des Systems. (Blasiak et al., Mukhopadhyay et al.)

3.4.3 Watson for Oncology

IBM Watson for Oncology (WFO) ist ein KI-gestütztes Entscheidungshilfesystem, das Ärzt*innen evidenzbasierte Empfehlungen für die Behandlung von Krebspatient*innen liefert. Die Plattform analysiert patient*innenspezifische Daten, wie diagnostische Befunde und genetische Profile, und bietet basierend auf umfangreicher medizinischer Literatur potenzielle Behandlungsstrategien an. Die Behandlungsmöglichkeiten werden dabei in Kategorien wie „empfohlen“, „in Erwägung zu ziehen“ und „nicht empfohlen“ eingeteilt, um die klinische Entscheidungsfindung zu unterstützen. Obwohl WFO insbesondere in ressourcenarmen Regionen ein großes Potenzial hat, bleiben wichtige methodologische und praktische Herausforderungen bestehen, die seine globale Anwendung erschweren (Jie et al., Tupasela and Di Nucci).

Eine Meta-Analyse dokumentiert eine durchschnittliche globale Konkordanz von 81,52 % zwischen den Empfehlungen von WFO und den Entscheidungen multidisziplinärer Teams (MDT). Diese Rate variiert jedoch stark je nach Krebsart, wobei sie bei Brustkrebs am höchsten (88,99 %) und bei Magenkrebs am niedrigsten (57,94 %) ausfällt. Der Nutzen von WFO wird

dabei vor allem in der Unterstützung der Entscheidungsfindung bei standardisierten Fällen gesehen, während es bei komplexeren Krankheitsbildern weniger hilfreich erscheint (Jie et al., 2021). Eine retrospektive Analyse in China zeigt eine noch größere Diskrepanz, mit einer Konkordanzrate von 96 % bei Eierstockkrebs, aber nur 12 % bei Magenkrebs. Diese Abweichungen werden vor allem auf Unterschiede in der Medikamentenverfügbarkeit und den Leitlinien zwischen den USA und China zurückgeführt (Zhou et al., 2018).

Kritik wird auch an der Abhängigkeit von US-amerikanischen Leitlinien und der algorithmischen Intransparenz von WFO geübt. Die mangelnde Anpassung an regionale Besonderheiten wird als eine wesentliche Einschränkung für die globale Anwendung betrachtet. Darüber hinaus fehlen randomisierte kontrollierte Studien (RCTs), die die Effektivität und Sicherheit von WFO unter klinischen Bedingungen validieren könnten. Ohne solche Studien bleibt die Evidenzbasis unzureichend, um die klinischen Vorteile von WFO zweifelsfrei zu belegen (Prochaska et al., Sheiner and Beal).

Die Ergebnisse der Studien zeigen, dass WFO regionale Anpassungen benötigt, um seine Wirksamkeit zu maximieren. Während globale Standards wie die NCCN-Leitlinien ein solides Fundament bieten, sind lokale medizinische Praktiken und kulturelle Faktoren entscheidend für die Akzeptanz und Effizienz des Systems. Gleichzeitig wird deutlich, dass Konkordanz allein kein verlässlicher Indikator für die Qualität der Empfehlungen ist, da sie die tatsächlichen klinischen Ergebnisse nicht vollständig widerspiegelt. Die Notwendigkeit, transparente Algorithmen zu entwickeln und RCTs durchzuführen, wird als zentrales Element hervorgehoben, um WFO weiter zu verbessern und seine Akzeptanz zu erhöhen (Choi, Tupasela and Di Nucci, Zhou et al.).

Zukünftige Studien sollten sich darauf konzentrieren, WFO regional zu optimieren, indem sie lokale pharmazeutische Märkte und klinische Gegebenheiten stärker berücksichtigen. Der Aufbau eines globalen Datennetzwerks könnte dazu beitragen, die Integration von internationalen und regionalen Erkenntnissen zu verbessern. Darüber hinaus sind RCTs erforderlich, um die Wirkung von WFO systematisch zu bewerten, insbesondere hinsichtlich der Verbesserung der Patient*innenversorgung und der langfristigen klinischen Ergebnisse. Solche Untersuchungen könnten dazu beitragen, die tatsächlichen Vorteile von WFO nachzuweisen und das System robuster und anpassungsfähiger zu gestalten (Choi, Jie et al., Tupasela and Di Nucci, Zhou et al.).

3.4.4 Woebot

Woebot ist ein KI-gestützter Chatbot, der entwickelt wurde, um Menschen bei der Bewältigung psychischer Herausforderungen wie Stress, Angstzuständen und Depressionen zu unterstützen. Woebot verwendet eine regelbasierte Entscheidungsbaumstruktur, die speziell für strukturierte, klinisch fundierte Konversationen entwickelt wurde. Dabei bietet er keine Diagnosen oder medizinische Ratschläge und betont die Eigenverantwortung der/des Nutzer*in. Die Effektivität von Woebot wurde in mehreren Studien untersucht, die seine Wirksamkeit in verschiedenen Kontexten sowie seine Herausforderungen beleuchten.

Eine randomisierte kontrollierte Studie untersuchte den Einsatz von Woebot bei Frauen mit postpartaler Depression (PPD). Die Ergebnisse zeigten eine signifikante Reduktion der Depressionswerte (gemessen mit dem PHQ-9) und eine hohe Zufriedenheit der Benutzer*in, wobei 91 % mittlere bis hohe Zufriedenheitswerte angaben. Allerdings fehlt es an Langzeitstudien, um die Nachhaltigkeit dieser Ergebnisse zu bewerten. (Darcy et al., 2023)

Eine weitere Untersuchung analysierte den Nutzen von Woebot bei Stressbewältigung und Resilienzförderung über eine 8-wöchige Intervention. Dabei konnte eine Reduktion von Stress (Cohen's $d = -0,391$) und eine Erhöhung der Resilienz ($d = 0,436$) festgestellt werden. Diese positiven Ergebnisse traten insbesondere bei Teilnehmer*innen mit klinisch erhöhten Symptomen auf. Allerdings wiesen die Autoren auf methodologische Einschränkungen hin, da das Studiendesign nicht randomisiert war und Langzeiteffekte nicht bewertet wurden. (Durden et al.)

Im Bereich problematischer Substanzgebrauch zeigte Woebot in einer explorativen Studie ebenfalls vielversprechende Ergebnisse. Die Nutzung des Chatbots führte zu einer Reduktion des Substanzgebrauchs (OR = 0,48) sowie einer Abnahme von Angst- und Depressionssymptomen (GAD-7 um -2,3; PHQ-8 um -2,1). Dennoch stellte die hohe Abbruchrate von 50,5 % eine signifikante Herausforderung dar. Zudem basierte die Studie auf einem explorativen Design ohne Kontrollgruppe, was die Aussagekraft einschränkt. (Prochaska et al., 2021)

Eine Studie zu Nutzer*innenbindung und Engagement identifizierte drei unterschiedliche Nutzer*innengruppen: typische Nutzer*innen, frühe Nutzer*innen und effiziente Teilnehmer*innen. Während typische Nutzer*innen mittlere Engagement-Werte zeigten, zeichnete sich die Gruppe der effizienten Teilnehmer*innen durch eine geringe Verweildauer, aber eine hohe Anwendung therapeutischer Prinzipien aus. Diese Gruppe zeigte die größten Verbesserungen bei Depressions- und Stresssymptomen. Die Studie verdeutlichte jedoch, dass

die Ergebnisse unter stark kontrollierten Bedingungen erzielt wurden, was ihre Generalisierbarkeit einschränkt. (Hoffman et al., 2023)

Die Effektivität von Woebot beruht auf seiner allgemeinen Wirksamkeit, der Skalierbarkeit und der Fähigkeit, therapeutische Bindungen aufzubauen. Besonders wirksam war Woebot bei spezifischen Zielgruppen wie Frauen mit PPD oder Personen mit moderatem Stress. (Darcy et al., Durden et al.) Durch seine digitale Natur ist der Chatbot kosteneffizient und bietet eine 24/7-Verfügbarkeit, wodurch physische und soziale Barrieren überwunden werden können. (Durden et al., Hoffman et al.) Studien zeigen zudem, dass Benutzer*innen eine starke therapeutische Bindung zu Woebot entwickeln, was die Therapieergebnisse positiv beeinflusst (Hoffman et al., 2023)

Trotz dieser Stärken bestehen auch Herausforderungen. Woebot ist nicht für den Einsatz bei schweren klinischen Fällen geeignet und kann in Krisensituationen trotz implementierter „Safety Net“-Funktionen nur eingeschränkt Unterstützung bieten. (Hoffman et al., Prochaska et al.) Zudem erschweren methodologische Herausforderungen wie explorative Designs und hohe Abbruchraten die Validität der Ergebnisse. (Hoffman et al., Prochaska et al.) Digitale Ungleichheiten stellen ein weiteres Hindernis dar, da stabile Internetverbindungen und digitale Kompetenz für die Nutzung erforderlich sind. (Hoffman et al., 2023) Schließlich fehlt Woebot trotz seines FDA-„Breakthrough Device“-Status eine umfassende regulatorische Zulassung, was seine klinische Anwendung begrenzt. (Darcy et al., Prochaska et al.)

In Zukunft könnte die Integration generativer KI die Personalisierung und Empathie von Woebot weiter verbessern, birgt jedoch Risiken wie unvorhersehbare Antworten. (Darcy et al., Hoffman et al.) Langzeitstudien sind erforderlich, um die Nachhaltigkeit und spezifische Wirksamkeit einzelner Module zu bewerten. (Durden et al., Prochaska et al.) Darüber hinaus sollten Maßnahmen ergriffen werden, um den Zugang für unterrepräsentierte Gruppen zu erweitern und so die Akzeptanz und Wirksamkeit des Chatbots zu erhöhen. (Hoffman et al., 2023)

3.5 Verzernte Ausgaben und Halluzinationen: Grenzen medizinischer KI-Systeme

Die Integration von Künstlicher Intelligenz (KI) in die medizinische Praxis eröffnet neue diagnostische und therapeutische Möglichkeiten, bringt jedoch auch substantielle Herausforderungen mit sich. Eine davon ist das Phänomen sogenannter „KI-Halluzinationen“. Darunter versteht man die Erzeugung von Informationen durch KI-Modelle, die zwar grammatikalisch korrekt und überzeugend formuliert erscheinen, inhaltlich jedoch falsch, irreführend oder frei erfunden sind. Dieses Problem tritt insbesondere bei großen Sprachmodellen auf, die auf probabilistischer Textgenerierung basieren – ein Mechanismus, der mit zunehmender Textlänge auch die Wahrscheinlichkeit semantischer Fehler erhöht (Humbusch et al.).

In medizinischen Kontexten hat diese Problematik besondere Relevanz. Wenn beispielsweise Chatbots oder Entscheidungsunterstützungssysteme auf unzureichenden oder verzerrten Trainingsdaten beruhen, können sie fehlerhafte Empfehlungen geben, die in sensiblen klinischen Entscheidungssituationen zu ernsthaften Konsequenzen führen. Studien zeigen, dass selbst fortgeschrittene KI-Modelle wie ChatGPT in der Lage sind, medizinische Fachfragen korrekt zu beantworten, bei der Generierung komplexerer Inhalte jedoch zu Halluzinationen neigen, insbesondere wenn sie außerhalb klar definierter Faktenbereiche agieren (Schubert and et al.).

Die Auswirkungen solcher Fehlaussagen reichen von falschen Diagnosevorschlägen bis hin zu riskanten Therapieempfehlungen. In der Praxis bedeutet dies, dass medizinisches Fachpersonal mit KI-generierten Inhalten konfrontiert wird, deren Wahrheitsgehalt ohne zusätzliche Kontrolle oft nicht zuverlässig eingeschätzt werden kann. Dies kann zu Verunsicherung führen – sowohl auf Seiten der Ärzt*innenschaft als auch bei Patient*innen, die zunehmend mit KI-basierten Informationen in Kontakt kommen. Die Gefahr liegt unter anderem darin, dass Patient*innen die von KI-Systemen gelieferten Aussagen nicht als probabilistische Schätzungen, sondern als gesicherte Fakten interpretieren (Haverkamp et al.).

Die wissenschaftliche Diskussion betont daher die Notwendigkeit eines kritischen und reflektierten Umgangs mit KI-Systemen in der medizinischen Versorgung. Das betrifft sowohl die Entwicklung robuster Prüfmechanismen für KI-Ausgaben als auch die Schulung von Anwender*innen, um Halluzinationen erkennen und bewerten zu können. Zudem wird zunehmend auf die Bedeutung von Trainingsdatenqualität, Modelltransparenz und

menschlicher Supervision hingewiesen, um das Risiko fehlerhafter Aussagen zu minimieren (Leible et al.).

Besondere Verantwortung kommt den Ärzt*innen zu, die trotz wachsender technischer Unterstützung weiterhin als zentrale Instanz für die korrekte Einordnung von Informationen agieren müssen. Ihre Aufgabe besteht nicht nur darin, fehlerhafte Ausgaben zu identifizieren, sondern auch Patient*innen über die Grenzen solcher Systeme aufzuklären. KI-Systeme sollten als Hilfsmittel verstanden werden, nicht jedoch als Ersatz für klinisches Fachwissen, Erfahrung und persönliche Beurteilung. Gerade bei komplexen oder seltenen Erkrankungen ist es unerlässlich, dass Entscheidungen nicht allein auf KI-generierten Vorschlägen basieren, sondern im interdisziplinären Dialog und unter Berücksichtigung der individuellen Patient*innensituation getroffen werden (Humbusch et al., Leible et al.).

