MASTERARBEIT

Künstliche Intelligenz im Krankenhaus

Gesetzliche Rahmenbedingungen und ethische Aspekte

ausgeführt am

CAMPUS 02 GRAZ

FACHHOCHSCHULE DER WIRTSCHAFT

Studiengang

Informationstechnologien und Wirtschaftsinformatik

Von: DI Georg Strahlhofer

Personenkennzeichen: 2310320017

Graz, am 01. Juli 2025 .....

Unterschrift

EHRENWÖRTLICHE ERKLÄRUNG

Ich erkläre ehrenwörtlich, dass ich

- die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst,
- andere als die angegebenen Quellen nicht benutzt,
- die den Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht,
- den Einsatz von generativen KI-Modellen (z.B. ChatGPT) kenntlich gemacht,
- oder nur zur Unterstützung bei der Paraphrasierung sowie zur Überprüfung von Rechtschreibung und Grammatik verwendet,
- und mich sonst keiner unerlaubten Hilfsmittel bedient habe.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht

Graz, am 01. Juli 2025 .....

Unterschrift

KURZFASSUNG

Diese Masterarbeit setzt sich mit den gesetzlichen Rahmenbedingungen sowie den ethischen Fragestellungen im Zusammenhang mit dem Einsatz Künstlicher Intelligenz (KI) im Krankenhauswesen, insbesondere im Kontext österreichischer Gesundheitseinrichtungen, auseinander. Hintergrund ist die zunehmende Integration von KI-Systemen in medizinische Prozesse, wie etwa in der Radiologie, der Pflege oder in Entscheidungsunterstützungssystemen. Besonderes Augenmerk wird auf den EU AI Act gelegt, der sich als zentrales Regulierungsinstrument etabliert. Zahlreiche dieser Anwendungen gelten als hochriskant, was erweiterte Anforderungen in Hinblick auf Sicherheit, Transparenz und Kontrolle zur Folge hat.

Ziel der Arbeit ist es, auf Grundlage einer Analyse des EU AI Act konkrete Handlungsempfehlungen für österreichische Krankenhäuser zu formulieren, um deren rechtskonformen und ethisch verantwortungsvollen Umgang mit KI zu fördern. Hierfür wurde ein gestaltungsorientierter Forschungsansatz (Design Science Research) gewählt. Das entwickelte Artefakt, in Form eines Leitfadens zur Verwendung und Entwicklung von KI, wurde durch qualitative Experteninterviews überprüft.

Der theoretische Teil der Arbeit beleuchtet sowohl technologische Grundlagen als auch organisationale Aspekte des Krankenhausbetriebs. Ergänzt um Perspektiven aus der Praxis zeigt sich, dass die erfolgreiche Umsetzung regulatorischer Anforderungen maßgeblich vom bestehenden Daten- und Prozessverständnis innerhalb der Organisation sowie von einem funktionierenden interdisziplinären Austausch abhängt.

Die abschließend formulierten Handlungsempfehlungen sollen Entscheidungsträger:innen im Gesundheitswesen dabei unterstützen, die Potenziale von KI zu nutzen und gleichzeitig konform nach dem EU AI ACT zu handeln.

ABSTRACT

This master's thesis examines the legal and ethical challenges associated with the application of Artificial Intelligence (AI) in hospital environments, with a particular focus on Austrian healthcare institutions. Against the backdrop of increasing integration of AI technologies into medical processes, such as radiology, nursing, and clinical decision support, the study analyzes the European Union's AI Act as a central regulatory framework. Due to the classification of many AI systems in healthcare as high-risk applications, the EU AI Act imposes extensive requirements regarding safety, transparency, and accountability.

The objective of this thesis is to develop practical, legally compliant, and ethically sound recommendations for the use of AI technologies in hospitals. To achieve this, a Design Science Research (DSR) methodology was applied, resulting in the development of a structured guideline (artifact) for the implementation of AI systems. This guideline was reviewed and refined through interviews with experts from both the healthcare and technology sectors.

The theoretical part provides an overview of AI technologies and the organizational structures of hospitals, while the practical part includes insights based on real-world experience. The findings emphasize the importance of interdisciplinary collaboration, a solid understanding of data processes, and ongoing monitoring to ensure the responsible use of AI. The resulting recommendations are intended to support hospital decision-makers in making the most of AI, while complying with legal requirements.

INHALTSVERZEICHNIS

1 EINLEITUNG 1

1.1 Forschungsfragen und Zielsetzung 3
1.2 Methodik und Vorgehensweise 4

2 INTELLIGENZ IM GESUNDHEITSWESEN 6

2.1 Begriffsdefinitionen Künstliche Intelligenz 6
2.1.1 Maschinelles Lernen 7
2.1.2 Tiefes Lernen und künstliche neuronale Netzwerke 8
2.1.3 Lebenszyklus eines KI-Systems 10

2.2 Begriffsdefinitionen Gesundheitswesen 12
2.2.1 Aufbauorganisation 15
2.2.2 Ablauforganisation 16

2.3 Anwendungen Künstlicher Intelligenz im Krankenhaus 16
2.3.1 Künstliche Intelligenz in der Pflege 17
2.3.2 ChatGPT im Krankenhaus 19
2.3.3 Künstliche Intelligenz in der Radiologie 20
2.3.4 Intelligente Triagesysteme 21
2.3.5 Spezifische Risiken der Künstlichen Intelligenz im Gesundheitswesen 21

3 GRUNDLAGEN DES EU AI ACT 24

3.1 Rollen und Zuständigkeiten 26
3.2 Risikoklassifizierungssystem 27
3.2.1 Verbotene Künstliche Intelligenz 28
3.2.2 Hochriskante Künstliche Intelligenz 29
3.2.3 Künstliche Intelligenz mit geringem Risiko 29
3.2.4 Allgemeine Künstliche Intelligenz 30

3.3 Verpflichtungen 31

4 AUFSTELLEN DER HANDLUNGSEMPFEHLUNGEN 34

4.1 Phase 1: Geschäfts- und Datenverständnis 35
4.2 Phase 2: Datenvorbereitung 35
4.3 Phase 3: Modellierung 36
4.4 Phase 4: Evaluation 37
4.5 Phase 5: Bereitstellung 37
4.6 Phase 6: Überwachung und Wartung 38

5 EVALUIERUNG DER HANDLUNGSEMPFEHLUNGEN 40

5.1 Erhebungsmethode – das Experteninterview 41
5.1.1 Erstellung des Interviewleitfadens 42
5.1.2 Auswahl der Experten und Expertinnen 46

5.2 Transkription und Auswertungsmethode 50

6 ERGEBNISSE 54

6.1 Phase 1: Geschäfts- und Datenverständnis 56
6.2 Phase 2: Datenvorbereitung 59
6.3 Phase 3: Modellierung 61
6.4 Phase 4: Evaluation 63
6.5 Phase 5: Bereitstellung 65
6.6 Phase 6: Überwachung und Wartung 68
6.7 Querschnittsfragen 69

7 DISKUSSION 72

8 FAZIT 75

ABBILDUNGSVERZEICHNIS 77

TABELLENVERZEICHNIS 78

LITERATURVERZEICHNIS 79

ANHANG 87

1 EINLEITUNG

Künstliche Intelligenz (KI) hat im täglichen Leben Einzug gefunden, und der Mensch ist ständig damit konfrontiert. Das Verwenden einer Suchmaschine, das Auswählen von Serien und Filmen, das Lenken eines Fahrzeugs oder das Steuern des eigenen Hauses – dies alles sind Bereiche, in denen KI bereits im Einsatz ist (Heinrichs, 2023). Aber nicht nur der gewöhnliche Alltag jeder Person ist von der Verbreitung der KI betroffen, sondern auch in der Medizin hat sich KI bereits etabliert oder wird erprobt. Das durch den Einsatz dieser Technologie entstehende Potenzial wird dabei als erheblich eingestuft (Arge ITA-AIT Parlament, 2020; Bures et al., 2023; Fraunhofer-Institut für Kognitive Systeme IKS, 2024b; Küster & Schultz, 2023). Eine im Jahr 2019 durchgeführte Studie, die sich auf die primäre Gesundheitsversorgung konzentrierte, stellte im Rahmen einer dreirundigen Delphi-Studie mit insgesamt vierzig Experten und Expertinnen aus verschiedenen Gremien fest, dass der Einsatz von KI in diesem Bereich grundsätzlich zu Verbesserungen führt. Die Meinungen der Experten und Expertinnen unterscheiden sich jedoch in der Fragestellung, ob KI-Systeme im Betrieb lernen sollen, da das Ausmaß der Veränderung nur bedingt abschätzbar ist. Weiters ist festzuhalten, dass besonders in diesem Bereich ein sorgfältiges Design und eine entsprechende Einschätzung anhand unterschiedlichster Kriterien der KI-Systeme sicherzustellen ist (Liyanage et al., 2019).

Die potenziellen Anwendungsbereiche von KI in Gesundheitseinrichtungen sind vielseitig und reichen von der Erstellung von Diagnosen, über die Optimierung von Prozessen, bis hin zur Aufklärung und Schulung des Gesundheitsfachpersonals (Fraunhofer-Institut für Kognitive Systeme IKS, 2024b). Die Arbeitsgemeinschaft ITA-AIT des österreichischen Parlaments führt ebenfalls vergleichende Beispiele an, wobei speziell der Einsatz in der Diagnostik und im Krankenhausmanagement hervorgehoben wird (Arge ITA-AIT Parlament, 2020).

Basierend auf den Ergebnissen der zuvor erwähnten Studie, würde eine ausschließliche Fokussierung auf die Potenziale und Möglichkeiten der Technologie eine unzulässige Vereinfachung darstellen, da der Einsatz von KI inhärente Risiken birgt, die insbesondere in der medizinischen Anwendung aufgrund der überwiegend kritischen Natur der Einsatzgebiete deutlich werden. Die Mehrheit der vorliegenden wissenschaftlichen Quellen stimmt weitgehend mit dieser Einschätzung überein (Arge ITA-AIT Parlament, 2020; Bures et al., 2023; Fraunhofer-Institut für Kognitive Systeme IKS, 2024b; Gama et al., 2022; Küster & Schultz, 2023). Das potenzielle Ende der Privatsphäre durch den missbräuchlichen Umgang mit Daten sowie die Rolle der KI als Katalysator für Ungleichheit und Vorurteile sind hier nur ein Ausschnitt an konkreten Beispielen (Bajema et al., 2022; Küster & Schultz, 2023). Ein weiterer nennenswerter Aspekt in diesem Kontext ist meist die fehlende Transparenz in der Entscheidungsfindung (Mitchell et al., 2023). Aber auch unvorhergesehene Fehler und deren Auswirkungen dürfen nicht unerwähnt bleiben (Keskinbora, 2019).

Ein erster wesentlicher Schritt im Umgang mit KI in diesem Bereich besteht daher in der Erlangung von Grundkompetenzen der Entscheidungsträger:innen im Gesundheitswesen (Chen & Decary, 2020). Die zuvor beschriebenen Probleme werden dadurch nicht gelöst, sondern in erster Linie nur in das Bewusstsein der Personen gerufen.

Zusammengefasst ist eine Regulierung von KI speziell, aber nicht ausnahmslos, im Gesundheitswesen notwendig. Diese Regulierung ist auch als Chance zu sehen, den Wert von KI-Systemen entsprechend zu steigern (Pesapane et al., 2018). Vor diesem Hintergrund wurden im Auftrag der Europäischen Kommission im April 2019 von einer unabhängigen, hochrangigen Expertengruppe für künstliche Intelligenz (HEG-KI) Ethik-Leitlinien für eine vertrauenswürdige KI auf Basis der EU-Verträge, der EU-Grundrechtecharta und der internationalen Menschenrechte veröffentlicht. Diese Leitlinien betonen die Notwendigkeit, dass KI-Systeme sicher, transparent und diskriminierungsfrei sein müssen, um das Vertrauen der Nutzer:innen zu gewinnen und zu erhalten (Europäische Kommission, Generaldirektion Kommunikationsnetze, Inhalte und Technologien, 2019).

Im April 2024 markierte die Verabschiedung des EU-Gesetz zur künstlichen Intelligenz, kurz EU AI ACT, einen bedeutenden nächsten Schritt für KI-Anwendungen innerhalb der Europäischen Union. Dieses Gesetz adressiert unter anderem die Einschränkung der Nutzung biometrischer Fernidentifizierungssysteme durch Strafverfolgungsbehörden sowie das Verbot des Einsatzes von KI-Systemen zur Bewertung sozialen Verhaltens, Beeinflussung von Personen oder Ausnutzung ihrer Schwächen. Darüber hinaus umfasst es Regelungen für den Bereich der sogenannten Hochrisikosysteme, die von wesentlicher Bedeutung für das Gesundheitswesen sind (Europäisches Parlament, 2024). Es gibt bereits erste Publikationen, die den EU AI ACT in Hinblick auf die wesentlichsten Anforderungen analysieren (Wagner et al., 2024). In älteren Publikationen wurde der Entwurf des EU AI ACT analysiert und es wurde damals festgestellt, dass Schlupflöcher für bestimmte Bereiche existierten. Weiters wurden zum damaligen Zeitpunkt strengere rechtliche Anforderungen gefordert. Als Beispiele werden die Bewertung der Konformität durch Dritte, die Bewertung der Auswirkungen auf die Grundrechte, Transparenzpflichten sowie die bestehenden EU-Datenschutzgesetze und die Rechte und Rechtsmittel, die den Einzelpersonen zur Verfügung stehen, genannt (Barkane, 2022). Zu einem ähnlichen Ergebnis gelangt auch Raposo (2022). In dieser Publikation werden die Gründe für die Schaffung als lobenswert bezeichnet, jedoch griff der Entwurf in vielerlei Hinsicht zu kurz. Zudem wird angemerkt, dass das Gesetz zu Lasten der digitalen Innovation gehen könnte. Diese Kritikpunkte sind insofern von Relevanz, da sie die Grenzen des EU AI ACT aufzeigen. Solaiman und Malik (2025) üben Kritik am EU AI Act, da dieser ihrer Ansicht nach die Arzt-Patient-Beziehung nicht hinreichend berücksichtigt und die Autonomie sowie die Partizipationsmöglichkeiten der Patienten und Patientinnen bei Entscheidungen über ihre Versorgung, potenziell einschränken könnte.

KI-Systeme in Einrichtungen im Gesundheitswesen, insbesondere Krankenhäuser, werden mit hoher Wahrscheinlichkeit gemäß dem EU AI Act als Hochrisikosysteme eingestuft oder sind von den daraus resultierenden regulatorischen Anforderungen betroffen. Vor diesem Hintergrund stehen solche Einrichtungen vor der Aufgabe, sich eingehend mit dem EU AI ACT und seinen Implikationen zu beschäftigen.

1.1 Forschungsfragen und Zielsetzung

Vor diesem Hintergrund widmet sich die vorliegende wissenschaftliche Arbeit der Beantwortung der folgenden zentralen Forschungsfrage:

„Wie sollen österreichische Krankenhäuser mit ihren spezifischen Gegebenheiten, KI-Systeme entwickeln bzw. einsetzen, um den Vorgaben des EU AI ACT gerecht zu werden?“

Weiters ergeben sich folgende drei Teilfragen:

„Welche spezifischen Anforderungen des EU AI ACT sind für KI-Anwendungen in österreichischen Krankenhäusern relevant?“

„Wie beeinflusst der EU AI ACT die Auswahl und Entwicklung von KI-Systemen in österreichischen Krankenhäusern?“

„Welche Maßnahmen sollen österreichische Krankenhäuser ergreifen, um Compliance mit dem EU AI Act sicherzustellen?“

Im weiteren Verlauf dieser Untersuchung zeigt sich, dass die Implementierung von KI-Systemen in Krankenhäusern nicht ausschließlich unter technischen Gesichtspunkten betrachtet werden kann. Vielmehr steht sie gleichermaßen im Spannungsfeld ethischer Verantwortung und rechtlicher Anforderungen. Der EU AI Act nimmt dabei eine zentrale Rolle ein, da er einen verbindlichen Rahmen für den sicheren und ethisch vertretbaren Einsatz von KI im Gesundheitswesen vorgibt.

Die Beantwortung der zentralen Forschungsfrage sowie der daraus abgeleiteten Teilfragen ist entscheidend, um nachvollziehen zu können, unter welchen Voraussetzungen Krankenhäuser neue KI-Technologien verantwortungsvoll einsetzen können, und gleichzeitig regulatorischen Vorgaben entsprechen. Diese Herausforderungen sind vielschichtig und erfordern eine sorgfältige Abwägung zwischen dem Potenzial von KI zur Verbesserung der medizinischen Versorgung und dem Schutz individueller Grundrechte sowie der Sicherheit sensibler Gesundheitsinfrastrukturen.

Das Ziel dieser wissenschaftlichen Arbeit besteht darin, die Inhalte des EU AI ACT entsprechend aufzuarbeiten, sodass sich daraus möglichst konkrete Handlungsempfehlungen für Krankenhäuser ableiten lassen. Anhand dieser Aufarbeitung wird es möglich sein, die zuvor genannten Forschungsfragen zu beantworten. Im Gesundheitswesen gibt es weitere Standards, wie spezifische ethische Richtlinien oder die EU-Medizinprodukte-Verordnung (Palmer & Schwan, 2024). Diese werden in dieser wissenschaftlichen Arbeit nur am Rande behandelt und es handelt sich somit um ein Nicht-Ziel andere Richtlinien oder Gesetze zu untersuchen.

1.2 Methodik und Vorgehensweise

In der Wirtschaftsinformatik gibt es zwei vorherrschende Forschungsströmungen. In der verhaltensorientierten Forschung wird versucht, Theorien zu entwickeln bzw. zu verifizieren, die das Verhalten von Menschen und Organisationen beschreiben. Im Gegensatz dazu werden in der gestaltungsorientierten Forschung Artefakte entwickelt. Für diese zweite Forschungsströmung findet sich im deutschsprachigen Raum der Begriff Design Science Research (DSR) (Hevner et al., 2004).

In der gestaltungsorientierten Forschung wird weiters zwischen zwei grundlegenden Strategien unterschieden. Einerseits ist es möglich, Artefakte auf einer Metaebene zu entwickeln, um dadurch eine Vielzahl an Problemen zu lösen. Andererseits besteht die Möglichkeit, ein spezifisches Problem zu lösen und die gewonnenen Erkenntnisse auf eine Vielzahl von Problemen zu übertragen (Iivari, 2015). Bezugnehmend auf die zuvor beschriebene Forschungsfrage und Zielsetzung sowie die dargelegten Grundlagen zur verhaltensorientierten und gestaltungsorientierten Forschung, wird für diese wissenschaftliche Arbeit der DSR-Ansatz gewählt. Ziel ist die Entwicklung eines Artefakts in Form von Handlungsempfehlungen für die Implementierung des EU AI Act in Krankenhäusern auf Metaebene. Obwohl die gewählte Anwendungsdomäne „KI-Anwendungen in Krankenhäusern“ bereits einen konkreten Rahmen vorgibt, kann aufgrund der stark variierenden Einsatzszenarien nicht von einem Einzelfall ausgegangen werden. Daher muss die Entwicklung auf einer Metaebene erfolgen.

Die Entwicklung dieser Handlungsempfehlungen erfolgt nach einem klar definierten und wissenschaftlich fundierten Prozess, der übergeordnet in vier Schritte unterteilt wird. Im ersten Schritt wird basierend auf einer Literaturrecherche das Problem identifiziert und seine Relevanz verdeutlicht. Im zweiten Schritt wird eine weitere Literaturrecherche durchgeführt und das Artefakt basierend auf den gewonnenen Erkenntnissen entwickelt. Im dritten Schritt wird anhand von empirischer Sozialforschung das Artefakt evaluiert und im vierten Schritt finalisiert. Dieser Prozess wurde von Offermann et al., 2009 publiziert und stellt eine Kombination mehrerer anderer DSR-Prozesse dar.

In der empirischen Sozialforschung wird eine Befragung von Experten und Expertinnen entweder in Form von Einzelinterviews durchgeführt (Häder, 2015, S. 194). Die Auswertung dieser Befragung erfolgt anhand einer qualitativen Inhaltsanalyse (Mayring, 2020).

Umgelegt auf diese wissenschaftliche Arbeit bedeutet das konkret, dass nach dieser Einleitung, die die Bedeutung des Themas und die zentralen Forschungsfragen darstellt, im zweiten Kapitel die theoretischen Grundlagen zum Thema KI im Gesundheitswesen behandelt werden. Damit werden die zuvor angesprochenen Grundkompetenzen, die zum Verständnis der nachfolgenden Absätze notwendig sind, geschaffen. Das dritte Kapitel ist dem EU AI ACT gewidmet. Dabei werden die spezifischen Anforderungen, die der EU AI ACT an KI-Anwendungen in Krankenhäusern stellt, detailliert erörtert und die daraus resultierenden Herausforderungen für diese Einrichtungen analysiert. Dieses Kapitel bildet den Abschluss des ersten Schrittes der zuvor beschriebenen vierstufigen Vorgehensweise. Auf Grundlage dieser Analyse werden im vierten Kapitel konkrete Handlungsempfehlungen für Krankenhäuser entwickelt. Diese sollen als

Leitfaden dienen, wie KI-Systeme gemäß den Vorgaben des EU AI ACT implementiert werden können. Dieses Kapitel stellt den zweiten Schritt des theoretischen Modells dar. Im fünften Kapitel, das dem dritten Schritt entspricht, werden die erarbeiteten Handlungsempfehlungen durch Interviews validiert. Der Austausch mit Experten und Expertinnen dient der Überprüfung der Praktikabilität und Vollständigkeit der Empfehlungen. Abschließend werden im sechsten Kapitel die Anmerkungen und Erkenntnisse der Experten und Expertinnen in die Handlungsempfehlungen integriert und finalisiert. Abschließend werden im siebten Kapitel die Ergebnisse diskutiert und im letzten Kapitel das Fazit gebildet.

2 INTELLIGENZ IM GESUNDHEITSWESEN

In diesem Kapitel werden zunächst die allgemeinen Grundlagen zum Thema KI dargestellt. Der Fokus liegt dabei auf der Vermittlung der Basiskenntnisse, um das Verständnis zu schärfen und das Bewusstsein der Entscheider:innen in Krankenhäusern zu stärken, wie in der Einleitung als notwendig dargelegt. Im nächsten Schritt werden die spezifischen Aspekte des Gesundheitswesens, insbesondere von Krankenhäusern in Österreich, ausgearbeitet. Im letzten Absatz dieses Kapitels werden ausgewählte Anwendungsbeispiele von KI-Systemen in der Anwendungsdomäne vorgestellt.

2.1 Begriffsdefinitionen Künstliche Intelligenz

Am Beginn eines Grundlagenkapitels erscheint es sinnvoll, mit entsprechenden Begriffsdefinitionen zu beginnen. Jedoch erweist sich bereits die erste Definition als herausfordernd, denn der Begriff KI hat keine einheitliche Definition und eine Vereinheitlichung ist aufgrund der stetigen Entwicklung nicht absehbar (Ertel, 2021; Mockenhaupt, 2021)

Aufgrund des Schwerpunkts dieser wissenschaftlichen Arbeit werden die Definitionen seitens der Europäischen Kommission und des Europäischen Parlaments als sinnvoll erachtet und herangezogen. Die Europäische Kommission definierte 2019 KI als ein System, das intelligentes Verhalten aufweist. Konkret analysieren diese Systeme ihre Umwelt und handeln mit einem bestimmten Grad an Autonomie, um definierte Ziele zu erreichen. Es handelt sich bei diesen Systemen einerseits um rein softwaregestützte Lösungen und andererseits kann das intelligente Verhalten auch in Hardware eingebettet sein. Ein weiterer Aspekt ist die Fähigkeit dieser Systeme, ihre Entscheidungen durch Lernen zu verbessern (Europäische Kommission, Generaldirektion Kommunikationsnetze, Inhalte und Technologien, 2019)

Auch das Europäische Parlament spricht im EU AI ACT von KI-Systemen und verwendet den Begriff „maschinenbasiert“, da KI-Systeme von Maschinen betrieben werden. In dieser Definition wird jedoch nicht von Hard- und Software gesprochen, sondern davon, ob sie im Produkt eingebettet oder nicht eingebettet sind. Dabei ist es irrelevant, ob es sich um ein physisches Produkt handelt. Die Hauptfähigkeit eines KI-Systems in dieser Definition ist das Ableiten, also das Erstellen von Ergebnissen wie Vorhersagen, Inhalten, Empfehlungen oder Entscheidungen. Dazu gehört auch das Erstellen von Modellen oder Algorithmen aus Eingaben oder Daten sowie Methoden des maschinellen Lernens. Außerdem wird in dieser Definition auf die Arbeit mit expliziten und impliziten Zielen hingewiesen und der Begriff Umgebung oder Kontext eingeführt. Umgebung beschreibt das System mit seinen Grenzen, indem das KI-System operiert. Das letzte Merkmal eines KI-Systems in dieser Definition ist die Anpassungsfähigkeit bzw. die Fähigkeit,

nach der Inbetriebnahme zu lernen und sich im Laufe des Lebenszyklus zu verändern (Europäisches Parlament, 2024).

Im Zusammenhang mit KI finden sich häufig die Begriffe maschinelles Lernen und tiefes Lernen. Bei beiden Begriffen handelt es sich um Teilgebiete bzw. Schlüsselkonzepte von KI und sie werden deshalb im Folgenden näher beschrieben. Zunächst wird das maschinelle Lernen behandelt, da das Deep Learning als dessen Teilbereich gilt und anschließend beschrieben wird (Fraunhofer-Institut für Kognitive Systeme IKS, 2024a; Schreyögg et al., 2022).

2.1.1 Maschinelles Lernen

KI-Systeme, die auf dem Konzept von maschinellem Lernen basieren, erkennen selbständig Muster und Gesetzmäßigkeiten. Es wird kein Modell zugrunde gelegt und durch Vorhandensein einer Rückkopplung der Resultate wird das System befähigt sich entweder mittels Trainingsdaten zeitpunktbezogen oder stetig weiterzuentwickeln (Auer et al., 2019; Babel, 2024; Mockenhaupt, 2021; Schreyögg et al., 2022). Eine Abgrenzung zur klassischen Programmierung ist insofern möglich, dass bei der klassischen Programmierung Regel und Daten als Input und Ergebnisse als Output fungieren. Im Gegensatz ist es beim maschinellen Lernen so, dass Daten und Ergebnisse als Input zu sehen sind und die Regeln den Output bilden (Wahl & Wilhelm, 2024). Beim Lernen wird übergeordnet zwischen überwachtem (z. B. Prädiktion), unüberwachtem (z. B. Klassifikation) und bestärkendem Lernen (engl. Reinforcement Learning) unterschieden (Klein et al., 2024; Schreyögg et al., 2022).

Das **überwachte Lernen** (engl. Supervised Learning) beschreibt einen Vorgang, wo gelabelte Daten verwendet werden, um den Algorithmus bzw. das Programm zu trainieren. Gelabelte Daten stehen´ für Datensätze, die Eingaben und Ausgaben umfassen. Während des Trainings versucht der Algorithmus anhand der Eingaben, die dazugehörigen Ausgaben vorherzusagen und adaptiert dabei seine freien Parameter. Nach dem iterativen Training mit Trainingsdaten und der späteren Validierung bzw. Testung mit abweichenden, aber gelabelte Daten, kann das Modell für unbekannte Eingaben herangezogen werden und die Ausgaben mit einer gewissen Genauigkeit vorhersagen. Eine Anwendung findet dieser Ansatz überall dort, wo bereits beschriftete Daten beispielsweise eine Klassifizierung vorhanden sind (Granigg & Lichtenegger, 2024; Lang, 2023). Ein Beispiel aus dem Gesundheitswesen für überwachtes Lernen stellt die Wundbehandlung dar. Ein Modell wird mit Wundbildern trainiert, um im Betrieb festzustellen, ob eine Infektion vorliegt. Dafür werden dem Modell zahlreiche Bilder verschiedener Wunden präsentiert, die vorab von medizinischen Fachpersonal manuell als infiziert oder nicht infiziert kategorisiert wurden. Das Modell lernt daraufhin, welche Merkmale auf den Bildern auf eine Infektion hinweisen. Nach ausreichendem Training könnte der Algorithmus in der klinischen Praxis verwendet werden, um den Diagnoseprozess von Wundinfektionen zu unterstützen (Wahl & Wilhelm, 2024).

Im Gegensatz zum überwachten Lernen werden beim **unbeaufsichtigten Lernen** (engl. Unsupervised Learning) keine beschrifteten Daten verwendet. Das Ziel bei dieser Form des Lernens ist es, dass das Modell selbstständig Muster und Strukturen ohne vordefinierte Merkmale erkannt und basierend auf Wahrscheinlichkeiten die beste Kategorisierung vornimmt. Dieser Ansatz wird dort verwendet, wo nicht klar ist, welche Merkmale in einem Datensatz

beschriftungswürdig sind und erst Ähnlichkeiten und Muster gefunden werden müssen (Granigg & Lichtenegger, 2024; Lang, 2023). Ein Anwendungsbeispiel für diese Form des Lernens im Gesundheitswesen wäre die Erkennung von Mustern in Vitalparametern wie Herzfrequenz, Blutdruck und Atmung, um auf eine kritische Entwicklung hinzuweisen (Wahl & Wilhelm, 2024).