Die kritische Auseinandersetzung mit Halluzinationen in medizinischen KI-Systemen zeigt eindrucksvoll, dass technologische Innovationen stets begleitet sein müssen von strukturierten Evaluationsmechanismen, kontinuierlicher Qualitätssicherung und einem klaren ethischen sowie professionellen Rahmen. Nur so kann das Vertrauen in KI-Anwendungen langfristig gesichert und ihr Nutzen für die medizinische Versorgung voll ausgeschöpft werden.

3.6 Potential von KI und NLP in der medizinischen Datenverarbeitung

3.6.1 Veränderungen im Gesundheitssystem durch Digitalisierung: Einsatzmöglichkeiten für NLP

Die fortschreitende Digitalisierung im Gesundheitswesen hat in den letzten Jahren tiefgreifende Veränderungen in der Art und Weise bewirkt, wie medizinische Daten generiert, gespeichert und verarbeitet werden. Durch die zunehmende elektronische Erfassung und Speicherung von Patient*innendaten, klinischen Berichten und Forschungsergebnissen stehen heute große Mengen unstrukturierter Daten zur Verfügung. Diese Entwicklung bildet die Grundlage für den erfolgreichen Einsatz von Natural Language Processing (NLP) in der Medizin, da solche Algorithmen in der Lage sind, Informationen aus textbasierten Datenquellen effizient zu extrahieren und für verschiedene Anwendungen nutzbar zu machen. (Mijwil et al., 2022)

Ein wesentlicher Aspekt der Digitalisierung im Gesundheitswesen ist die wachsende Verfügbarkeit elektronischer Gesundheitsakten (Electronic Health Records, EHRs). Diese Dokumente enthalten eine Vielzahl wertvoller Informationen wie Diagnoseberichte, Laborergebnisse, ärztliche Befundbriefe und Behandlungshistorien, die häufig in unstrukturierter Textform vorliegen. Ohne den Einsatz moderner Technologien wären diese Daten nur schwer zugänglich und könnten nicht im vollen Umfang genutzt werden. NLP-Algorithmen ermöglichen es, solche Textdaten zu strukturieren, zu analysieren und gezielt nach relevanten Informationen zu durchsuchen. Dadurch wird es beispielsweise möglich, medizinische Dokumente automatisch auf bestimmte Krankheitsbilder, Symptome oder Medikamentenwechselwirkungen zu prüfen, was die Diagnose und Therapieentscheidung erheblich beschleunigen kann. (Arora, 2020)

Ein konkretes Beispiel für die Auswirkungen der Digitalisierung ist die Möglichkeit, Patient*innenakten automatisiert auszuwerten, um Muster und Trends zu erkennen, die zuvor manuell nicht ersichtlich gewesen wären. So können durch NLP-Verfahren Risikogruppen identifiziert, Krankheitsverläufe besser prognostiziert und gezielte Präventionsmaßnahmen entwickelt werden. Die Analyse großer Datenmengen führt zu einer evidenzbasierten Entscheidungsfindung, die es Ärzt*innen ermöglicht, personalisierte Therapien für ihre Patient*innen zu entwickeln. Ein Beispiel hierfür ist die Verwendung von NLP zur Erkennung von seltenen Erkrankungen anhand von zuvor unzusammenhängenden Symptomen, die in verschiedenen klinischen Berichten dokumentiert wurden. (World Economic, 2023)

Ein weiterer Vorteil der Digitalisierung ist die Verbesserung der Kommunikation und Zusammenarbeit im Gesundheitswesen. Die Möglichkeit, Informationen schnell zwischen

verschiedenen Institutionen auszutauschen, erleichtert es Fachärzt*innen, Krankenhäusern und anderen medizinischen Einrichtungen, auf die relevanten Daten zuzugreifen und fundierte Entscheidungen zu treffen. NLP spielt in diesem Zusammenhang eine entscheidende Rolle, da es die Verarbeitung großer Mengen an textbasierten Informationen beschleunigt und sicherstellt, dass wichtige Inhalte zuverlässig erkannt und hervorgehoben werden. (Dicuonzo et al., 2023) Trotz dieser Fortschritte bringt die Digitalisierung auch Herausforderungen mit sich. Insbesondere der Datenschutz und die Datensicherheit sind essenzielle Aspekte, die berücksichtigt werden müssen, um das Vertrauen von Patient*innen und medizinischem Personal in digitale Technologien zu stärken. Die sensiblen Informationen in elektronischen Gesundheitsakten müssen vor unbefugtem Zugriff geschützt werden. Gleichzeitig muss sichergestellt werden, dass die eingesetzten Algorithmen transparent und nachvollziehbar sind, um potenzielle Verzerrungen und Fehlinterpretationen zu vermeiden. (Sætra and Fosch-Villaronga, 2021)

Die fortschreitende Digitalisierung hat somit sowohl Potenziale als auch Herausforderungen für den Einsatz von NLP in der Medizin geschaffen. Während die Verarbeitung großer Datenmengen neue Möglichkeiten zur Verbesserung der Diagnose und Therapie eröffnet, erfordert die sichere und ethische Nutzung dieser Daten ein hohes Maß an Verantwortung. Die Integration von NLP in digitale Gesundheitsanwendungen zeigt jedoch eindrucksvoll, dass die Digitalisierung ein entscheidender Faktor für die zukünftige Weiterentwicklung der medizinischen Versorgung ist. (Dicuonzo et al., 2023)

3.6.2 Fallstudien und Projekte

Anwendung von GPT-basierten Modellen in der medizinischen Diagnostik

Die Anwendung von GPT-basierten Modellen, wie GPT-3 und GPT-4, in der medizinischen Diagnostik hat in den letzten Jahren erheblich an Bedeutung gewonnen. Große Sprachmodelle (Large Language Models, LLMs) zeichnen sich durch ihre Fähigkeit aus, komplexe medizinische Texte zu verstehen, zu verarbeiten und präzise Antworten auf diagnostische Fragestellungen zu liefern. Die Integration dieser Modelle in den medizinischen Alltag hat das Potenzial, die klinische Entscheidungsfindung zu verbessern, administrative Prozesse zu automatisieren und die Qualität der Patient*innenversorgung zu steigern.

Ergebnisse aktueller Fallstudien

Eine Vielzahl von Studien zeigt vielversprechende Ergebnisse im Hinblick auf die Anwendung von GPT-Modellen in der Medizin. Eine Untersuchung von Horiuchi et al. (2024) analysierte den Einsatz von GPT-4 zur Interpretation von radiologischen Befunden. Dabei wurde

festgestellt, dass das Modell in der Lage war, auffällige Muster in Bilddaten wie Röntgenaufnahmen oder MRT-Scans präzise zu identifizieren und mögliche Diagnosen vorzuschlagen. In 22 % der Fälle deckten sich die Diagnosen des Modells mit den Beurteilungen erfahrener Radiologen. Besonders beeindruckend war die Fähigkeit des Modells, seltene Erkrankungen zu erkennen, die in den klinischen Praxisleitlinien nicht immer ausführlich beschrieben sind. (Horiuchi et al., 2024a)

Darüber hinaus zeigte eine Studie von Hirano et al. (2024), dass GPT-4 in der Lage ist, komplexe medizinische Dokumentationen zu analysieren und potenzielle diagnostische Fehler zu erkennen. In einem Test mit 324 anonymisierten Patient*innenakten identifizierte das Modell in 41 % der Fälle Diagnosen korrekt, die von den behandelnden Ärzt*innen übersehen wurden. Diese Erkenntnisse unterstreichen das Potenzial von GPT-Modellen, als unterstützendes Werkzeug in der Diagnostik zu fungieren und mögliche Fehldiagnosen zu reduzieren. (Sonoda et al., 2024)

Auch bei der Erstellung von Anamnese- und ärztlichen Briefen kommen GPT-Modelle zunehmend zum Einsatz. Studien zeigen, dass sie strukturierte und präzise Anamnesen formulieren können, die sowohl die Krankengeschichte des Patient*innen als auch aktuelle Beschwerden berücksichtigen. (Wang et al., 2024)

Herausforderungen

Trotz dieser Erfolge bestehen jedoch auch Herausforderungen. Eine zentrale Problematik ist die Frage der Haftung. Wenn ein KI-System eine fehlerhafte Diagnose vorschlägt, stellt sich die Frage, wer letztlich für die daraus resultierenden medizinischen Entscheidungen verantwortlich ist. Darüber hinaus gibt es Bedenken hinsichtlich der Transparenz und Nachvollziehbarkeit der von GPT-Modellen generierten Antworten. Da die Modelle auf umfangreichen Datenmengen trainiert werden, ist es oft schwierig nachzuvollziehen, auf welcher Grundlage eine bestimmte Diagnose vorgeschlagen wird. (Huppertz et al., 2024)

Ein weiterer kritischer Punkt ist die Notwendigkeit, GPT-Modelle kontinuierlich mit aktuellen medizinischen Daten zu versorgen. Medizinisches Wissen entwickelt sich ständig weiter, und es besteht die Gefahr, dass veraltete Informationen in die Entscheidungsfindung einfließen, wenn die Modelle nicht regelmäßig aktualisiert werden. Zudem muss sichergestellt werden, dass die Daten, auf denen die Modelle trainiert werden, frei von Verzerrungen (Bias) sind, um eine faire und objektive Diagnosestellung zu gewährleisten. (Brin et al., 2024)

Trotz dieser Limitationen eröffnen sich vielfältige Perspektiven für die Weiterentwicklung dieser Technologie.

Ausblick in die Zukunft

Der Einsatz von GPT-Modellen in der medizinischen Diagnostik steht zwar noch am Anfang, doch die bisherigen Ergebnisse lassen darauf schließen, dass diese Technologie die klinische Praxis nachhaltig verändern könnte. Künftig könnten GPT-Modelle verstärkt in elektronische Gesundheitsakten integriert werden, um Ärzt*innen in Echtzeit Entscheidungshilfen zu bieten – etwa durch das Vorschlagen potenzieller Diagnosen oder das Erkennen möglicher Wechselwirkungen von Medikamenten. (Horiuchi et al., 2024b)

Ein weiterer vielversprechender Ansatz ist die Kombination von GPT-Modellen mit anderen KI-Technologien wie der Computer Vision. Während GPT-Modelle Texte analysieren, könnten Bilderkennungsalgorithmen gleichzeitig radiologische Bilddaten auswerten. Diese multimodale Diagnostik würde nicht nur die Präzision erhöhen, sondern auch den Zeitaufwand für die Analyse reduzieren. (Dehdab et al., 2024)

Langfristig könnten GPT-Modelle zudem in der personalisierten Medizin eingesetzt werden. Durch die Analyse großer Mengen an Patient*innendaten lassen sich individuelle Risikoprofile erstellen und maßgeschneiderte Therapieempfehlungen ableiten. Dies könnte nicht nur die Effizienz der Versorgung steigern, sondern auch die Patient*innensicherheit erheblich verbessern. (Hu et al., 2023)

Abschließend lässt sich festhalten, dass GPT-basierte Modelle bedeutende Chancen zur Verbesserung der medizinischen Versorgung bieten – jedoch begleitet von der Notwendigkeit, ethische, rechtliche und technische Fragestellungen kontinuierlich zu adressieren.

Verbesserung der Kommunikation und Zusammenarbeit im Gesundheitswesen durch KI

Neben diagnostischen Aufgaben zeigt Künstliche Intelligenz (KI) auch großes Potenzial in der interpersonellen Kommunikation und organisatorischen Zusammenarbeit im Gesundheitswesen. Insbesondere im Zuge der fortschreitenden Digitalisierung eröffnen sich neue Möglichkeiten, den Informationsaustausch zwischen medizinischem Personal effizienter und patient*innenzentrierter zu gestalten. In diesem Abschnitt werden zentrale Anwendungsfelder, Ergebnisse aus Fallstudien sowie künftige Entwicklungsperspektiven beleuchtet.

Aktuelle Entwicklungen und Anwendungen

KI findet zunehmend Anwendung in der Verbesserung des Informationsflusses zwischen verschiedenen Akteuren im Gesundheitswesen. Ein Beispiel hierfür ist der Einsatz von Large Language Models (LLMs), die dazu beitragen, die Interoperabilität zwischen Gesundheitsdienstleistern und der pharmazeutischen Industrie zu optimieren. Durch die Strukturierung und Harmonisierung von Informationen ermöglichen LLMs eine konsistente und qualitativ hochwertige Patient*innenversorgung. (Isbary et al., 2024)

Auch in digitalen Gesundheits-Apps kommt KI verstärkt zum Einsatz. Diese Anwendungen unterstützen nicht nur beim Gesundheitsmonitoring und der Symptomerkennung, sondern fördern zugleich die Kommunikation zwischen Ärzt*innen und Patient*innen. Personalisierte Empfehlungen sowie optimierte Therapieabläufe können die Effizienz der Versorgung weiter verbessern. (Widmer, 2024)

Ergebnisse aus Fallstudien

Dass KI im klinischen Alltag konkrete Verbesserungen ermöglicht, zeigen verschiedene Studien. An der Universitätsmedizin Essen wurde beispielsweise untersucht, wie KI-basierte Anwendungen das medizinische Personal entlasten können, indem sie Routineaufgaben automatisieren und somit mehr Zeit für die direkte Patient*innenversorgung schaffen. Die Ergebnisse deuten darauf hin, dass KI zur Fehlervermeidung, Effizienzsteigerung und Qualitätsverbesserung beiträgt. (Bures et al., 2023)

Eine weitere Untersuchung beschäftigte sich mit der Nutzer*innenakzeptanz von KI im Gesundheitswesen. Dabei zeigte sich, dass KI-Anwendungen von medizinischem Fachpersonal als wertvolle Assistenzsysteme wahrgenommen werden. Auf Patient*innenseite bestehen jedoch weiterhin Bedenken – insbesondere im Hinblick auf den Datenschutz und das Bedürfnis nach einer persönlichen ärztlichen Betreuung. (Dahm et al., 2020)

Herausforderungen und zukünftige Perspektiven

Trotz der positiven Entwicklungen stellt die Integration von KI in die medizinische Praxis eine komplexe Aufgabe dar. Datenschutz und ethische Aspekte nehmen dabei eine zentrale Rolle ein, ebenso wie die Gewährleistung der technischen Zuverlässigkeit und Systemintegrität. Darüber hinaus erfordert eine erfolgreiche Implementierung eine enge Zusammenarbeit zwischen verschiedenen Akteuren im Gesundheitssystem – von IT-Entwicklern über Kliniken bis hin zu politischen Entscheidungsträgern. (Isbary et al., 2024)

Ein vielversprechender Ausblick ergibt sich im Bereich der KI-gestützten Ressourcenplanung in Krankenhäusern. Zukünftig könnten intelligente Systeme helfen, Prozesse zu optimieren und Verschwendung zu vermeiden. Durch die Integration von Steuerungsfunktionen ließe sich eine neue Qualität der Ressourcenallokation erreichen, die über bisherige Methoden hinausgeht. (Goerke et al., 2024)

Insgesamt zeigt sich, dass KI über ein enormes Potenzial zur Verbesserung der Kommunikation und Zusammenarbeit im Gesundheitswesen verfügt. Durch den gezielten Einsatz entsprechender Technologien können Informationsflüsse optimiert, die Versorgungsqualität gesteigert und das Gesundheitspersonal entlastet werden. Damit dies gelingt, bedarf es jedoch

nicht nur technischer Innovation, sondern auch gesellschaftlicher und institutioneller Akzeptanz..

Strukturierung von Patient*innendaten durch Verarbeitung natürlicher Sprache

Während sich der vorherige Abschnitt auf zwischenmenschliche und institutionelle Kommunikationsprozesse konzentrierte, steht in diesem Teil die semantische Aufbereitung medizinischer Informationen im Vordergrund. Die Strukturierung von Patient*innendaten mittels Natural Language Processing (NLP) hat in den letzten Jahren erheblich an Bedeutung gewonnen. NLP-Technologien ermöglichen es, unstrukturierte medizinische Texte wie ärztliche Befundberichte, Befunde oder Patient*innenakten in strukturierte Daten zu überführen. Dies erleichtert die Analyse, Speicherung und den Austausch von Gesundheitsinformationen und trägt damit zu einer effektiveren Patient*innenversorgung bei.

Ergebnisse aktueller Studien

Eine Studie von Sarmiento und Deroncourt (2016) untersuchte den Einsatz von NLP zur Identifizierung spezifischer Patient*innenkohorten aus großen klinischen Datenbanken. Die Forschenden verglichen die Extraktion strukturierter Daten mit der Analyse unstrukturierter klinischer Notizen mittels NLP. Die Ergebnisse zeigten, dass NLP-basierte Ansätze eine höhere Genauigkeit bei der Identifizierung von Patient*innengruppen aufwiesen, insbesondere in Fällen, in denen relevante Informationen ausschließlich in Freitextnotizen vorhanden waren. (Sarmiento and Deroncourt, 2016)

Auch eine systematische Überprüfung durch Hossain et al. (2023) beleuchtete den Einsatz von NLP in elektronischen Gesundheitsakten (EHR). Die Studie identifizierte vielfältige Anwendungsfelder, darunter die Klassifizierung medizinischer Notizen, die Erkennung klinischer Entitäten sowie die automatische Textzusammenfassung. Die Autoren unterstrichen, dass NLP-Technologien einen wertvollen Beitrag leisten können, um aus umfangreichen EHR-Daten klinisch relevante Erkenntnisse zu extrahieren und die Entscheidungsfindung im Gesundheitswesen zu unterstützen. (Hossain et al., 2023)

Zudem beleuchtete eine Publikation von Moslmani und Thielscher (2024) die Anwendung von Large Language Models (LLMs) wie ChatGPT in der Onkologie. Die Autoren diskutierten, wie LLMs bei der Auswertung von Patient*innendaten helfen können, insbesondere bei der Analyse und Zusammenfassung von Patient*innenakten. Sie hoben hervor, dass diese Technologien das Potenzial haben, die Effizienz in der onkologischen Versorgung zu steigern, warnten jedoch auch vor möglichen Auswirkungen auf die Beziehung zwischen Ärzt*innen und Patient*innen. (Moslmani and Thielscher, 2024)

Herausforderungen

Trotz dieser Fortschritte bestehen mehrere Herausforderungen. Eine zentrale Schwierigkeit liegt in der Vielfalt und Komplexität medizinischer Terminologie. Synonyme, Abkürzungen und unterschiedliche Schreibweisen erschweren die präzise Extraktion relevanter Informationen. Hinzu kommen regionale Unterschiede in der medizinischen Dokumentationspraxis, die die Entwicklung universell einsetzbarer NLP-Modelle zusätzlich erschweren.

Ein weiteres Problem stellt die Qualität der zugrundeliegenden Daten dar. Unvollständige oder fehlerhafte Dokumentationen können die Leistung von NLP-Systemen erheblich beeinträchtigen. Hinzu kommt die Notwendigkeit, Datenschutz- und Sicherheitsaspekte zu beachten, da Patient*innendaten als besonders schützenswert gelten und vor unbefugtem Zugriff geschützt werden müssen.

Ausblick in die Zukunft

Die kontinuierliche Weiterentwicklung von NLP-Technologien sowie der Einsatz leistungsfähiger LLMs eröffnen vielversprechende Perspektiven für die strukturierte Verarbeitung medizinischer Texte. Künftige Systeme könnten dazu in der Lage sein, nicht nur Inhalte zu analysieren, sondern auch Kontextinformationen einzubeziehen – etwa den zeitlichen Verlauf einer Erkrankung oder implizite Zusammenhänge zwischen Symptomen. Dies würde die Generierung umfassender Patient*innenprofile erleichtern und die personalisierte Medizin maßgeblich voranbringen.

Zudem bietet sich Potenzial in der Integration von NLP-Systemen in elektronische Gesundheitsakten, wodurch Echtzeit-Analysen und Entscheidungsunterstützung in der klinischen Praxis möglich würden. Gleichzeitig könnten administrative Aufgaben – etwa die Kategorisierung von Diagnosen oder die Dokumentation von Behandlungsverläufen – automatisiert werden. Dies hätte das Potenzial, medizinisches Personal erheblich zu entlasten und die Effizienz im Gesundheitswesen insgesamt zu steigern.

Langfristig könnten NLP-Technologien auch dazu beitragen, Gesundheitsdaten über Institutionen und Ländergrenzen hinweg zu vernetzen. Die Analyse solcher globaler Datenströme würde es ermöglichen, epidemiologische Trends frühzeitig zu erkennen und entsprechende Präventionsstrategien zu entwickeln – mit potenziell weitreichenden Effekten für die öffentliche Gesundheit.

Identifizierung von Mustern und Trends in Patient*innenakten für eine evidenzbasierte Entscheidungsfindung

Aufbauend auf der Strukturierung medizinischer Informationen rückt dieser Abschnitt die analytische Nutzung von Patient*innendaten zur Unterstützung evidenzbasierter

Entscheidungen in den Fokus. Die Erkennung von Mustern und Trends in klinischen Daten spielt eine zentrale Rolle für eine moderne, datengestützte Medizin. Durch den Einsatz von Künstlicher Intelligenz (KI) und Maschinellem Lernen (ML) lassen sich große Datenmengen systematisch auswerten, um fundierte Rückschlüsse auf Krankheitsverläufe und Therapiewirkungen zu ziehen.

Ergebnisse aktueller Studien

Eine Studie von Schneeweiss (2023) betont die Bedeutung von Real-World-Evidence (RWE) für die Bewertung der Wirksamkeit von Behandlungen im klinischen Alltag. Im Gegensatz zu randomisierten kontrollierten Studien ermöglichen RWE-Analysen Rückschlüsse auf die tatsächliche Wirksamkeit therapeutischer Maßnahmen unter Alltagsbedingungen. Dies erleichtert eine bessere Anpassung von Therapien an individuelle Patient*innenbedürfnisse und fördert somit die personalisierte Medizin. (Schneeweiss, 2023)

Auch in der Onkologie finden KI-basierte Verfahren zunehmend Anwendung. Laut einer Veröffentlichung von Kather und Niehues (2019) können KI-gestützte Systeme dazu beitragen, die wachsende Datenflut zu bewältigen. Durch das Erkennen klinisch relevanter Muster und Trends in Parametern wie Tumormarkern oder Therapieverläufen wird die Entscheidungsfindung unterstützt und die Behandlung von Krebspatient*innen optimiert. (Neumuth, 2020)

Darüber hinaus zeigt eine Untersuchung von Schmücker (2024), dass eine strukturierte medizinische Dokumentation eine wichtige Grundlage für das Health Data Management bildet. Durch die gezielte Erfassung und Analyse von Gesundheitsdaten lassen sich Informationen zur richtigen Zeit am richtigen Ort bereitstellen. Dies unterstützt die diagnostische und therapeutische Entscheidungsfindung erheblich. (Schmücker, 2024)

Herausforderungen

Trotz vielversprechender Fortschritte bestehen auch in diesem Bereich zentrale Herausforderungen. Die Qualität und Vollständigkeit der zugrunde liegenden Daten ist entscheidend für die Aussagekraft der Analysen. Unvollständige, fehlerhafte oder inkonsistente Einträge können zu falschen Interpretationen führen und die klinische Entscheidungsgrundlage verfälschen.

Ein weiteres Hindernis stellt die Integration heterogener Datenquellen dar. Da Patient*innendaten häufig aus verschiedenen Systemen stammen – etwa Laborinformationssystemen, Radiologiebefunden oder hausärztlichen Dokumentationen – sind standardisierte Schnittstellen und Interoperabilität unerlässlich, um diese Informationen sinnvoll zusammenzuführen und auszuwerten.

Ausblick in die Zukunft

Die fortlaufende Entwicklung von KI- und ML-Methoden eröffnet vielfältige Möglichkeiten, um Muster und Trends noch effizienter und präziser zu identifizieren. Zukünftig könnten prädiktive Modelle entwickelt werden, die Krankheitsverläufe antizipieren und frühzeitige Interventionen ermöglichen – etwa durch die automatische Erkennung von Risikofaktoren oder sich anbahnenden Komplikationen. Dadurch ließe sich nicht nur die Qualität der Versorgung verbessern, sondern auch die Effizienz des Gesundheitssystems steigern.

Zudem könnten Echtzeit-Analysen dazu beitragen, aktuelle Trends und Ausbrüche von Krankheiten zu identifizieren. Dies wäre besonders in Pandemie-Situationen von Vorteil, um schnell reagieren und entsprechende Maßnahmen ergreifen zu können.

Langfristig besteht die Möglichkeit, dass durch die Integration von Patient*innendaten aus verschiedenen Quellen ein umfassendes Bild des individuellen Gesundheitszustands entsteht. Dies würde eine echte personalisierte Medizin ermöglichen, bei der Behandlungen exakt auf die Bedürfnisse und Voraussetzungen des einzelnen Patient*innen abgestimmt sind.

3.7 Regulierungen

Regulierungen bilden die Grundlage für den sicheren und ethisch vertretbaren Einsatz von Künstlicher Intelligenz (KI) in der Medizin. Sie umfassen rechtliche und ethische Anforderungen, die sicherstellen sollen, dass KI-Systeme den Ansprüchen an Sicherheit, Transparenz und Datenschutz gerecht werden. Auf europäischer Ebene setzen Verordnungen wie der AI Act und die Datenschutz-Grundverordnung (DSGVO) den rechtlichen Rahmen für den Einsatz von KI. Innerhalb dieses Rahmens ermöglichen nationale Gesetze – wie etwa in Österreich – spezifische Umsetzungen und Anpassungen an das jeweilige Gesundheitssystem. Neben den rechtlichen Rahmenbedingungen spielen ethische Fragen eine zentrale Rolle. Dazu gehören Aspekte wie Verantwortung, Diskriminierung und Transparenz in KI-Systemen. Die ethischen Leitlinien der EU fordern, dass KI-Systeme rechtmäßig, ethisch und robust sein müssen, um den Schutz der Patient*innen und die Qualität der medizinischen Versorgung zu gewährleisten.

Unterschiede in der Haftung zwischen europäischen und nationalen Regelungen beeinflussen die Integration von KI-Systemen in die Praxis. Während auf EU-Ebene die Herstellerhaftung betont wird, liegt in Österreich die Verantwortung häufig bei der behandelnden Person. Solche Unterschiede haben Auswirkungen auf die Akzeptanz und den Einsatz von KI im klinischen Alltag.

Eine verantwortungsvolle Regulierung muss sowohl technische als auch ethische Aspekte umfassen, um eine sichere und innovationsfördernde Nutzung von KI im Gesundheitswesen zu ermöglichen.

3.7.1 Überblick über die bestehenden Rechtsrahmen in der EU und in Österreich

Der Einsatz von KI im Gesundheitswesen wird durch umfangreiche rechtliche Vorgaben sowohl auf EU-Ebene als auch auf nationaler Ebene in Österreich geregelt. Dabei zeigt sich ein dualer Ansatz: Während die EU harmonisierte Vorschriften für alle Mitgliedstaaten schafft, bietet die österreichische Gesetzgebung Raum für spezifische nationale Anpassungen. Dieser Ansatz hat praktische Relevanz, da nationale Besonderheiten in der Gesundheitsversorgung berücksichtigt werden müssen.