Das **bestärkende Lernen** (engl. Reinforcement Learning) unterscheidet sich gänzlich von den zuvor beschriebenen Formen des maschinellen Lernens. Bei dieser Form des Lernens führt ein sogenannter Agent Handlungen in seiner Umwelt durch. Das Ziel dieser Art des Lernens ist eine Maximierung der Belohnung. Eine Belohnung erhält der Agent im Zuge seiner Handlungen, wenn sie ihn der Erreichung des Endziels näherbringt. Ist dies nicht der Fall, so gibt es einen Malus. Verwendung findet diese Form des Lernens, wenn der ideale Zustand des Agenten im System nicht definiert werden kann und wenn eine Entwicklung nur über Interaktion mit der Umgebung möglich ist (Granigg & Lichtenegger, 2024; Lang, 2023). Im Gesundheitswesen wird das Trainieren eines Pflegeroboters als Anwendungsbeispiel betrachtet. Der Roboter kann mithilfe eines implementierten Feedbacksystems aus positiven und negativen Rückmeldungen für seine Handlungen lernen. Voraussetzung dafür ist, dass der Roboter seine Umgebung sowie die Bedürfnisse der zu pflegenden Person versteht und seine Aktionen entsprechend ausrichtet. Durch das iterative Feedback kann der Roboter seine Handlungen kontinuierlich verbessern und somit eine individuelle Pflege gewährleisten (Wahl & Wilhelm, 2024). Ohne an dieser Stelle detaillierter in das vorliegende Problem einzutauchen, muss darauf hingewiesen werden, dass das genannte Beispiel auch Risiken mit sich bringt. Theoretisch wäre es möglich, dass ein Roboter Fehler macht, die der zu pflegenden Person schaden und somit zu einem Malus für den Roboter führen. Dieses Beispiel verdeutlicht, dass je nach Einsatzbereich des KI-Systems unterschiedliche Lernmethoden angewendet werden können.

Eine weitere, eher generelle Unterscheidung betrifft das Lernen mit Trainingsdaten bzw. in einem Trainingssetting und das kontinuierliche Lernen. Welche Lernmethode für den jeweiligen Anwendungsfall geeignet ist, kann durch die Beantwortung der folgenden Fragen ermittelt werden (Mockenhaupt, 2021):

- Erlaubt die Anwendung das Lernen durch Fehler?
- Müssen Ergebnisse im Betrieb nachvollziehbar sein?
- Ist eine Überwachung der Ergebnisse notwendig?
- Gibt es definierte Systemgrenzen, die im Betrieb nicht überschritten werden dürfen?

Auf Grundlage der Anforderungen an das KI-System, der Beantwortung dieser Fragestellungen und der zugrundeliegenden Datenstruktur wird ein geeignetes Modell ausgewählt. Eine bedeutende Modellform sind hierbei sogenannte künstliche neuronale Netzwerke, die eng mit dem Begriff des Deep Learning verknüpft sind.

2.1.2 Tiefes Lernen und künstliche neuronale Netzwerke

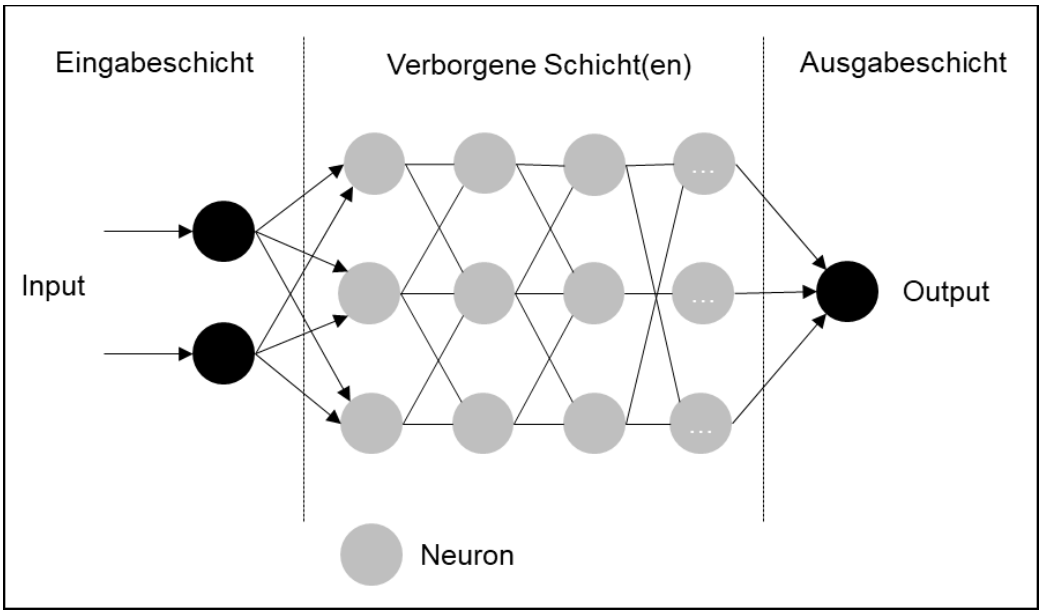
Es wird vom tiefen Lernen (engl. Deep Learning) gesprochen, wenn anstatt klassischer Verfahren des maschinellen Lernens, wie Lineare oder Logistische Regression, künstliche neuronale

Netzwerke als logisches Konzept verwendet werden. Anwendung findet das tiefe Lernen bei komplexen Aufgaben, wie Bild- und Objekterkennung oder Verarbeitung natürlicher Sprache (Klein et al., 2024; Lang, 2023; Schreyögg et al., 2022).

Grundsätzlich wird mit künstlichen neuronalen Netzwerken versucht, den menschlichen Lernprozess im Gehirn nachzubilden. Babel (2024); Ertel (2021) beschreiben das Gehirn von Menschen als Zusammenschluss aus Milliarden von Nervenzellen, sogenannten Neuronen zu einem Netzwerk. Diese Netzwerke und ihre teilweise beschriebene Funktionsweise bilden die Grundlage für menschliche Intelligenz. Intelligenz wird in diesem Kontext über die Lernfähigkeit und die Anpassungsfähigkeit an eine sich verändernde Umwelt definiert. Taulli (2023) beschreibt den Zusammenhang zwischen den Lernprozess im Gehirn und künstlichen neuronalen Netzwerken als schwach. Als konkretes Beispiel wird in dieser Publikation der Prozess für die Datenaufnahme und -verarbeitung über die Netzhaut durch ein komplexes Netzwerk, das auf der Zuweisung von Gewichtungen beruht, genannt. Dieses konkrete Beispiel kann mit künstlichen neuronalen Netzwerken verglichen werden. Jedoch ist damit nicht die ganzheitliche Funktionsweise des menschlichen Gehirns erklärt und durch weitere Gehirnforschung besteht die Chance, neue Modelle für die KI basierend auf den Erkenntnissen zu entwickeln.

Künstliche neuronale Netzwerke sind im mathematischen Sinne verschachtelte Funktionen. Jede Schicht eines neuronalen Netzes kann als eine Funktion angesehen werden, deren Eingabe entweder die ursprüngliche Eingabe in das Netz (wie ein Bild oder ein Text) oder das Ergebnis der vorherigen Schicht ist (Babel, 2024). Auf der grundlegendsten Ebene ist ein künstliches neuronales Netzwerk eine Funktion, die aus sogenannten Neuronen, Perzeptronen und Knoten besteht. Jede dieser Einheiten hat einen Wert und eine Gewichtung, die die relative Wichtigkeit angibt, und wird in die versteckte Schicht eingefügt. Die versteckte Schicht verwendet eine Funktion, deren Ergebnis die finale Ausgabe darstellt (Puppe et al., 2024; Taulli, 2023). Abbildung 1 zeigt einen schematischen Aufbau eines neuronalen Netzwerkes.

Abbildung 1: Aufbau eines künstlichen neuronalen Netzwerkes



Anmerkung: Eigene Darstellung in Anlehnung an Babel (2024); Kirste und Schürholz (2019); Mockenhaupt (2021).

Aus der Abbildung 1 ist zu entnahmen, dass ein künstliches neuronales Netzwerk aus zumindest drei Schichten zusammengesetzt ist. Die Eingabe wird erstmalig in der Eingabeschicht verarbeitet und entsprechend an in die verborgene Schicht, die selbst aus unterschiedlichen Schichten besteht, weitergeleitet. Die Weiterleitung erfolgt dabei wie zuvor beschreiben anhand der Gewichtung und des Wertes der einzelnen Neuronen. Erst wenn alle relevanten Schichten abgearbeitet wurden, wird über die Ausgabeschicht ein Ergebnis dargestellt. Aus dieser Beschreibung lässt sich der Begriff des tiefen Lernens ableiten, da die Tiefe des Netzwerks proportional zur Anzahl der verwendeten Schichten steigt: Je mehr Schichten integriert sind, desto "tiefer" ist das Netzwerk (Babel, 2024; Kirste & Schürholz, 2019; Klein et al., 2024; Mockenhaupt, 2021).

Im nächsten Absatz werden die Grundlagen der KI durch die Darstellung eines beispielhaften Lebenszyklus eines KI-Systems abgeschlossen. Dieser Lebenszyklus ist von entscheidender Bedeutung für die spätere Ableitung relevanter Handlungsempfehlungen.

2.1.3 Lebenszyklus eines KI-Systems

In dieser wissenschaftlichen Arbeit wird der Lebenszyklus eines KI-Systems anhand des Modells **CR**oss-**I**ndustry **S**tandard **P**rocess model for the development of **M**achine **L**earning applications with **Q**uality assurance methodology kurz CRISP-ML(Q) beschrieben. Das Modell gliedert die Entwicklung eines KI-Systems in sechs bzw. sieben Phasen. Theoretisch durchläuft jedes KI-System diese Phasen, jedoch nicht zwingend in der vorgegebenen Reihenfolge oder Systematik, da es in diesem Entwicklungsbereich noch an allgemein anerkannten Standards fehlt. CRISP-ML(Q) ist ein speziell für ML-Anwendungen entwickeltes Modell, das branchen- und anwendungsneutral ist. Es wurde primär für die Entwicklung von ML-Anwendungen konzipiert und integriert Methoden der Qualitätssicherung. Ein besonderer Fokus des Modells liegt auf dem Monitoring und der Wartung von ML-Anwendungen in einer sich verändernden Umgebung (Studer et al., 2021). Im Sinne der Vollständigkeit sei an dieser Stelle das Lebenszyklusmodell aus den Ethik-Leitlinien für eine vertrauenswürdige KI erwähnt. Dieses Modell umfasst fünf Phasen (Europäische Kommission, Generaldirektion Kommunikationsnetze, Inhalte und Technologien, 2019). Tabelle 1 stellt beide Modelle gegenüber, und im weiteren Verlauf dieses Absatzes wird CRISP-ML(Q) näher erläutert und verwendet.

Tabelle 1: Modelle eines KI-System-Lebenszyklus

Ethik-Leitlinien für eine vertrauenswürdige KI	CRISP-ML(Q)
Analyse	Geschäfts- und Datenverständnis
Design bzw. Re-Design	Datenvorbereitung
Entwicklung	Modellierung
Nutzung	Evaluierung
Bewertung und Begründung	Bereitstellung
	Wartung- und Monitoringphase

Anmerkung: Eigene Darstellung (Czvetkó et al., 2022; Europäische Kommission, Generaldirektion Kommunikationsnetze, Inhalte und Technologien, 2019; Studer et al., 2021)

Ergänzend zur vorhergehenden Tabelle ist zu erwähnen, dass es sich bei diesen Modellen um zyklische Modelle handelt, bei denen die Phasen je nach Anpassung des Systems mehrfach durchlaufen werden. Dies ist besonders im Kontext eines KI-Systems relevant, da beispielsweise durch das Training mit neuen Daten das Systemverhalten verändert wird, obwohl der Algorithmus selbst unverändert bleibt.

In der ersten Phase **Geschäfts- und Datenverständnis** werden die Ziele, die mit dem Use Case erreicht werden sollen, definiert und in Ziele für das KI-System übersetzt. Gleichzeitig wird die Qualität der verfügbaren Daten geprüft und bewertet, ob das Projekt umsetzbar ist. Bei Unsicherheiten kann ein Proof of Concept (PoC) helfen, die Machbarkeit zu prüfen. In der **Datenvorbereitung** werden die wichtigsten Merkmale und Stichproben ausgewählt. Dabei wird darauf geachtet, die Daten von Fehlern oder Rauschen zu befreien und fehlende Werte zu ergänzen. Neue Merkmale können durch Feature Engineering oder durch die künstliche Erweiterung bestehender Datensätze mittels veränderter Kopien erzeugt werden. Eine Standardisierung der Datenstruktur hilft, Verzerrungen zu vermeiden. Während der **Modellierung** werden passende Algorithmen ausgewählt und KI-Modelle entwickelt, die den Projektanforderungen entsprechen. Oft wird mit einfachen Modellen begonnen, die im Verlauf weiter verbessert werden. Techniken wie Regularization und Cross-Validation helfen, Over Fitting zu vermeiden. Die Reproduzierbarkeit des Systems wird durch genaue Dokumentation und Vergleichstests sichergestellt. In der **Evaluierung** wird die Leistung des KI-Systems mit einem separaten Testdatensatz überprüft. Ziel ist es, sicherzustellen, dass das System robust und nachvollziehbar ist. Die Ergebnisse werden mit den vorher festgelegten Erfolgskriterien verglichen. In der **Bereitstellung** wird das KI-System in der Praxis eingesetzt. Um Fehler zu vermeiden, erfolgt die Einführung idealerweise schrittweise und wird unter realen Bedingungen laufend getestet. Feedback der Nutzer:innen wird genutzt, um das System weiter zu verbessern. Nach der Bereitstellung wird das KI-System in der **Wartungs- und Monitoringphase** regelmäßig überprüft. Da sich Daten und Systemumgebungen verändern können, ist es notwendig, das System kontinuierlich anzupassen. Neue Daten werden gesammelt, und das System wird bei

Bedarf aktualisiert. Automatisierte Prozesse helfen dabei, diesen Ablauf effizient zu gestalten und die Qualität langfristig zu sichern.

Das Modell veranschaulicht den idealen Lebenszyklus eines KI-Systems, der von der Entwicklung über die Bereitstellung bis hin zur Wartung reicht. Anbieter können dieses Modell nutzen, um KI-Systeme systematisch zu entwickeln, während es Betreibern als Orientierungshilfe dient (Studer et al., 2021). Mit der Erklärung des Lebenszyklus eines KI-Systems werden die Grundlagen zum Thema KI abgeschlossen. Im weiteren Verlauf dieses Kapitels werden die grundlegenden Begriffe im Gesundheitswesen definiert. Zunächst wird das Gesundheitswesen als Ganzes beschrieben, und im nächsten Schritt wird das Krankenhaus als zentrale Einrichtung näher erläutert.

2.2 Begriffsdefinitionen Gesundheitswesen

Neben KI erfordert auch das Gesundheitswesen als Ganzes Begriffsdefinitionen, um die später beschriebenen Szenarien besser einordnen zu können. Das Gesundheitssystem ist ein wesentlicher Teil des Sozialsystems (Wiemschulte, 2024). Die Organisation des Gesundheitssystems weist eine starke länderspezifische Ausprägung auf, weshalb die folgenden Beschreibungen auf das österreichische Gesundheitssystem fokussiert sind (Halwachs-Baumann, 2024).

In Österreich ist das Gesundheitssystem öffentlich strukturiert und basiert auf dem Prinzip der solidarischen Finanzierung. Dadurch wird allen Menschen ein gleichberechtigter Zugang zu medizinischer Versorgung ermöglicht. Bund, Länder, Gemeinden, Sozialversicherungsträger sowie gesetzliche Interessensvertretungen teilen sich die Aufgaben der Organisation und Steuerung. Die Finanzierung setzt sich aus Beiträgen zur Sozialversicherung, öffentlichen Steuermitteln und zusätzlichen privaten Zahlungen zusammen. Die Gesundheitsversorgung umfasst ambulante und stationäre Leistungen, Prävention und Gesundheitsförderung. Für diese Arbeit ist im speziellen die Unterteilung in ambulante und stationäre Versorgung relevant. Die ambulante medizinische Versorgung in Österreich erfolgt auf drei Wegen:

- Niedergelassene Allgemeinmediziner:innen, Fachärzte und Fachärztinnen und Zahnärzte und Zahnärztinnen in Praxen, von denen etwa die Hälfte Verträge mit Krankenkassen haben.
- Ambulatorien der Krankenkassen und privat geführte Ambulatorien.
- Spitalsambulanzen in Krankenhäusern.

Zusätzlich werden Heilmittel und Heilbehelfe sowie Dienstleistungen von weiteren Gesundheitsberufen, wie Ergo- und Physiotherapeuten und Ergo- und Physiotherapeutinnen, zur ambulanten Versorgung gezählt. Die stationäre medizinische Versorgung wird von öffentlichen, privat-gemeinnützigen und rein privaten Krankenhäusern bereitgestellt, betrieben von Ländern, Gemeinden, Glaubensgemeinschaften, Sozialversicherungsträgern oder privaten Trägern (Gesundheitsportal, 2024).

Bezugnehmend auf die Einleitung ist es erforderlich, die Einrichtung Krankenhaus näher zu erläutern. Synonym für den Begriff Krankenhaus werden die Begriffe Spital und Krankenanstalten verwendet. Laut Angaben des österreichischen Bundesministeriums für Arbeit, Soziales, Gesundheit und Konsumentenschutz gab es im Jahr 2019 in Österreich 271 Spitäler, verteilt auf 301 Standorte, mit insgesamt rund 64.800 Betten. Von diesen 271 Spitälern waren 45 % öffentlich finanzierte gemeinnützige Akutspitäler (Landesfondsspitäler und Unfallkrankenhäuser) mit etwa 45.600 Betten und fast 90 % aller Krankenhausaufenthalte. Zum damaligen Zeitpunkt waren rund 24.600 Ärzte und Ärztinnen und 93.300 Personen in nichtärztlichen Gesundheitsberufen (ohne Verwaltungspersonal) beschäftigt (Bundesministerium für Arbeit, Soziales, Gesundheit und Konsumentenschutz, 2019).

Krankenhäuser können weiter differenziert werden. In dieser wissenschaftlichen Arbeit erfolgt die Unterteilung gemäß § 2 Abs. 1 des Krankenanstalten- und Kuranstaltengesetzes.

In **allgemeinen Krankenhäusern** werden alle Personen ohne Unterschied des Geschlechts, des Alters oder der Art der ärztlichen Betreuung versorgt. Gemeinnützige Krankenhäuser, die Allgemeinversorgung erbringen, werden diesem Typ zugeordnet. Im Gegensatz werden in **Sonderkrankenhäusern** nur Personen mit bestimmten Krankheiten oder Personen bestimmter Altersstufen oder für bestimmte Zwecke behandelt. Zu dieser Gruppe zählen alle entsprechenden Spitäler nach dieser Definition außer Sanatorien. **Sanatorien** unterscheiden sich durch ihre besondere Ausstattung im Sinne von höheren Ansprüchen hinsichtlich Verpflegung und Unterbringung. In Sanatorien wird entweder Allgemeinversorgung im Akutsektor oder eine Spezialversorgung erbracht. Diese Häuser führen in der Regel die Bezeichnung im Namen. Der letzte Typ der Kranken- und Kuranstalten sind **Pflegeanstalten für chronisch Kranke.** In diesen Einrichtungen werden Personen betreut, die ärztliche Betreuung und besondere Pflege in der nicht-akuten Spezialversorgung bedürfen. Einrichtungen, die nicht dem Krankenanstaltengesetz unterliegen, fallen nicht in diese Kategorie. Beispiele hierfür sind Alten- und Pflegeheime bzw. geriatrische Zentren.

Eine weitere Form der Unterteilung ist nach Eurostat. Hier wird zwischen General Hospitals, Mental Health and Substance Abuse Hospitales und Speciality Hospitals unterschieden. Die Zuordnung erfolgt anhand einer vergleichbaren Logik wie im Krankenanstalten- und Kuranstaltengesetz (Bundesministerium für Arbeit, Soziales, Gesundheit, Pflege und Konsumentenschutz, 2024).

Krankenhäuser sind zentrale Bestandteile des Gesundheitssystems und erfüllen daher spezifische Funktionen. Diese Funktionen lassen sich übergeordnet in vier Kategorien einteilen: Versorgung von Patienten und Patientinnen, Forschung und Lehre, Koordination der Versorgungsaktivitäten und gesellschaftliche Aufgaben. Eine nähre Ausführung findet sich in Tabelle 2 (Busse et al., 2022).

Tabelle 2: Funktionen eines Krankenhauses

Versorgung von Patienten und Patientinnen
Stationäre und teilstationäre Betreuung ergänzt durch vor- sowie nachgelagerte Behandlungsangebote
Notfallversorgung in enger organisatorischer Abstimmung mit dem Rettungsdienst
Rehabilitative Maßnahmen in Kooperation mit entsprechenden Einrichtungen sowie gezielte Überleitung in solche
Übergang in die ambulante oder stationäre Pflegeversorgung
Hospizdienste als Ergänzung bei unzureichender Versorgung Schwerstkranker außerhalb von Kliniken
Ambulante ärztliche Versorgung im Rahmen strukturierter Behandlungsprogramme (z. B. DMP)
Erbringung spezialisierter ambulanter Leistungen, u. a. bei seltenen oder komplexen Krankheitsbildern
Forschung und Lehre
Aus-, Weiter- und Fortbildungsangebote, insbesondere für medizinisches Fachpersonal und Pflegekräfte
Klinische Forschung sowie Transfer und Anwendung neuer wissenschaftlicher Erkenntnisse
Koordination der Versorgungsaktivitäten
Abstimmung zwischen ambulanter, stationärer und rehabilitativer Versorgung zur Sicherstellung durchgängiger Behandlungsprozesse
Strukturierte Zusammenarbeit mit anderen Gesundheitsdienstleistern im Sinne einer gezielten Auftragsvergabe und Kooperationskultur
Gesellschaftliche Aufgaben
Beitrag zur staatlich gesicherten Gesundheitsversorgung als Teil des sozialen Sicherungssystems
Stärkung der regionalen Standortattraktivität für Städte und Kommunen
Rolle als Arbeitgeber sowie als wirtschaftlicher Partner für Zulieferunternehmen

Die Erfüllung dieser Funktionen wird durch das Zusammenspiel der Aufbau- und Ablauforganisation sichergestellt.

2.2.1 Aufbauorganisation

Die Aufbauorganisation beschreibt das statische System einer Einrichtung, in dem organisatorische Einheiten und deren Zuständigkeiten für die arbeitsteilige Erfüllung der Unternehmensaufgaben festgelegt sind. Dabei werden Aufgaben entweder nach Funktion oder nach Leistungseinheit strukturiert, um eine Spezialisierung zu ermöglichen. Zudem regelt die Aufbauorganisation den Rang innerhalb der Aufgabenverteilung und koordiniert die einzelnen Handlungsbeziehungen. Diese Struktur manifestiert sich sichtbar in einem Organigramm und bildet das Fundament für institutionelle Fragestellungen. In Krankenhäusern wird typischerweise eine funktionale Gliederung angewendet, bei der ähnliche Tätigkeiten auf den nachfolgenden Hierarchieebenen in spezifischen Abteilungen zusammengefasst werden (Hensen, 2022; Schreyögg et al., 2022).

Das LKH-Univ. Klinikum Graz, als Teil der Steiermärkischen Krankenanstaltengesellschaft m.b.H., bietet ein praktisches Beispiel für eine strukturierte Aufbauorganisation. Das Direktorium setzt sich aus der Betriebsdirektion, der Pflegedirektion und der ärztlichen Direktion zusammen. Die Betriebsdirektion verantwortet die wirtschaftlichen, administrativen und technischen Angelegenheiten und koordiniert diese zwischen den Organisationseinheiten des Klinikums. Die Pflegedirektion ist zuständig für die Organisation und Überwachung aller Aufgaben in der Gesundheits- und Krankenpflege. Die ärztliche Direktion steuert die medizinischen Aufgaben, erteilt Weisungen im Rahmen ihres Wirkungsbereichs und gewährleistet eine einwandfreie medizinische Versorgung. Dem Direktorium untergeordnet sind Stabstellen, das Bereichsmanagement und Abteilungen. Beispiele für Stabstellen sind die Medizintechnik und Informationstechnologie. Ebenso trägt die Abteilung Qualitäts- und Risikomanagement zur Sicherheit und Qualität der Behandlungsprozesse bei, indem es qualitätssichernde Maßnahmen begleitet und deren Umsetzung unterstützt. Das Bereichsmanagement übernimmt die Steuerung der Kliniken gemeinsam mit den medizinisch und pflegerisch Verantwortlichen. Es berät Führungspersonen, optimiert den Ressourceneinsatz und erstellt Leistungs- sowie Investitionsbudgets. Zudem entlastet es die Anstaltsleitung, indem es auf Klinikebene die wirtschaftliche Geschäftsführung wahrnimmt. Die Anstaltsleitung umfasst einen Klinikvorstand aus ärztlicher Sicht und eine Pflegedienstleitung. Die Abteilungen bilden die letzte Organisationseinheit. Ein Beispiel ist die Abteilung Wirtschaft/Logistik, die für die Betriebslogistik sowie die Beschaffung von Materialien zuständig ist. Diese Struktur orientiert sich am Aufbau des Direktoriums (Steiermärkische Krankenanstaltengesellschaft m.b.H, 2025a, 2025b). Wird diese Organisationsstruktur als Grundlage für die Zielgruppendefinition dieser wissenschaftlichen Arbeit herangezogen, zeigt sich, dass der Begriff der Entscheidungsträger:innen eine breite Spannweite aufweist. Aus diesem Grund wurden die Grundlagen zu KI sowie deren Anwendungsfällen umfassender dargestellt, um nicht nur technisches Fachpersonal, sondern auch Pflegedienstleitungen und leitendes ärztliches Personal gezielt anzusprechen.

2.2.2 Ablauforganisation

Die Ablauforganisation befasst sich mit der strukturierten Gestaltung und Steuerung von Tätigkeitsabfolgen unter Berücksichtigung zeitlicher, räumlicher, personeller und sachlicher Ressourcen. Ziel ist es, komplexe Prozesse durch Standardisierung beherrschbar zu machen, die Effizienz zu steigern, Wartezeiten und Kosten zu senken sowie Durchlaufzeiten zu verkürzen und gleichzeitig die Qualität der Leistungserbringung zu erhöhen. Im Krankenhauskontext bedeutet dies eine Abkehr von der klassischen Abteilungsstruktur zugunsten einer prozessorientierten Ausrichtung entlang des Behandlungsverlaufs. Dieser wird interdisziplinär gestaltet und stellt die behandelte Person in den Mittelpunkt (Busse et al., 2022; Hensen, 2022).

Auch Behar et al. (2022) rücken den klinischen Leistungsprozess in den Fokus, indem sie ihn als zentrales Element des Krankenhausbetriebs bezeichnen. Alle weiteren Abläufe und Aktivitäten sind diesem Kernprozess untergeordnet und dienen seiner Unterstützung. Durch eine gut organisierte Diagnostik und Therapie soll ein möglichst zügiger und komplikationsarmer Heilungsverlauf erreicht werden. Daraus ergeben sich unter anderem eine nachhaltige Verbesserung der medizinischen Versorgungsqualität, eine effizientere Nutzung von Personal- und Sachressourcen sowie eine höhere Zufriedenheit der Patienten und Patientinnen.

Auf Basis dieser organisatorischen Grundlagen lassen sich die im folgenden Abschnitt dargestellten Anwendungsfälle für den Einsatz von KI im klinischen Umfeld gezielter einordnen.

2.3 Anwendungen Künstlicher Intelligenz im Krankenhaus

Am Beginn dieses Absatzes erscheint es sinnvoll, beschreibende Merkmale für entsprechende Anwendungen unabhängig vom EU AI Act einzuführen. Eine erste allgemeine Unterscheidung ist die Einteilung in patientennahe und -ferne Anwendungen. Der Fokus dieser Unterscheidung liegt nicht ausschließlich auf KI-Systemen, sondern allgemein auf digitalen Lösungen. Unter patientennahen Anwendungen fallen beispielsweise Lösungen für das Patientendatenmanagement, einschließlich digitaler Fieberkurven, Labore, Bildarchiv- und Kommunikationssysteme sowie Entscheidungsunterstützungssysteme. Zusammengefasst stehen solche Anwendungen im direkten Zusammenhang mit der zu behandelnde Person und es sind deshalb weitreichendere technische und organisatorische Maßnahmen zu berücksichtigen. Im Gegensatz dazu stehen patientenferne Anwendungen, die ebenfalls im klinischen Alltag präsent und notwendig sind, jedoch in keinem direkten Zusammenhang mit Patienten und Patientinnen stehen. Beispiele hierfür sind Anwendungen, die innerbetriebliche Prozesse betreffen, wie das Erstellen von Dienst- und Urlaubsplänen, die digitale Erfassung von Bewerbungen, Patientenportale und Aufnahme-Apps. Ein weiteres Einsatzgebiet von KI ist die Kodierung und Rechnungslegung oder die Übersetzung von Texten in andere Sprachen (Blase et al., 2023). Diese Unterscheidung dient als erster Indikator für das Risiko einer KI-Anwendung im Sinne des EU AI Acts. Sie ermöglicht jedoch keine eindeutigen Schlussfolgerungen, weshalb jede Anwendung individuell bewertet werden muss. Bei der Auswahl, der in den folgenden Absätzen beschriebenen Beispiele, wurde versucht, patientennahe Anwendungen zu

berücksichtigen und die spezifischen Risiken von KI im Gesundheitswesen aufzuzeigen. Bei der Beschreibung der Anwendungsfälle wurde nicht auf länderspezifische Unterschiede, wie beispielsweise die Verantwortungsteilung zwischen Pflegepersonal und Ärzteschaft, eingegangen.

2.3.1 Künstliche Intelligenz in der Pflege

Die Einsatzszenarien von KI in der Pflege sind vielfältig und werden im Verlauf dieses Absatzes näher erläutert. Insgesamt ist eine Zunahme der Forschung und Entwicklung von KI-Anwendungen für die Pflege erkennbar. Ein Beispiel hierfür ist die Pflegedokumentation. Domhoff et al. (2024) führten eine Rapid-Review-Studie durch, um einen Überblick über vielversprechende KI-Anwendungen in der Pflege zu gewinnen. Analysiert wurden 292 relevante englischsprachige Publikationen aus verschiedenen wissenschaftlichen Datenbanken bis Juni 2020. Der Fokus lag auf Originalstudien, Übersichtsarbeiten und Beiträgen zu ethischen, rechtlichen und sozialen Aspekten.

Krankenhäuser stellten dabei den häufigsten Anwendungsort dar. Maschinelle Lernverfahren kamen insbesondere in der Bild- und Signalverarbeitung zur Überwachung von Aktivitäten und Gesundheitszuständen zum Einsatz. Weitere Anwendungen betrafen Pflegekoordination, Kommunikation und Sturzerkennung.

Trotz des identifizierten Potenzials mangelt es an Evidenz zur tatsächlichen Wirksamkeit in der Praxis. Ethische und rechtliche Fragen wurden nur am Rande behandelt. Vor diesem Hintergrund ist die folgende Zusammenfassung kritisch zu betrachten. Dennoch wird der Überblick von Wolf-Ostermann et al. (2021) insgesamt als sinnvoll erachtet. In dieser Publikation wurde anhand einer systematischen Literaturrecherche eine tabellarische Übersicht über verschiedene Szenarien erstellt (siehe Tabelle 3).

Tabelle 3: KI in der Pflege

Setting	Anwendungskontext des maschinellen Lernens
Stationäre Langzeitpflege	Vermeidung von Stürzen Unterstützung bei der Kontrolle der Blasen- und Darmfunktion Förderung individueller, bewohnerorientierter Pflege Beobachtung von Aktivität, Gesundheitszustand und Wohlbefinden Lokalisierung und Orientierung von Bewohner:innen Analyse von Bewegungsmustern Organisation und Unterstützung im Notfall Nutzung intelligenter Mobilitätshilfen Einschätzung individueller Versorgungsbedarfe Planung und Dokumentation pflegerischer Abläufe

Setting	Anwendungskontext des maschinellen Lernens
	Schutz des Pflegepersonals vor physischen Belastungen
Ambulante Pflege	Prognose des Erholungspotenzials Umgang mit Belastungssituationen bei demenziell Erkrankten Überwachung von Gesundheitsparametern im Alltag
Akutstationäre Pflege/ Intensivpflege	Vorbeugung von Druckgeschwüren Sturzrisikoreduktion Bedarfsgerechte Versorgungserfassung Prävention und Kontrolle von Harnwegsinfektionen Steuerung der Nahrungsaufnahme und -versorgung Erkennung und Linderung von Schmerz Überwachung und Management von Beatmungsprozessen Optimierung der Alarmbearbeitung Unterstützung bei der Blasen- und Darmregulation Einsatz zur besseren Personaleinsatzplanung Verbesserung der Abstimmung und Kommunikation im Team Automatisierte Analyse freier Texte in Pflegedokumentationen Förderung der physischen und psychischen Gesundheit von Pflegekräften Analyse des Einflusses von Teammerkmalen auf klinische Ergebnisse

Mit Verweis auf genannten Studien lässt sich die Dokumentation als eine zentrale Aufgabe in der Pflege hervorheben (Domhoff et al., 2024; Wolf-Ostermann et al., 2021). Korrespondierend dazu umfasst der nächste Absatz zwei Anwendungsfälle, die den Einsatz von ChatGPT bzw. allgemein Large Language Models (LLMs) in unterschiedlichen Szenarien beschreiben und in Studien untersucht wurden.

2.3.2 ChatGPT im Krankenhaus

In einer Studie von Dos Santos et al. (2024) wurde das Ziel verfolgt, die Wirksamkeit eines validierten Rahmens zur Erstellung von ChatGPT-Prompt, die gültige Pflegepläne für einen hypothetischen älteren Patienten und Patientinnen mit Lungenkrebs generieren, zu demonstrieren. Die Untersuchung wurde anhand eines Fallbeispiels mit einem Lungenkrebspatienten durchgeführt. Nach Eingabe einer pflegerischen Beurteilung des Patienten in ChatGPT wurden Pflegeplanvorschläge generiert. Diese Pflegepläne wurden anschließend hinsichtlich ihrer Qualität bewertet. Die Ergebnisse der Studie zeigen, dass der von ChatGPT generierte Pflegeplan in Bezug auf Umfang und Art weitgehend dem Standard entsprach. Besonders positiv wurde hervorgehoben, dass ChatGPT bestimmte Bedürfnisse (Sauerstoffversorgung und Beatmung) korrekt priorisierte. Die Schlussfolgerung lautet, dass die Verwendung einer validierten Rahmen-Prompts zur Erstellung von Pflegeplänen durch ChatGPT vielversprechend ist. Es zeigt das Potenzial von ChatGPT als Entscheidungshilfetool, um die Pflegedokumentation zu optimieren (Dos Santos et al., 2024).

Eine weitere Studie bietet interessante Einblicke in die Anwendung von KI im Bereich der Krankenhausdokumentation. Es handelt sich per Definition nicht um eine Anwendung in der Pflege, jedoch ähnelt sich die technische Umsetzung und wird deshalb an dieser Stelle angeführt. In dieser Studie wurde der Einsatz von ChatGPT bei der Erstellung von Entlassungsberichten untersucht. Diese Berichte sind entscheidend für die kontinuierliche Versorgung von Patienten und Patientinnen. Ziel der Machbarkeitsstudie war es, das Potenzial von KI zur Erstellung von qualitativ hochwertigen Entlassungsberichten zu bewerten. Dazu wurden 25 simulierte Patientenvignetten erstellt, basierend auf denen sowohl ChatGPT als auch Assistenzärzte und Assistenzärztinnen Entlassungsberichte verfassten. Die Qualität der Berichte wurde von unabhängigen Hausärzten und Hausärztinnen und anhand eines Mindestdatensatzes bewertet. Die Ergebnisse der Studie zeigen, dass 100 % der KI-generierten Entlassungsberichte von Hausärzten und Hausärztinnen als qualitativ akzeptabel eingestuft wurden, im Vergleich zu 92 % der von Assistenzärzten und Assistenzärztinnen erstellten Berichte. Beide Gruppen erreichten eine durchschnittliche Übereinstimmung von 97 % mit dem vorgegebenen Mindestdatensatz. Darüber hinaus hatten die Hausärzte und Hausärztinnen Schwierigkeiten, den Unterschied zwischen von KI und von Menschen erstellten Berichten zu erkennen, mit einer Trefferquote von nur 60 %. Auch ein KI-Erkennungstool bewertete alle Berichte als sehr unwahrscheinlich von KI erstellt. Die Studie kommt zu dem Schluss, dass KI in der Lage ist, Entlassungsberichte in vergleichbarer Qualität zu denen von Assistenzärzten und Assistenzärztinnen zu erstellen. Größere Studien mit realen Daten von erkrankten Personen und zugelassenen KI-Tools sind jedoch erforderlich, um das Potenzial weiter zu untersuchen (Clough et al., 2024).

Der nächste Anwendungsbereich für den Einsatz von KI liegt in der Radiologie, insbesondere zur Unterstützung bei der Diagnosestellung. Ziel dieses Anwendungsfeldes ist die Verbesserung der diagnostischen Qualität sowie die Steigerung der Effizienz, unter anderem durch eine Reduktion des Ressourcenaufwands in verschiedenen Bereichen der Gesundheitsversorgung (Auer et al., 2019).

2.3.3 Künstliche Intelligenz in der Radiologie

In vorhergegangenen Absätzen wurden bereits beschrieben, dass KI-Systeme je nach Anwendungszweck unterschiedliche Datenstrukturen als Input verarbeiten können. Dies umfasst auch CT- oder MRT-Bildgebung. Die visuelle Begutachtung dieser Bilder ist generell ressourcenintensiv. KI-Systeme unterstützen, indem sie die Informationen aus den Bildern und anderen Daten filtern, diese zusammenführen, einen Kontext herstellen und relevante Muster aufzeigen (Auer et al., 2019). Lakhani und Sundaram (2017) haben beispielsweise bereits im Jahr 2017 Modelle zur Erkennung von pulmonaler Tuberkulose basierend auf künstlichen neuronalen Netzwerken entwickelt. Als Input wurden in diesem Anwendungsfall 1.007 Röntgenaufnahmen der Brust verwendet. Das Datenset wurde in Trainings- (68,0 %), Validierungs- (17,1 %) und Testdaten (14,9 %) aufgeteilt. Das Ergebnis zeigt, dass künstliche neuronale Netzwerke pulmonaler Tuberkulose auf Thoraxröntgenbildern mit einer Area under the curve (AUC) von 0,99 klassifizieren. Ein menschenunterstützter Ansatz für Fälle, in denen zwischen Modellen Uneinigkeit bestand, verbesserte die Genauigkeit weiter. Ein weiteres Beispiel aus dem Jahr 2017 wird von Rajpurkar et al. (2017) beschrieben. Hier wurde ein konvolutionales neuronales Netzwerk (engl. Convolutional Neural Network), bestehend aus 121 Schichten und trainiert mit 100.000 Lungenaufnahmen, verwendet, um Lungenembolien zu erkennen. Die Entscheidungsfindung des Modells war für Menschen nicht mehr nachvollziehbar, jedoch übertragen die Ergebnisse der maschinellen Klassifizierung die der Expertenklassifizierung.

Im Hinblick auf den Fokus dieser wissenschaftlichen Arbeit ist festzuhalten, dass der Einsatz von KI in der Radiologie bedeutsame rechtliche Fragestellungen in Bezug auf die Haftung bei Fehlern aufwirft. In diesem Zusammenhang wird die Art der Implementierung von entscheidender Bedeutung sein, sei es als eigenständiges Diagnosewerkzeug oder als unterstützendes Instrument für Radiologen und Radiologinnen. In Bezug, auf den im nächsten Kapitel detailliert erläuterten EU AI Act kann bereits an dieser Stelle darauf hingewiesen werden, dass KI-Systeme in der Radiologie als sogenannte Hochrisiko-Anwendungen eingestuft werden (Martín-Noguerol et al., 2024).

Ein weiterer Anwendungsfall für KI findet sich mit dem Triageprozess, der als entscheidender erster Schritt im Behandlungsablauf gilt und eine effektive sowie effiziente Gestaltung erfordert.

2.3.4 Intelligente Triagesysteme

Allgemein beschreibt eine Triage den Prozess der Festlegung der Behandlungspriorität von Patienten und Patientinnen basierend auf ihrem Gesundheitszustand. Durch den Einsatz von Triage wird es möglich, knappe Ressourcen im Behandlungsprozess effizient zu nutzen, da eine gleichzeitige Behandlung aller Patienten und Patientinnen in der Praxis meist nicht möglich ist. Historisch betrachtet beschreibt der Begriff Triage in erster Linie den Prozess zur Festlegung der Reihenfolge und Priorität der Notfallbehandlung, des Notfalltransports oder der Bestimmung des Transportziels. Heutzutage findet der Triageprozess Anwendung in klinischen Notaufnahmen (Madigan). Die Triage in Notaufnahmen wird vom medizinischem Fachpersonal meist manuell mit standardisierten Triage-Richtlinien und Scoring-Systemen durchgeführt. Dieser Prozess bindet knappe Ressourcen und weist einen bestimmten repetitiven Faktor auf, auch wenn die Krankheitsbilder variieren mögen (Behar et al., 2022). Mithilfe von KI-gestützten Triagesystemen kann der Zustand der zu triagierenden Person automatisch bewertet werden und die Person dementsprechend an die richtige versorgende Stelle weitergeleitet werden. Die vorhandenen Systeme basieren dabei auf unterschiedlichen Methoden wie Expertenregeln, Bayes'sche Inferenz (ML) und Deep Learning (Semigran et al., 2015). Im Detail errechnet das KI-System den aktuellen gesundheitlichen Zustand der Patienten und Patientinnen anhand individueller demographischer Daten und den vorliegenden Symptomen. Basierend auf dem aktuellen gesundheitlichen Zustand wird automatisch vorgeschlagen, wo und bis wann medizinische Versorgung in Anspruch genommen werden sollte. Eine weitere erforschte Variante ist die Vorhersage von Krankheiten basierend auf vergangenen Aufnahmeprotokollen. Diese Variante erreichte in der vorliegenden Untersuchung eine Genauigkeit von 96,1 % (Auer et al., 2019; Li et al., 2017; Priyanka, 2016).

2.3.5 Spezifische Risiken der Künstlichen Intelligenz im Gesundheitswesen

Abschließend lässt sich festhalten, dass im klinischen Alltag zahlreiche Entscheidungen mit unmittelbarem Einfluss auf die Gesundheit von Patienten und Patientinnen getroffen werden. Aktuelle Forschungsergebnisse zeigen, dass der Einsatz von KI-Systemen zur Entscheidungsunterstützung sinnvoll sein kann, jedoch nicht ohne menschliche Mitwirkung erfolgen sollte. Insbesondere bei gesundheitsrelevanten Entscheidungen ist eine ausschließliche Automatisierung kritisch zu bewerten. Zudem müssen sowohl die Bedürfnisse der zu behandelnden Personen als auch jene des medizinischen Fachpersonals in die Bewertung und Gestaltung solcher Systeme einfließen. Wahl und Wilhelm (2024) haben hierzu sieben Punkte formuliert, die zu berücksichtigen sind (siehe Tabelle 4).

Tabelle 4: Anforderungen an KI-Systeme aus Sicht des Gesundheitswesens

Kategorie	Beschreibung
Datenschutz und Vertraulichkeit	Gesundheitsdaten sind hochsensibel und müssen gemäß den Datenschutzbestimmungen besonders geschützt werden. Die Daten dürfen nur für legitime Zwecke verwendet und die Betroffenen müssen über den Einsatz der Technologie informiert werden sowie ihre Einwilligung geben.
Transparenz und Erklärbarkeit	Die Entscheidungen, die von KI-Systemen getroffen werden, müssen für alle Beteiligten nachvollziehbar sein. Die Algorithmen sollten so gestaltet sein, dass erklärt werden kann, wie die Algorithmen zu einer bestimmten Schlussfolgerung gekommen sind.
Menschliche Interaktion	KI sollte die menschliche Interaktion in der Pflege ergänzen und nicht ersetzen. Pflegebedürftige müssen weiterhin Zugang zu menschlicher Betreuung und emotionaler Unterstützung haben.
Verantwortlichkeit	Pflegekräfte sollen in der Lage sein, Empfehlungen der KI kritisch zu überprüfen und die finale Entscheidung zu treffen.
Ethik und Gerechtigkeit	KI-Systeme haben ethische und rechtliche Standards einzuhalten.
Schulung und Ausbildung	Pflegekräfte sollten angemessen geschult und ausgebildet werden, um den Einsatz von KI zu verstehen und sicherzustellen, dass KI korrekt und ethisch eingesetzt wird.
Bewertung der Wirksamkeit	Die Auswirkungen von KI müssen regelmäßig in Hinblick auf die Bedürfnisse pflegebedürftiger Personen als auch den Anforderungen des Pflegepersonals und den Beitrag zur Qualitätsverbesserung bewertet werden.

Auf Grundlage dieser Auflistung können die spezifischen Risiken des Einsatzes von KI im Gesundheitswesen, insbesondere im klinischen Kontext, analysiert und herausgearbeitet werden.

Im Kontext von KI-Anwendungen im Gesundheitswesen, und damit auch für Krankenhäuser, lassen sich zumindest drei übergeordnete Risikokategorien identifizieren: fehlerhafte Diagnostik und Therapieentscheidungen, Verzerrungen (Biases) sowie Datenschutz und -missbrauch.

Das erste Risiko bezieht sich auf die potenziellen Fehler in den Entscheidungen von KI-Systemen, die die Gesundheit und das Leben von Patienten und Patientinnen unmittelbar negativ beeinflussen können. Die zweite Risikokategorie bezieht sich auf Verzerrungen (Biases), die beispielsweise zu Diskriminierung führen können. Dies hat zur Folge, dass bestimmte demografische Gruppen, wie Minderheiten oder Personen mit seltenen Erkrankungen, schlechter diagnostiziert oder behandelt werden. Ein konkretes Beispiel ist die fehlerhafte Diagnosestellung

von Hautkrankheiten bei dunkelhäutigen Personen, falls der Trainingsdatensatz überwiegend Daten von hellhäutigen Personen enthielt. Die dritte Risikokategorie umfasst den Datenschutz und die Gefahr des Datenmissbrauchs. In diesem Zusammenhang gilt der Grundsatz „Trust is key!“, da es nur wenige gesellschaftliche Bereiche gibt, in denen die Sensibilität für datenschutzrechtliche Risiken, Ängste und teilweise irrational erscheinende Bedenken so stark ausgeprägt sind. Dies führt häufig zu der Sorge, unbeabsichtigt strenge Datenschutzvorgaben zu verletzen. Gleichzeitig ist es wichtig, auf die Risiken einer übermäßigen Erfüllung der Datenschutzanforderungen hinzuweisen, da dies die Entwicklung neuer Technologien, insbesondere im europäischen Raum, hemmen kann (Wendland, 2024). In vergleichbarer Form werden diese Risiken von García-Gómez et al. (2024) ebenfalls beschrieben. Zusätzlich wird auf die fehlende Transparenz und die Regelung der Verantwortlichkeit verwiesen.

Abschließend zu diesem Kapitel und bezugnehmend auf die Anwendungsdomäne dieser wissenschaftlichen Arbeit ist darauf hinzuweisen, dass das klinische Setting im Fokus der Forschung und Entwicklung von KI im Gesundheitsweisen liegt. Bereiche wie die stationäre und ambulante Langzeitpflege, die Pflegeausbildung oder gar andere Anwendungskontexte wie Prävention von Pflegebedürftigkeit oder Rehabilitation sind bislang seltener Gegenstand von Forschung und Entwicklung. In Summe fehlen nach Wolf-Ostermann et al. (2021) Erkenntnisse in unter anderem folgenden Bereichen:

- Einsatz und Effekten von KI-Systemen unter Realbedingungen
- Vergleich von KI-Lösungen zu herkömmlichen oder alternativen Lösungen
- Perspektiven und Erfahrungen von Pflegebedürftigen, Pflegefachpersonen oder informellen Pflegepersonen
- Möglichkeiten der Generierung und Nutzung von qualitativ hochwertigen Daten aus Pflegeeinrichtungen

Dieser Auszug soll nochmals verdeutlichen, dass rund um das Thema KI im Gesundheitswesen noch Forschungsbedarf besteht, jedoch das klinische Setting aufgrund des gesetzten Fokus sich als Anwendungsdomäne für die Analyse des EU AI ACT anbietet.

3 GRUNDLAGEN DES EU AI ACT

Der EU AI ACT regelt übergreifend den Lebenszyklus von KI-Systemen. Der Ansatz orientiert sich an einer Bewertung der Risiken, die von KI-Systemen ausgehen können, wobei diese je nach potenzieller Auswirkung auf den Menschen in unterschiedliche Risikostufen eingeteilt werden. Besonderes Augenmerk gilt dabei Anwendungen, die als hochriskant eingestuft werden. Ergänzend werden grundlegende Vorgaben definiert, um den verantwortungsvollen Einsatz von Künstlicher Intelligenz sicherzustellen (Barenkamp, 2024).

Am Beginn dieses Kapitels werden die wichtigsten und praxisrelevanten Punkte des Gesetzes zusammengefasst und in weiterer Folge näher erläutert. Es werden dabei folgende Fragen einleitend beantwortet:

- Wen wird das Gesetz betreffen?
- Was sind die Hauptmerkmale des Gesetzes?
- Wann tritt das Gesetz in Kraft?
- Welche Maßnahmen sollten Unternehmen und andere Organisationen von Anfang an ergreifen?

Das Gesetz schützt alle Menschen in der EU und umfasst deshalb alle KI-Systeme, die Menschen in der EU betreffen. Es ist dabei nicht relevant, ob diese KI-Systeme innerhalb der EU entwickelt und betrieben werden oder aus nicht EU-Ländern stammen. Weiters werden alle Akteure und Akteurinnen im Lebenszyklus eines KI-Systems in die Pflicht genommen und mit Verpflichtungen versehen. Wichtig an dieser Stelle ist auch darauf hinzuweisen, dass in bestimmen Fällen auch KI-Systeme vom Gesetz betroffen sind, die vor Inkrafttreten des Gesetzes auf den Markt gebracht wurden.

Inhaltich basiert das Gesetz, wie bereits einleitend erwähnt, auf einem risikobasierten Ansatz mit dem Fokus auf Anwendungsfälle. Die zu setzenden Maßnahmen und geltenden Verpflichtungen sind im jeweiligen Anwendungsfall vom Risikolevel abhängig. Die zugrundeliegende Technologie spielt in der Regel keine Rolle. Das Risikolevel des jeweiligen KI-Systems ergibt sich anhand einer Einschätzung mittels eines definierten Rahmenwerks. Folglich müssen alle KI-Systeme inventarisiert und bewertet werden, um ihre Risikoeinstufung und die damit verbundenen Verantwortlichkeiten zu bestimmen. Das Rahmenwerk spricht dabei übergeordnet von folgenden Risikolevel: Verbotene Systeme, Hochrisiko-KI-Systeme und KI-Systeme mit geringem Risiko (Visengeriyeva, 2025).

Verbotene KI-Systeme stellen ein untragbares Risiko für die Sicherheit und Grundrechte der Menschen dar und werden in der EU deshalb verboten. Im Gegensatz sind **Hochrisiko-KI-Systeme** am stärksten von den kommenden Änderungen betroffen, da sie den Großteil der

Maßnahmen erfordern. **KI-Systeme mit geringem Risiko** erfordern im ersten Schritt eine anfängliche Risikobewertung und die Erfüllung bestimmter Anforderungen zur Sicherstellung der geforderten Transparenz. Eine weitere Kategorie sind allgemeine KI-Modelle (engl. General Purpose AI Models). Diese Modelle und die darauf aufbauenden Systeme werden durch einen separaten gestuften Ansatz reguliert, mit zusätzlichen Verpflichtungen für Modelle, die als „systemisches Risiko“ eingestuft werden. Sie unterliegen ähnlichen Anforderungen, wie Hochrisiko-KI-Systeme. Das Gesetz trat am 1. August 2024 in Kraft, jedoch erstreckt sich die konkrete Umsetzung aller Teilbereiche planmäßig bis zum Jahr 2030. Im ersten Schritt und direkt nach Verabschiedung wird mit der Errichtung des Europäischen Amts für künstliche Intelligenz begonnen. Zusätzlich sind die Mitgliedstaaten mit der Schaffung von KI-Regulierungs-Sandboxes beauftragt. Diese sollen es insbesondere kleinen und mittleren Unternehmen ermöglichen, in einer sicheren Umgebung zu innovieren, experimentieren, testen und die Konformität ihrer KI-Systeme mit dem KI-Gesetz zu validieren. Sechs Monate nach Inkrafttreten werden die Verbote, die mit dem EU AI ACT verbunden sind, durchgesetzt. Regulierung zu allgemeinen KI-Modellen werden zwölf Monate nach Inkrafttreten des ursprünglichen Gesetzes schlagend. Allgemeine KI-Modelle, die vor diesem Datum auf den Markt gebracht wurden, haben zusätzliche 24 Monate Zeit. In Summe treten die meisten Verpflichtungen 24 Monate nach Inkrafttreten des Gesetzes (2. August 2026) in Kraft. Im Hinblick auf Hochrisiko-KI-Systeme sind weitere zwei Zeithorizonte zu berücksichtigen. Erstens sind KI-Systeme, die als hochriskant eingestuft werden, weil sie eine Sicherheitskomponente eines Systems darstellen, das der Harmonisierungsrechtsvorschrift der EU unterliegt, erst 36 Monate nach Inkrafttreten des Gesetzes davon betroffen. Zweitens sind Hochrisiko-KI-Systeme, die für die Verwendung durch öffentliche Behörden bestimmt sind und vor Inkrafttreten des KI-Gesetzes auf dem Markt waren, erst ab dem 31. Dezember 2030 vom Gesetz betroffen. Bezugnehmend auf diese zeitliche Einordung des Inkrafttretens des EU AI ACT werden folgend Sofortmaßnahmen für Unternehmen und andere Organisationen beschrieben, die direkt nach Inkrafttreten abgehandelt werden sollten. Zuerst sind alle KI-Systeme zu inventarisieren. Es wird dabei nicht unterschieden, ob es sich um eine Eigenentwicklung handelt oder ob dieses System nur eingesetzt wird. Darüber hinaus sollte auch der zukünftige Einsatz und geplante Entwicklungen in die Betrachtung miteinbezogen werden. Im Anschluss erfolgt eine Prüfung, ob diese Systeme vom Anwendungsbereich des EU AI ACT betroffen sind. Hierfür ist eine Bewertung und Kategorisierung nach dem zuvor und im Absatz 3.2 detaillierter beschriebenen Risikolevel notwendig. An dieser Stelle ist zu erwähnen, dass ein Fokus auf potenziell verbotene KI-Systemen zu legen ist, da diese bereits mit Februar 2025 außer Betrieb zu nehmen sind. Neben diesen Sofortmaßnahmen sollten allgemein die Position der Organisation in relevanten KI-Wertschöpfungsketten, die damit verbundenen Compliance-Verpflichtungen und die Erfüllung dieser Verpflichtungen verstanden werden. Es ist notwendig, dass Compliance in alle Funktionen integriert wird, die entlang der Wertschöpfungskette für die KI-Systeme verantwortlich sind (EY, 2024; KPMG, 2024).

Wie bereits erwähnt, legt der EU AI Act klare Regelungen zu den Verantwortlichkeiten fest. Dabei finden die im folgenden Absatz beschriebenen Rollen Anwendung.

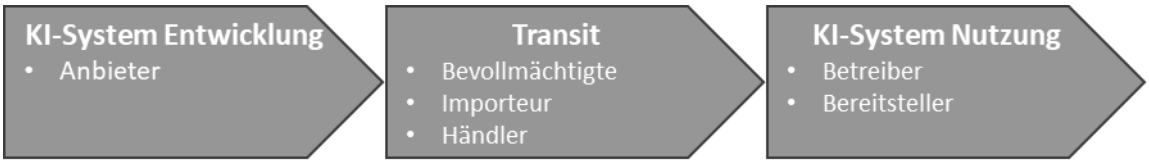
3.1 Rollen und Zuständigkeiten

Im Rahmen des EU AI Act werden verschiedene Rollen und deren Verantwortlichkeiten definiert. Im weiterem Verlauf dieser wissenschaftlichen Arbeit wird auf eine geschlechtergerechte Sprache für die Rollen aus dem EU AI Act verzichtet, da die verwendeten Begriffe aus dem Englischen stammen und in ihrer ursprünglichen Form beibehalten werden.

Ein **Anbieter (engl. provider)** ist eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System oder ein KI-Modell für allgemeine Zwecke entwickelt oder entwickeln lässt und es unter ihrem eigenen Namen oder ihrer eigenen Marke in Verkehr bringt bzw. in Betrieb nimmt. Es ist dabei nicht relevant, ob dies entgeltlich oder unentgeltlich geschieht. Der **Bereitsteller (engl. deployer)** hingegen ist eine natürliche oder juristische Person, Behörde, Einrichtung oder sonstige Stelle, die ein KI-System unter ihrer Aufsicht einsetzt, es sei denn, das KI-System wird im Rahmen einer persönlichen, nicht beruflichen Tätigkeit verwendet. Weiterhin wird der **Bevollmächtigte (engl. authorised representative)** als eine natürliche oder juristische Person definiert, die in der EU ansässig oder niedergelassen ist und von einem eines KI-Systems oder eines KI-Modells für allgemeine Zwecke ein schriftliches Mandat erhalten hat, um in dessen Namen die in dieser Verordnung festgelegten Verpflichtungen und Verfahren zu erfüllen. Der **Importeur (engl. importer)** ist eine in der EU ansässige oder niedergelassene natürliche oder juristische Person, die ein KI-System in Verkehr bringt, welches den Namen oder die Marke einer in einem Drittland ansässigen natürlichen oder juristischen Person trägt. Ein **Händler (engl. distributor)** ist wiederum eine natürliche oder juristische Person in der Lieferkette, die ein KI-System auf dem Unionsmarkt bereitstellt, ohne Anbieter oder Einführer zu sein. Der **Betreiber (engl. operator)** umfasst schließlich alle zuvor genannten Akteure, also Anbieter, Produkthersteller, Verteiler, Bevollmächtigte, Importeure oder Händler, und beschreibt somit eine Sammelbezeichnung für diejenigen, die in den verschiedenen Phasen der Wertschöpfungskette eines KI-Systems tätig sind (EU AI ACT, 2024).