Auf EU-Ebene regelt die Verordnung (EU) 2017/745 über Medizinprodukte die Anforderungen an KI-gestützte Software als Medizinprodukt (Europäische Kommission, 2017). (DeepMind, 2020) Besonders relevant ist die Datenschutz-Grundverordnung (DSGVO), die klare Regeln für den Schutz personenbezogener Daten aufstellt und damit eine wichtige Grundlage für den Einsatz von KI im Gesundheitsbereich bildet (Europäische Union, 2016). (Europäische, 2017)

Ergänzt wird dieser Rahmen durch den geplanten AI Act, der einen risikobasierten Ansatz verfolgt und KI-Systeme im Gesundheitswesen als Hochrisiko-Anwendungen einstuft. (Europäische, 2021)

In Österreich wird der Datenschutz durch das Gesundheitstelematikgesetz (GTelG 2012) konkretisiert, das spezielle Anforderungen an die Verarbeitung von Gesundheitsdaten stellt. (Republik, 2012) Das Medizinproduktegesetz (MPG) regelt darüber hinaus die Zulassung und Überwachung von KI-Systemen als Medizinprodukte. (Republik, 2021) Das Ärztegesetz betont, dass klinische Entscheidungen weiterhin von medizinisch qualifiziertem Personal getroffen werden müssen, auch wenn KI-Systeme zur Unterstützung eingesetzt werden. (Republik, 1998) Ein Beispiel für den Unterschied zwischen EU- und österreichischen Regelungen zeigt sich in der Zulassung von KI-Systemen. Während die EU eine zentrale Harmonisierung anstrebt, müssen KI-Systeme in Österreich zusätzlich nationale Anforderungen erfüllen. Dies kann etwa bei der Entwicklung eines Diagnoseassistenzsystems relevant sein, das sowohl den europäischen Anforderungen an Sicherheitsprüfungen als auch den österreichischen Vorgaben zur Datenverarbeitung gerecht werden muss.

3.7.2 Ethische Aspekte des Einsatzes von KI in der Medizin

Ethische Fragen im Zusammenhang mit KI in der Medizin betreffen Themen wie Verantwortlichkeit, Bias und Diskriminierung, Transparenz und Patient*innenrechte. Die Ethik-Leitlinien der EU betonen die Notwendigkeit, dass KI-Systeme rechtmäßig, ethisch und robust sein müssen. (Europäische, 2019) Transparenz ist ein zentrales Prinzip, um sicherzustellen, dass medizinische Entscheidungen für Patient*innen nachvollziehbar bleiben. Ein Beispiel für eine ethische Herausforderung ist die potenzielle Diskriminierung durch KI-Systeme. Ein KI-gestütztes Diagnosewerkzeug könnte ungewollt Vorurteile reproduzieren, wenn die Trainingsdaten verzerrt sind. In Österreich legt das Gesundheitsqualitätsgesetz (GQG) besonderen Wert auf Patient*innensicherheit und fordert Maßnahmen zur Qualitätskontrolle und Transparenz. (Republik, 2004) Hier könnten verpflichtende Qualitätsstandards für KI-Systeme eingeführt werden, um solche Risiken zu minimieren.

Ein weiteres Beispiel ist die Frage der Verantwortung bei Fehlentscheidungen. Während die EU durch den AI Act eine klare Haftung der Hersteller von KI-Systemen vorsieht, bleibt in Österreich die Verantwortung bei der behandelnden Person. Dies könnte dazu führen, dass Ärzt*innen zögern, KI-Systeme in ihre Entscheidungsprozesse zu integrieren, da sie befürchten, für Fehler haftbar gemacht zu werden.

Um die zuvor dargestellten ethischen Fragestellungen greifbarer zu machen, werden im Folgenden konkrete KI-Systeme vorgestellt, die bereits in unterschiedlichen Bereichen des Gesundheitswesens eingesetzt werden. Anhand dieser Beispiele lassen sich potenzielle Herausforderungen – etwa hinsichtlich Verantwortung, Fairness, Transparenz oder Datenschutz – exemplarisch nachvollziehen und kritisch einordnen.

- **Unterstützende Diagnosetools** wie IBM Watson Health, das medizinische Daten analysiert und Behandlungsoptionen vorschlägt (IBM, 2023).
- **Organisatorische Hilfsprogramme** wie KI-gestützte Terminplanungssysteme, die den Patient*innenfluss in Krankenhäusern optimieren. (Nuance, 2022)
- **Bildanalyse-Algorithmen** wie DeepMind Health von Google, das in der Lage ist, Augenkrankheiten frühzeitig zu erkennen. (DeepMind, 2020)
- **Sprachbasierte Assistenzsysteme** wie Nuance Dragon Medical One, die Ärzt*innen bei der Dokumentation und Kommunikation helfen. (Nuance, 2022)

Diese Systeme zeigen, wie KI bereits heute in verschiedenen Bereichen des Gesundheitswesens eingesetzt wird und sowohl Chancen als auch Risiken birgt.

3.7.3 Analyse von Haftungsfragen und Sicherheitsrisiken

Die Haftung für Schäden, die durch KI-Systeme verursacht werden, ist sowohl auf EU- als auch auf österreichischer Ebene ein zentrales Thema. Der AI Act sieht vor, dass Hersteller von KI-Systemen für Schäden haften, die durch Fehlfunktionen ihrer Systeme verursacht werden (Europäische Kommission, 2021). In Österreich regelt das Medizinproduktegesetz die Haftung für Schäden durch Medizinprodukte und legt fest, dass Hersteller für Sicherheitsmängel ihrer Produkte verantwortlich sind (Republik Österreich, 2021).

Ein konkretes Beispiel ist der Einsatz eines KI-basierten Systems zur Analyse medizinischer Bilder. Sollte das System eine falsche Diagnose stellen, haftet der Hersteller gemäß den EU-Vorgaben. In Österreich könnte die behandelnde Person jedoch ebenfalls haftbar gemacht werden, wenn er sich blind auf das System verlässt. Dies zeigt, dass Haftungsfragen noch klarer geregelt werden müssen, um Ärzt*innen Sicherheit im Umgang mit KI-Systemen zu bieten.

Sicherheitsrisiken wie Cyberangriffe sind eine weitere Herausforderung. Das GTelG 2012 fordert strikte Sicherheitsmaßnahmen, um den Missbrauch von Gesundheitsdaten zu verhindern (Republik Österreich, 2012). Ein Angriff auf ein KI-System könnte schwerwiegende Folgen haben, etwa wenn ein System falsche Medikationsvorschläge macht oder kritische Daten manipuliert werden.

3.7.4 Vergleich verschiedener Regulierungsrahmen

Der Vergleich zwischen EU-weiten Vorgaben und österreichischen Regelungen zeigt wesentliche Unterschiede und Gemeinsamkeiten in der Art und Weise, wie KI-Systeme im Gesundheitswesen reguliert werden. Während die EU auf eine harmonisierte und standardisierte Regulierung abzielt, verfolgt Österreich eine stärker lokalisierte Anpassung an nationale Besonderheiten. Beide Ansätze bieten spezifische Vorteile und Herausforderungen.

Ein zentrales Beispiel ist die Zulassung von KI-Systemen. In der EU wird der Zulassungsprozess durch die Verordnung (EU) 2017/745 über Medizinprodukte geregelt, die einheitliche Anforderungen an Sicherheitsprüfungen und Leistungsnachweise stellt. In Österreich müssen diese Systeme zusätzlich den Anforderungen des Medizinproduktegesetzes (MPG) und des Gesundheitstelematikgesetzes (GTelG 2012) entsprechen. Diese beinhalten spezifische Regelungen zur Verarbeitung von Gesundheitsdaten.

Ein weiteres Beispiel betrifft die Haftung. Während der AI Act in der EU eine klare Herstellerhaftung vorsieht, bleibt in Österreich die Haftung häufig bei der behandelnden Person. Dies hat direkte Auswirkungen auf die Bereitschaft, KI-Systeme in die Praxis zu integrieren. Ärzt*innen könnten zögern, neue Technologien einzusetzen, wenn unklar ist, wer im Schadensfall verantwortlich ist.

Zusammenfassend lässt sich sagen, dass beide Regulierungsansätze spezifische Herausforderungen adressieren. Während die EU auf eine einheitliche Regulierung abzielt, um den Binnenmarkt zu stärken, stellt Österreich sicher, dass nationale Besonderheiten berücksichtigt werden. Dieser duale Ansatz kann zu praktischen Hürden führen, etwa wenn ein in der EU zugelassenes KI-System in Österreich zusätzliche Anforderungen erfüllen muss.

Thema	EU-Ebene	Österreichische Ebene
Zulassung von KI-Systemen	Verordnung (EU) 2017/745	Medizinproduktegesetz (MPG)
Datenschutz	DSGVO	Gesundheitstelematikgesetz (GTelG 2012)
Ethische Leitlinien	Ethik-Leitlinien für vertrauenswürdige KI	Gesundheitsqualitätsgesetz (GQG)
Haftung	AI Act	Ärztegesetz (ÄrzteG)

TABELLE 3: VERGLEICH UNTERSCHIEDLICHER REGULATORISCHER RAHMENBEDINGUNGEN

4 Diskussion

Die Entwicklung von Künstlicher Intelligenz (KI) im medizinischen Bereich hat das Potenzial, die Art und Weise, wie medizinische Diagnosen gestellt und Behandlungen durchgeführt werden, grundlegend zu verändern. Daher ist es für Mediziner*Innen essenziell, sich mit KI-Anwendungen auseinanderzusetzen. Dies betrifft sowohl den praktischen Umgang mit KI-Systemen in der klinischen Praxis als auch die aktive Teilnahme am Entwicklungsprozess solcher Technologien. Die Fähigkeit, KI-Modelle kritisch zu hinterfragen, ihre Ergebnisse zu interpretieren und sie im Arbeitsalltag sinnvoll einzusetzen, wird zunehmend eine Kernkompetenz für ÄrztInnen.

Ein zentraler Aspekt in diesem Kontext ist die Wahl der geeigneten Entwicklungsstrategie für KI-Systeme. Der Erfolg solcher Systeme hängt maßgeblich davon ab, wie gut sie auf die spezifischen Anforderungen der medizinischen Praxis abgestimmt sind. Hierbei kommen zwei grundlegende Ansätze zum Einsatz: der Top-to-Bottom-Ansatz, bei dem Expertenwissen in den Entwicklungsprozess einfließt, und der Bottom-to-Top-Ansatz, der datengetriebene Methoden nutzt. Beide Ansätze bieten unterschiedliche Potenziale und Herausforderungen für Mediziner*Innen.

Mediziner*Innen sind im Entwicklungsprozess von KI-Systemen unverzichtbar, da sie das notwendige Domänenwissen einbringen, um die Modelle praxisrelevant zu gestalten. Der Top-to-Bottom-Ansatz erfordert insbesondere ihre Expertise, um bestehende medizinische Klassifikationen und Ontologien sinnvoll einzubinden. Gleichzeitig können ÄrztInnen beim Bottom-to-Top-Ansatz durch die Bereitstellung und Annotation von Daten dazu beitragen, die Qualität der Modelle zu verbessern und Verzerrungen in den Daten zu minimieren.

Der richtige Entwicklungsansatz hat auch direkte Auswirkungen auf die spätere Nutzung von KI-Systemen in der Praxis. Ein Top-to-Bottom-Ansatz bietet zunächst hohe Verlässlichkeit, könnte jedoch zu starr sein, um neue Entwicklungen schnell zu integrieren. Ein datengetriebener Ansatz hingegen bietet mehr Flexibilität, birgt jedoch Risiken hinsichtlich der Nachvollziehbarkeit der Ergebnisse. Die Kombination beider Ansätze – der sogenannte Hybridansatz – könnte eine zukunftsweisende Lösung sein, um die Vorteile beider Welten zu vereinen.

4.1 Warum ist es für Mediziner*Innen relevant, sich mit KI-Anwendungen auseinanderzusetzen?

Mediziner*Innen spielen eine zentrale Rolle im Gesundheitswesen und sind sowohl in der klinischen Praxis als auch in der Forschung mit neuen Technologien konfrontiert. KI-Systeme können nicht nur die Diagnosegenauigkeit erhöhen, sondern auch administrative Aufgaben automatisieren. Das Verständnis für die Funktionsweise solcher Systeme ermöglicht es ÄrztInnen, deren Stärken gezielt einzusetzen und mögliche Schwächen kritisch zu hinterfragen. Die aktive Auseinandersetzung mit KI kann zudem dazu beitragen, die Kommunikation zwischen medizinischem Fachpersonal und Patient*innen zu verbessern, indem technische Sachverhalte verständlich vermittelt werden.

4.2 Wie können Mediziner*Innen im Prozess der KI-Entwicklung teilnehmen?

Eine aktive Beteiligung von Mediziner*Innen im Entwicklungsprozess von KI-Systemen ist essenziell, um sicherzustellen, dass diese Systeme den praktischen Anforderungen im klinischen Alltag gerecht werden. Dies kann durch die Bereitstellung und Annotation medizinischer Daten, durch die Formulierung klinischer Fragestellungen und durch die Validierung der Modelle geschehen. Mediziner*Innen bringen das notwendige Domänenwissen ein, um KI-Algorithmen auf reale Szenarien anzupassen und die Qualität der Ergebnisse zu gewährleisten. Dies gilt insbesondere für den Top-to-Bottom-Ansatz, bei dem Expertenwissen eine entscheidende Rolle spielt.