Eine vereinfachte Darstellung für diesen Sachverhalt bietet die Abbildung 2. In dieser Darstellung wird ein vereinfachter Lebenszyklus eines KI-Systems über drei Phasen dargestellt und die betroffenen Rollen aus dem EU AI ACT zugeordnet.

Abbildung 2: Verantwortlichkeiten in einem vereinfachten Lebenszyklus eines KI-Systems



Anmerkung: Eigene Darstellung in Anlehnung an Visengeriyeva (2025).

Neben der Festlegung von Verantwortlichkeiten stellt die Risikoklassifizierung ein zentrales Konzept des EU AI Acts dar und wird aus diesem Grund im folgenden Absatz ausführlich erläutert.

3.2 Risikoklassifizierungssystem

Einleitend zu diesem Kapitel wurde bereits die Klassifizierung in verbotene Systeme, Hochrisiko-KI-Systeme und KI-Systeme mit geringem Risiko genannt. An dieser Stelle ist zu ergänzen, dass die letzte Kategorie je nach Darstellung nochmals in Systeme mit geringem und minimalem Risiko unterteil wird. Eine nähere Beschreibung dieser Klassifizierungen wird in diesem Absatz durchgeführt.

Um ein besseres Verständnis der Risikoklassifizierung von KI-Systemen zu fördern, wird es als hilfreich angesehen, konkrete Zahlen heranzuziehen. Eine Studie von Liebl und Klein (2023) liefert hierfür, trotz einiger relevanten Limitationen, eine Grundlage. In dieser Untersuchung wurden über 100 KI-Systeme aus unterschiedlichen Unternehmensbereichen gemäß dem Entwurf der KI-Verordnung der Europäischen Kommission von 2021 sowie dem Diskussionsstand im Parlament Anfang 2022 klassifiziert. Die Ergebnisse zeigten, dass 19 % der Systeme als hochriskant und 41 % als gering riskant eingestuft wurden. Für 39 % der Systeme war die Risikoeinstufung unklar, wobei die betroffenen Bereiche kritische Infrastruktur, Beschäftigung, Strafverfolgung und Produktsicherheit umfassten. Ein KI-System könnte basierend auf dieser Klassifizierung möglicherweise verboten werden. Zusammengefasst liegt die theoretische Anzahl der Hochrisiko-KI-Systeme zwischen 19 % und 58 %. Als Einschränkung dieser Studie und mit Verweis auf die wissenschaftliche Arbeit muss erwähnt werden, dass KI-Systeme für den medizinischen Bereich und das Gesundheitswesen nicht berücksichtigt wurden.

Um noch ein klareres Bild zu vermittelt, dient Tabelle 5. Sie enthält jeweils eine Beschreibung und Anwendungsbeispiele pro Risikolevel (EY, 2024; Gilbert, 2024; KPMG, 2024).

Tabelle 5: Risikoklassifizierung

Klassifizierung	Beschreibung	Anwendungsbeispiele
Verbotene Systeme	Verboten, da deren Nutzung ein untragbares Risiko für die Sicherheit, den Schutz und die Grundrechte der Menschen darstellt	Soziale Bewertung, die zu nachteiliger Behandlung führen könnte; Erkennungssysteme für Emotionen am Arbeitsplatz; biometrische Kategorisierung zur Ableitung sensibler Daten und prädiktiver Polizeiarbeit bei Einzelpersonen
Hochrisiko-KI-Systeme	Erlaubt, vorbehaltlich der Einhaltung der Anforderungen (einschließlich Konformitätsbewertungen vor dem Inverkehrbringen)	Recruiting; biometrische Identifikationsüberwachungssysteme; Sicherheitskomponenten von Systemen, die durch harmonisierte Gesetze abgedeckt sind (z. B. medizinische Geräte, Automobil); Zugang zu wesentlichen privaten und öffentlichen Dienstleistungen (z. B. Kreditwürdigkeit, Sozialleistungen, Kranken- und

Klassifizierung	Beschreibung	Anwendungsbeispiele
		Lebensversicherung); Sicherheit kritischer Infrastrukturen (z. B. Energie, Transport)
KI-Systeme mit geringem Risiko	Erlaubt, vorbehaltlich spezifischer Transparenz- und Offenlegungspflichten, wenn deren Nutzung ein begrenztes Risiko darstellt	Bestimmte KI-Systeme, die direkt mit Menschen interagieren (z. B. Chatbots); visuelle oder auditive Inhalte, die von einem KI-System manipuliert wurden (z.B. „Deepfakes“)
	Erlaubt, ohne zusätzliche Anforderungen, wenn deren Nutzung ein minimales Risiko darstellt	Standardmäßig alle anderen KI-Systeme, die nicht in die obigen Kategorien fallen (z. B. Bildbearbeitungssoftware, Produktempfehlungssysteme, Spam-Filtersoftware, Planungssoftware)

Bezugnehmend auf diese Tabelle werden in den folgenden Absätzen die jeweiligen Risikoklassen näher beschrieben und ihre Relevanz für die Anwendungsdomäne in dieser wissenschaftlichen Arbeit erläutert. Begonnen wird mit der höchsten Risikoklasse, den verbotenen Systemen und am Ende wird noch auf die Sonderbestimmungen rund um allgemeine KI-Modelle eingegangen.

3.2.1 Verbotene Künstliche Intelligenz

KI-Systeme, die als verboten klassifiziert werden, stehen per Definition im direkten Widerspruch zu den zentralen Werten der EU, wie der Achtung der Menschenwürde, der Freiheit, Gleichheit, Demokratie und Rechtsstaatlichkeit. Um die in der EU-Grundrechtecharta verankerten Rechte, darunter das Recht auf Nichtdiskriminierung, Datenschutz, Privatsphäre sowie den Schutz der Kinderrechte, zu wahren, verbietet die EU den Einsatz von KI-Systemen, die diese Prinzipien verletzen. Konkret betroffen sind KI-Systeme,

- die Entscheidungen von Menschen manipulieren oder deren Schwachstellen ausnutzen.
- die Menschen basierend auf ihrem sozialen Verhalten oder ihren persönlichen Eigenschaften bewerten oder klassifizieren.
- die das Risiko einer Person vorhersagen, ein Verbrechen zu begehen.
- die Gesichtsbilder aus dem Internet oder Überwachungskameras (CCTV) sammeln, Emotionen am Arbeitsplatz oder in Bildungseinrichtungen erfassen und Menschen basierend auf ihren biometrischen Daten kategorisieren.

In Hinblick auf den letzten Aspekt bietet das Gesetz Ausnahmen für Zwecke der Strafverfolgung, etwa bei der Suche nach vermissten Personen oder der Verhinderung von Terroranschlägen. Weiters regelt das Gesetz Praktiken bzw. Anwendungen, die bereits durch anderes geltendes Unionsrecht verboten sind, nicht (EU AI ACT, 2024). Angesichts der inhärenten Verknüpfung der vier Kategorien und der zahlreichen Möglichkeiten ihrer wechselseitigen Kombination und

Verflechtung, wird diese Unterscheidung durchaus problematisch gesehen und kritisiert. Es muss deshalb ein Augenmerk auf die spezifische Rechtsprechung gelegt werden und es wird eine breite interdisziplinäre Debatte notwendig sein (Neuwirth, 2023).

3.2.2 Hochriskante Künstliche Intelligenz

Im Kontext von Hochrisiko-KI-Systemen wird erstmalig von zu erfüllenden Anforderungen gesprochen, die es ermöglichen, dass solche Systeme auf dem Unionsmarkt in Verkehr gebracht, in Betrieb genommen oder allgemein verwendet werden. Die Einhaltung dieser Anforderungen soll sicherstellen, dass durch den Einsatz von KI-Systemen keine unzumutbaren Risiken für zentrale Interessen der EU entstehen, wie sie im Unionsrecht definiert sind (van Kolfschooten & van Oirschot, 2024). Im Gesundheitswesen bedeutet dies, dass jedes KI-System, das bei der Diagnose oder Behandlung verwendet wird, unabhängig davon, ob es nur unterstützende oder beratende Vorschläge liefert, grundsätzlich als Hochrisiko-KI-System einzustufen ist. Eine Ausnahme besteht nur, wenn das System bereits im Vorfeld ausdrücklich als verboten klassifiziert wurde (Gilbert, 2024). Gleichzeitig wurden Kriterien definiert, anhand derer bestimmte Anwendungsfälle von der Hochrisikoklassifikation ausgenommen werden können. Dazu zählen insbesondere Systeme, die auf die Erfüllung eng definierter, prozeduraler Aufgaben ausgerichtet sind, wie etwa die Umwandlung unstrukturierter Daten in strukturierte Formate, das Kategorisieren eingehender Dokumente oder das Erkennen von Duplikaten in großen Datenmengen. Solche Anwendungen haben keine direkten Auswirkungen auf Sicherheit oder Schutz. Zudem gelten KI-Systeme nicht als hochriskant, wenn sie dazu verwendet werden, das Ergebnis einer bereits abgeschlossenen menschlichen Aktivität zu verbessern. In diesen Fällen stellt die KI lediglich eine zusätzliche Schicht zur Qualitätsverbesserung dar, ohne eigenständig in sicherheitsrelevante Entscheidungen einzugreifen. Ein weiteres Kriterium ist der Einsatz von KI-Systemen, die Entscheidungsmuster oder Abweichungen von vorherigen Mustern erkennen, um Inkonsistenzen aufzudecken. Diese Systeme beeinflussen keine Entscheidungen, sondern unterstützen lediglich die menschliche Überprüfung. Ebenso sind Anwendungen, die eine vorbereitende Aufgabe für eine spätere menschliche Bewertung ausführen, von der Hochrisikoeinstufung ausgeschlossen. Ein Beispiel hierfür wäre der Einsatz zur Erkennung von finanziellen Betrugsfällen, ohne eigenständig Entscheidungen zu treffen (EY, 2024; Gilbert, 2024).

3.2.3 Künstliche Intelligenz mit geringem Risiko

Beispiele für KI-Systeme mit geringem Risiko im Gesundheitswesen reichen von Wohlfühl-Apps bis hin zur Krankenhausverwaltung. Vereinfacht kann festgehalten werden, dass die meisten KI-Systeme, die keinem rein medizinischen Zweck dienen und somit nicht unter die Definition der EU-Verordnung für Medizinprodukte fallen, als KI mit geringem Risiko eingestuft werden können. Medizinprodukte, die keine externe Konformitätsbewertung benötigen, sind ebenfalls eingeschlossen. Beispiele für solche Produkte sind Apps, die das Wohlbefinden und die Stimmung überwachen oder personalisierte Ernährungsempfehlungen basierend auf Nutzerdaten geben. Diese Systeme müssen lediglich bestimmte Transparenzvorgaben erfüllen.

Es gilt, dass Anbieter und Betreiber gewährleisten müssen, dass ihre Mitarbeitenden sowohl im Umgang mit risikoarmen als auch risikobehafteten KI-Systemen umfassend geschult sind. Für KI-Systeme mit minimalem Risiko, die keinen direkten Kontakt mit Personen haben gibt es keine speziellen Vorschriften. Mögliche Beispiele sind hierfür Produkte in der Krankenhausverwaltung (van Kolfschooten & van Oirschot, 2024). Die Sicherstellung der erforderlichen Kompetenzen kann durch die Implementierung eines freiwilligen Verhaltenskodexes (engl. Code of Conduct) gewährleistet werden (Stuurman & Lachaud, 2022).

3.2.4 Allgemeine Künstliche Intelligenz

Allgemein einsetzbare KI, Foundation-Modelle und generative KI spielen eine zunehmend wichtige Rolle in der KI-Regulierung. Während sie im ursprünglichen Vorschlag nicht explizit definiert waren, wurden sie in der letzten Version des KI-Gesetzes aufgenommen, um Szenarien abzudecken, in denen diese Systeme für verschiedene Zwecke verwendet oder in hochriskante Anwendungen integriert werden. Allgemein einsetzbare KI-Systeme (GPAI), wie generative KI-Anwendungen (z.B. ChatGPT), sind darauf ausgelegt, vielseitige Funktionen wie Bild- und Spracherkennung, Audio-/Videogenerierung und Musterdetektion zu erfüllen. Foundation-Modelle werden auf umfangreichen, breit gefächerten Daten trainiert und können für unterschiedliche Aufgaben angepasst werden. Da diese Systeme, insbesondere LLMs, in der Regel in andere KI-Systeme integriert werden und bedeutenden Einfluss auf deren Funktionsweise haben, fordert der AI Act besondere Transparenzanforderungen. Dazu gehört die Offenlegung technischer Dokumentationen und die Beschreibung der verwendeten Trainingsdaten. Da einige dieser Systeme aufgrund ihrer hohen Leistungsfähigkeit systemische Risiken bergen können, gelten zusätzliche Auflagen wie Modellevaluierungen und die Pflicht, schwerwiegende Vorfälle zu dokumentieren und zu melden (KPMG, 2024; van Kolfschooten & van Oirschot, 2024).

Da die Einstufung eines Systems, als risikoreich auf dessen Verwendungszweck basiert, ist es nachvollziehbar, dass allgemein einsetzbare KI-Systeme (GPAI) eine eigene Kategorie innerhalb des KI-Gesetzes erhalten haben. Diese Systeme können potenziell eine Vielzahl von Zwecken erfüllen, oft ohne das Wissen des Anbieters. Selbst wenn der Anbieter bestimmte hochriskante Anwendungsfälle ausdrücklich ausschließt, bleibt es technisch möglich, dass das System in einem gesellschaftlich risikobehafteten Bereich verwendet wird (Bayer, 2024).

3.3 Verpflichtungen

Dieser Absatz befasst sich mit den spezifischen Verpflichtungen von Deployern und Providern im Gesundheitswesen, die sich aus dem EU AI Act ableiten. Einen Überblick dieser Verpflichtungen bietet hierzu die Tabelle 6, die auf den Publikationen von van Kolfschooten und van Oirschot (2024) und Wagner et al. (2024) basiert.

Tabelle 6: Verpflichtungen für Anbieter und Betreiber nach Risikolevel

Klassifizierung	Verpflichtungen Provider	Verpflichtungen Deployer
Verbotene Systeme	Die Inverkehrbringung, Bereitstellung und Nutzung sind verboten (Artikel 5).	Die Inverkehrbringung, Bereitstellung und Nutzung sind verboten (Artikel 5).
Hochrisiko-KI-Systeme	Maßnahmen zur KI-Kompetenz (Artikel 4) Risikomanagementsystem (Artikel 9) Datenqualität und -governance (Artikel 10) Technische Dokumentation (Artikel 11) Dokumentation und Protokollierung (Artikel 12 und 18) Transparenz- und Informationspflichten (Artikel 13) Menschliche Überwachungsmaßnahmen (Artikel 14) Korrektheit, Robustheit und Cybersicherheit (Artikel 15) Qualitätssicherungssystem (Artikel 17) Abhilfemaßnahmen bei Nichtkonformität (Artikel 20) Durchführung der Konformitätsbewertung; CE-Kennzeichnung und Registrierung (Artikel 43 - 49)	Maßnahmen zur KI-Kompetenz (Artikel 4) Nutzung von Systemen gemäß den Anweisungen (Artikel 26(1)) Zuweisung der menschlichen Überwachung an qualifizierte natürliche Personen (Artikel 26(2)) Sicherstellung relevanter und ausreichend repräsentativer Eingangsdaten (Artikel 26(3)) Überwachung des Betriebs und Information der Stakeholder über schwerwiegende Vorfälle (Artikel 26(5)) (Artikel 72) Durchführung einer Datenschutz-Folgenabschätzung (Artikel 26(9)) Durchführung einer Folgenabschätzung der Auswirkungen auf Grundrechte (Artikel 27)

Klassifizierung	Verpflichtungen Provider	Verpflichtungen Deployer
KI-Systeme mit geringem Risiko	Maßnahmen zur KI-Kompetenz (Artikel 4) Transparenzpflichten (Artikel 50)	Maßnahmen zur KI-Kompetenz (Artikel 4) Transparenzpflichten (Artikel 50)
Allgemeine KI-Modelle	Technische Dokumentation und Beschreibung der Trainingsdaten (Artikel 53 - 54) Transparenz- und Informationspflichten Implementierung von Richtlinien zu Urheber- und verwandten Rechten Durchführung von Modellauswertungen und Risikobewertungen (Artikel 55) Meldung von Vorfällen an das AI Office (Artikel 55) Kennzeichnung maschinell generierter Inhalte	Keine Verpflichtungen

Zusammenfassend lässt sich feststellen, dass die größten Änderungen für KI-Systeme mit hohem Risiko sowie für allgemeine KI-Modelle zu erwarten sind. Ein überwiegender Teil der Verpflichtungen betrifft die Anbieter von KI-Systemen. Der Schwerpunkt dieser Verpflichtungen liegt auf Technologieunternehmen und nicht auf Krankenhäusern. Allerdings haben auch Krankenhäuser KI-Systeme für den eigenen Gebrauch entwickelt und werden dies auch in Zukunft tun. In der Regel treten Krankenhäuser jedoch als Betreiber von KI-Systemen auf (van Kolfschooten & van Oirschot, 2024). Bezugnehmend auf die im Absatz 2.3.5 beschriebenen spezifischen Risiken der KI im Gesundheitswesen bedeutet dieser Umstand, dass in Hinblick auf das Risiko **fehlerhafter Diagnostik und Therapieentscheidungen** KI-Systeme als hochriskant eingestuft werden und besonders die Maßnahmen nach Art. 9, Art. 15 und Art. 14 greifen. Letzterer rückt in den Vordergrund, da damit sichergestellt wird, dass klinische Diagnosen trotz KI-Unterstützung nur durch ärztliches Personal gestellt werden bzw. eine menschliche Überprüfung stattfindet. Medizinische KI-Systeme müssen, wie alle Medizinprodukte, eine klinische Konformitätsbewertung nach der Medizinprodukteverordnung (MDR) durch eine benannte Stelle durchlaufen. Dies stellt sicher, dass ihre Sicherheit und Leistungsfähigkeit nachgewiesen werden. Dadurch wird gewährleistet, dass medizinische KI-Anwendungen auf Genauigkeit und Zuverlässigkeit geprüft werden, bevor sie in der klinischen Praxis verwendet werden dürfen. Weiters kommt in diesem Kontext auch die EU-Datenschutz-Grundverordnung (DSGVO) zur Anwendung. Diese formuliert nochmals allgemeiner, dass Entscheidungen, die

ausschließlich auf einer automatisierten Verarbeitung beruhen, menschlicher Bewertung unterzogen werden müssen, da Patienten und Patientinnen das Recht haben, einer nicht ausschließlich maschinellen Entscheidung unterworfen zu werden. Auch in Hinblick auf das Risiko von **Biases** wird auf Art. 14 und die menschliche Überwachung verwiesen. Genannt wird im Kontext von Entwicklungsprozessen Art. 10, der entsprechende Datenqualität und -governance fordert. An dieser Stelle wird nochmals auf Absatz 2.1.3 verwiesen.

Abschließend wird dieses Risiko auch durch Art. 9, der die Einführung eines Risikomanagementsystems adressiert, behandelt. Eine vergleichbare bzw. in Teilen strengere Anforderung für eine Risikobewertung fordert die MDR und die Verordnung über In-vitro-Diagnostika (IVDR). Im Bereich **Datenschutz und Datenmissbrauch** ergänzt und konkretisiert der EU AI Act die DSGVO. Zentrale Elemente sind die Risikobewertung und -klassifizierung von Hochrisiko-KI-Systemen nach Art. 9, die Daten-Governance, der Grundsatz der Datensparsamkeit und in diesem Kontext das Prinzip Privacy by Design nach Art. 10, die Transparenz der Datenverwendung nach Art. 13 und die Cybersicherheit nach Art. 15 (Wendland, 2024).

Zusammenfassend bringt die Integration von KI-Systemen in bestehende regulatorische Rahmenwerke eine zusätzliche Komplexitätsebene mit sich. Obwohl Gesundheitsdienstleister mit solchen Rahmenwerken vertraut sind, wird berichtet, dass es Herausforderungen bei der Vereinbarkeit der Anforderungen des EU AI Act mit den bestehenden Regelungen für Medizinprodukte gibt. Im Vergleich zu etablierten Standards wie ISO 27001 im Bereich der Cybersicherheit fehlt es dem EU AI Act bislang an klaren, objektiven Nachweisen zur Erfüllung seiner Vorgaben. Darüber hinaus bestehen Unstimmigkeiten zwischen verschiedenen regulatorischen Rahmenbedingungen, die in der Praxis kritisch bewertet werden. Dennoch wird der EU AI Act in Summe von vielen Anbietern im Gesundheitswesen als ein Beitrag zur Stärkung des Marktvertrauens gesehen. Einrichtungen, die bereits mit bestehenden Regelwerken vertraut sind, zeigen sich besser vorbereitet und nutzen ihre Erfahrungen, um die neuen Anforderungen zu erfüllen. Untersuchungen bestätigen, dass insbesondere größere Organisationen von dieser Vorerfahrung profitieren. In diesem Zusammenhang wird der Nutzen standardisierter Vorgaben betont, etwa zur Verbesserung der Interoperabilität und zur reibungslosen Integration von KI in bestehende klinische Systeme. Es wird deshalb gefordert, dass Standardisierungsorganisationen spezifische technische Standards für KI-Systeme im Gesundheitswesen entwickeln, um bestehende Inkonsistenzen zu adressieren. Ebenso erscheint eine enge Abstimmung zwischen europäischen und internationalen Standards erforderlich, um den regulatorischen Aufwand für Unternehmen möglichst gering zu halten (Kilian et al., 2025).

Bezugnehmend auf diese Verpflichtungen und Besonderheiten werden im nächsten Kapitel Handlungsempfehlungen aufgestellt, die im Zuge der Experteninterviews umfassend geprüft, bewertet und auf ihre praktische Umsetzbarkeit sowie ihren potenziellen Nutzen für die Zielsetzung hin validiert werden.

4 AUFSTELLEN DER HANDLUNGSEMPFEHLUNGEN

Der Schwerpunkt dieser wissenschaftlichen Arbeit liegt in der Entwicklung von Handlungsempfehlungen für Entscheidungsträger:innen in Krankenhäusern im Umgang mit den Implikationen des EU AI Act und den erforderlichen Maßnahmen.

Die folgenden Handlungsempfehlungen basieren auf ausgewählten Quellen und der Fokus lag unter anderem auf praxisrelevanten Aspekten, weshalb auch graue Literatur in die Entwicklung einbezogen wurde. Es zeigte sich, dass spezifische und relevante Quellen, die sich explizit auf den EU AI Act beziehen, rar sind. Ein Kritikpunkt an der Vorgehensweise ist, dass keine systematische Literaturrecherche (Datenbanksuche mit Ein- und Ausschlusskriterien) oder systematischen Textanalysen durchgeführt wurden, um zusätzliche Quellen zu identifizieren und eine entsprechende wissenschaftliche Auswertung zu erhalten. Dennoch sollte die Vorgehensweise im Sinne der intersubjektiven Nachvollziehbarkeit des Forschungsprozesses gegeben sein, da die Handlungsempfehlungen im weiteren Verlauf einer Validierung durch Experten und Expertinnen unterzogen werden. Eine abschließende Limitation ist der zeitliche Rahmen in zweierlei Hinsicht. Einerseits sind aufgrund der Neuheit des Themas die Quellen, wie bereits erwähnt, unvollständig mit spezifischem Verweis auf den EU AI Act. Andererseits handelt es sich um eine Abschlussarbeit, die einem zeitlichen Rahmen unterliegt, der nur ein bestimmtes Ausmaß an unterschiedlichen Methoden zulässt.

Als Rahmenwerk für die Handlungsempfehlungen für Provider und Deployer wird das Modell CRIPS-ML(Q) verwendet, das bereits im Absatz 2.1.3 beschrieben wurde. Die Grundlage bzw. der Ansatz für diese Verwendung liefert Visengeriyeva (2025). In dieser Publikation wurde ein detaillierter und umsetzbarer Prozess beschrieben, wie Provider die Einhaltung des EU AI Acts aus einer Entwicklungsperspektive sicherstellen können.

Es werden an dieser Stelle lediglich die zentralen Konzepte erläutert. Für weiterführende Informationen und eine detaillierte Vorgehensweise wird auf die entsprechende Publikation verwiesen. Nach Ansicht des Autors ist eine modifizierte Anwendung der Konzepte auch für Deployer möglich, weshalb diese im Folgenden näher beschrieben wird.

Abschließend sei zur Einführung dieses Kapitels darauf hingewiesen, dass es sich hierbei nicht um rechtliche Empfehlungen handelt. Im Zweifelsfall wird daher empfohlen, eine Beratung durch die zuständigen Fachstellen oder Institutionen in Anspruch zu nehmen.

4.1 Phase 1: Geschäfts- und Datenverständnis

- Alle KI-Systeme in der Organisation identifizieren und prüfen, ob diese unter das KI-Gesetz fallen (Provider und Deployer) (EY, 2024; KPMG, 2024).
- Die Rolle der Organisation im gesamten Lebenszyklus der KI-Systeme verstehen (Provider und Deployer) (EY, 2023, 2024).
- Ein Risikomanagementsystem für alle Hochrisiko-KI-Systeme über deren gesamten Lebenszyklus hinweg etablieren (Provider und Deployer) (Aboy et al., 2024; Bayer, 2024; EY, 2023, 2024; KPMG, 2024). Ein exemplarisches Risikomanagementsystem wird durch das NIST AI Risk Management Framework (AI RMF 1.0) bereitgestellt. Dieses Framework bietet einen detaillierten und strukturierten Ansatz für das Management von Risiken im Zusammenhang mit KI. Insbesondere beschreibt dieses Framework ausführlich die Phasen der Risikoanalyse, der Risikobewältigung und der Risikokontrolle im Lebenszyklus eines KI-Systems (Capparelli et al., 2024).
- Risiken für Gesundheit, Sicherheit und grundlegende Rechte in Bezug auf den vorgesehenen Einsatzzweck und über den gesamten Produktlebenszyklus berücksichtigen (Provider und Deployer) (Bayer, 2024; Gilbert, 2024; KPMG, 2024).
- Registrierungsanforderungen und Grundrechts-Folgenabschätzungen durchführen, wenn es sich beim Deployer um eine öffentliche Behörde handelt (Deployer) (EY, 2024).
- Ein Überwachungssystem für Provider und deren KI-Systeme etablieren, dass die Prüfung der technischen Dokumentation und die Identifikation von nachgelagerten Risiken umfasst (Deployer) (KPMG, 2024).

Weitere Best Practices (Visengeriyeva, 2025):

- Workshops mit Stakeholdern organisieren, um mögliche Risiken zu brainstormen (Provider und Deployer).
- Spezifische Risikorahmenwerke (z. B. IEEE AI Ethics) nutzen, um Fragen zu Datenrisiken, operationellen Risiken und regulatorischen Risiken zu beantworten (Provider und Deployer).
- Eine Failure Mode and Effects Analysis (FMEA) durchführen, um potenzielle Ausfälle systematisch zu identifizieren, zu bewerten und zu mindern (Provider und Deployer).

4.2 Phase 2: Datenvorbereitung

- Praktiken zur Datenverwaltung und -steuerung anwenden, um sicherzustellen, dass die Datensätze vollständig und fehlerfrei sind (Provider) (Aboy et al., 2024; Bayer, 2024; Gilbert, 2024; KPMG, 2024; Ranjbar & Ravn, 2024).
- Maßnahmen ergreifen, um die Relevanz der Eingabedaten für den Verwendungszweck sicherzustellen (Deployer) (EY, 2024).

Weitere Best Practices (Visengeriyeva, 2025):

- Prozesse zur Prüfung der Datenqualität etablieren, potenzielle Risiken wie fehlende Werte, Inkonsistenzen oder Verzerrungen analysieren und identifizieren (Provider und Deployer).
- Bias-Erkennungstools einsetzen, um Verzerrungen im Datensatz zu erkennen und zu quantifizieren (Provider und Deployer).
- Datenvalidierung und -profilierung mithilfe von Frameworks automatisieren (Provider und Deployer).
- Echtzeitüberwachung der Datenqualität durch Integration von Unit-Tests in Datenpipelines gewährleisten (Provider und Deployer).
- Sicherstellen, dass Datensätze notwendige Diversität enthalten, um verzerrte Ergebnisse zu vermeiden (Provider und Deployer).
- Ethik-Checks in Pipelines integrieren und durch spezifische Tools unterstützen (Provider und Deployer).