4.3 Welche Auswirkungen hat/wird es auf die Arbeit von Mediziner*Innen in der Praxis und Theorie haben?

KI-Systeme können die Arbeit von Mediziner*Innen in vielfacher Hinsicht beeinflussen. In der Praxis ermöglichen sie eine effizientere Organisation von Arbeitsabläufen, eine schnellere Diagnostik sowie eine personalisierte Patient*innenversorgung. In der Theorie eröffnen KI-Technologien neue Forschungsfelder, etwa in der Datenanalyse, der Bildverarbeitung oder der Entwicklung von Entscheidungshilfesystemen. Gleichzeitig verändert sich das Berufsbild der ÄrztInnen: Neben medizinischem Wissen wird zunehmend auch technisches Verständnis gefordert, um die Interaktionen mit KI-Systemen kritisch zu begleiten.

4.4 Vergleich der Strategien: Top-to-Bottom und Bottom-to-Top

4.4.1 Top-to-Bottom-Ansatz

Der Top-to-Bottom-Ansatz zeichnet sich dadurch aus, dass die Anforderungen und Strukturen von KI-Modellen von vornherein durch Expertenwissen definiert werden. Dabei wird das Modell auf Grundlage bestehender medizinischer Klassifikationen, Ontologien und Terminologien entwickelt.

Vorteile:

- **Hohe Präzision:** Durch den Einsatz von etablierten medizinischen Standards wie ICD oder SNOMED-CT wird eine hohe Genauigkeit bei der Erkennung und Verarbeitung von medizinischen Begriffen erreicht.
- **Schneller Einstieg:** Da bereits strukturierte Daten und Kategorien vorhanden sind, kann die Entwicklung von KI-Modellen effizienter erfolgen.
- **Verlässlichkeit:** Expertenwissen reduziert die Wahrscheinlichkeit von Fehlinterpretationen durch das KI-Modell.

Nachteile:

- **Hoher manueller Aufwand:** Die Erstellung und Pflege von medizinischen Klassifikationen erfordern erhebliche Ressourcen.
- **Begrenzte Flexibilität:** Der Ansatz kann Schwierigkeiten haben, neue medizinische Entwicklungen oder seltene Krankheitsbilder schnell zu integrieren.
- **Bias durch bestehende Standards:** Vorurteile und Einschränkungen in bestehenden medizinischen Systemen können sich auf die Leistung des Modells auswirken.

4.4.2 Bottom-to-Top-Ansatz

Der Bottom-to-Top-Ansatz basiert auf einer datengetriebenen Herangehensweise. Hierbei werden große Mengen unstrukturierter Daten genutzt, um das Modell durch maschinelles Lernen zu trainieren.

Vorteile:

- **Höhere Anpassungsfähigkeit:** Das Modell kann schnell neue Muster und Trends erkennen.
- **Reduzierter manueller Aufwand:** Automatisierte Prozesse minimieren die Notwendigkeit einer manuellen Annotation.
- **Dynamische Skalierbarkeit:** Der Ansatz kann große Datenmengen effizient verarbeiten und sich an Veränderungen im medizinischen Bereich anpassen.

Nachteile:

- **Qualitätsrisiken:** Ohne ausreichende Kontrolle können fehlerhafte oder unvollständige Daten die Modellleistung negativ beeinflussen.
- **Black-Box-Problem:** Die Entscheidungen des Modells sind oft schwer nachvollziehbar, was in medizinischen Anwendungen problematisch ist.
- **Erhöhtes Risiko von Bias:** Verzerrungen in den Trainingsdaten können zu unerwünschten Ergebnissen führen.

4.5 Hybridansatz: Top-to-Bottom-to-Top

Ein möglicher Hybridansatz, der die Vorteile beider Strategien kombiniert, könnte sich als praktikabler erweisen. Dieser Ansatz würde zunächst Expertenwissen nutzen, um eine Grundstruktur zu schaffen (Top-to-Bottom), anschließend jedoch datengetriebene Verfahren einsetzen, um diese Struktur dynamisch zu erweitern und zu verbessern (Bottom-to-Top).

Potenziale:

- **Verbesserte Genauigkeit:** Die Kombination von strukturierten Daten und maschinellem Lernen kann die Modellleistung erheblich verbessern.
- **Flexibilität:** Der Ansatz ermöglicht es, sowohl bestehende medizinische Standards zu nutzen als auch auf neue Entwicklungen zu reagieren.
- **Erhöhte Akzeptanz:** Die Nachvollziehbarkeit der Entscheidungen könnte durch den Einsatz von Expertenwissen in der Initialphase verbessert werden.

Risiken:

- **Komplexität:** Die Implementierung eines solchen Hybridansatzes erfordert umfangreiche Ressourcen und interdisziplinäre Zusammenarbeit.
- **Datenschutz:** Die Nutzung großer Datenmengen könnte Datenschutzrisiken erhöhen.
- **Potenzielle Konflikte:** Unterschiede zwischen den strukturierten Vorgaben und den datenbasierten Erkenntnissen können zu Inkonsistenzen führen.

4.6 Sicherheits- und Qualitätsaspekte

Bei der Wahl der Entwicklungsstrategie spielen **Datensicherheit** und **Qualität der Metadaten** eine zentrale Rolle.

Datensicherheit

- **Schutz sensibler Daten:** Medizinische Daten sind besonders sensibel und erfordern daher umfassende Schutzmaßnahmen. Dies gilt insbesondere bei datengetriebenen Ansätzen.

- **Regulatorische Anforderungen:** In Europa gelten strenge Datenschutzvorschriften wie die DSGVO, die sicherstellen sollen, dass personenbezogene Daten geschützt werden.
- **Anonymisierung:** Eine effektive Anonymisierung der Daten ist essenziell, um den Datenschutz zu gewährleisten.

Qualität der Metadaten

- **Relevanz hochwertiger Annotationen:** Die Qualität der Annotationen hat einen direkten Einfluss auf die Leistungsfähigkeit des Modells.
- **Standardisierung:** Der Einsatz etablierter Standards wie ICD oder SNOMED-CT verbessert die Interoperabilität und Vergleichbarkeit der Daten.
- **Kontinuierliche Aktualisierung:** Metadaten müssen regelmäßig überprüft und aktualisiert werden, um die Relevanz und Genauigkeit der KI-Modelle sicherzustellen.

4.7 Ausblick

Zukünftige Ansätze könnten verstärkt auf **adaptive Modelle** setzen, die sich kontinuierlich an neue Daten anpassen. Ein weiteres vielversprechendes Konzept ist der Einsatz von **Explainable AI (XAI)**, die es ermöglicht, die Entscheidungen von KI-Modellen nachvollziehbar zu machen. Dies könnte die Akzeptanz und das Vertrauen in KI-Systeme im medizinischen Bereich weiter erhöhen.

Darüber hinaus könnte der Einsatz von **föderiertem Lernen** dazu beitragen, Datenschutzprobleme zu minimieren. Bei diesem Ansatz verbleiben die Daten vor Ort, und nur die Modelle werden zentralisiert trainiert. Dies reduziert das Risiko von Datenschutzverletzungen erheblich.

5 Zusammenfassung

Die vorliegende Arbeit hat sich intensiv mit der Bedeutung von Klassifikationen, Terminologien und Ontologien für die Entwicklung von KI-Systemen im medizinischen Bereich auseinandergesetzt. Diese grundlegenden Strukturen bilden das Fundament für die Effizienz und Genauigkeit von KI-Anwendungen, die in der Medizin zunehmend eine zentrale Rolle einnehmen. Vor dem Hintergrund der stetig wachsenden Menge medizinischer Daten und der steigenden Anforderungen an personalisierte Diagnosen und Therapien bietet der Einsatz von KI-Systemen ein enormes Potenzial. Doch die Implementierung solcher Systeme bringt auch Herausforderungen mit sich — ethische, regulatorische und technologische Aspekte müssen dabei berücksichtigt werden. Diese Schlussfolgerung greift zentrale Themen der Arbeit auf und bietet weiterführende Überlegungen zu den spezifischen Herausforderungen und Möglichkeiten von KI in der Medizin.

5.1 Medizinischer Fokus: Klinische Anwendungen und Auswirkungen auf die Patient*innenversorgung

5.1.1 Diagnostik und Frühwarnsysteme

Eines der bedeutendsten Einsatzgebiete von KI in der Medizin ist die Diagnostik. KI-basierte Systeme können große Mengen an medizinischen Bilddaten analysieren und Muster erkennen, die für das menschliche Auge kaum wahrnehmbar sind. In der Radiologie, Pathologie und Dermatologie kommen solche Systeme bereits zum Einsatz, um Krankheiten wie Krebs frühzeitig zu erkennen und diagnostische Fehler zu reduzieren. Dabei liegt der Vorteil insbesondere in der Geschwindigkeit und Genauigkeit der Analyse, was die Behandlungsqualität erheblich verbessern kann.

Ein weiteres Einsatzgebiet ist die Entwicklung von Frühwarnsystemen. KI-Systeme können auf Basis von Patient*innendaten Risiken frühzeitig erkennen und vor potenziell kritischen Entwicklungen warnen, beispielsweise bei der Vorhersage von Herzinfarkten oder der Verschlechterung chronischer Krankheiten. Diese Systeme entlasten Ärzt*innen und Pflegekräfte, indem sie Routineaufgaben automatisieren und den Fokus auf komplexere medizinische Entscheidungen lenken.

Allerdings birgt der Einsatz von KI in der Diagnostik auch Risiken. Eine der größten Herausforderungen ist die sogenannte Black-Box-Problematik: Viele KI-Modelle liefern zwar präzise Ergebnisse, doch ihre Entscheidungswege bleiben für Ärzt*innen oft intransparent. Dies

kann das Vertrauen in die Technologie beeinträchtigen und ethische Fragen zur Verantwortung aufwerfen.

5.1.2 Personalisierte Medizin und KI-gestützte Therapie

Ein weiterer zentraler Aspekt des Einsatzes von KI in der Medizin ist die personalisierte Medizin. KI-Systeme können große Mengen an Patient*innendaten analysieren und daraus individualisierte Therapieempfehlungen ableiten. Besonders in der Onkologie zeigt sich das Potenzial dieser Anwendungen: Durch die Analyse genetischer Daten lassen sich spezifische Tumoreigenschaften identifizieren, die bei der Auswahl der optimalen Behandlungsmethode helfen.

Auch in der Pharmakogenetik spielt KI eine wichtige Rolle. Hierbei werden genetische Informationen genutzt, um vorherzusagen, wie eine/ein Patient*in auf bestimmte Medikamente reagieren wird. Dies ermöglicht eine individuell angepasste Medikation und reduziert das Risiko unerwünschter Nebenwirkungen.

Die Herausforderung besteht jedoch darin, dass personalisierte Therapien auf qualitativ hochwertigen Daten basieren müssen. Die Qualität der zugrunde liegenden Daten entscheidet über die Genauigkeit der Vorhersagen. Fehlerhafte oder unvollständige Daten können zu falschen Therapieempfehlungen führen, was die Patient*innensicherheit gefährden kann.

5.1.3 Administrative Prozesse und Effizienzsteigerung

Neben der direkten Patient*innenversorgung können KI-Systeme auch administrative Prozesse im Gesundheitswesen optimieren. Sie unterstützen beispielsweise bei der Terminplanung, der Patient*innenaufnahme und der Ressourcentverteilung in Krankenhäusern. Dies entlastet das medizinische Personal von Routineaufgaben und schafft mehr Freiraum für die eigentliche Patient*innenversorgung.

Ein Beispiel für die Effizienzsteigerung durch KI ist die Vermeidung von Medikationsfehlern. KI-Systeme können Dosierungsfehler erkennen und vermeiden, indem sie Patient*innendaten mit den verordneten Medikamenten abgleichen und potenzielle Risiken anzeigen.

Allerdings besteht auch hier die Gefahr der Überautomatisierung. Eine zu starke Abhängigkeit von KI-Systemen könnte dazu führen, dass die zwischenmenschliche Kommunikation zwischen Ärzt*innen und Patient*innen vernachlässigt wird. Es gilt daher, ein Gleichgewicht zwischen technischer Unterstützung und menschlicher Interaktion zu finden.

5.1.4 Auswirkungen auf die Beziehung zwischen Ärzt*innen und Patient*innen

Der Einsatz von KI-Systemen in der Medizin hat das Potenzial, die Beziehung zwischen Ärzt*innen und Patient*innen grundlegend zu verändern. Während Ärzt*innen bislang die Hauptverantwortung für Diagnosen und Therapieentscheidungen tragen, könnte diese Verantwortung durch KI-Systeme zunehmend geteilt werden. Dies führt zu neuen Fragen hinsichtlich der Haftung und der Rolle der/des Ärzt*in als Entscheidungsmanager*in.

Zudem können KI-Systeme das Vertrauen der Patient*innen beeinflussen. Während einige Patient*innen die Technologie als Hilfsmittel zur Verbesserung ihrer Behandlung ansehen, könnten andere skeptisch gegenüber automatisierten Entscheidungen sein. Eine transparente Kommunikation über den Einsatz von KI und deren Grenzen ist daher essenziell, um das Vertrauen der Patient*innen zu gewinnen.

5.2 Zugang zu qualitativ hochwertigen Daten

Der Zugang zu qualitativ hochwertigen Daten ist eine Grundvoraussetzung für die Entwicklung effektiver KI-Systeme in der Medizin. Daten sind der Treibstoff, der KI-Modelle antreibt, und die Qualität dieser Daten bestimmt die Genauigkeit und Zuverlässigkeit der Ergebnisse.

Ein zentrales Problem ist der Datenschutz. Aufgrund strenger Datenschutzgesetze wie der DSGVO sind medizinische Daten oft nur eingeschränkt zugänglich. Dies stellt eine Herausforderung für die Entwicklung von KI-Systemen dar, die auf große Datenmengen angewiesen sind.

Darüber hinaus besteht die Gefahr, dass Daten aus unterschiedlichen Quellen nicht interoperabel sind. Unterschiedliche Krankenhaussysteme verwenden oft verschiedene Standards, was den Austausch von Daten erschwert. Hier sind gemeinsame Standards und Schnittstellen notwendig, um die Interoperabilität zu gewährleisten.

Ein weiteres Problem ist der sogenannte Bias in den Datensätzen. Wenn bestimmte Patient*innengruppen in den Daten unterrepräsentiert sind, kann dies zu unausgewogenen und potenziell diskriminierenden Ergebnissen führen. Daher ist es wichtig, dass die verwendeten Datensätze repräsentativ sind und eine breite Bevölkerungsgruppe abdecken.

Strategien zur Sicherstellung der Datenqualität umfassen die Erstellung umfangreicher, kuratierter Datensätze sowie die Zusammenarbeit zwischen Krankenhäusern, Forschungseinrichtungen und der Industrie. Solche Kooperationen können dazu beitragen, qualitativ hochwertige und umfassende Datensätze zu generieren, die für die Entwicklung von KI-Systemen unerlässlich sind.

5.3 Regulatorische Anforderungen und deren Integration

Der rechtliche Rahmen für den Einsatz von KI in der Medizin stellt eine bedeutende Herausforderung dar, die jedoch für die Gewährleistung von Sicherheit und Transparenz unerlässlich ist. In der Europäischen Union bilden die „Medical Device Regulation“ (MDR) und der geplante AI Act zwei zentrale regulatorische Vorschriften, die die Entwicklung und den Einsatz von KI-Systemen im medizinischen Bereich betreffen.

Die MDR regelt die Zulassung und das Inverkehrbringen von medizinischen Geräten, zu denen auch KI-gestützte Systeme gehören. Diese Vorschrift fordert umfassende Dokumentationen, Risikobewertungen und Nachweise über die Sicherheit und Wirksamkeit der Systeme. Der AI Act, der sich derzeit in der Entwurfsphase befindet, legt darüber hinaus spezifische Anforderungen an KI-Systeme fest, insbesondere hinsichtlich Transparenz, Datenqualität und Risikomanagement.

Eine zentrale Herausforderung bei der Einhaltung dieser Vorschriften ist der hohe Aufwand an Dokumentation und Compliance, der bereits in der Entwicklungsphase berücksichtigt werden muss. Der Ansatz des „Compliance by Design“, bei dem regulatorische Anforderungen von Beginn an in den Entwicklungsprozess integriert werden, kann dabei helfen, den Aufwand zu reduzieren und die Sicherheit der Systeme zu erhöhen.

Eine kontinuierliche Überwachung und Nachbesserung der Systeme ist ebenfalls erforderlich, um sicherzustellen, dass die KI-Systeme auch nach ihrer Zulassung den gesetzlichen Anforderungen entsprechen. Dies umfasst unter anderem die Überwachung von Updates, das Sammeln von Nutzungsdaten und die Anpassung der Systeme an neue regulatorische Vorgaben. Darüber hinaus spielen ethische Überlegungen eine wichtige Rolle bei der regulatorischen Bewertung von KI-Systemen. Regulierungsbehörden müssen sicherstellen, dass KI-Anwendungen nicht nur sicher und effektiv, sondern auch ethisch vertretbar sind. Dies umfasst Aspekte wie Datenschutz, Bias-Vermeidung und die Sicherstellung der menschlichen Kontrolle über kritische Entscheidungen.

Insgesamt erfordert die Integration regulatorischer Anforderungen in den Entwicklungsprozess von KI-Systemen eine enge Zusammenarbeit zwischen Entwicklern, medizinischen Fachkräften und Regulierungsbehörden. Nur so kann sichergestellt werden, dass die Systeme sowohl den rechtlichen als auch den praktischen Anforderungen gerecht werden.

5.4 Ethische Herausforderungen

Die Implementierung von KI-Systemen in der Medizin wirft grundlegende ethische Fragen auf, die eng mit dem Vertrauen in die Technologie und der Sicherheit der Patient*innen verbunden sind. Zu den zentralen ethischen Herausforderungen gehören die Patient*innenautonomie, die potenzielle Diskriminierung durch algorithmische Verzerrungen und die Transparenz der Entscheidungsprozesse.

Ein wesentliches ethisches Prinzip in der Medizin ist die Wahrung der Patient*innenautonomie. Der Einsatz von KI in Diagnostik und Therapie könnte jedoch die Entscheidungsfreiheit von Patient*innen und Ärzt*innen einschränken. Wenn KI-Systeme beispielsweise Therapieempfehlungen aussprechen, besteht die Gefahr, dass sich Ärzt*innen stärker auf die Technologie verlassen und weniger eigene Entscheidungen treffen. Dies könnte zu einem Verlust der menschlichen Kontrolle führen und die Beziehung zwischen Ärzt*innen und Patient*innen beeinträchtigen.

Darüber hinaus besteht das Risiko von Diskriminierung durch algorithmische Verzerrungen. Wenn die Trainingsdaten von KI-Systemen nicht repräsentativ sind, können bestimmte Patient*innengruppen benachteiligt werden. Ein Beispiel dafür ist die Unterrepräsentation von Frauen oder ethnischen Minderheiten in medizinischen Datensätzen, was zu ungenauen Diagnosen oder Therapieempfehlungen für diese Gruppen führen kann. Es ist daher essenziell, dass die Datensätze divers und inklusiv sind, um solche Verzerrungen zu vermeiden.

Ein weiteres ethisches Problem ist die Transparenz der Entscheidungsprozesse von KI-Systemen. Die sogenannte Black-Box-Problematik beschreibt die Schwierigkeit, die Entscheidungswege von KI-Systemen nachzuvollziehen. Dies kann das Vertrauen in die Technologie beeinträchtigen, insbesondere wenn es um lebenswichtige Entscheidungen geht. Es bedarf daher technischer und rechtlicher Maßnahmen, um sicherzustellen, dass die Entscheidungsprozesse von KI-Systemen transparent und nachvollziehbar sind.

Zusammenfassend ist es wichtig, ethische Leitlinien und Standards zu entwickeln, die den Einsatz von KI in der Medizin regeln. Dies umfasst die Gewährleistung der Patient*innenautonomie, die Vermeidung von Diskriminierung und die Sicherstellung der Transparenz von Entscheidungsprozessen. Nur durch die Berücksichtigung dieser ethischen Aspekte kann sichergestellt werden, dass KI-Systeme sicher und verantwortungsvoll eingesetzt werden.

5.5 Zukunftsaussichten

Die Zukunft des Einsatzes von KI in der Medizin birgt sowohl große Chancen als auch Herausforderungen. Zu den wichtigsten zukünftigen Entwicklungen zählt die Weiterentwicklung von KI-Modellen hin zu sogenannten „Explainable AI“ (XAI)-Systemen, die eine bessere Nachvollziehbarkeit von Entscheidungen bieten sollen. Diese Systeme könnten das Vertrauen in die Technologie stärken und die Integration in den klinischen Alltag erleichtern. Darüber hinaus wird die Personalisierung von Medizin weiter an Bedeutung gewinnen. Durch die zunehmende Verfügbarkeit von genomischen und biomedizinischen Daten könnten KI-Systeme zukünftig noch präzisere und individuellere Therapieempfehlungen geben. Dies erfordert jedoch auch eine Weiterentwicklung der rechtlichen und ethischen Rahmenbedingungen, um den sicheren Einsatz dieser Systeme zu gewährleisten.

Ein weiterer wichtiger Aspekt ist die gesellschaftliche Akzeptanz von KI im Gesundheitswesen. Es wird entscheidend sein, Patient*innen und medizinisches Fachpersonal umfassend über die Möglichkeiten und Grenzen der Technologie aufzuklären. Nur so kann das Vertrauen in die Systeme gestärkt und eine breite Akzeptanz erreicht werden.

Zusammenfassend bleibt festzuhalten, dass KI-Systeme das Potenzial haben, die Medizin grundlegend zu verändern und die Patient*innenversorgung nachhaltig zu verbessern. Gleichzeitig sind jedoch umfassende Maßnahmen erforderlich, um die ethischen, regulatorischen und technischen Herausforderungen zu bewältigen. Die Weiterentwicklung und verantwortungsvolle Nutzung von KI in der Medizin wird entscheidend dazu beitragen, die Gesundheitsversorgung der Zukunft zu gestalten.

5.6 Schlusswort

Die Implementierung von KI im medizinischen Bereich stellt eine der vielversprechendsten Entwicklungen der modernen Medizin dar. Die Möglichkeit, große Datenmengen effizient zu analysieren und daraus Erkenntnisse abzuleiten, bietet enorme Chancen zur Verbesserung der Patient*innenversorgung und zur Optimierung klinischer Abläufe. Gleichzeitig dürfen jedoch die ethischen und regulatorischen Herausforderungen nicht vernachlässigt werden.

Es ist essenziell, dass der Einsatz von KI im Gesundheitswesen stets von menschlicher Expertise begleitet wird. Die menschliche Kontrolle und das Vertrauen in die Technologie sind entscheidend, um die Sicherheit der Patient*innen zu gewährleisten und die Akzeptanz neuer Systeme zu fördern.

Zukünftig wird es darauf ankommen, dass alle relevanten Akteure — Entwickler, medizinisches Personal, Patient*innen und Regulierungsbehörden — eng zusammenarbeiten, um die Potenziale der KI verantwortungsvoll zu nutzen. Nur so kann sichergestellt werden, dass KI-Systeme nicht nur innovativ, sondern auch sicher, effektiv und ethisch vertretbar eingesetzt werden.

Die vorliegende Arbeit zeigt auf, dass eine fundierte Grundlage durch Klassifikationen, Terminologien und Ontologien wesentlich ist, um den Entwicklungsprozess von KI-Systemen zu unterstützen. Die Integration von rechtlichen und ethischen Aspekten wird dabei eine entscheidende Rolle spielen, um die Patient*innenversorgung langfristig zu verbessern und die Medizin der Zukunft zu gestalten.

Quellenverzeichnis

- AAPC. 2025. *Medical Terminology* [Online]. Available: <https://www.aapc.com/resources/medical-terminology> [Accessed].
- AGHAMIRI, S. S., AMIN, R. & HELIKAR, T. 2022. Recent applications of quantitative systems pharmacology and machine learning models across diseases. *J Pharmacokinet Pharmacodyn*, 49, 19-37.
- AIWAKFORCE. 2025. *The Future of Natural Language Processing: Trends and Predictions* [Online]. Available: <https://www.aiwakforce.com/blog/the-future-of-natural-language-processing-trends-and-predictions/> [Accessed].
- ALAMGIR, A., MOUSA, O. & SHAH, Z. 2021. Artificial Intelligence in Predicting Cardiac Arrest: Scoping Review. *JMIR Med Inform*, 9, e30798.
- ALOWAIS, S. A., ALGHAMDI, S. S., ALSUHEBANY, N., ALQAHTANI, T., ALSHAYA, A. I., ALMOHAREB, S. N., ALDAIREM, A., ALRASHED, M., BIN SALEH, K., BADRELDIN, H. A., AL YAMI, M. S., AL HARBI, S. & ALBEKAIRY, A. M. 2023. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Medical Education*, 23, 689.
- AMINOLOLAMA-SHAKERI, S. & LÓPEZ, J. E. 2019. The Doctor-Patient Relationship With Artificial Intelligence. *American Journal of Roentgenology*, 212, 308-310.
- ARORA, A. 2020. Conceptualising Artificial Intelligence as a Digital Healthcare Innovation: An Introductory Review. *Medical Devices: Evidence and Research*, 13, 223-230.
- AZER, K. & BARRETT, J. S. 2022. Quantitative system pharmacology as a legitimate approach to examine extrapolation strategies used to support pediatric drug development. *CPT Pharmacometrics Syst Pharmacol*, 11, 797-804.
- BHAKAR, S., SINWAR, D., PRADHAN, N., DHAKA, V. S., CHERREZ-OJEDA, I., PARVEEN, A. & HASSAN, M. U. 2023. Computational Intelligence-Based Disease Severity Identification: A Review of Multidisciplinary Domains. *Diagnostics*, 13, 1212.
- BLASIAK, A., KHONG, J. & KEE, T. 2020. CURATE.AI: Optimizing Personalized Medicine with Artificial Intelligence. *SLAS TECHNOLOGY: Translating Life Sciences Innovation*, 25, 95-105.
- BRÄM, D. S., PARROTT, N., HUTCHINSON, L. & STEIERT, B. 2022. Introduction of an artificial neural network-based method for concentration-time predictions. *CPT Pharmacometrics Syst Pharmacol*, 11, 745-754.
- BRIN, D., SORIN, V., BARASH, Y., KONEN, E., GLICKSBERG, B. S., NADKARNI, G. N. & KLANG, E. 2024. Assessing GPT-4 multimodal performance in radiological image analysis. *European Radiology*.
- BURES, D., HOSTERS, B., REIBEL, T., JOVY-KLEIN, F., SCHRAMM, J., BRENDT-MÜLLER, J., SANDER, J. & DIEHL, A. 2023. Die transformative Wirkung von künstlicher Intelligenz im Krankenhaus. *Die Innere Medizin*, 64, 1025-1032.
- CADARETTE, S. M. & WONG, L. 2015. An Introduction to Health Care Administrative Data. *Can J Hosp Pharm*, 68, 232-7.
- CHOI, Y. S. 2017. Concepts, Characteristics, and Clinical Validation of IBM Watson for Oncology. *hmr*, 37, 49-60.
- ĆIRKOVIĆ, A. 2020. Evaluation of Four Artificial Intelligence–Assisted Self-Diagnosis Apps on Three Diagnoses: Two-Year Follow-Up Study. *J Med Internet Res*, 22, e18097.
- CROWELL, L. E., LU, A. E., LOVE, K. R., STOCKDALE, A., TIMMICK, S. M., WU, D., WANG, Y., DOHERTY, W., BONNYMAN, A., VECCHIARELLO, N., GOODWINE, C., BRADBURY, L., BRADY, J. R., CLARK, J. J., COLANT, N. A., CVETKOVIC, A., DALVIE, N. C., LIU, D., LIU, Y., MASCARENHAS, C. A., MATTHEWS, C. B., MOZDZIERZ, N. J., SHAH, K. A., WU, S.-L., HANCOCK, W. S., BRAATZ, R. D., CRAMER, S. M. & LOVE, J. C. 2018. On-demand manufacturing of clinical-quality biopharmaceuticals. *Nature Biotechnology*, 36, 988-995.
- CUOCOLO, R., CARUSO, M., PERILLO, T., UGGA, L. & PETRETTA, M. 2020. Machine Learning in oncology: A clinical appraisal. *Cancer Letters*, 481, 55-62.
- DAHM, M. H., BERGMOSER, S. & YOGENDIRAN, T. 2020. Nutzerakzeptanz und Potenziale von KI im Gesundheitswesen. In: BUCHKREMER, R., HEUPEL, T. & KOCH, O. (eds.) *Künstliche*