4.3 Phase 3: Modellierung

- KI-Systeme so entwickeln und nutzen, dass eine angemessene Rückverfolgbarkeit und Erklärbarkeit gewährleistet sind (Provider und Deployer) (Gilbert, 2024; KPMG, 2024).
- Systeme auf mögliche Verzerrungen testen, die die Gesundheit und Sicherheit von Personen beeinträchtigen oder zu Diskriminierung führen könnten (Provider und Deployer) (Aboy et al., 2024; Bayer, 2024).

Weitere Best Practices (Visengeriyeva, 2025):

- Eine Risikoanalyse des Modells durchführen, um Overfitting und Underfitting zu identifizieren (Provider und Deployer).
- Potenzielle Verzerrungen und unerwartete Ergebnisse prüfen (Provider und Deployer).
- Automatisierte Test-Frameworks nutzen, um Modelle zu testen, und Cross-Validation-Techniken implementieren, um die Generalisierbarkeit des Modells zu bewerten (Provider und Deployer).
- Algorithmen mit ethischen Überlegungen auswählen und interpretierbare Modelle für Entscheidungen mit hoher Tragweite verwenden (Provider und Deployer).
- Fairness-Metriken definieren und prüfen (z. B. demografische Parität, gleiche Chancen) (Provider und Deployer).
- Tools für kontinuierliche Ethik-Checks nutzen (Provider und Deployer).

4.4 Phase 4: Evaluation

- Das System vor der Markteinführung testen, um sicherzustellen, dass es konsistent für seinen vorgesehenen Zweck funktioniert (Provider und Deployer) (Bayer, 2024; KPMG, 2024).
- Eine technische Dokumentation für Hochrisiko-KI-Systeme erstellen, die eine allgemeine Beschreibung des KI-Systems, seinen Zweck, das Datum und die Version des Systems sowie Benutzeranweisungen, eine detaillierte Beschreibung der Systemelemente und des Entwicklungsprozesses umfasst (Provider) (Aboy et al., 2024; Bayer, 2024; KPMG, 2024).
- Hochrisiko-KI-Systeme in der EU-Datenbank vor der Markteinführung registrieren (Provider) (EY, 2024).
- Eine Drittanbieter-Konformitätsbewertung durchführen, wenn das KI-System Teil einer sicherheitsrelevanten Komponente ist (Provider) (EY, 2024).

Weitere Best Practices (Visengeriyeva, 2025):

- Tools wie SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations) oder InterpretML verwenden, um globale und lokale Interpretationen zu erstellen und unerwartete Merkmalsinteraktionen und Proxy-Variablen zu analysieren (Provider).
- Ergebnisse dokumentieren und Tests mit Frameworks wie der Adversarial Robustness Toolbox oder Cleverhans durchführen, um Angriffe wie Evasion, Poisoning oder Modell-Extraktion zu prüfen (Provider).
- Testarten (Unit-, Integrations- und Systemtests) implementieren, um einzelne Komponenten, deren Interaktionen und das Gesamtsystem zu prüfen (Provider).
- A/B-Tests durchführen, um Modelle schrittweise einzuführen und ihre Performance sowie Compliance-Metriken zu bewerten (Provider und Deployer).
- Fairness, Transparenz und Verantwortlichkeit durch die Bewertung von Ethik-Metriken sicherstellen und Tools wie Fairlearn oder AI Fairness 360 nutzen (Provider).

4.5 Phase 5: Bereitstellung

- Nutzer:innen angemessen über Risiken informieren, die nicht vollständig eliminiert oder gemindert werden können (Provider und Deployer) (Aboy et al., 2024; Bayer, 2024).
- Nutzer:innen darauf hinweisen, dass sie mit einem KI-System kommunizieren oder interagieren (Provider und Deployer) (EY, 2024; Gilbert, 2024).
- Ein System zur Überwachung etablieren, um sicherzustellen, dass das System eingebauten betrieblichen Beschränkungen unterliegt und auf menschliche Bedienung reagiert (Provider und Deployer) (Bayer, 2024; Botero Arcila, 2024; EY, 2024; Gilbert,

2024; Morales Santos et al., 2024; Ranjbar & Ravn, 2024; van Kolfschooten & van Oirschot, 2024).

- Hochrisiko-KI-Systeme so gestalten, dass die automatische Protokollierung von Ereignissen während ihres gesamten Lebenszyklus ermöglicht wird (Provider und Deployer) (Aboy et al., 2024; Bayer, 2024; EY, 2024).
- Systemprotokolle aufbewahren (Provider und Deployer) (EY, 2024).
- Schwerwiegende Vorfälle und Fehlfunktionen, die grundlegende Rechte verletzen, an die zuständigen Behörden und den Provider des KI-Systems melden (Provider und Deployer) (EY, 2024).
- Eine KI-Ethik- und Governancekommission einrichten, um bei Streitfällen zu entscheiden und die regulatorischen Anforderungen kontinuierlich zu überwachen (Provider und Deployer) (EY, 2024; KPMG, 2024).
- Prozesse etablieren, die eine interdisziplinäre Zusammenarbeit zwischen KI-Entwicklern, Fachkräften im Gesundheitswesen, Patientenorganisationen und Regulierungsbehörden sicherstellen (Provider und Deployer) (van Kolfschooten & van Oirschot, 2024).
- Mitarbeitende in Bezug auf die Vorteile und Risiken von KI und die entsprechenden Szenarien schulen (Provider und Deployer) (EY, 2024; KPMG, 2024).

Weitere Best Practices (Visengeriyeva, 2025):

- Sicherheitsaudits vor der Bereitstellung durchführen, Sicherheitslücken identifizieren und manuelle oder automatisierte Penetration-Tests durchführen (Provider und Deployer).
- Sicherheitsorientierte Code-Reviews für Bereitstellungsskripte und Infrastruktur durchführen (Provider und Deployer).
- Wichtige Metriken wie Modellgenauigkeit, Latenz und CPU-Nutzung überwachen und Alarmierungen einrichten (Provider und Deployer).
- Modelle regelmäßig anhand ethischer Leitlinien überprüfen und Frameworks wie den Foundational Value Finder nutzen (Provider und Deployer).

4.6 Phase 6: Überwachung und Wartung

- Ein Qualitätsmanagementsystem einführen oder in ein vorhandes System nach ISO/IEC 13485 integrieren (Provider) (Aboy et al., 2024; EY, 2024).
- Allgemeine Geschäftsbedingungen, Datenschutzrichtlinien und Einwilligungserklärungen prüfen und überarbeiten, wenn KI-Systeme mit Verbrauchern interagieren (Provider und Deployer) (KPMG, 2024).

Weitere Best Practices (Visengeriyeva, 2025):

- Daten- und Modell-Drift überwachen und technische, operationelle, ethische sowie regulatorische Risiken regelmäßig im Risikoregister aktualisieren (Provider und Deployer).
- Chaos-Engineering nutzen, um Szenarien wie Datenkorruption oder Netzwerkfehler zu simulieren (Provider und Deployer).
- Ethikleitlinien regelmäßig überprüfen, Stakeholder-Feedback einholen und dieses mit dem Systemverhalten abgleichen (Provider und Deployer).

5 EVALUIERUNG DER HANDLUNGSEMPFEHLUNGEN

Das fünfte Kapitel befasst sich mit der Validierung der zuvor entwickelten Handlungsempfehlungen. Ziel ist es, die Praxistauglichkeit, Relevanz und Vollständigkeit der theoretisch abgeleiteten Maßnahmen anhand von Experteninterviews systematisch zu überprüfen. Dieser Schritt ist von zentraler Bedeutung, um sicherzustellen, dass die erarbeiteten Strategien nicht nur den regulatorischen Anforderungen entsprechen, sondern auch praktikabel sind und den spezifischen Gegebenheiten im Krankenhausalltag Rechnung tragen. Der Mehrwert dieses Kapitels liegt in der gezielten Verknüpfung von Theorie und Praxis. Die Handlungsempfehlungen werden um die Perspektiven und Erfahrungen von Fachpersonen aus dem Gesundheitswesen ergänzt, wodurch eine belastbare Grundlage für die abschließende Überarbeitung und Finalisierung geschaffen wird. Darüber hinaus ermöglicht dieser Abschnitt, potenzielle Lücken, Missverständnisse oder unrealistische Annahmen in den bisherigen theoretischen Überlegungen zu identifizieren und gezielt zu adressieren. Im Rahmen der Interviews wird insbesondere untersucht, in welchem Maß die vorgeschlagenen Maßnahmen als umsetzbar eingeschätzt werden, welche Herausforderungen bei der Implementierung bestehen könnten und welche zusätzlichen Aspekte aus Sicht der Praxis berücksichtigt werden sollten. Die gewonnenen Erkenntnisse dienen im weiteren Verlauf als Grundlage für die Beantwortung der Forschungsfragen.

Zunächst ist zwischen qualitativer und quantitativer Forschung sowie einem Mixed-Method-Ansatz zu differenzieren. In der quantitativen Forschung werden theoretisch abgeleitete Hypothesen an einer großen Anzahl von Untersuchungseinheiten mittels strukturierter Datenerhebungsmethoden überprüft. Im Unterschied dazu konzentriert sich die qualitative Forschung auf eine geringe Anzahl von Fällen, die mithilfe unstrukturierter oder teilstrukturierter Methoden intensiv untersucht werden. Ziel ist es, den Untersuchungsgegenstand umfassend zu beschreiben und theoretische Zusammenhänge zu entwickeln. Die gewonnenen Daten werden dabei interpretativ ausgewertet. Aufgrund der Forschungsfragen, die keine Hypothesen ableiten lassen, und der Neuheit der Thematik wurde ein qualitativer Forschungsansatz gewählt. Es sei darauf hingewiesen, dass die entwickelten Handlungsempfehlungen in zukünftigen Studien statistisch auf ihre Verbreitung und Wirksamkeit untersucht werden könnten. Ein Mixed-Method-Ansatz, der beide Aspekte vereint, wurde aufgrund des Umfangs dieser wissenschaftlichen Arbeit nicht in Betracht gezogen. Im nächsten Schritt ist in der qualitativen Forschung und in Hinblick auf den Forschungsprozess zu argumentieren, wie die Datenerhebung und -analyse erfolgt (Döring, 2023). Diese Argumentation erfolgt in den nächsten beiden Absätzen.

5.1 Erhebungsmethode – das Experteninterview

Für die Datenerhebung wurden Experteninterviews gewählt, da sich dieses Format besonders gut eignet, um fachliche Einschätzungen strukturiert zu erfassen und gleichzeitig genügend Spielraum für spontane Vertiefungen im Gespräch zu lassen. Im Vergleich zu anderen Interviewformen ermöglichen Experteninterviews eine systematische Auswertung, ohne die Offenheit für neue Impulse einzuschränken. Aufgrund der Komplexität und Spezifik des Themenfelds war es erforderlich, die relevanten Experten und Expertinnen im Vorfeld gezielt auszuwählen. Die vorangegangene Literaturrecherche verdeutlichte, dass das vorhandene Wissen in diesem Bereich stark fragmentiert ist. Durch die Interviews sollen gezielt neue Erkenntnisse generiert werden, indem die fachlichen Einschätzungen und praktischen Erfahrungen der Befragten möglichst umfassend erfasst und nachvollzogen werden. Ein zusätzlicher Beweggrund für die Wahl dieser Methode liegt in ihrer etablierten Anwendung innerhalb wissenschaftlicher Abschlussarbeiten sowie in der breiten Verfügbarkeit methodischer Literatur, die eine fundierte konzeptionelle Orientierung ermöglicht (Berger-Grabner, 2022).

Um Experteninterviews erfolgreich durchzuführen, ist eine sorgfältige Planung unerlässlich. Die Vorbereitung gliedert sich in folgende vier Schritte (Bogner et al., 2014).

1. **Erstellung eines Interviewleitfadens**: Dieser dient als Checkliste und Orientierungshilfe für den Interviewer.
2. **Auswahl der Experten und Expertinnen**: Die Auswahl muss nachvollziehbar und gut begründet sein.
3. **Kontaktaufnahme und Terminvereinbarung**: Da die befragten Personen oft einen vollen Terminkalender haben, ist dieser Schritt besonders wichtig und darf nicht vernachlässigt werden.
4. **Entscheidung über das Dokumentationsmedium**: Es wird festgelegt, welches Medium zur Aufzeichnung des Interviews verwendet wird.

Aufgrund des Umfangs der Schritte werden der erste und zweite Schritt in den folgenden beiden Absätzen beschrieben. Der dritte Schritt wird an dieser Stelle nur kurz erwähnt. Experten und Expertinnen, die im entsprechenden Zeitfenster keine Verfügbarkeit gewähren konnten, wurden nicht interviewt, und dies wurde im zweiten Schritt entsprechend vermerkt. Der vierte Schritt wird im Absatz 5.2 behandelt.

Zur Schaffung einer fundierten Bewertungsgrundlage für die nachfolgenden detaillierten Erläuterungen der durchgeführten Schritte werden an dieser Stelle die drei Gütekriterien nach Kaiser (2021) vorgestellt, die im Kontext dieses Forschungsdesigns berücksichtigt wurden. Erstens muss das Verfahren der Datenerhebung und Datenauswertung intersubjektiv nachvollziehbar sein. Der zweite Aspekt umfasst die theoriegeleitete Vorgehensweise. Aus diesem Grund werden in diesem Kapitel auch Aspekte, wie die Erstellung des Interviewleitfadens und die Auswahl der Experten und Expertinnen näher beschrieben. Der letzte Punkt betrifft die Neutralität und Offenheit des Forschers gegenüber neuen Erkenntnissen.

5.1.1 Erstellung des Interviewleitfadens

Der Interviewleitfaden ist ein Instrument zur Datenerhebung, das oft als „halbstandardisiert“ bezeichnet wird. Es wird im Vorfeld ein Fragenkatalog, ohne vorgegebene Antwortmöglichkeiten, erstellt. Der Leitfaden kann und soll flexibel, beispielsweise in Bezug auf Wortwahl, Reihenfolge der Fragen und Zusatzfragen, angepasst werden. Die Standardisierung des Leitfadens liegt nicht in der exakten Formulierung der Fragen, sondern in der inhaltlichen Vergleichbarkeit des Fragenkatalogs. Ein typischer Interviewleitfaden umfasst 1-2 Seiten mit 8-15 Fragen (Gläser & Laudel, 2010). Diese lassen sich in Hauptfragen und Differenzierungsfragen unterteilen und können wörtlich oder stichpunktartig im Leitfaden stehen. Falls biografische und demografische Angaben nicht oder nur teilweise zu Beginn des qualitativen Interviews erhoben wurden, werden sie am Ende mündlich oder schriftlich erfasst (z.B. Nationalität, Bildungsstand, Beruf, Einkommen). Je nach Forschungsfrage können mit einem abschließenden Kurzfragebogen weitere relevante Aspekte erhoben werden. Dabei gilt jedoch das Gebot der Datensparsamkeit bei persönlichen Daten (Döring, 2023).

Der Interviewleitfaden beginnt nicht mit den ersten inhaltlichen Fragen, sondern mit einer Frage zur Einwilligung der Aufzeichnung des Interviews und der Verwendung der Aussagen in anonymisierter oder nicht anonymisierter Form. Zudem wird die befragte Person über das Ziel der wissenschaftlichen Arbeit und des Interviews informiert. Anschließend wird die erste sogenannte Aufwärmfrage gestellt, die als Eisbrecher fungiert und das Gesprächsklima, die Rollen der Gesprächspartner und das Kommunikationsniveau bestimmt. Diese Frage sollte daher gut durchdacht und leicht zu beantworten sein, um einen guten Start zu gewährleisten (Gläser & Laudel, 2010). Generell sollten Fragen zu gleichen Themengebieten in Fragenkomplexe organisiert und nacheinander gestellt werden. Obwohl es inhaltliche Brüche im Interview geben kann, können diese durch eine gute Organisation minimiert werden, was auch das Erinnerungsvermögen bei längeren Interviews unterstützt (Bogner et al., 2014). Neben diesen spezifischen Punkten zum Aufbau gibt es weitere Aspekte, die bei der Konstruktion des Interviewleitfadens berücksichtigt werden sollten. Der Leitfaden sollte klar strukturiert und möglichst auf zwei Seiten begrenzt sein, damit er während des Interviews eine schnelle Orientierung bietet. Es wird empfohlen, die Fragen auszuformulieren, damit alle Interviewpartner:innen die Fragen in gleicher Form hören. Dadurch wird die Vergleichbarkeit der Interviews verbessert und Sicherheit in schwierigen Situationen für die forschende Person geboten. Ein Nachteil vorformulierter Fragen ist, dass sie zum Festhalten am Leitfaden verleiten können. Es ist dennoch empfehlenswert, die Fragen zunächst vollständig aufzuschreiben, um sich mit ihrer Formulierung bewusst auseinanderzusetzen. Im Interview sollten die Fragen klar, verständlich und möglichst in Alltagssprache gestellt werden. Meinungsfragen sind nur dann sinnvoll, wenn die Einschätzungen und Beweggründe der befragten Person für das Verständnis sozialer Prozesse relevant sind. Auch hypothetische Fragen sollten nur eingesetzt werden, wenn subjektive Theorien von Bedeutung sind. Simulationsfragen eignen sich besonders für schwer kommunizierbare Sachverhalte oder als starke Erzählanregungen. Faktfragen sollten als offene Erzählanregungen gestellt werden. Erzählanregungen sind wichtiger als Detailfragen und sollten zuerst gestellt werden, in der Hoffnung, dass sich Detailfragen erübrigen. Im Interviewleitfaden

kann dies so aussehen, dass die Erzählanregung als Frage hervorgehoben wird und die Detailfragen als Checkliste darunter notiert werden. Während der Erzählung müssen dann nur noch gedanklich die Details abgehakt werden. Die Klarheit der Frage ist wichtiger als ihre Offenheit, damit die befragte Person genau versteht, was gefragt ist. Provokationen sollten sparsam und nur für spezielle Zwecke eingesetzt werden. Wenn in Interviewfragen Informationen aus anderen Interviews verwendet werden, muss die Anonymität gewahrt werden. Das bedeutet, dass der Name der ursprünglich befragten Person nicht genannt werden darf und die Information so formuliert werden muss, dass die Identität nicht preisgegeben wird. Im Zweifelsfall hat die Wahrung der Anonymität Priorität (Gläser & Laudel, 2010). Bezugnehmend auf diese allgemeinen Vorgaben wurde der Interviewleitfaden in thematische Phasen unterteilt, die aufeinander aufbauen. Diese Phasen spiegeln Fragenkomplexe wider und dienen als Analysedimensionen für die Auswertung. Der Hauptteil des Leitfadens gliedert sich in drei Phasen: (1) Allgemeine Erfahrungen mit KI im Gesundheitswesen und dem EU AI Act, (2) Validierung der Handlungsempfehlungen anhand des CRISP-ML(Q)-Modells sowie (3) Zukunftsaussichten. Die zweite Dimension wird zusätzlich in die einzelnen Phasen des Modells unterteilt, sodass jede Phase separat validiert werden kann. Die Gliederung des Interviewleitfadens mit entsprechenden Zielen und dem detaillierten Ablauf ist in der Tabelle 7 dargestellt.

Tabelle 7: Fragenkomplexe des Hauptteils des Interviewleitfadens

Phase	U-Phase	Ziel/ Ablauf
Allgemeine Erfahrungen mit KI im Gesundheitswesen und dem EU AI Act		**Ziel**: Einstieg ins Thema, Verständnis für vorhandene KI-Anwendungen und Wissen über EU AI Act **Ablauf**: Fragen zu bisherigen KI-Projekten, Erfolgen und Herausforderungen; Einschätzung der Akzeptanz bei Personal und Patienten und Patientinnen; Kenntnisstand EU AI Act
Validierung der Handlungsempfehlungen anhand des CRISP-ML(Q)-Modells	Phase 1: Geschäfts- und Datenverständnis	**Ziel**: Abprüfen, wie Teilnehmende KI-Systeme identifizieren, Verantwortlichkeiten klären und Risiken managen **Ablauf**: Fragen zum Umgang mit KI-Systemen, Rolle(n) im Lebenszyklus, Risikomanagement und rechtlichen Anforderungen- Erfassung praktischer Erfahrungen und möglicher Lücken
	Phase 2: Datenvorbereitung	**Ziel**: Prüfen von Datenqualität, Bias-Vermeidung, Relevanz der Daten für den Einsatzzweck **Ablauf**: Fragen zur Data Governance und konkreten Prozessen; Bewertung, wie systematisch Daten aufbereitet werden

Phase	U-Phase	Ziel/ Ablauf
	Phase 3: Modellierung	**Ziel**: Einschätzung zu Modellrückverfolgbarkeit, Erklärbarkeit, Fairness und Risikoprüfung **Ablauf**: Fragen zu Mechanismen der Fairness-Prüfung, Modellrisikoanalyse und Testverfahren
	Phase 4: Evaluation	**Ziel**: Feststellen, wie umfangreich die Evaluations- und Testverfahren sind (z. B. technische Dokumentation, Robustheit, Konformität) **Ablauf**: Fragen zu Pre-Market-Tests, Dokumentation, Registrierung und Robustheitsprüfungen
	Phase 5: Bereitstellung	**Ziel**: Verständnis für den Produktions-/Livegang der KI und Einhaltung von Transparenz, Sicherheitsvorkehrungen, Meldewesen **Ablauf**: Fragen zum Risikokommunikationsprozess, Kennzeichnungspflicht und menschlicher Kontrolle
	Phase 6: Überwachung und Wartung	**Ziel**: Einsicht in langfristige Qualitätssicherung, Umgang mit Daten-/Modell-Drift, Ethikleitlinien und Notfallplanung **Ablauf**: Fragen zu QM-System, Vertragsanpassungen, Monitoring des Systems im Betrieb; Einblicke in Kontinuität und Weiterentwicklung
	Querschnittsfragen	**Ziel**: Ergänzende Fragen, die mehrere Phasen betreffen (organisatorische Verankerung, strategische Bedeutung, Herausforderungen, Schulungsbedarf) **Ablauf**: Fragen ggf. am Ende oder verteilt stellen- Sammeln zusätzlicher Infos zu Hürden und Rahmenbedingungen
Zukunftsaussicht		**Ziel**: Blick in die Zukunft, Rolle von KI in Krankenhäusern, mögliche Kooperationen mit Gesetzgebern und Technologieentwicklern **Ablauf**: Fragen zu künftigen KI-Entwicklungen und strategischer Relevanz- Ideen zur Verbesserung der Zusammenarbeit zwischen Stakeholdern

Die erste Analysedimension hat das Ziel, einen thematischen Einstieg zu ermöglichen und ein gemeinsames Verständnis hinsichtlich bestehender KI-Anwendungen und des Wissens über den EU AI Act zu etablieren. Die dazugehörigen Fragestellungen adressieren bisherige KI-Projekte, erzielte Erfolge und bestehende Herausforderungen, die Akzeptanz bei Personal und Patienten

und Patientinnen sowie den Kenntnisstand zum EU AI Act. Wie eingangs erwähnt, bildet die zweite Analysedimension den zentralen Inhalt dieser Arbeit – die Validierung der Handlungsempfehlungen anhand des CRISP-ML(Q)-Modells. Diese Dimension wird in sechs Unterkategorien entsprechend den Phasen des Modells gegliedert. In der ersten Phase, dem Geschäfts- und Datenverständnis, erfolgt die Erhebung darüber, wie Teilnehmende KI-Systeme identifizieren, Verantwortlichkeiten klären und Risiken managen. An dieser Stelle werden Fragen zum Umgang mit KI-Systemen, zu den jeweiligen Rollen im Lebenszyklus sowie zum Risikomanagement und zu den rechtlichen Anforderungen gestellt, um praktische Erfahrungen und potenzielle Lücken zu erfassen. In der zweiten Phase, der Datenvorbereitung, liegt der Schwerpunkt auf der Bewertung der Datenqualität, der Vermeidung von Bias sowie der Relevanz der Daten für den jeweiligen Einsatzzweck. Hierbei wird der Aspekt der Data Governance und die systematische Aufbereitung der Daten konkret hinterfragt. Die dritte Phase, die Modellierung, dient der Einschätzung hinsichtlich der Rückverfolgbarkeit, Erklärbarkeit, Fairness und Risikoprüfung der entwickelten Modelle. Fragen zur Modellrisikoanalyse und zu den angewandten Testverfahren ermöglichen eine differenzierte Bewertung. Im Anschluss erfolgt in der vierten Phase, der Evaluation, die Erfassung des Umfangs der Evaluations- und Testverfahren, beispielsweise bezüglich technischer Dokumentation, Robustheit und Konformität. Hierbei werden Fragen zu Pre-Market-Tests, Dokumentationsstandards, Registrierungsvorgaben und Robustheitsprüfungen gestellt. Die fünfte Phase, die Bereitstellung, fokussiert auf das Verständnis für den Produktions- beziehungsweise Livegang der KI-Systeme sowie auf die Einhaltung von Transparenz-, Sicherheits- und Meldevorgaben. Hierzu werden Fragen zum Risikokommunikationsprozess, zur Kennzeichnungspflicht und zur Rolle der menschlichen Kontrolle gestellt. In der sechsten Phase, der Überwachung und Wartung, wird schließlich die langfristige Qualitätssicherung thematisiert. Der Fokus liegt hierbei auf dem Umgang mit Daten- oder Modell-Drift, der Implementierung von Ethikleitlinien. Fragen zu Qualitätsmanagementsystemen, zu Vertragsanpassungen und zum Monitoring des Betriebs geben Aufschluss über die Kontinuität und Weiterentwicklung der Systeme. Ergänzend zu diesen phasenspezifischen Fragestellungen werden Querschnittsfragen integriert, die mehrere Phasen betreffen und zusätzliche Informationen zur organisatorischen Verankerung, strategischen Bedeutung, zu Herausforderungen sowie zum Schulungsbedarf liefern. Abschließend richtet sich der Hauptteil des Leitfadens auf die Zukunftsaussicht, in der der Blick in die künftige Rolle von KI im Gesundheitswesen geschärft wird. Hierbei werden Fragestellungen zu potenziellen Entwicklungen, zu möglichen Kooperationen mit Gesetzgebern und Technologieentwicklern sowie zur strategischen Relevanz zukünftiger KI-Anwendungen erörtert. Im Zuge dieser Darstellung wurde die Einleitung und der Schluss des Interviewleitfadens nicht näher behandelt. Die Einleitung umfasst die Begrüßung, administrative Aspekte und eine Vorstellung der Masterarbeit. Der Leitfaden wurde bereits vorab der zu interviewende Person übermittelt. Der Schluss beinhaltet abschließend eine Zusammenfassung und die Klärung noch offener Fragen. Der vollständige Interviewleitfaden ist als Anhang 1 zu finden. Auf Grundlage des Interviewleitfadens wurden im Anschluss die geeigneten Experten und Expertinnen identifiziert und kontaktiert. Der Auswahl- und Kontaktprozess wird im folgenden Absatz detailliert erläutert.

5.1.2 Auswahl der Experten und Expertinnen

Am Beginn dieses Absatzes ist zu definieren, was unter dem Begriff "Expertentum" verstanden wird. Eine Abgrenzung findet hier auf der einen Seite zum Laien und auf der anderen Seite zum Spezialisten oder zur Spezialistin statt. Laien verfügen über Allgemein- oder Alltagswissen, im Gegensatz dazu verfügt ein Experte oder eine Expertin über Sonderwissen und Problemlösungskompetenzen, die an eine Funktion oder Berufsrolle gebunden sind. Die Abgrenzung zum Spezialistentum umfasst den Umstand, dass Spezialisten und Spezialistinnen nicht selbstständig entscheiden, ob das Wissen und die Kompetenzen zur Lösung eines Problems ausreichen. Ausgehend von dieser Aussage lässt sich schließen, dass für diese Art von Interviews sowohl Experten und Expertinnen als auch Spezialisten und Spezialistinnen geeignet sind, da ihr Wissen im Rahmen eines explorativen Ansatzes zur Lösung des Problems beitragen kann. In der Theorie wären auch Interviews mit Laien möglich, jedoch handelt es sich bei dieser wissenschaftlichen Arbeit um ein Thema, das mit Allgemein- und Alltagswissen nicht behandelt werden kann, und aus diesem Grund wird auf die beiden anderen Gruppen zurückgegriffen. Es ist zu vermerken, dass befragte Personen unterschiedliches Wissen in das Interview einbringen, und daraus ergibt sich die Fragestellung nach der Anzahl der Interviews. In diesem Kontext muss auf das zu gewinnende Wissen verwiesen werden. Hier wird zwischen Betriebswissen, Kontextwissen und Deutungswissen unterschieden. Jede dieser Dimensionen ist für sich gesehen wichtig, jedoch können diese durch den unterschiedlichen Background der Experten und Expertinnen unterschiedlich eingebracht werden. Das Betriebswissen beschäftigt sich mit Prozessen und Routinen. Bezogen auf diese wissenschaftliche Arbeit kann damit das Wissen über Prozesse und Routinen in Krankenhäusern sowie entsprechenden KI-Systemen gemeint sein. Die zweite Kategorie beschreibt das Kontextwissen. Bei dieser Form von Wissen geht es um die Rahmenbedingungen, und umgelegt auf diese Arbeit wären die Grundlagen zum EU AI Act als nennenswert zu sehen. Die letzte Dimension beschreibt das Deutungswissen. Hier stehen die subjektiven Relevanzen, Sichtweisen und Interpretationen, beispielsweise zum EU AI Act und seiner Beschaffenheit, im Vordergrund (Kaiser, 2021).