- Intelligenz in Wirtschaft & Gesellschaft: Auswirkungen, Herausforderungen & Handlungsempfehlungen.* Wiesbaden: Springer Fachmedien Wiesbaden.
- DARCY, A., BEAUDETTE, A., CHIAUZZI, E., DANIELS, J., GOODWIN, K., MARIANO, T. Y., WICKS, P. & ROBINSON, A. 2023. Anatomy of a Woebot® (WB001): agent guided CBT for women with postpartum depression. *Expert Review of Medical Devices*, 20, 1035-1049.
- DARWICH, A. S., POLASEK, T. M., ARONSON, J. K., OGUNGBENRO, K., WRIGHT, D. F. B., ACHOUR, B., RENY, J. L., DAALI, Y., EIERMANN, B., COOK, J., LESKO, L., MCLACHLAN, A. J. & ROSTAMI-HODJEGAN, A. 2021. Model-Informed Precision Dosing: Background, Requirements, Validation, Implementation, and Forward Trajectory of Individualizing Drug Therapy. *Annu Rev Pharmacol Toxicol*, 61, 225-245.
- DEEPMIND. 2020. *DeepMind Health: Using AI to Detect Eye Diseases* [Online]. Available: <https://deepmind.com/health> [Accessed].
- DEHDAB, R., BRENDLIN, A., WERNER, S., ALMANSOUR, H., GASSENMAIER, S., BRENDDEL, J. M., NIKOLAOU, K. & AFAT, S. 2024. Evaluating ChatGPT-4V in chest CT diagnostics: a critical image interpretation assessment. *Japanese Journal of Radiology*, 42, 1168-1177.
- DHIMAN, G., JUNEJA, S., VIRIYASITAVAT, W., MOHAFAEZ, H., HADIZADEH, M., ISLAM, M. A., EL BAYOUMY, I. & GULATI, K. 2022. A Novel Machine-Learning-Based Hybrid CNN Model for Tumor Identification in Medical Image Processing. *Sustainability*, 14, 1447.
- DICUONZO, G., DONOFRIO, F., FUSCO, A. & SHINI, M. 2023. Healthcare system: Moving forward with artificial intelligence. *Technovation*, 120, 102510.
- DURDEN, E., PIRNER, M. C., RAPOPORT, S. J., WILLIAMS, A., ROBINSON, A. & FORMAN-HOFFMAN, V. L. 2023. Changes in stress, burnout, and resilience associated with an 8-week intervention with relational agent “Woebot”. *Internet Interventions*, 33, 100637.
- ESTEVA, A., KUPREL, B., NOVOA, R. A., KO, J., SWETTER, S. M., BLAU, H. M. & THRUN, S. 2017. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542, 115-118.
- EUROPÄISCHE, K. 2017. Verordnung (EU) 2017/745 über Medizinprodukte.
- EUROPÄISCHE, K. 2019. Ethik-Leitlinien für vertrauenswürdige KI.
- EUROPÄISCHE, K. 2021. Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union.
- FEDERAL INSTITUTE FOR, D. & MEDICAL, D. 2025. *History and Outlook of OPS* [Online]. Available: https://www.bfarm.de/EN/Code-systems/Classifications/OPS-ICHI/OPS/History-and-outlook/_node.html [Accessed].
- GIL-RIOS, M. A., CRUZ-ACEVES, I., HERNANDEZ-AGUIRRE, A., MOYA-ALBOR, E., BRIEVA, J., HERNANDEZ-GONZALEZ, M. A. & SOLORIO-MEZA, S. E. 2024. High-Dimensional Feature Selection for Automatic Classification of Coronary Stenosis Using an Evolutionary Algorithm. *Diagnostics (Basel)*, 14.
- GOERKE, M., SKRZIPALE, S., SONG, H. & ZACHERL, C. 2024. Chancen und Herausforderungen von KI-basierter Ressourcenplanung und -optimierung im Krankenhaus. In: HENKE, V., HÜLSKEN, G., SCHNEIDER, H. & VARGHESE, J. (eds.) *Health Data Management: Schlüsselfaktor für erfolgreiche Krankenhäuser*. Wiesbaden: Springer Fachmedien Wiesbaden.
- HARITHA, C. K., SHAIKH, A., GADIA, V. & MULLICK, A. 2024. Enhancing Patient Care With AI Chatbots And Virtual Assistants. *International Journal of Pharmaceutical Sciences*, 2, 10-13.
- HAVERKAMP, W., STRODTHOFF, N., TENNENBAUM, J. & ISRAEL, C. 2023. Großer Hype um ChatGPT in der Medizin. *Herzschrittmachertherapie + Elektrophysiologie*, 34, 240-245.
- HOFFMAN, V., FLOM, M., MARIANO, T. Y., CHIAUZZI, E., WILLIAMS, A., KIRVIN-QUAMME, A., PAJARITO, S., DURDEN, E. & PERSKI, O. 2023. User Engagement Clusters of an 8-Week Digital Mental Health Intervention Guided by a Relational Agent (Woebot): Exploratory Study. *J Med Internet Res*, 25, e47198.
- HOLFORD, N. H. G. & SHEINER, L. B. 1982. Kinetics of pharmacologic response. *Pharmacology & Therapeutics*, 16, 143-166.
- HORIUCHI, D., TATEKAWA, H., OURA, T., OUE, S., WALSTON, S. L., TAKITA, H., MATSUSHITA, S., MITSUYAMA, Y., SHIMONO, T., MIKI, Y. & UEDA, D. 2024a. Comparing the Diagnostic

- Performance of GPT-4-based ChatGPT, GPT-4V-based ChatGPT, and Radiologists in Challenging Neuroradiology Cases. *Clinical Neuroradiology*, 34, 779-787.
- HORIUCHI, D., TATEKAWA, H., SHIMONO, T., WALSTON, S. L., TAKITA, H., MATSUSHITA, S., OURA, T., MITSUYAMA, Y., MIKI, Y. & UEDA, D. 2024b. Accuracy of ChatGPT generated diagnosis from patient's medical history and imaging findings in neuroradiology cases. *Neuroradiology*, 66, 73-79.
- HOSSAIN, E., RANA, R., HIGGINS, N., SOAR, J., BARUA, P. D., PISANI, A. R. & TURNER, K. 2023. Natural Language Processing in Electronic Health Records in Relation to Healthcare Decision-making: A Systematic Review.
- HU, X., RAN, A. R., NGUYEN, T. X., SZETO, S., YAM, J. C., CHAN, C. K. M. & CHEUNG, C. Y. 2023. What can GPT-4 do for Diagnosing Rare Eye Diseases? A Pilot Study. *Ophthalmology and Therapy*, 12, 3395-3402.
- HUGHES, J. H. & KEIZER, R. J. 2021. A hybrid machine learning/pharmacokinetic approach outperforms maximum a posteriori Bayesian estimation by selectively flattening model priors. *CPT Pharmacometrics Syst Pharmacol*, 10, 1150-1160.
- HUMBSCH, P., HORN, E., BOHM, K. & GINTROWICZ, R. 2024. ChatGPT im Einsatz für „technology-enhanced learning“ in Anästhesiologie und Notfallmedizin und potenzielle klinische Anwendung von KI-Sprachmodellen. *Die Anaesthesiologie*, 73, 324-335.
- HUPPERTZ, M. S., SIEPMANN, R., TOPP, D., NIKOUBASHMAN, O., YÜKSEL, C., KUHL, C. K., TRUHN, D. & NEBELUNG, S. 2024. Revolution or risk?—Assessing the potential and challenges of GPT-4V in radiologic image interpretation. *European Radiology*.
- HUSER, V. Evaluation of LOINC as a reference terminology for clinical document types: A case report of an outpatient EHR. *Academia*.
- HUSER, V., TAFT, L. M. & CIMINO, J. J. Suitability of LOINC Document Ontology as a reference terminology for clinical document types: A case report of a research-oriented EHR. *Academia*.
- HUTCHINS, J. 2014. Hutchins 2014 - Machine Translation Archive.
- IMAI, S., TAKEKUMA, Y., MIYAI, T. & SUGAWARA, M. 2020. A New Algorithm Optimized for Initial Dose Settings of Vancomycin Using Machine Learning. *Biological and Pharmaceutical Bulletin*, 43, 188-193.
- INTERNATIONAL, S. 2025a. *SNOMED in Action: Australian Medicines Terminology* [Online]. Available: <https://www.snomed.org/snomed-in-action/australian-medicines-terminology> [Accessed].
- INTERNATIONAL, S. 2025b. *SNOMED in Action: B2i Healthcare* [Online]. Available: <https://www.snomed.org/snomed-in-action/b2i-healthcare> [Accessed].
- INTERNATIONAL, S. 2025c. *SNOMED in Action: St John, New Zealand* [Online]. Available: <https://www.snomed.org/snomed-in-action/st-john%2C-new-zealand> [Accessed].
- ISBARY, G., ZIMMER, E., DETTMAR, K. & ZAIAC, M. 2024. KI zur Verbesserung des Informationsflusses zwischen verschiedenen Akteuren der Gesundheitsbranche und Pharmaindustrie. *Forum*, 39, 309-311.
- IVANOVIĆ, M. & BUDIMAC, Z. 2014. An overview of ontologies and data resources in medical domains. *Expert Systems with Applications*, 41, 5158-5166.
- JELLIFFE, R. W., SCHUMITZKY, A., BAYARD, D., MILMAN, M., VAN GUILDER, M., WANG, X., JIANG, F., BARBAUT, X. & MAIRE, P. 1998. Model-Based, Goal-Oriented, Individualised Drug Therapy. *Clinical Pharmacokinetics*, 34, 57-77.
- JETTÉ, N., QUAN, H., HEMMELGARN, B., DROSLER, S., MAASS, C., MOSKAL, L., PAOIN, W., SUNDARARAJAN, V., GAO, S., JAKOB, R., ÜSTÜN, B., GHALI, W. A. & INVESTIGATORS, F. T. I. 2010. The Development, Evolution, and Modifications of ICD-10: Challenges to the International Comparability of Morbidity Data. *Medical Care*, 48, 1105-1110.
- JIE, Z., ZHIYING, Z. & LI, L. 2021. A meta-analysis of Watson for Oncology in clinical application. *Scientific Reports*, 11, 5792.
- KOROT, E., PONTIKOS, N., LIU, X., WAGNER, S. K., FAES, L., HUEMER, J., BALASKAS, K., DENNISTON, A. K., KHAWAJA, A. & KEANE, P. A. 2021. Predicting sex from retinal fundus photographs using automated deep learning. *Sci Rep*, 11, 10286.
- KRENZER, A., MAKOWSKI, K., HEKALO, A., FITTING, D., TROYA, J., ZOLLER, W. G., HANN, A. & PUPPE, F. 2022. Fast machine learning annotation in the medical domain: a semi-

- automated video annotation tool for gastroenterologists. *BioMedical Engineering OnLine*, 21, 33.
- LEE, S., SONG, M., HAN, J., LEE, D. & KIM, B. H. 2022. Application of Machine Learning Classification to Improve the Performance of Vancomycin Therapeutic Drug Monitoring. *Pharmaceutics*, 14.
- LEIBLE, S., GÜCÜK, G.-L., SIMIC, D., VON BRACKEL-SCHMIDT, C. & LEWANDOWSKI, T. 2024. Zwischen Forschung und Praxis: Fähigkeiten und Limitationen generativer KI sowie ihre wachsende Bedeutung in der Zukunft. *HMD Praxis der Wirtschaftsinformatik*, 61, 344-370.
- LESLIE, D. 2019. Understanding Artificial Intelligence Ethics and Safety: A Guide for the Responsible Design and Implementation of AI Systems in the Public Sector.
- LI, Y., ESTEVA, A., KUPREL, B., NOVOA, R., KO, J. & THRUN, S. 2016. Skin Cancer Detection and Tracking using Data Synthesis and Deep Learning.
- LITJENS, G., KOOI, T., BEJNORDI, B. E., SETIO, A. A. A., CIOMPI, F., GHAFORIAN, M., VAN DER LAAK, J. A. W. M., VAN GINNEKEN, B. & SÁNCHEZ, C. I. 2017. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60-88.
- LOINC. 2025a. *LOINC (Logical Observation Identifiers Names and Codes)* [Online]. Available: <https://loinc.org/> [Accessed].
- LOINC. 2025b. *LOINC and Other Standards - FAQ* [Online]. Available: <https://loinc.org/kb/faq/loinc-and-other-standards/> [Accessed].
- MAAD, M. M., ABDEL-HAMEED, A.-M., MOSTAFA, A., EL-SAYED, M. E.-K., ABDELHAMEED, I., ABDELAZIZ, A. A. & MARWA, M. E. 2023. From Pixels to Diagnoses: Deep Learning's Impact on Medical Image Processing-A Survey. *Wasit Journal of Computer and Mathematics Science*, 2, 9-15.
- MALIK, I., IQBAL, A., GU, Y. H. & AL-ANTARI, M. A. 2024. Deep Learning for Alzheimer's Disease Prediction: A Comprehensive Review. *Diagnostics*, 14, 1281.
- MANOHARAN, G., PANT, L. M., TEWARI, R., VICHARE, S., SARKAR, P. & DHANRAJ, J. A. AI-Powered Chatbots for Mental Health Support. 2024 7th International Conference on Contemporary Computing and Informatics (IC3I), 18-20 Sept. 2024 2024. 436-440.
- MARQUET, P., CROS, F., MICALLEF, L., JACQZ-AIGRAIN, E., WOILLARD, J. B., MONCHAUD, C., SAINT-MARCOUX, F. & DEBORD, J. 2021. Tacrolimus Bayesian Dose Adjustment in Pediatric Renal Transplant Recipients. *Ther Drug Monit*, 43, 472-480.
- MIJWIL, M. M., FAIEQ, A. K. & AL-MISTAREHI, A.-H. 2022. The Significance of Digitalisation and Artificial Intelligence in The Healthcare Sector: A Review. *Asian Journal of Pharmacy, Nursing and Medical Sciences*.
- MIOTTO, R., WANG, F., WANG, S., JIANG, X. & DUDLEY, J. T. 2017. Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, 19, 1236-1246.
- MOSLMANI, A. & THIELSCHER, C. 2024. Large Language Models zur Auswertung von Patientendaten in der Onkologie. *InFo Hämatologie + Onkologie*, 27, 16-18.
- MOULD, D. R. & UPTON, R. N. 2012. Basic concepts in population modeling, simulation, and model-based drug development. *CPT Pharmacometrics Syst Pharmacol*, 1, e6.
- MUKHOPADHYAY, A., SUMNER, J., LING, L. H., QUEK, R. H. C., TAN, A. T. H., TENG, G. G., SEETHARAMAN, S. K., GOLLAMUDI, S. P. K., HO, D. & MOTANI, M. 2022. Personalised Dosing Using the CURATE.AI Algorithm: Protocol for a Feasibility Study in Patients with Hypertension and Type II Diabetes Mellitus. *International Journal of Environmental Research and Public Health*, 19, 8979.
- NATIONAL LIBRARY OF, M. 2009a. *Metathesaurus*, Bethesda (MD), National Library of Medicine (US).
- NATIONAL LIBRARY OF, M. 2009b. Semantic Network. *In: UPDATED, A. (ed.)*. Bethesda (MD): National Library of Medicine (US).
- NATIONAL LIBRARY OF, M. 2009c. SPECIALIST Lexicon and Lexical Tools. *In: UPDATED, A. (ed.)*. Bethesda (MD): National Library of Medicine (US).
- NATIONAL LIBRARY OF, M. 2024. *MeSH Highlights 2024* [Online]. Available: <https://www.nlm.nih.gov/oet/ed/mesh/2024/mesh-highlights.html> [Accessed].
- NATIONAL LIBRARY OF, M. 2025a. *Introduction to MeSH* [Online]. Available: <https://www.nlm.nih.gov/mesh/introduction.html> [Accessed].

- NATIONAL LIBRARY OF, M. 2025b. *Medical Subject Headings (MeSH)* [Online]. Available: <https://www.nlm.nih.gov/mesh/meshhome.html> [Accessed].
- NATIONAL LIBRARY OF, M. 2025c. *RxNorm - National Library of Medicine (NLM)* [Online]. Available: <https://www.nlm.nih.gov/research/umls/rxnorm/index.html> [Accessed].
- NATIONAL LIBRARY OF, M. 2025d. *UMLS Metathesaurus Release Statistics* [Online]. Available: https://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/release/statistics.html [Accessed].
- NATIONAL LIBRARY OF, M. 2025e. *UMLS Source Release Documentation* [Online]. Available: <https://www.nlm.nih.gov/research/umls/sourcereleasedocs/> [Accessed].
- NEUMUTH, T. 2020. Künstliche Intelligenz – Anwendungsbereiche in der Onkologie. *Forum*, 35, 104-108.
- NUANCE, C. 2022. *Dragon Medical One: AI-Powered Clinical Documentation* [Online]. Available: <https://www.nuance.com> [Accessed].
- OPENSTAX. 2025. *Anatomical Terminology* [Online]. Available: <https://openstax.org/books/anatomy-and-physiology/pages/1-6-anatomical-terminology> [Accessed].
- OTERO VARELA, L., DOKTORCHIK, C., WIEBE, N., QUAN, H. & EASTWOOD, C. 2021. Exploring the differences in ICD and hospital morbidity data collection features across countries: an international survey. *BMC Health Services Research*, 21, 308.
- PROCHASKA, J. J., VOGEL, E. A., CHIENG, A., KENDRA, M., BAIOCCHI, M., PAJARITO, S. & ROBINSON, A. 2021. A Therapeutic Relational Agent for Reducing Problematic Substance Use (Woebot): Development and Usability Study. *J Med Internet Res*, 23, e24850.
- PUSTEJOVSKY, J. & STUBBS, A. 2012. *Natural Language Annotation for Machine Learning: A Guide to Corpus-Building for Applications*, O'Reilly Media.
- RAJPURKAR, P., IRVIN, J., ZHU, K., YANG, B., MEHTA, H., DUAN, T., DING, D., BAGUL, A., LANGLOTZ, C., SHPANSKAYA, K., LUNGREN, M. P. & NG, A. Y. 2017. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning.
- REPUBLIK, Ö. 1998. Bundesgesetz über die Ausübung des ärztlichen Berufes und die Standesvertretung der Ärzte (Ärztegesetz 1998 – ÄrzteG 1998).
- REPUBLIK, Ö. 2004. Bundesgesetz zur Qualität von Gesundheitsleistungen (Gesundheitsqualitätsgesetz – GQG).
- REPUBLIK, Ö. 2012. Gesundheitstelematikgesetz 2012 (GTelG 2012).
- REPUBLIK, Ö. 2021. Bundesgesetz betreffend Medizinprodukte 2021 (Medizinproduktegesetz 2021 – MPG 2021).
- SÆTRA, H. S. & FOSCH-VILLARONGA, E. 2021. Healthcare Digitalisation and the Changing Nature of Work and Society. *Healthcare*, 9, 1007.
- SARMIENTO, R. F. & DERNONCOURT, F. 2016. Improving Patient Cohort Identification Using Natural Language Processing. *Secondary Analysis of Electronic Health Records*. Cham: Springer International Publishing.
- SCHMÜCKER, P. 2024. Medizinische Dokumentation als Basis für das Health Data Management. In: HENKE, V., HÜLSKEN, G., SCHNEIDER, H. & VARGHESE, J. (eds.) *Health Data Management: Schlüsselfaktor für erfolgreiche Krankenhäuser*. Wiesbaden: Springer Fachmedien Wiesbaden.
- SCHNEEWEISS, S. 2023. Von Real-World-Daten zur Real-World-Evidenz: eine praktische Anleitung. *Prävention und Gesundheitsförderung*.
- SCHORK, N. J. 2019. Artificial Intelligence and Personalized Medicine. In: VON HOFF, D. D. & HAN, H. (eds.) *Precision Medicine in Cancer Therapy*. Cham: Springer International Publishing.
- SCHUBERT, M. C. & ET AL. 2024. Besteht ChatGPT die neurologische Facharztprüfung? Eine kritische Analyse. *SpringerLink*.
- SHEINER, L. B. 1969. Computer-aided long-term anticoagulation therapy. *Computers and Biomedical Research*, 2, 507-518.
- SHEINER, L. B. & BEAL, S. L. 1982. Bayesian individualization of pharmacokinetics: simple implementation and comparison with non-Bayesian methods. *J Pharm Sci*, 71, 1344-8.
- SIBIEUDE, E., KHANDELWAL, A., HESTHAVEN, J. S., GIRARD, P. & TERRANOVA, N. 2021. Fast screening of covariates in population models empowered by machine learning. *J Pharmacokinet Pharmacodyn*, 48, 597-609.

- SOEORG, H., SVERRISDÓTTIR, E., ANDERSEN, M., LUND, T. M. & SESSA, M. 2022. Artificial Neural Network vs. Pharmacometric Model for Population Prediction of Plasma Concentration in Real-World Data: A Case Study on Valproic Acid. *Clin Pharmacol Ther*; 111, 1278-1285.
- SONODA, Y., KUROKAWA, R., NAKAMURA, Y., KANZAWA, J., KUROKAWA, M., OHIZUMI, Y., GONOI, W. & ABE, O. 2024. Diagnostic performances of GPT-4o, Claude 3 Opus, and Gemini 1.5 Pro in “Diagnosis Please” cases. *Japanese Journal of Radiology*, 42, 1231-1235.
- STACEY, D., LÉGARÉ, F., LEWIS, K., BARRY, M. J., BENNETT, C. L., EDEN, K. B., HOLMES-ROVNER, M., LLEWELLYN-THOMAS, H., LYDDIATT, A., THOMSON, R. & ET AL. 2017. Decision aids for people facing health treatment or screening decisions. *Cochrane Database of Systematic Reviews*.
- STANFORD UNIVERSITY - ERIC, R. 2004. An Overview of Natural Language Processing: History and Current State.
- SURUSHA, T. 2023. *Emerging Natural Language Processing Technologies of 2023* [Online]. Available: <https://medium.com/@surushatechnology53/emerging-natural-language-processing-technologies-of-2023-8816718a3cfb> [Accessed].
- THE ECONOMIST, I. 2025. Artificial Intelligence in the Real World.
- TOPOL, E. J. 2019. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25, 44-56.
- TUPASELA, A. & DI NUCCI, E. 2020. Concordance as evidence in the Watson for Oncology decision-support system. *AI & SOCIETY*, 35, 811-818.
- VICINI, P. & VAN DER GRAAF, P. H. 2013. Systems pharmacology for drug discovery and development: paradigm shift or flash in the pan? *Clin Pharmacol Ther*, 93, 379-81.
- WANG, Z., GUO, R., SUN, P., QIAN, L. & HU, X. 2024. Enhancing Diagnostic Accuracy and Efficiency with GPT-4-Generated Structured Reports: A Comprehensive Study. *Journal of Medical and Biological Engineering*, 44, 144-153.
- WIDMER, A. 2024. Digitale Gesundheits-Apps. *Die Innere Medizin*, 65, 1261-1265.
- WOILLARD, J. B., LABRIFFE, M., DEBORD, J. & MARQUET, P. 2021. Tacrolimus Exposure Prediction Using Machine Learning. *Clin Pharmacol Ther*; 110, 361-369.
- WORLD ECONOMIC, F. 2023. *World Health Day: How AI and Digital Health Are Transforming Healthcare* [Online]. Available: <https://www.weforum.org/stories/2023/04/world-health-day-how-ai-and-digital-health-healthcare/> [Accessed].
- WORLD HEALTH, A. 1967. WHO Nomenclature Regulations 1967. Geneva: World Health Organization.
- WORLD HEALTH, O. 2004. *WHO Classifications* [Online]. Available: <https://web.archive.org/web/20041003044904/http://www.who.int/classifications/en/> [Accessed].
- WORLD HEALTH, O. 2022. International Classification of Diseases (ICD) 11 Reference Guide.
- WORLD HEALTH, O. 2025. *ICHI Reference Guide*.
- ZHANG, F., WU, S., ZHANG, C., CHEN, Q., YANG, X., JIANG, K. & ZHENG, J. 2019. Multi-domain features for reducing false positives in automated detection of clustered microcalcifications in digital breast tomosynthesis. *Medical Physics*.
- ZHANG, T., ANDROULAKIS, I. P., BONATE, P., CHENG, L., HELIKAR, T., PARIKH, J., RACKAUCKAS, C., SUBRAMANIAN, K. & CHO, C. R. 2022. Two heads are better than one: current landscape of integrating QSP and machine learning : An ISoP QSP SIG white paper by the working group on the integration of quantitative systems pharmacology and machine learning. *J Pharmacokinet Pharmacodyn*, 49, 5-18.
- ZHOU, N., ZHANG, C. T., LV, H. Y., HAO, C. X., LI, T. J., ZHU, J. J., ZHU, H., JIANG, M., LIU, K. W., HOU, H. L., LIU, D., LI, A. Q., ZHANG, G. Q., TIAN, Z. B. & ZHANG, X. C. 2018. Concordance Study Between IBM Watson for Oncology and Clinical Practice for Patients with Cancer in China. *The Oncologist*, 24, 812-819.
- ZHU, X., HUANG, W., LU, H., WANG, Z., NI, X., HU, J., DENG, S., TAN, Y., LI, L., ZHANG, M., QIU, C., LUO, Y., CHEN, H., HUANG, S., XIAO, T., SHANG, D. & WEN, Y. 2021. A machine learning approach to personalized dose adjustment of lamotrigine using noninvasive clinical parameters. *Sci Rep*, 11, 5568.