Neben diesen Dimensionen des Wissens stehen bei der Wahl der Experten und Expertinnen vier allgemeine Fragen im Zentrum (Gläser & Laudel, 2010; Kaiser, 2021):

- Verfügt die Person über das notwendige Wissen?
- Ist die Person qualifiziert, das Wissen weiterzugeben?
- Zeigt die Person die Bereitschaft, das Wissen weiterzugeben?
- Ist die Person terminlich im Untersuchungszeitraum verfügbar?

Zur Sicherstellung der intersubjektiven Nachvollziehbarkeit wird die Auswahl der Experten und Expertinnen tabellarisch in der Tabelle 8 anhand dieser Fragestellungen und den zuvor beschriebenen Wissensdimensionen dargelegt. Insgesamt wurden 13 Experten und Expertinnen per E-Mail kontaktiert. Aufgrund ausbleibender Rückmeldungen sowie weiterer Ausschlussgründe konnten letztlich Interviews mit fünf der kontaktierten Fachpersonen durchgeführt werden.

Tabelle 8: Auswahl der Experten und Expertinnen

I-Nr	Allgemeine Beschreibung	Besonderes Wissen über die relevanten Informationen	Fähigkeit und Bereitschaft der präzisen Weitergabe
1	Leitung einer klinischen Abteilung an einer Universitätsklinik im Bereich der Radiologie	**Betriebswissen:** Expertise in Prozessen und Routinen der pädiatrischen Radiologie, insbesondere in der Trauma- und Herzbildgebung. **Kontextwissen:** Forschung zu Strahlenexposition bei pädiatrischen CTs sowie zu unbeabsichtigten Folgen von KI in der Radiologie. **Deutungswissen:** Eigene Publikationen zur Abwägung zwischen Leistungsfähigkeit und Interpretierbarkeit von KI-Systemen; differenzierte Sicht auf den Einfluss technischer Innovationen.	Wissenschaftlich aktiv (Herausgeberschaften, Lehre, Fachgesellschaften); hohe Bereitschaft und Fähigkeit zur fundierten Wissensweitergabe.
2	Leitung einer interdisziplinären Forschungsgruppe mit Fokus auf medizinische Bildgebung	**Betriebswissen:** Tiefe Kenntnisse radiologischer Diagnoseprozesse sowie technischer Abläufe in der Bildgebung (insb. MRT, funktionelle Bildgebung). **Kontextwissen:** Forschung zu KI-gestützter Bildverarbeitung; Kenntnisse zu ethischen und regulatorischen Aspekten. **Deutungswissen:** Fachliche Einordnung der Rolle von KI bei klinischen Entscheidungen und Verbesserung der Diagnosequalität.	Breite fachliche und didaktische Expertise durch internationale Publikationstätigkeit, Lehre und aktive Fachnetzwerke.

I-Nr	Allgemeine Beschreibung	Besonderes Wissen über die relevanten Informationen	Fähigkeit und Bereitschaft der präzisen Weitergabe
3	Leitung der IT-Steuerung und CISO in einem österreichischen Krankenhausverbunds	**Betriebswissen:** Langjährige Erfahrung mit IT-Prozessen und IT-Sicherheitsstrategien im Krankenhauskontext; Umsetzung der DSGVO und NIS-Richtlinien. **Kontextwissen:** Umfassende Kenntnisse der regulatorischen Anforderungen, insbesondere des EU AI Act. **Deutungswissen:** Einschätzung der Umsetzbarkeit von EU-Vorgaben im Spitalsbetrieb; Bewertung von Herausforderungen und Auswirkungen auf Datenschutz und Sicherheit.	Hohe Fachkompetenz durch Leitungstätigkeit, aktive Beteiligung an Audits und Know-how-Transfer in internen Gremien.
4	Leitung einer medizininformatischen Fachabteilung in einem österreichischen Krankenhausverbund	**Betriebswissen:** Erfahrung mit Aufbau, Steuerung und Optimierung von IT-Infrastrukturen im Krankenhaus, insbesondere SAP-Systeme. **Kontextwissen:** Tiefe Auseinandersetzung mit regulatorischen Anforderungen und deren Integration in strategische IT-Planung; Mitwirkung in internationalen SAP-Gremien. **Deutungswissen:** Subjektive Einschätzung zu Auswirkungen des EU AI Act auf Krankenhaus-IT, Softwarearchitekturen und organisatorische Abläufe.	Nachgewiesene Vermittlungskompetenz durch Gremienarbeit, Projektleitung und aktives Engagement im Wissensaustausch.

I-Nr	Allgemeine Beschreibung	Besonderes Wissen über die relevanten Informationen	Fähigkeit und Bereitschaft der präzisen Weitergabe
5	Geschäftsführende Leitung (CEO) und Mitbegründung eines Unternehmens im Bereich klinischer KI-Anwendungen; Lehrbeauftragte*r für Clinical Decision Support an einer Hochschule	**Betriebswissen:** Über 15 Jahre Erfahrung in der Arbeit mit Gesundheitsdaten und dem Aufbau von Machine-Learning-Modellen zur klinischen Entscheidungsunterstützung; Leitung interdisziplinärer Projekte und Anwendungen im klinischen Betrieb. **Kontextwissen:** Forschung zu KI-Anwendungen im Gesundheitswesen mit Fokus auf Risikovorhersagen (z. B. Delirium, Mangelernährung); fundiertes Verständnis regulatorischer Anforderungen, insbesondere in Bezug auf Patientensicherheit und Systemintegration. **Deutungswissen:** Analysen zur Akzeptanz von KI-gestützten Entscheidungsunterstützungssystemen bei medizinischem Personal; differenzierte Perspektiven zur Wirkung von KI auf Entscheidungsprozesse im Krankenhaus.	Nachweisbare Transferkompetenz durch Lehrtätigkeit, zahlreiche Vorträge auf Fachkonferenzen, Publikationstätigkeit sowie praxisnahe Umsetzung durch unternehmerische Verantwortung im KI-Bereich.

Im Sinne der angewandten Methodik wurde jede Beschreibung gemeinsam mit dem entsprechenden Interviewtranskript an die jeweilige Person übermittelt und eine Freigabe zur Verwendung eingeholt. Dadurch konnten die Transkripte nochmals überprüft und zugleich die Anonymität der Beteiligten gewährleistet werden.

5.2 Transkription und Auswertungsmethode

Jedes geführte Interview erfordert eine Transkription, um aufbereitete qualitative Daten zu erhalten. Die Grundlage bilden Video- und Audioaufzeichnungen, die in dieser wissenschaftlichen Arbeit mit Microsoft Teams erstellt werden. Es wird zwischen Volltranskription (vollständige Verschriftlichung) und Teiltranskription (nur relevante Passagen) unterschieden. Volltranskriptionen sind zeit- und kostenintensiv, während Teiltranskriptionen Zeit sparen können, jedoch das Risiko bergen, relevante Passagen zu übersehen. Bei Qualifikationsarbeiten, bei denen Transkription und Datenanalyse in einer Hand liegen, kann eine begründete Teiltranskription Zeit sparen. Für diese wissenschaftliche Arbeit wurde entschieden, dass ein KI-gestütztes deutsches Transkript mit Microsoft Teams während der Interviews erstellt und anschließend zusammen mit den Video- bzw. Audioaufzeichnungen in die Software MAXQDA hochgeladen und überarbeitet wird. MAXQDA ist eine QDA-Software, die das Transkript synchron zum Ton darstellt, um Vergleich und Verbesserung zu erleichtern (Döring, 2023). Somit wird ein vollständiges Transkript erstellt, wodurch die intersubjektive Nachvollziehbarkeit sichergestellt ist. Eine Teiltranskription würde an dieser Stelle einen subjektiven und nicht reproduzierbaren Interpretationsschritt darstellen (Gläser & Laudel, 2010). Die Überprüfung der Inhalte erfolgt manuell, während erforderliche Umformulierungen mithilfe von ChatGPT vorgenommen werden.

Im nächsten Schritt werden die aufbereiteten Daten einer qualitativen Inhaltsanalyse, eine häufig in Abschlussarbeiten verwendete Form der Textanalyse, unterzogen. Das Ziel ist es, das Datenmaterial so zu reduzieren und zu analysieren, dass die Forschungsfragen beantwortbar sind und die Zielsetzung erreicht werden kann. Jeder Analyseschritt muss auf begründeten und getesteten Regeln basieren, um die Nachvollziehbarkeit der Interpretationen zu gewährleisten.

Die qualitative Inhaltsanalyse umfasst vier Phasen: Planungs-, Entwicklungs-, Test- und Haupterhebungsphase mit anschließender Auswertung. In der Planungsphase werden die Forschungsfragen, sowie die Analyseeinheiten festgelegt. In der Entwicklungsphase wird das Kategoriensystem erstellt und Kodierregeln formuliert. Die Testphase beinhaltet einen Pretest des Kategoriesystems und der Kodierregeln. In der Haupterhebungsphase erfolgt die tatsächliche Auswertung des Materials, wobei in dieser wissenschaftlichen Arbeit eine zusammenfassende Inhaltsanalyse angewendet wird. Die zusammenfassende Inhaltsanalyse verdichtet das Material, um wesentliche Inhalte zu erhalten. Abschließend werden die Kategorien überprüft und das Kategoriensystem überarbeitet. Konkret umfasst die Hauptphase folgende Schritte:

1. Der transkribierte Text wird aufmerksam gelesen und relevante Stellen werden markiert. Nach diesem ersten Durchgang ergibt sich eine grobe Vorstellung vom Inhalt.
2. Im zweiten Durchgang werden nur die markierten Stellen betrachtet und für jede eine Kategorie formuliert, die den Inhalt prägnant zusammenfasst.
3. Die gebildeten Kategorien werden separat erfasst.

4. Zu jeder Kategorie werden die entsprechenden Textstellen mit Kurzbezeichnung des Textes und Zeilennummer angegeben.
5. Auf dem Kategorienblatt werden die Textstellen den Kategorien zugeordnet, was das schnelle Auffinden erleichtert.
6. Die Kategorien werden geordnet, indem zusammengehörige Begriffe gruppiert werden. Dabei wird darauf geachtet, dass sich die Kategorien klar voneinander unterscheiden. Jede Kategorie wird genau definiert. So entsteht ein Kategoriensystem, das die Inhalte strukturiert und Zusammenhänge aufzeigt. Sofern eine direkte Zusammenfassung auf Basis der codierten Textstellen nicht möglich ist, erfolgt eine Paraphrasierung der entsprechenden Inhalte, wobei mindestens ein inhaltlicher Reduktionsschritt vorgenommen wird.

Für die Bearbeitung dieser sechs Schritt wird auch das zuvor genannte System MAXQDA verwendet (Berger-Grabner, 2022; Gläser & Laudel, 2010). Der Einsatz dieser Software zur Datenanalyse steigert die Effizienz durch eine strukturierte qualitative Inhaltsanalyse. Zu den unterstützenden Funktionen gehören das Management der Transkriptionen, das Erstellen von Memos, das Management von Codes und Kategorien sowie die Nutzung quantitativer Variablen, wie beispielsweise Statistiken über die Häufigkeit verwendeter Wörter (Döring, 2023; Mayring, 2020). Ergänzend ist an dieser Stelle auch der AI Assist von MAXQDA zu erwähnen, dieser wurde für das Zusammenfassen der codierten Segmente verwendet. Das entwickelte Kategoriensystem sowie die zugehörigen Kodierregeln wurden deduktiv aus den im Interviewleitfaden operationalisierten Handlungsempfehlungen abgeleitet. Die Kodierung erfolgte direkt in der Software. Zur Strukturierung des Materials wurde ein hierarchisches Codierschema mit drei Ebenen verwendet. Auf der ersten Ebene wurden Hauptcodes definiert, die den einzelnen Phasen bzw. Fragenkomplexen des Leitfadens entsprechen. In Anlehnung an das CRISP-ML(Q)-Prozessmodell wurden diesen Hauptcodes spezifische Subcodes zugeordnet, die eine inhaltliche Differenzierung innerhalb der jeweiligen Phase ermöglichen. Für eine weiterführende Feinkodierung wurde eine dritte Ebene eingeführt, auf der ausgewählte Subcodes in differenzierte Unterkategorien aufgeschlüsselt wurden. Beispielhaft wird an dieser Stelle der Subcode „Vielversprechende KI-Anwendungen“ genannt. Das finale dreistufige Kategoriensystem mit Haupt- und Subcodes ist in der Tabelle 9 dargestellt.

Tabelle 9: Codierschema MAXQDA

Ebene 0	Ebene 1	Ebene 2	Beschreibung
Allgemeine Erfahrungen mit KI und EU AI Act	Vielversprechende KI-Anwendungen	Sonstige	Besonders vielversprechende KI-Anwendungen in Pflege, Diagnostik und Verwaltung und Beispiele für erfolgreichen KI-Einsatz in der Einrichtung.
		Triagesysteme	
		Extraktion von Informationen aus großen Datenmengen	
		KI-Anwendungen in der Radio- und Pathologie	
		Unterstützung in der Dokumentation	
	Herausforderungen	-	Herausforderungen bei Implementierung von KI-Systemen.
	Akzeptanz	Akzeptanz von KI-Systemen bei Patienten und Patientinnen	Akzeptanz von KI-Systemen bei Fachpersonal und Patienten und Patientinnen.
		Akzeptanz von KI-Systemen im klinischen Personal	
	Kenntnisstand EU AI Act	-	Vertrautheit der Befragten mit dem EU AI Act.
CRISP-ML(Q)	Geschäfts- und Datenverständnis	Identifikation von KI-Systemen	Sicherstellung, dass KI-Systeme erfasst und geprüft werden.
		Rollen im KI-Lebenszyklus	Verantwortlichkeiten gemäß EU AI Act (Provider/Deployer).
		Risikomanagementsystem	Existenz und Nutzung eines Risikomanagementsystems.
		Registrierungsanforderungen	Umsetzung von Registrierungsanforderungen und Grundrechts-Folgenabschätzungen.
		Überwachungssystem Provider	Überwachung der Provider hinsichtlich technischer Dokumentation und Standards.
	Datenvorbereitung	Datenqualität und Verwaltung	Prozesse zur Sicherstellung von Vollständigkeit und Konsistenz der Daten.
		Relevanz der Daten	Gewährleistung der Relevanz der Eingabedaten für den Anwendungsfall.
		Bias-Erkennung und Vermeidung	Prozesse zur Bias-Erkennung und -Vermeidung.
		Automatisierte Datenvalidierung	Kontinuierliche Überprüfung der Datenqualität.

	Modellierung	Rückverfolgbarkeit und Erklärbarkeit	Sicherstellung der Nachvollziehbarkeit von KI-Modellen im klinischen Umfeld.
		Verzerrungen und Diskriminierung	Mechanismen zur Erkennung diskriminierender Ergebnisse.
		Modellrisikoanalyse	Nutzung und Dokumentation formaler Risikoanalysen.
		Fairness-Metriken	Anwendung und Interpretation von Fairness-Kennzahlen.
	Evaluation	Vor-Markteinführungstests	Prüfung der Eignung des Systems vor Markteinführung.
		Technische Dokumentation	Erstellung vollständiger technischer Dokumentationen.
		Registrierung und Konformitätsbewertung	Prozesse bei Registrierung und Konformitätsbewertungen für Hochrisiko-KI-Systeme.
		Testarten und -methoden	Durchführung strukturierter Tests zur Robustheit.
	Bereitstellung	Risikokommunikation	Kommunikation verbleibender Risiken an medizinisches Personal und Patienten und Patientinnen.
		Kennzeichnung von KI-Systemen	Sichtbarkeit der Interaktion mit KI-Systemen für Patienten und Patientinnen.
		Überwachung und menschliche Kontrolle	Existenz von Not-Aus und Interventionsmöglichkeiten.
		Protokollierung und Meldung von Vorfällen	Dokumentation und Meldung kritischer Vorfälle.
		Ethik- und Governancekommission	Existenz eines Gremiums zur ethischen Prüfung.
		Interdisziplinäre Zusammenarbeit und Schulung	Organisation der Zusammenarbeit und Durchführung von Schulungen.
	Überwachung und Wartung		Langfristige Qualitätssicherung und Umgang mit Daten-/Modell-Drift.
	Querschnittsfragen	Organisatorische Verankerung	Zuständigkeiten für KI-Regulierungen und deren Organisation.
		Strategische Bedeutung	Bedeutung des EU AI Acts auf Führungsebene und Ressourcenverfügbarkeit.
		Herausforderungen und Verbesserungen	Identifizierung der größten Hindernisse und Unterstützungsbedarf.
		Wissensstand und Schulungsbedarf	Bedarf an zusätzlichen Schulungen und Informationen.
Zukunftsaussicht			Zukunftsvision für KI in Krankenhäusern und Zusammenarbeit mit Stakeholdern.

6 ERGEBNISSE

In diesem Kapitel werden die Experteninterviews aufbereitet und die entsprechenden Ergebnisse ausformuliert. Wie in Tabelle 7 und Tabelle 9 ersichtlich, wurde im Rahmen der Interviews zu Beginn auf Erfahrungen mit KI im Gesundheitswesen sowie den EU AI Act eingegangen. Deshalb beginnt auch die Aufbereitung der Ergebnisse mit diesem Fragenkomplex.

Vielversprechende KI-Anwendungen in Pflege, Diagnostik und Verwaltung

KI-Anwendungen in der Medizin umfassen Bereiche wie Triage-Systeme zur Priorisierung von Patienten und Patientinnen, Unterstützung bei der Wundbeurteilung in der Pflege sowie die Extraktion und Zusammenfassung von Informationen aus großen Datenmengen. Bildverarbeitende KI-Systeme finden bereits breite Anwendung in der Radiologie und Pathologie, etwa zur Erkennung von Auffälligkeiten in Bildaufnahmen. Auch im Bereich der Sprachverarbeitung bietet KI Potenzial, beispielsweise für die automatische Erstellung von Arztbriefen und Befunden. Insgesamt zeigt sich, dass KI-Technologien in der Medizin zunehmend an Bedeutung gewinnen und vielfältige Möglichkeiten zur Effizienzsteigerung und Unterstützung des medizinischen Personals bieten. Eine Übersicht zu Häufigkeit der Nennung bietet Tabelle 10.

Tabelle 10: Statistik zu vielversprechende KI-Anwendungen

Vielversprechende KI-Anwendungen	Nennung in Interviews	Prozent (%)
Extraktion von Informationen aus großen Datenmengen	5	100,00
KI-Anwendungen in der Radio- und Pathologie	5	100,00
Unterstützung in der Dokumentation	4	80,00
Triagesysteme	2	40,00
Sonstige	4	80,00

Herausforderungen von KI-Systemen in Krankenhäusern

Ein zentrales Problem betrifft die aktuelle Reife dieser Anwendungen. In hochsensiblen medizinischen Bereichen wie der Notaufnahme müssen KI-Systeme nicht nur präzise, sondern auch zuverlässig und nachvollziehbar arbeiten. Fehler, die in weniger kritischen Anwendungsfeldern hinnehmbar wären, können im klinischen Umfeld schwerwiegende Folgen haben. Darüber hinaus zeigen KI-gestützte Anwendungen wie automatische Transkriptionsdienste häufig Ungenauigkeiten. Besonders kritisch ist, dass KI- Systeme möglicherweise auch dann Ergebnisse liefern, wenn keine belastbare Evidenz vorliegt. Dies erfordert eine fortlaufende Überprüfung der Resultate durch fachlich geschultes Personal. Ein

weiterer zentraler Aspekt ist die Qualität der verwendeten Daten. Verzerrungen in den Trainingsdaten sowie die große Vielfalt medizinischer Vorgehensweisen, etwa zwischen verschiedenen Ländern oder Einrichtungen, erschweren die Entwicklung allgemein einsetzbarer KI-Lösungen erheblich.

Nicht zuletzt ist die Akzeptanz solcher Systeme im klinischen Alltag entscheidend. Studien zeigen, dass medizinisches Fachpersonal KI-Empfehlungen teils mit Zurückhaltung begegnet, selbst wenn diese objektiv bessere Ergebnisse erzielen. Ein solches Vertrauensdefizit kann die Integration von KI in die Praxis erheblich erschweren.

Ein interessanter Aspekt in diesem Zusammenhang ist die Frage, ob Krankenhäuser nicht dazu verpflichtet sein sollten, kostenintensivere, jedoch leistungsstärkere KI-Systeme einzusetzen, um eine Benachteiligung von Patienten und Patientinnen durch qualitativ minderwertige Technologien zu vermeiden.

Insgesamt wird deutlich, dass der erfolgreiche Einsatz von KI in Krankenhäusern eine ganzheitliche Herangehensweise erfordert, die technologische, datenspezifische, ethische und soziale Aspekte gleichermaßen berücksichtigt.

Akzeptanz von KI-Systemen bei medizinischem Fachpersonal und bei Patienten und Patientinnen

Die Akzeptanz von KI-Systemen im klinischen Personal ist insgesamt gegeben, jedoch differenziert zu betrachten. Während ein Großteil des Personals offen und neugierig gegenüber dem Einsatz von KI ist und diese als hilfreiche Unterstützung, insbesondere in belasteten Bereichen wie der Radiologie oder Pathologie, wahrnimmt, bestehen auch vereinzelt Vorbehalte. Besonders gut angenommen werden Systeme, die im Hintergrund automatisiert arbeiten, eine hohe Genauigkeit aufweisen und deren Entscheidungen nachvollziehbar sind. Auch technikfernere Mitarbeitende zeigen Begeisterung, wenn die Systeme zuverlässig funktionieren und Fehlalarme minimiert werden. Gleichzeitig ist das Bewusstsein für bereits bestehende KI-Anwendungen oft gering, und es herrscht Skepsis gegenüber möglichen Fehlentscheidungen. Daher ist es entscheidend, KI-Ergebnisse stets kritisch zu prüfen und ihre Plausibilität sicherzustellen. Die Akzeptanz von KI-Systemen bei Patienten und Patientinnen ist noch nicht vollständig geklärt. Viele Patienten und Patientinnen sind sich nicht bewusst, dass KI bereits im medizinischen Bereich eingesetzt wird. Wenn KI direkt eine Diagnose stellen würde, müsste dies für die Patienten und Patientinnen klar erkennbar sein. Derzeit liegt die Verantwortung noch vollständig bei den Ärzten und Ärztinnen, und die die Befunde erstellen. KI dient bisher lediglich als unterstützendes Tool. In manchen Fällen wird bereits vermerkt, dass ein KI-Tool bei der Befunderstellung geholfen hat. Wenn KI in der Routine verwendet wird, beispielsweise für automatische Volumenmessungen des Gehirns, wird dies in der Regel im Befundtext erwähnt.

Vertrautheit mit dem EU AI ACT

Die Auseinandersetzung mit dem EU AI Act stellt für Unternehmen und Institutionen im Gesundheitsbereich eine komplexe, aber notwendige Aufgabe dar. Insbesondere medizinische KI-Anwendungen gelten vermutlich als hochriskant und unterliegen strengen Anforderungen. Viele Einrichtungen befinden sich noch in einer frühen Phase der Umsetzung und analysieren derzeit, wie die neuen Vorgaben mit bestehenden Regularien wie der DSGVO und der MDR zusammenwirken. Für Anbieter von Medizinprodukten ergeben sich dadurch zusätzliche Problemstellungen. Unterstützt wird die Umsetzung durch Leitfäden und Fragenkataloge, etwa vom Johner Institut. Krankenhäuser bereiten sich aktiv vor, indem sie sich in nationalen Foren austauschen, gemeinsame Lösungen erarbeiten und Doppelarbeit vermeiden wollen – insbesondere bei der Bewertung identischer Systeme und Cloud-Lösungen. Trotz einiger noch offener Detailfragen wird das Gesetz grundsätzlich als positive und notwendige Entwicklung wahrgenommen, die langfristig für Klarheit im europäischen Rechtsraum sorgen soll.

Zukunftsausblick

Die Zukunftsaussichten für den Einsatz von KI-Systemen im medizinischen Bereich sind vielversprechend. Zum einen können Erklärbarkeit und Transparenz der Entscheidungsfindung in neuronalen Netzwerken verbessert werden, was für den medizinischen Kontext wichtig ist. Zum anderen können KI-basierte Entscheidungsunterstützungssysteme in Zukunft eine übergreifende Unterstützung bieten, insbesondere bei der Dokumentation. Darüber hinaus sehen Experten und Expertinnen die Notwendigkeit, Patientendaten öffentlich zugänglich zu machen und zum Wohl der Allgemeinheit einzusetzen. Allerdings stellen die Komplexität und der hohe Ressourcenaufwand für die Erfüllung eine Herausforderung dar, die es zu bewältigen gilt.

6.1 Phase 1: Geschäfts- und Datenverständnis

In der ersten Phase des Interviews lag der Schwerpunkt auf dem grundlegenden Verständnis der befragten Experten und Expertinnen hinsichtlich des Einsatzes von KI-Systemen in ihrem jeweiligen beruflichen Kontext. Im Fokus der Befragung standen insbesondere Kenntnisse zur Identifikation von KI-Anwendungen, zur Zuordnung von Verantwortlichkeiten über den gesamten Lebenszyklus hinweg sowie zum Umgang mit Risiken und bestehenden Rahmenbedingungen. Ziel war es, einen Einblick in bestehende Strukturen und Rollenverständnisse zu gewinnen.

Identifikation von KI-Systemen

Die Identifikation und Prüfung von KI-Systemen in Organisationen, insbesondere im Gesundheitswesen, stellt eine vielschichtige Herausforderung dar. Um sicherzustellen, dass alle eingesetzten KI-Systeme vollständig erfasst und überprüft werden, ist eine strukturierte und strategisch ausgerichtete Vorgehensweise erforderlich. Die befragten Personen betonen in diesem Zusammenhang die Bedeutung einer Erweiterung bestehender Strukturen, wie etwa des Verzeichnisses der Verarbeitungstätigkeiten (VVT), anstelle der Schaffung paralleler neuer Systeme. Dies fördert eine bessere Integration in bestehende Prozesse und reduziert zusätzlichen administrativen Aufwand. Ein zentrales Element im Umgang mit KI im

Gesundheitswesen ist die systematische Bewertung sämtlicher eingesetzter Softwarelösungen, einschließlich solcher, die nur indirekt als KI gelten oder in umfassendere Systeme integriert sind. Im Krankenhauskontext ist dabei insbesondere zu prüfen, ob KI-Produkte regulatorische Anforderungen erfüllen und beispielsweise über eine gültige CE-Kennzeichnung verfügen. Auch eingebettete KI-Module innerhalb zugelassener Medizinprodukte sind in die Risikobewertung einzubeziehen, vor allem dann, wenn sie funktional eng mit medizinischen Anwendungen verbunden sind. Eine wesentliche Herausforderung ergibt sich aus der Vielzahl eingesetzter Systeme sowie aus der fortlaufenden Veränderung der Softwarelandschaft durch Updates oder Erweiterungen, die zusätzliche KI-Funktionalitäten einführen können. Vor diesem Hintergrund ist es für Organisationen unerlässlich, geeignete Mechanismen zu etablieren, mit denen auch solche dynamischen Entwicklungen zuverlässig erfasst und bewertet werden. Als zielführend hat sich eine koordinierte Vorgehensweise auf nationaler Ebene erwiesen, wie sie etwa in österreichischen Krankenhausverbünden praktiziert wird. Einrichtungen stimmen sich hier in gemeinsamen Gremien ab, um eine einheitliche Bewertung vorzunehmen. Dies erscheint besonders sinnvoll, da häufig identische Systeme im Einsatz sind. Eine zentrale Steuerung der Softwareinstallationen durch die IT-Abteilungen unterstützt zudem dabei, unkoordinierte oder nicht autorisierte Anwendungen zu vermeiden.

Insgesamt zeigt sich, dass die Identifikation und Bewertung von KI-Systemen als fortlaufender, organisationsübergreifender Prozess zu verstehen ist, der strategische, technische und regulatorische Aspekte miteinander verbindet.

Rolle eines Krankenhauses im KI-Lebenszyklus

Im Rahmen des EU AI Acts übernehmen Krankenhäuser in der Regel die Rolle des Deployers, also derjenigen Organisation, die KI-Systeme in Betrieb nimmt, ohne aktiv in deren Entwicklung involviert zu sein. In dieser Funktion liegt ihre Verantwortung vor allem darin, sicherzustellen, dass die eingesetzten Systeme entsprechend ihrer Zweckbestimmung verwendet werden und zentrale Anforderungen wie Datenschutz, Ausfallsicherheit und Datensicherheit erfüllt sind. Dies schließt auch die Pflicht ein, Patienten und Patientinnen über die Datenverarbeitung aufzuklären. Es handelt sich hierbei um eine Aufgabe, die beim Betreiber liegt und für Anbieter weniger relevant ist. Die Rollenzuweisung im Lebenszyklus von KI-Systemen ist in der Praxis nicht immer eindeutig. Krankenhäuser treten nicht nur als Nutzer solcher Systeme auf, sondern können auch als Anbieter im Sinne des EU AI Acts gelten. Dies ist beispielsweise dann der Fall, wenn eigene Algorithmen entwickelt, bestehende Systeme funktional erweitert oder patientenbezogene Daten zur Entwicklung neuer Anwendungen herangezogen werden. In diesen Konstellationen übernehmen die Einrichtungen auch die damit verbundenen Pflichten, etwa im Hinblick auf die Konformitätsbewertung und die Dokumentation gemäß den gesetzlichen Vorgaben. Deutlich wird diese Problematik, wenn KI-Systeme in sensiblen medizinischen Bereichen eingesetzt werden, etwa zur automatisierten Klassifizierung von Fieberkurven. Ein ursprünglich als Verwaltungslösung konzipiertes System kann dadurch in den Anwendungsbereich eines Hochrisikosystems fallen und unterliegt somit erweiterten regulatorischen Anforderungen. Herausfordernd sind zudem Graubereiche, in denen Einrichtungen kleinere Softwareanpassungen vornehmen oder eigene, weniger komplexe KI-Lösungen entwickeln. In solchen Fällen verschwimmt die Grenze zwischen bloßer Nutzung (Deployer) und aktiver

Entwicklung (Provider), was die rechtliche Einordnung erschwert. Zusätzliche Komplexität entsteht in forschungsnahen Umgebungen wie Universitätskliniken, in denen sich wissenschaftliche Tätigkeit und klinischer Betrieb überschneiden. Die Nutzung anonymisierter Daten in Studien fällt nicht automatisch unter die Verantwortlichkeit als Provider oder Deployer, es sei denn, die entwickelten Systeme gelangen tatsächlich in die klinische Anwendung.

In der Praxis ist festzuhalten, dass auch Einrichtungen, die KI-Systeme lediglich anwenden, den Anforderungen des EU AI Acts unterliegen. Die Unschärfe zwischen Nutzung und Eigenentwicklung bleibt jedoch eine Herausforderung.

Einführung eines Risikomanagementsystems

Ein umfassendes Risikomanagementsystem, das den gesamten Lebenszyklus von KI-Systemen abdeckt und spezifisch auf deren Risiken im Hinblick auf Gesundheit, Sicherheit und Grundrechte eingeht, ist in vielen Einrichtungen bisher nur in Ansätzen vorhanden oder wird derzeit aufgebaut bzw. erweitert. Organisationen verfügen bereits über etablierte Risikomanagementsysteme, die auf allgemeinen Normen wie ISO 13485 (Qualitätsmanagement für Medizinprodukte), ISO 14971 (Risikomanagement für Medizinprodukte) oder ISO 27001 (Informationssicherheit) basieren. Diese Systeme bilden eine solide Grundlage für die Integration von KI-bezogenen Anforderungen, insbesondere in hochregulierten Bereichen wie der Gesundheitsversorgung, wo mit sensiblen Patientendaten gearbeitet wird. So etwa wird bei der Entwicklung geriatrischer Softwaremodule eine gemeinsame Risikobewertung angewendet, solange diese Module ähnliche medizinische Settings betreffen. Bei Erweiterungen, beispielsweise durch neue, deutlich abweichende Module, wäre hingegen eine separate Risikobewertung erforderlich. Andere Einrichtungen setzen ebenfalls auf bestehende, etablierte Systeme und streben an, diese, um KI-spezifische Komponenten zu erweitern, anstatt parallele Systeme aufzubauen. Dabei zeigt sich ein pragmatischer Ansatz: Die Unternehmen nutzen bereits vorhandene Risikostrukturen, etwa im Rahmen der Netzwerksicherheitsrichtlinie (NIS), und adaptieren diese auf die neuen Anforderungen durch KI-Anwendungen. Insbesondere bei Bild- und Kurvenanalysen ist dies bereits gängige Praxis. Für komplexere KI-Systeme wie Sprachmodelle ist eine genauere, möglicherweise separate Bewertung notwendig. Gleichzeitig wird deutlich, dass in der Praxis auch Grauzonen bestehen. Die Verantwortung für den sicheren und reflektierten Einsatz der Systeme liegt dort oft bei Anwender:innen selbst, die die Resultate kritisch hinterfragen müssen.

Insgesamt lässt sich festhalten, dass zwar wichtige Grundlagen für ein Risikomanagement über den gesamten Lebenszyklus von KI-Systemen existieren, etwa durch ISO-Normen und etablierte Informationssicherheits-Frameworks. Jedoch fehlt es vielerorts noch an einer konsequenten, KI-spezifischen Weiterentwicklung dieser Systeme, um den besonderen Problemstellungen systematisch und umfassend begegnen zu können.

Registrierungsanforderungen

Die Registrierungsanforderungen für medizinische Software umfassen verschiedene Verantwortlichkeiten für Anbieter und Betreiber. Anbieter müssen die Konformitätsbewertungen korrekt durchführen und einhalten. Betreiber haben wiederum andere Anforderungen, insbesondere im Zusammenhang mit der DSGVO. Dazu zählt auch die Durchführung einer Datenschutz-Folgenabschätzung bei der Inbetriebnahme der Software. In der Regel werden die

Betreiber hierbei vom Anbieter unterstützt, da dieser über entsprechende Erfahrung in der Umsetzung datenschutzrechtlicher Anforderungen verfügt.

Überwachungssystem Provider

Die Kontrolle und Überwachung der Einhaltung technischer Dokumentationen und Standards durch Deployer von KI-Systemen in der Gesundheitsversorgung befindet sich derzeit im Aufbau. Sie wird maßgeblich durch bestehende regulatorische Rahmenwerke wie die MDR, einschlägige ISO-Normen geprägt. Ein zentrales Anliegen ist dabei die Erwartung, dass benannte Stellen im Rahmen der Konformitätsbewertung etablierte Qualitäts- und Risikomanagementnormen wie ISO 13485 und ISO 14971 berücksichtigen. Es besteht die Hoffnung, dass der EU AI Act keine zusätzlichen parallelen Qualitätsanforderungen schafft, sondern vielmehr bestehende Standards erweitert oder konkretisiert. Die technische Dokumentation und deren Überprüfung sollen sich eng an diese bewährten Normen anlehnen.

Insgesamt lässt sich feststellen, dass die Kontrolle durch Deployer derzeit einem dynamischen Anpassungsprozess unterliegt. Zwar stehen in der Medizintechnik etablierte Normen und Verfahren zur Verfügung, doch bedarf es gezielter Erweiterungen und Konkretisierungen, um den besonderen Anforderungen von KI-Systemen im klinischen Kontext gerecht zu werden.

6.2 Phase 2: Datenvorbereitung

Der zweite Abschnitt des Interviews widmete sich der systematischen Aufbereitung und Qualitätssicherung von Daten für KI-Systeme. Im Fokus standen hierbei Fragen zur Data Governance, zur Relevanz und Aussagekraft der verwendeten Daten sowie zur Vermeidung von Verzerrungen (Bias). Diese Phase diente der Einschätzung, wie bewusst und strukturiert der Umgang mit Trainingsdaten erfolgt und welche Standards in der Praxis Anwendung finden.

Datenqualität und Datenverwaltung

Im Gesundheitsbereich stellen Datenqualität und -verwaltung zentrale Herausforderungen bei der Entwicklung und Anwendung von KI-Systemen dar. Die verfügbaren Daten sind in der Praxis oft unvollständig, unstrukturiert oder enthalten zum Schutz der Privatsphäre gerundete Angaben wie Alter oder Geschlecht. Um dennoch verlässliche Vorhersagen treffen zu können, müssen KI-Systeme lernen, mit diesen Einschränkungen umzugehen. Dabei kommen unterschiedliche Strategien zum Einsatz, etwa das Filtern seltener Diagnosen, das Zusammenfassen von Messwerten oder das Anonymisieren von Patientendaten. Zugleich ist es wichtig, dass die Datenverfügbarkeit auf individueller Ebene ausreichend gegeben ist. Die Datenanalyse verläuft im Zuge der Entwicklung meist in mehreren Stufen: Zunächst erfolgt eine interne Prüfung der vorhandenen Daten und ihrer Aussagekraft, darauf aufbauend wird ein Modell entwickelt und in einem Proof-of-Concept getestet. Bei Erfolg folgt ein Piloteinsatz, der durch die jeweiligen Anwendergruppen evaluiert wird. Besonders betont wird der Bedarf an domänenspezifischem Expertenwissen, um Verzerrungen in den Daten zu erkennen und potenzielle Biases zu minimieren. Ein standardisierter Prozess ist in diesem Zusammenhang nicht ausreichend, da jede KI-Anwendung individuell bewertet werden muss. Der Fokus richtet sich daher zunehmend auf

die strukturierte und nachvollziehbare Verwaltung der verwendeten Trainingsdaten. Anbieter sind gefordert, die Qualität und Genauigkeit dieser Daten transparent offenzulegen, da sie die Grundlage für die Leistungsfähigkeit und Verlässlichkeit der Modelle bilden. Aus fachlicher Sicht ist ein ausgewogenes Verhältnis zwischen Datenqualität, Datenschutz und praktischer Anwendbarkeit anzustreben. Der Aufwand ist in Relation zum erwarteten Nutzen jeweils anwendungsbezogen abzuwägen.

Die Verfügbarkeit von Daten für KI-Systeme kann unter anderem über bestehende Strukturen sichergestellt werden. In Krankenhäusern ist die Anonymisierung von Patientendaten beispielsweise durch etablierte Abrechnungsprozesse bereits gut umgesetzt. Diese bestehende Infrastruktur lässt sich auch für die Integration von KI-Systemen nutzen.

Relevanz der Eingabedaten (Deployer)

Die Relevanz der verwendeten Daten ist ein zentraler Faktor bei der Entwicklung und Anwendung von KI-Systemen im Gesundheitswesen. Auch wenn die Qualität medizinischer Daten in der Praxis nicht immer optimal ist, kommt es entscheidend darauf an, dass diese Daten die reale klinische Versorgungssituation abbilden. KI-Modelle sollten daher nicht ausschließlich auf idealisierten Datensätzen beruhen, sondern auf tatsächlich verfügbaren Informationen basieren, um eine hohe Praxistauglichkeit zu gewährleisten. Dabei kommt auch eine Feedback-Funktion zum Einsatz, die Daten für das Retraining generiert, um unterrepräsentierte Fälle besser zu erfassen und die Modelle kontinuierlich zu verbessern. Zugleich müssen regulatorische Anforderungen wie das Datensparsamkeitsprinzip der DSGVO sowie Vorgaben aus dem EU AI Act berücksichtigt werden, insbesondere bei der Nutzung von Cloud-Systemen. Hierbei stellt sich die Frage, wie stark Daten pseudonymisiert oder anonymisiert werden müssen, was im Gesundheitsbereich besonders sensibel ist. Ein weiterer zentraler Punkt ist die Ausgewogenheit und Breite des Datensatzes. Während ein enges gefasstes Modell, das sich nur auf wenige Merkmale konzentriert, zu Overfitting neigen und schlecht generalisieren kann, benötigt ein breit angelegtes Modell mehr Daten, um Vielfalt und unerwartete Muster abbilden zu können. Insbesondere bei seltenen Krankheiten rückt dieser Aspekt in der Vordergrund, da die geringe Fallzahl eine Herausforderung für das Training darstellt. Trotz dieser Limitationen ist der Einsatz von KI in der Regel sinnvoll. Ergänzend ist zu betonen, dass die kontinuierliche Weiterentwicklung von KI-Systemen nicht außer Acht gelassen werden darf. Viele Anwendungen benötigen Zeit sowie geeignete Entwicklungsbedingungen in einem geschützten Umfeld, um verlässlich funktionieren zu können.

Bias-Erkennung und Vermeidung

Generell erforderten die Erkennung und anschließende Vermeidung von Bias ein Zusammenspiel aus technischem und domänenspezifischem Wissen. Für die Entwicklung verlässlicher und fairer Modelle ist es wichtig, dass die zugrunde liegenden Trainingsdaten die tatsächliche Patientenstruktur möglichst gut abbilden. Dabei werden auch Merkmale wie Geschlecht und Alter angemessen berücksichtigt. So ist beispielsweise bekannt, dass medizinische Symptome wie bei Herzinfarkten geschlechtsspezifisch unterschiedlich auftreten können und diese Unterschiede entsprechend in den Modellen berücksichtigt werden müssen. Auch das Alter als Risikofaktor darf nicht pauschal bewertet werden, sondern muss differenziert in das Modell einfließen.

Im Hinblick auf das beschriebene Zusammenspiel ist es wesentlich, in dieser Phase ein interdisziplinäres Team einzubinden. Nur so können mögliche Verzerrungen aus unterschiedlichen Perspektiven erkannt und angemessen berücksichtigt werden. Dabei ist zu beachten, dass es keine universell anwendbare Lösung zur Vermeidung von Bias gibt, da jeder Anwendungsfall eigene Anforderungen und spezifische Risikopotenziale mit sich bringt.

Automatisierte Datenvalidierung

Automatisierte Datenvalidierung ist ein komplexes Thema, das vom jeweiligen Anwendungsfall abhängt. Einerseits kann man Datensätze nicht in einen starren Prozess packen, da jede Anwendung individuell bewertet werden muss. Andererseits wird es bei großen und vielfältigen Datenmengen deutlich komplizierter: Die Daten müssen strukturiert werden, es braucht Kooperationen mit anderen Institutionen und einheitliche Schnittstellen sowie Systeme müssen geschaffen werden. Dies macht den Entwicklungsprozess insgesamt viel schwieriger. Daher gibt es keine generelle Antwort, sondern es hängt immer vom spezifischen Anwendungsfall ab.

6.3 Phase 3: Modellierung

In der dritten Phase wurde der Prozess der Modellentwicklung näher betrachtet. Die Fragen zielten darauf ab, wie die Nachvollziehbarkeit, Fairness und Erklärbarkeit von KI-Modellen sichergestellt werden. Darüber hinaus wurde thematisiert, inwieweit Risikoprüfungen und Testmechanismen bereits im Entwicklungsprozess integriert sind.

Rückverfolgbarkeit und Erklärbarkeit

Die Rückverfolgbarkeit und Erklärbarkeit von KI-Systemen sind grundlegende Voraussetzungen für deren Einsatz im medizinischen Kontext, da sie das Vertrauen der Anwender:innen unmittelbar beeinflussen. Um die Entscheidungen von Algorithmen nachvollziehbar zu machen, ist eine transparente Darstellung der Entscheidungswege erforderlich. Technische Verfahren wie SHAP kommen hierbei zum Einsatz, da sie aufzeigen können, welche Eingabedaten zu bestimmten Vorhersagen geführt haben und wie stark diese jeweils gewichtet wurden. Gleichzeitig zeigt sich, dass der Grad der Erklärbarkeit je nach Fachgebiet unterschiedlich ausgeprägt ist. In der Radiologie beispielsweise bieten KI-Systeme visuelle Hinweise wie Heatmaps, die auf mögliche Gewebeveränderungen hinweisen. Die abschließende Beurteilung liegt jedoch weiterhin beim medizinischen Fachpersonal, wobei die KI als unterstützendes Instrument fungiert. Herausfordernd bleibt die Tatsache, dass KI-Modelle in Einzelfällen plausibel wirkende, aber nicht fundierte Ergebnisse liefern können. Dies unterstreicht die Notwendigkeit einer kritischen Auseinandersetzung mit der jeweiligen Systemlogik und deren Einbettung in den klinischen Entscheidungsprozess. Besonders bei DL-Anwendungen ist die Nachvollziehbarkeit der Entscheidungsprozesse eingeschränkt, da diese die relevanten Merkmale automatisch lernen und sich dabei auch auf unerwünschte oder irrelevante Muster stützen können, wie etwa ein Beispiel bei Lungenröntgenbildern zeigte, in denen nicht die Krankheit, sondern eine bestimmte Körperhaltung erkannt wurde. Während klassische ML-Verfahren durch manuelle Feature-Extraktion noch eine gewisse Transparenz erlauben, erschweren es komplexe neuronale Netzwerke, die internen Entscheidungsprozesse zu auditieren. Diese liefern zwar

Wahrscheinlichkeiten für bestimmte Befunde, aber in der Regel keine nachvollziehbaren Erklärungen, wie diese zustande kommen. Das Fehlen eines echten „Reasonings“ stellt somit eine wesentliche Hürde für die Validierung und den verantwortungsvollen Einsatz von KI in der Medizin dar.

Verzerrungen und Diskriminierung

Die Sicherstellung vollständiger und konsistenter Datensätze setzt fundiertes Fachwissen innerhalb der jeweiligen Domäne voraus. Ohne diese inhaltliche Expertise ist es kaum möglich, Verzerrungen in den Daten zu erkennen oder potenzielle Bias-Risiken angemessen zu bewerten.

Modellrisikoanalyse

In der Modellrisikoanalyse werden potenzielle Schwachstellen wie Fehlentscheidungen, Overfitting oder eine eingeschränkte Generalisierbarkeit systematisch identifiziert und bewertet. Besonders bei Anwendungen, die auf Daten aus Krankenhaus-Informationssystemen basieren, ist Sorgfalt geboten, da diese Daten häufig heterogen und stark vom jeweiligen Kontext abhängig sind. Ein methodischer Schwerpunkt liegt auf der sauberen Trennung von Trainings- und Testdaten sowie dem Einsatz von Verfahren wie der Kreuzvalidierung. Diese Maßnahmen tragen dazu bei, Modellverzerrungen zu erkennen und sowohl Overfitting als auch Underfitting zu vermeiden. Ergänzend sollte geprüft werden, inwieweit sich die entwickelten Modelle auf andere klinische Anwendungsbereiche übertragen lassen. Ein „Plug-and-Play“-Ansatz ist aufgrund unterschiedlicher Datenstrukturen in verschiedenen Krankenhäusern als gefährlich einzustufen. Stattdessen braucht es individuelle Anpassungen und Daten-Audits, um sicherzustellen, dass ein Modell in einem spezifischen klinischen Umfeld zuverlässig arbeitet. Hier unterscheiden sich Anbieter.

Ein weiteres Risiko besteht darin, dass Modelle auf Trainingsdaten basieren, die nicht repräsentativ für bestimmte Patientengruppen sind. Ein typisches Beispiel ist, dass das Modell mit Daten von Erwachsenen trainiert wurde, jedoch für Kinder zum Einsatz kommen soll. Solche Diskrepanzen können zu Fehldiagnosen führen, weshalb eine gründliche Validierung vor der produktiven Nutzung notwendig ist. Dabei müssen Modelle nicht nur besser sein als bestehende Verfahren, sondern auch deren Schwächen ausgleichen können. Ein Beispiel in der Kommunikation der Befundergebnisse ist der Einsatz eines sogenannten Consumer Health Vocabulary, das medizinische Fachbegriffe in allgemein verständliche Sprache überträgt. Dadurch kann das Risiko von Missverständnissen verringert werden, insbesondere im Hinblick auf potenziell unzutreffende oder erfundene Inhalte, wie sie bei großen Sprachmodellen gelegentlich auftreten. Insgesamt zeigt sich, dass eine sorgfältige und kontextspezifische Risikoanalyse, inklusive Qualitätsprüfung, Datenstrukturierung, Validierung und Nachjustierung, unabdingbar ist, um die Sicherheit und Verlässlichkeit von KI-Systemen im medizinischen Einsatz zu gewährleisten.

6.4 Phase 4: Evaluation

Die vierte Phase des Interviews befasste sich mit der systematischen Evaluation von KI-Systemen vor deren Einsatz. Thematisiert wurden unter anderem die technische Dokumentation, die Durchführung von Robustheitstests sowie die Einhaltung regulatorischer Anforderungen. Im Fokus stand dabei, wie Evaluationsprozesse im Pre-Market-Kontext verstanden und umgesetzt werden.

Vor-Markteinführungstests

Vor der Einführung von KI-Systemen im medizinischen Umfeld erfolgen in der Regel umfangreiche Tests, um Aspekte wie Wirksamkeit, Sicherheit und praktische Anwendbarkeit zu überprüfen. Eine wichtige Maßnahme in diesem Zusammenhang ist die Anpassung oder das Nachtrainieren der Modelle auf retrospektive, einrichtungsspezifische Daten. Idealerweise geschieht dies über einen längeren Zeitraum, um eine ausreichende Sensitivität und Spezifität zu erzielen und die Verlässlichkeit der Vorhersagen zu stärken. Auch wenn es sich bei vielen Anwendungen um Inhouse-Lösungen handelt, werden dennoch systematische Vorabtests durchgeführt. Diese beinhalten unter anderem Studien, in denen die neuen KI-basierten Modelle mit etablierten Verfahren wie manuellen Scoring-Systemen verglichen werden. Dabei wird überprüft, ob die KI-Modelle entweder überlegen oder zumindest gleichwertig sind und ob durch ihren Einsatz ein konkreter Mehrwert in Form von höherer Genauigkeit, Zeitersparnis oder reduziertem Aufwand entsteht. Die Einbindung der späteren Nutzer:innen, insbesondere des medizinischen Fachpersonals, sollte frühzeitig erfolgen, da deren Perspektiven wesentlich dazu beitragen, Anforderungen realitätsnah zu erfassen und den praktischen Nutzen der Systeme zu bewerten. Dabei steht nicht allein die technische Funktionsfähigkeit im Vordergrund, sondern auch der Beitrag zur Verbesserung des klinischen Alltags.

Für den Einsatz von KI-Systemen gelten keine gänzlich neuen Regelwerke. Vielmehr orientieren sich die Anforderungen an etablierten Bewertungskriterien, wie sie bereits seit Jahren bei der Einführung von IT-Lösungen im Gesundheitswesen zur Anwendung kommen. Dazu zählen unter anderem Zweckmäßigkeit, Verlässlichkeit und ein erkennbarer Mehrwert gegenüber bestehenden Verfahren.

Technische Dokumentation (Provider)

Die technische Dokumentation stellt einen zentralen Bestandteil bei der Entwicklung, Evaluierung und Implementierung von KI-Systemen im medizinischen Bereich dar. Aufgrund der strengen regulatorischen Anforderungen, etwa durch die europäische MDR sowie die einschlägigen Normen DIN EN ISO 20304 und DIN EN ISO 82304, ist für Medizinprodukteanbieter bereits heute ein hoher Dokumentationsstandard etabliert. Es wird deshalb nicht davon ausgegangen, dass in diesem Bereich zusätzliche regulatorische Anforderungen entstehen, da die bestehenden Vorgaben eine detaillierte und engmaschige Dokumentation bereits abdecken. Die technische Dokumentation muss sowohl qualitativ hochwertig als auch wissenschaftlich fundiert und praxisrelevant sein. Sie sollte unter anderem Informationen zu verwendeten Daten, eingesetzten Algorithmen und Modellversionen sowie Datenschutzmaßnahmen enthalten. Ein gut strukturierter Datenmanagement-Plan ist essenziell, um die Herkunft, Nutzung, Speicherung und

Sicherheit der Daten transparent darzulegen. Die Versionierung der Trainingsdatensätze und Modelle trägt zur Rückverfolgbarkeit bei und erleichtert die Identifikation möglicher Schwachstellen. Besonders in der Evaluierungsphase spielt die Dokumentation eine entscheidende Rolle, etwa zur Prüfung, ob die Trainingsdaten repräsentativ sind und welche Risiken im Modell berücksichtigt wurden. Die Erstellung interner Referenzdatensätze kann hier sinnvoll sein, erfordert jedoch eine genaue Auseinandersetzung mit deren Verarbeitung und Nutzung. Dies ist vor allem für Anbieter von KI-Systemen relevant, während sich für Deployer die Frage stellt, inwieweit eine eigene Evaluierung erforderlich ist. Darüber hinaus ist es wichtig zu verstehen, dass viele KI-Modelle, insbesondere von großen Anbietern, als vortrainierte Versionen geliefert werden, die nur durch gezielte Updates verändert werden. Ein automatisches Weiterlernen im Einsatzkontext findet aus Sicherheitsgründen nicht statt. Vor einer Implementierung erfolgt bei bestimmten Anbietern ein standortspezifischer Data Audit, sodass nicht zwingend alle Standorte mit demselben Software- oder Datenstand arbeiten müssen. Insgesamt bildet die technische Dokumentation die Grundlage für Transparenz, Sicherheit und Nachvollziehbarkeit im Einsatz medizinischer KI-Systeme. Dies gilt sowohl intern bei der Entwicklung als auch extern bei der Bewertung und Zulassung.

Registrierung und Konformitätsbewertung (Provider)

Die Registrierung und Konformitätsbewertung von KI-Systemen im Gesundheitswesen erfolgt im Rahmen der bestehenden regulatorischen Strukturen für Medizinprodukte, insbesondere der MDR. KI-Systeme, die als Medizinprodukte eingestuft werden, unterliegen somit denselben Anforderungen wie klassische medizinische Geräte, einschließlich der Pflicht zur CE-Kennzeichnung. Diese Zertifizierung stellt sicher, dass die Systeme den geltenden Sicherheits- und Leistungsstandards entsprechen. Auch unter dem neuen EU AI Act wird erwartet, dass bestehende benannte Stellen, die bereits für die CE-Zertifizierung von Medizinprodukten zuständig sind, künftig auch für die Bewertung und Überwachung von KI-Systemen verantwortlich sind. Dies umfasst sowohl Erstprüfungen als auch Rezertifizierungen und laufende Überwachungsaudits und soll den Krankenhäusern die Handhabung solcher Systeme erleichtern. In der Praxis ist es dabei weniger relevant, ob die KI-Komponente automatisierte Vorschläge macht oder lediglich unterstützende Funktionen wie Kontrastanpassung übernimmt. Entscheidend ist, dass das Gesamtsystem zuverlässig funktioniert und die medizinische Aufgabe erfüllt, unabhängig davon, ob einzelne Prozesse manuell oder automatisiert ablaufen. Für KI-Systeme, die sich noch in der Erprobungsphase befinden und noch keine vollständige Zulassung besitzen, gelten klare Regelungen zur Nutzung im Studienkontext. Solche Systeme müssen vorab registriert und bei den zuständigen Behörden gemeldet werden. Während der Studienlaufzeit werden auftretende Vorkommnisse und mögliche Funktionsmängel systematisch dokumentiert und bewertet, um den regulatorischen Anforderungen zu entsprechen und eine spätere Konformitätsbewertung fundiert vorzubereiten.

Testarten und -methoden

Die Testarten und -methoden zur Evaluierung von KI-Systemen im medizinischen Bereich folgen einem mehrstufigen, sorgfältig abgestimmten Prozess, der auf höchste Genauigkeit und Patientensicherheit ausgerichtet ist. Im Anschluss an die zuvor beschriebenen Maßnahmen

werden Pilotversuche mit Testpatienten und -patientinnen durchgeführt, bei denen Ärzte und Ärztinnen die von der KI generierten Ergebnisse prüfen und mit eigenen Einschätzungen vergleichen. Im Anschluss erfolgt eine Validierung dieser Befunde durch interdisziplinäre Teams. Um auch die Perspektive von Patienten und Patientinnen angemessen zu berücksichtigen, wird ihr Feedback eingeholt. Dabei stehen Aspekte wie Verständlichkeit, Akzeptanz und Vertrauen in die Systeme im Fokus. Dieser Prozess wird schrittweise weiterentwickelt, bis ein qualitativ zufriedenstellendes Ergebnis erreicht wird. Gerade im medizinischen Kontext sind hohe Anforderungen an die Genauigkeit unerlässlich. Selbst eine Treffergenauigkeit von 95 Prozent gilt häufig als unzureichend, da bereits geringe Abweichungen gravierende Folgen nach sich ziehen können. Darüber hinaus sind Datenschutzaspekte besonders relevant, insbesondere im Hinblick auf die sensible Kommunikation mit Patienten und Patientinnen.

Ein weiterer kritischer Punkt betrifft die Aussagekraft von Testverfahren. Anbietergeführte klinische Studien werden oft dafür kritisiert, nicht ausreichend repräsentativ zu sein und die tatsächliche Leistungsfähigkeit der Systeme im klinischen Alltag zu überschätzen. Um realistischere Bewertungen zu ermöglichen, wird der Aufbau unabhängiger und standardisierter Testdatensätze gefordert – vergleichbar mit Crashtests in der Automobilindustrie. Diese müssten jedoch vor missbräuchlicher Nutzung, etwa zum gezielten Training durch Anbieter, geschützt werden. Aus diesem Grund wird die Einrichtung einer neutralen Instanz, beispielsweise einer EU-Behörde oder einer wissenschaftlichen Gesellschaft, empfohlen, die solche Testdatensätze verwaltet, kontrolliert und objektive Benchmarks ermöglicht.

6.5 Phase 5: Bereitstellung

Im fünften Teil des Leitfadens stand die Phase der Inbetriebnahme und des produktiven Einsatzes von KI-Systemen im Vordergrund. Dabei wurden Fragen zur Transparenz gegenüber Nutzer:innen, zu Sicherheitsvorkehrungen sowie zu Meldepflichten und menschlicher Kontrolle gestellt. Die Auswertung dieses Abschnitts gibt Aufschluss darüber, wie KI-Systeme in Organisationen verantwortungsvoll eingeführt werden.

Risikokommunikation

Die Risikokommunikation sowie der verantwortungsvolle Umgang mit sicherheitsrelevanten Aspekten zählen zu den grundlegenden Anforderungen beim Einsatz von KI-Systemen im Gesundheitswesen. Vor dem Hintergrund hoher Anforderungen an Datenschutz und Informationssicherheit setzen viele Krankenhäuser auf On-Premise-Installationen. Anbieter reagieren darauf mit dem Einsatz klar definierter Sicherheitszertifizierungen, deren Einhaltung von den betreibenden Einrichtungen schriftlich bestätigt werden muss. Auf diese Weise wird die Verantwortung für die Datensicherheit formell auf die Betreiber übertragen – eine pragmatische Lösung, um regulatorischen und haftungsrechtlichen Vorgaben zu entsprechen.

In der Anwendungspraxis wird häufig betont, dass KI-Systeme derzeit keine autonomen Entscheidungen treffen, sondern lediglich Empfehlungen aussprechen, die vom medizinischen Fachpersonal überprüft und verantwortet werden. Dies verringert zwar das unmittelbare Risiko für Patienten und Patientinnen, macht potenzielle Fehlfunktionen jedoch nicht weniger relevant.

Kritisch angemerkt wird insbesondere, dass Fehler oder Auffälligkeiten nur selten systematisch an die Anbieter zurückgemeldet werden und dass anwendungsbezogene Daten bislang kaum in die Weiterentwicklung der Modelle einfließen.

Ein weiteres Problemfeld stellt die unzureichende Risikokommunikation gegenüber Patienten und Patientinnen dar. Viele sind sich der potenziellen Risiken und Funktionsweisen von KI-Systemen nicht bewusst. Um dem entgegenzuwirken, werden in Forschungsprojekten Ärzte und Ärztinnen in Annotierungsaufgaben eingebunden und Patienten und Patientinnen gezielt zu ihrer Wahrnehmung, zum Verständnis von Befunden sowie zu ihrer Bereitschaft, persönliche Daten bereitzustellen oder für solche Leistungen zu zahlen, befragt. Ziel ist es, die Transparenz zu erhöhen und die Kommunikation über Risiken und Nutzen solcher Systeme zu verbessern. In den Einrichtungen der befragten Personen gibt es aktuell noch keine KI-Systeme im Einsatz, die in direktem Kontakt mit Patienten und Patientinnen stehen. Viele Anwendungen sind ausschließlich für medizinisches Fachpersonal konzipiert. Dennoch ist klar: Je näher KI-Systeme an den Patientenkontakt rücken, desto wichtiger wird eine klare, verständliche und vertrauensbildende Risikokommunikation.

Kennzeichnung von KI-Systemen

Die Kennzeichnung von KI-Systemen im medizinischen Bereich spielt eine wichtige Rolle für Transparenz, Sicherheit und regulatorische Konformität. Unternehmen sind verpflichtet, ihre Systeme gemäß der MDR korrekt zu kennzeichnen. Dazu gehört insbesondere die CE-Kennzeichnung sowie der explizite Hinweis darauf, dass es sich um ein Software-Medizinprodukt handelt. Zusätzlich wird geprüft, ob ein gesonderter Hinweis auf den Einsatz von KI erforderlich ist, da dieser für Nutzer:innen oft nicht ohne weiteres erkennbar ist. Anbieter warten hierzu noch auf Rückmeldungen der zuständigen Behörden, um ihre Kennzeichnung entsprechend anzupassen. Seit dem 1. August besteht außerdem eine generelle Verpflichtung, Nutzer:innen darauf hinzuweisen, wenn KI-Systeme eingesetzt werden. In bestimmten Anwendungsbereichen wie der medizinischen Bildverarbeitung ist dies meist gegeben, da die KI-Funktionalitäten bewusst aktiviert werden. Dennoch wird betont, dass die Kennzeichnung teils noch stärker hervorgehoben werden sollte, um die Transparenz weiter zu erhöhen und das Bewusstsein für den KI-Einsatz zu schärfen. Gleichzeitig gibt es auch kritische Stimmen, die bezweifeln, dass zusätzliche Hinweise tatsächlich zu einem besseren Verständnis führen. Bereits in der Vergangenheit sei nicht jede eingesetzte Technologie oder deren Funktionsweise für die Nutzer:innen vollständig nachvollziehbar gewesen.

Überwachungssysteme und menschliche Kontrolle

Die eingesetzten Systeme sind explizit als Entscheidungsunterstützung und nicht als autonome Entscheidungsträger gedacht. Sie sollen Pflegekräfte und Ärzte sowie Ärztinnen dabei unterstützen, Risiken frühzeitig zu erkennen und fundierte Entscheidungen zu treffen. Die finale Diagnose und therapeutische Entscheidung bleiben jedoch stets in menschlicher Hand.

Die menschliche Expertise wird dabei nicht ersetzt, sondern durch technologische Unterstützung ergänzt. Insbesondere im Pflegealltag, wo nicht immer eine lückenlose Betreuung gewährleistet ist, kann die KI als „Sicherheitsnetz“ dienen. Gleichzeitig muss sichergestellt sein, dass das medizinische Fachpersonal die Plausibilität der Empfehlungen nachvollziehen kann. Daher ist die

Transparenz der Entscheidungsprozesse essenziell, auch wenn diese bei komplexeren bildgebenden Verfahren technisch anspruchsvoller umzusetzen ist.

Ein wichtiger Bestandteil der Qualitäts- und Sicherheitskontrolle ist die Einhaltung einschlägiger Zertifizierungen wie ISO 27001 (Informationssicherheit) und ISO 13485 (Qualitätsmanagement für Medizinprodukte). Diese tragen dazu bei, dass selbst entwickelte Softwarelösungen robust, sicher und kontrollierbar bleiben.

Ursprünglich war der Zugang zu medizinischen Befunden ausschließlich den behandelnden Ärzten und Ärztinnen vorbehalten. Inzwischen ist es jedoch gängige Praxis, dass auch Patienten und Patientinnen direkten Zugriff erhalten. Dies kann jedoch zu Missverständnissen oder Verunsicherung führen, insbesondere bei komplexen oder schwerwiegenden Diagnosen. Um dem entgegenzuwirken, besteht die Möglichkeit, bestimmte Befunde vorübergehend zu sperren und erst nach einem erläuternden Gespräch mit dem medizinischen Fachpersonal freizugeben. Die Verantwortung für medizinische Entscheidungen verbleibt unverändert beim Menschen. Die ethische Auseinandersetzung mit Risiken wie Fehldiagnosen und die kontinuierliche Diskussion über den verantwortungsvollen Einsatz von KI verdeutlichen, wie wichtig eine reflektierte und kontrollierte Nutzung solcher Systeme im Gesundheitswesen ist.

Protokollierung und Meldung von Vorfällen

Die Protokollierung und Meldung von Vorfällen bei medizinischen KI-Systemen ist ein zentraler Bestandteil der regulatorischen Überwachung und Qualitätssicherung. Anbieter von Medizinprodukten sind verpflichtet, jedes relevante Vorkommnis an die zuständigen Behörden zu melden. In Österreich wäre es beispielsweise das Bundesamt für Sicherheit im Gesundheitswesen. Diese Meldungen dienen der kontinuierlichen Marktüberwachung durch die Behörden. Gleichzeitig sind die Anbieter dazu verpflichtet, auch Vorfälle mit ähnlichen Produkten anderer Anbieter zu beobachten, um mögliche Risiken frühzeitig zu erkennen und gegebenenfalls präventive Maßnahmen zu ergreifen.

Historisch betrachtet gab es für die Protokollierung und das Feedback aus der Praxis lange Zeit keine klar geregelten Prozesse. Viele Entscheidungen darüber, ob und wie Probleme in neue Produktgenerationen einfließen, lagen bei den Unternehmen selbst. Diese Praxis ist teilweise noch heute zu beobachten, insbesondere aufgrund der langen Entwicklungszyklen von der Forschung bis zum klinischen Einsatz. Bei nicht-medizinischen KI-Systemen wie großen Sprachmodellen bestehen aktuell noch erhebliche Lücken im Umgang mit Fehlverhalten oder Fehlfunktionen. Ein standardisierter Prozess zur Meldung von Vorfällen existiert in diesen Bereichen oft nicht.

Insgesamt zeigt sich, dass für medizinisch eingesetzte KI-Systeme ein klar regulierter Rahmen für die Protokollierung und Meldung von Vorfällen existiert, der jedoch in der Praxis weiterentwickelt und konsequenter umgesetzt werden muss. Für nicht-medizinische KI-Anwendungen besteht hingegen noch deutlicher Nachholbedarf bei der Entwicklung geeigneter Melde- und Reaktionsmechanismen.

Ethik- und Governancekommission

Die Ethik- und Governancekommission befasst sich mit ethischen Fragen, die sich ergeben, wenn neue, vielversprechende Technologien entwickelt werden, die deutlich besser sind als bisherige menschengemachte Lösungen. Es stellt sich die Frage, ob Krankenhäuser und andere Einrichtungen verpflichtet sind, diese kommerziellen Lösungen anzuschaffen, oder ob dies den Patienten und Patientinnen vorenthalten werden darf. Die Kommission muss diese komplexen ethischen Fragestellungen beantworten, sollte eine solche Situation tatsächlich eintreten.

Interdisziplinäre Zusammenarbeit und Schulung

Die Zusammenarbeit zwischen Anbietern, Betreibern und weiteren Beteiligten ist ein entscheidender Erfolgsfaktor für den sicheren und effektiven Einsatz von KI-Systemen im Gesundheitswesen. Anbieter bieten hierfür meist einen mehrstufigen Support an, bei dem der Betreiber der Software als erster Ansprechpartner fungiert und Anfragen bei Bedarf an den Second-Level-Support des Anbieters weiterleitet. Zusätzlich werden interaktive Erklärkomponenten, etwa über die Unternehmenswebsite, bereitgestellt, um den Anwender:innen den Umgang mit der Software zu erleichtern. Zentral ist dabei auch die interdisziplinäre Zusammenarbeit zwischen verschiedenen medizinischen Einrichtungen. Der Austausch über Erfahrungen und Bewertungen mit bestimmten KI-Systemen kann dazu beitragen, einheitlichere Einschätzungen zur Wirksamkeit, Sicherheit und Praxistauglichkeit zu entwickeln. Ebenso bedeutsam ist die gezielte Schulung des medizinischen Personals. Die Nutzer und Nutzerinnen müssen in der Lage sein, die Systeme sicher und korrekt anzuwenden, um die Patientensicherheit nicht zu gefährden. Nur durch fundierte und praxisorientierte Schulungsmaßnahmen lässt sich sicherstellen, dass KI-Systeme nicht nur technisch funktionieren, sondern auch im klinischen Alltag verlässlich und sinnvoll genutzt werden können.

Zukünftig sollen auch die Perspektiven der Patienten und Patientinnen stärker eingebunden werden. Geplant sind Plattformen, über die Patienten und Patientinnen Rückfragen stellen und Feedback geben können. Dieses soll in die Weiterentwicklung der Systeme einfließen, etwa in neue Softwareversionen, und somit langfristig zur Verbesserung der Nutzererfahrung und Systemakzeptanz beitragen.

6.6 Phase 6: Überwachung und Wartung

Die sechste Phase fokussierte auf die langfristige Betreuung und Qualitätssicherung von KI-Systemen im laufenden Betrieb. Behandelt wurden Themen wie Monitoring, Umgang mit Daten- und Modell-Drift, Notfallplanung sowie die Einbindung ethischer Leitlinien. Ziel war es, zu verstehen, wie nachhaltig KI-Systeme gemanagt und weiterentwickelt werden.

Qualitätsmanagementsystem (Provider)

Die Leistungsfähigkeit der Modelle muss über den gesamten Lebenszyklus hinweg erhalten bleiben, auch im Hinblick auf Veränderungen im medizinischen Umfeld. Zu diesen Veränderungen zählen unter anderem neue Medikamente oder Therapieansätze. Aus diesem Grund ist eine kontinuierliche Wartung erforderlich. Einige Anbieter ermöglichen daher ein

regelmäßiges Nachtrainieren der Modelle, beispielsweise in halbjährlichen Abständen, sofern sich dadurch nachweisbare Verbesserungen erzielen lassen.

Zur kontinuierlichen Qualitätssicherung können Provider zusätzlich eine quartalsweise Auswertung der Patientenerkennungsraten etablieren, deren Ergebnisse an die jeweiligen Krankenhäuser zurückgemeldet werden. Ergänzend dazu können einmal jährlich Feedbackgespräche mit den Standorten helfen, um auch qualitative Erfahrungen und Verbesserungsvorschläge in die Weiterentwicklung einfließen zu lassen.

Rechtlich gesehen ist der Einsatz selbstlernender Systeme im laufenden Betrieb derzeit nicht zulässig. Zwar besteht kein ausdrückliches Verbot, doch lassen die Vorgaben des MDR und weiterer regulatorischer Rahmenbedingungen darauf schließen, dass adaptive Systeme, die sich ohne erneute Prüfung und Zulassung verändern, nicht rechtskonform eingesetzt werden dürfen. Entsprechend werden KI-Systeme in der Regel als fertig trainierte Produkte mit klar definierter Versionsnummer ausgeliefert. Sie lernen nach der Markteinführung nicht selbstständig weiter.

Gemäß den EU-Vorgaben sollten bzw. müssen zudem Informationen zu den verwendeten Trainingsdatensätzen, etwa zu deren Umfang und Verteilung, offengelegt werden, insbesondere im Rahmen wissenschaftlicher Studien oder Zulassungsprozesse. Dies dient der Transparenz und ermöglicht eine fundierte Bewertung der Systemleistung sowie der zugrundeliegenden Datenbasis.

6.7 Querschnittsfragen

Ergänzend zu den phasenspezifischen Themen wurden sogenannte Querschnittsfragen gestellt, die über mehrere Phasen hinweg von Bedeutung sind. Diese Fragen zielten unter anderem auf die organisatorische Verankerung von KI, strategische Überlegungen, Schulungsbedarfe und übergreifende Themen ab. Sie bieten eine zusätzliche Perspektive auf strukturelle und kulturelle Rahmenbedingungen in der Umsetzung des EU AI ACT.

Organisatorische Verankerung

Die organisatorische Verankerung von KI- und IT-gestützten Systemen im Gesundheitswesen ist komplex und erfordert eine enge Zusammenarbeit verschiedener Akteure innerhalb einer Einrichtung. In der Praxis zeigt sich, dass die Zuständigkeiten für IT-gestützte Medizinsysteme nicht immer eindeutig geregelt sind. Es handelt sich um ein Querschnittsthema, bei dem sowohl IT-Abteilungen als auch Bereiche der Medizintechnik involviert sind. Dieser Umstand ist vergleichbar zum Risikomanagement. Nur durch ein gemeinsames Verständnis und geteilte Verantwortung kann sichergestellt werden, dass Systeme effektiv, sicher und regelkonform eingeführt und betrieben werden.

In vielen Fällen liegt die konkrete Entscheidung über den Einsatz eines bestimmten Systems bei den jeweiligen medizinischen Fachabteilungen. Diese beurteilen die fachliche Relevanz und Praxistauglichkeit auf Basis ihres Arbeitsalltags. Die rechtliche und vertragliche Abwicklung, etwa im Hinblick auf Datenschutz, Konformität und Vertragsgestaltung, wird anschließend von der Beschaffungsabteilung übernommen. Dabei steht die medizinische Notwendigkeit der Systeme

im Vordergrund. Die IT-Abteilung übernimmt in der Regel die Verantwortung für die technische Umsetzung, insbesondere wenn es um Systemintegration, Datenverarbeitung und Infrastrukturfragen geht. Allerdings kann es auch zu Spannungsfeldern kommen, insbesondere in Situationen, in denen Sicherheitsbedenken gegenüber bestimmten Technologien wie Cloud-Lösungen bestehen. Solche Vorbehalte können Einführung und Nutzung moderner Lösungen erheblich einschränken.

Insgesamt zeigt sich, dass eine erfolgreiche organisatorische Verankerung von KI-gestützten Medizinsystemen eine strukturierte und übergreifende Abstimmung erfordert.

Strategische Bedeutung

Die Bedeutung regulatorischer Vorgaben entwickelt sich zunehmend zu einer strategischen Fragestellung für Unternehmen und Gesundheitseinrichtungen im medizinischen Bereich, insbesondere im Zusammenhang mit dem EU AI Act. Anbieter von KI-basierten Medizinprodukten sehen sich bereits heute mit einem erheblichen Ressourcenaufwand konfrontiert, um bestehende Anforderungen wie jene der MDR oder der DSGVO zu erfüllen. Ein Unternehmen gibt an, derzeit rund 20 Prozent seiner Ressourcen für regulatorische Aufgaben einzusetzen. Mit dem Inkrafttreten des EU AI Acts wird erwartet, dass dieser Anteil auf etwa 25 Prozent ansteigt.

Obwohl vielfach bereits etablierte Compliance-Strukturen bestehen, rechnen viele Akteure mit einem zusätzlichen Bedarf an personellen und organisatorischen Kapazitäten, um den neuen rechtlichen Vorgaben gerecht zu werden. Auch aufseiten der Gesundheitseinrichtungen wird der EU AI Act zunehmend als strategisch relevantes Thema wahrgenommen. Insbesondere in Österreich setzen Krankenhäuser verstärkt auf koordinierte Ansätze, um die Umsetzung effizient und ressourcenschonend zu gestalten. Übergreifende Gremien wie das IT-Forum der österreichischen Krankenhausträger dienen dabei als Plattform für Erfahrungsaustausch, Zuständigkeitsklärung und die gemeinsame Interpretation der neuen Anforderungen.

Der EU AI Act ist somit nicht nur als rechtliche Verpflichtung zu verstehen, sondern rückt zunehmend in den Fokus strategischer Überlegungen rund um den Einsatz von KI im Gesundheitswesen. Auch wenn die Umsetzung mit einem erheblichen Aufwand verbunden ist, wird dieser als notwendig angesehen, um langfristig sichere, nachvollziehbare und vertrauenswürdige KI-Anwendungen zu gewährleisten.

Herausforderungen und Verbesserungsmöglichkeiten

Eine der derzeit größten Hürden liegt in der fehlenden Rechtssicherheit. Die Auslegung und praktische Umsetzung bestehender Regelwerke, insbesondere im Zusammenspiel zwischen der MDR und dem EU AI Act, ist bislang uneinheitlich. Unterschiedliche Interpretationen durch Prüfstellen und Auditoren führen zu Unklarheiten, die sich unmittelbar auf die Produktentwicklung und Zulassungsverfahren auswirken. Diese Unsicherheiten erschweren nicht nur die Planung, sondern auch die Verlässlichkeit regulatorischer Prozesse, was den Marktzugang für neue KI-Systeme deutlich verzögern kann. Der Wunsch nach einer offiziellen, verbindlichen Klärung seitens der gesetzgebenden Stelle ist daher groß, um Einrichtungen eine klare Orientierung und mehr Planungssicherheit zu geben. Auch bei der praktischen Umsetzung neuer

Sicherheitsstandards zeigen sich Konflikte. So müssen Systeme mit bekannten Schwachstellen zwar rasch ersetzt werden, jedoch lassen gesetzliche Regelungen dies oft nicht ohne weiteres zu. Gleichzeitig ergibt sich die Frage, wie die hohen Umsetzungskosten gegen potenzielle Strafandrohungen abgewogen werden sollen, was dazu führt, dass andere wichtige Aufgaben zurückgestellt werden müssen. Technisch gesehen liegt eine große Herausforderung in der Sicherstellung von Datenschutz und Datenhoheit. Es muss verhindert werden, dass sensible Patientendaten zweckentfremdet werden oder unkontrolliert in Drittsysteme gelangen.

Insbesondere große Institutionen und der öffentliche Sektor geraten organisatorisch häufig an ihre Grenzen, da aufwendige Dokumentationspflichten, komplexe Risikomanagementprozesse und die Angst vor Fehlern nicht selten zu einem lähmenden Zögern führen. Ein flexibler und pragmatischer Ansatz, der Herausforderungen im Verlauf des Umsetzungsprozesses aufgreift, anstatt bereits im Vorfeld sämtliche potenzielle Risiken vollständig lösen zu wollen, wäre hier deutlich zielführender.

Wissensstand und Schulungsbedarf

Im Bereich der Schulungsangebote wird sich voraussichtlich bald ein Markt etablieren, da bestehende Berater:innen im Medizinproduktebereich ihre Expertise auf den EU AI Act ausweiten, wobei die Qualität dieser Beratungen erst die Praxis zeigen wird.

7 DISKUSSION

In diesem Kapitel werden die in Kapitel 4 theoretisch abgeleiteten Handlungsempfehlungen kritisch mit den im Rahmen der Experteninterviews gewonnenen Erkenntnissen verglichen. Ziel ist es, Übereinstimmungen, Ergänzungen sowie etwaige Widersprüche zwischen Theorie und Praxis systematisch darzustellen. Die Diskussion ermöglicht somit eine fundierte Reflexion über die Anwendbarkeit und Relevanz der Handlungsempfehlungen im konkreten Kontext. Auf eine erneute Zitation der Handlungsempfehlungen wird bewusst verzichtet, da deren theoretische Fundierung bereits im vorangegangenen Kapitel erfolgt ist. Die Handlungsempfehlungen basieren auf einem mehrphasigen, strukturierten Modell zur Umsetzung der regulatorischen Anforderungen des EU AI Acts und adressieren Provider und Deployer gleichermaßen. Sie sind umfangreich und orientieren sich an international etablierten Rahmenwerken. Die Experteninterviews zeigen, dass diese Empfehlungen grundsätzlich als relevant und sinnvoll anerkannt werden, ihre praktische Umsetzung jedoch in hohem Maße von lokalen Rahmenbedingungen, bestehenden Infrastrukturen, personellen Ressourcen und rechtlichen Graubereichen beeinflusst wird.

Bereits in der ersten Phase zeigt sich eine weitgehende Übereinstimmung zwischen Theorie und Praxis. Die Handlungsempfehlungen betonen die Notwendigkeit einer systematischen Identifikation aller eingesetzten KI-Systeme sowie der Rollenzuweisung im Lebenszyklus. Dies wird durch die Interviewaussagen gestützt, wonach viele Einrichtungen bestehende Prozesse, etwa das Verzeichnis der Verarbeitungstätigkeiten (VVT), als Ausgangspunkt nutzen und erweitern, anstatt parallele Systeme aufzubauen. Besonders betont wird dabei der dynamische Charakter von KI-Systemen: Neue Funktionalitäten können durch Softwareupdates eingeführt werden, was die vollständige Erfassung erschwert und ein kontinuierliches Monitoring erforderlich macht.

Auch das in den Handlungsempfehlungen adressierte Rollenverständnis (Deployer vs. Provider) wird durch die Interviews problematisiert. In der Praxis ist eine klare Trennung nur bedingt möglich. Viele Krankenhäuser modifizieren bestehende Systeme, entwickeln interne Prototypen oder nutzen Forschungskontexte zur KI-Erprobung. Dadurch entsteht eine Grauzone, in der regulatorische Pflichten teils unklar sind – eine Herausforderung, die in den Empfehlungen zwar erkannt, aber nicht abschließend aufgelöst wird.

Die Etablierung eines umfassenden Risikomanagements speziell für KI-Systeme wird in der Praxis zwar als notwendig erkannt, befindet sich aber vielerorts erst im Aufbau. Die Handlungsempfehlungen fordern einen strukturierten Risikoprozess entlang des gesamten Lebenszyklus. Die Interviews zeigen, dass bislang überwiegend bestehende generische Systeme, etwa auf Basis der ISO 13485 oder ISO 27001, genutzt werden. Diese werden um KI-spezifische Elemente ergänzt, jedoch nicht vollständig auf KI-Risiken ausgerichtet.

Die Datenvorbereitung ist ein weiteres Element der Handlungsempfehlungen und findet auch in den Interviews breite Bestätigung als kritischer Erfolgsfaktor für den KI-Einsatz. Theoretisch wird eine strukturierte Data Governance, die Prüfung der Datenqualität sowie der gezielte Einsatz von Bias-Detection-Tools gefordert. In der Praxis erweist sich jedoch bereits die Sicherstellung einer ausreichenden Datenqualität als herausfordernd. Die Interviews zeigen, dass medizinische Daten häufig unvollständig, heterogen und durch Anonymisierungsverfahren entwertet sind. Der Umgang mit solchen Daten erfordert domänenspezifisches Wissen und manuelle Eingriffe. Automatisierte Validierungstools sind derzeit nur bedingt einsatzfähig, insbesondere bei unstrukturierten Daten. Besonders hervorgehoben wird im Interviewkontext die Problematik seltener Diagnosen, die häufig nicht in ausreichender Fallzahl zur Verfügung stehen. In den Handlungsempfehlungen findet sich die Forderung nach Datendiversität, ein Anspruch, der in der Realität aus statistischen und infrastrukturellen Gründen oft nur bedingt erfüllt werden kann. Die Notwendigkeit einer kontinuierlichen Rückkopplung zwischen Datenanalyse und klinischer Relevanz wird von den Interviewpartner:innen betont, ebenso wie der sensible Umgang mit Pseudonymisierung und Datenschutzanforderungen im Einklang mit DSGVO und EU AI Act.

In der Phase der Modellierung wird in den Handlungsempfehlungen eine hohe Verantwortung auf technische Nachvollziehbarkeit, Fairness und kontinuierliche Risikoanalysen gelegt. Diese Ziele werden auch in der Praxis anerkannt, jedoch zeigen die Interviews die Grenzen der operativen Umsetzung auf. Besonders DL-Modelle gelten als schwierig hinsichtlich ihrer Erklärbarkeit. Weiters finden Tools wie SHAP oder LIME Anwendung. Die Interviews zeigen auch, dass typische Validierungsverfahren wie Cross-Validation und Modell-Risikoanalysen bekannt sind und Anwendung finden. Die in den Handlungsempfehlungen geforderte Anwendung von Fairness-Metriken wurde in Summe nicht adressiert und es kann somit keine finale Aussage getroffen werden. Verzerrungen lassen sich basierend auf den Aussagen der befragten Personen durch interdisziplinäre Teams erkennen und interpretieren. Die theoretischen Anforderungen an eine fundierte Evaluation, einschließlich technischer Dokumentation, Robustheitstests und Konformitätsbewertung, spiegeln sich deutlich in den Aussagen der befragten Personen wider. Besonders bei Systemen, die als Medizinprodukte gelten, sind durch die MDR, ISO 13485 und andere Normen bereits umfassende Prüfverfahren etabliert. In der Praxis erfolgt eine umfangreiche Vortestung, häufig mit hausinternen Daten, ergänzt durch Vergleichsstudien und Nutzerfeedback. Allerdings zeigen die Interviews auch, dass es derzeit keine neutralen und standardisierten Testdatensätze gibt, die eine objektive Bewertung von KI-Systemen ermöglichen würden – ein Aspekt, der in den Handlungsempfehlungen nicht explizit adressiert wurde.

Im Rahmen der fünften Phase steht die sichere und transparente Integration von KI-Systemen in den klinischen Alltag im Fokus. Die formulierten Empfehlungen betonen unter anderem die Bedeutung klarer Kommunikation gegenüber Nutzer:innen, die Einführung eines strukturierten Überwachungssystems sowie die systematische Protokollierung von Ereignissen während des Einsatzes. Diese Anforderungen werden grundsätzlich auch in der Praxis umgesetzt, etwa durch Sicherheitszertifizierungen, On-Premise-Installationen und begleitende Schulungen. Allerdings zeigen die Interviews, dass Rückmeldungen zu Fehlfunktionen von KI-Systeme nicht systematisch erfasst werden, wodurch die kontinuierliche Verbesserung der Systeme einschränkt wird.

Ein weiteres Spannungsfeld ergibt sich aus der Nutzerkennzeichnung. Während in den Handlungsempfehlungen eine explizite Information über den Einsatz von KI gefordert wird, zeigt sich in der Praxis, dass diese Information oft nicht erfolgt oder als wenig hilfreich eingeschätzt wird. Die Interviewten betonen vielmehr die Bedeutung einer funktionalen, risikoorientierten Kommunikation statt technischer Transparenz allein. Die Einrichtung einer Ethik- und Governancekommission, wie sie theoretisch empfohlen wird, ist in der Praxis meist erledigt, jedoch nicht mit speziellem Fokus auf KI-Systeme. Hier ist zu prüfen, ob bestehende Strukturen übernommen werden können.

Auch die letzte Phase erfährt in den Interviews weitgehende Bestätigung. Die Handlungsempfehlungen fordern ein aktives Qualitätsmanagement, Monitoring und regelmäßige Updates. Viele Organisationen setzen auf geplante Nachtrainingszyklen sowie kontinuierliche Rückmeldeschleifen mit Anwender:innen, wobei insbesondere Feedbackgespräche mit Krankenhäusern als sinnvolle Ergänzung zur rein technischen Evaluation genannt werden. Begrenzend wirkt sich jedoch aus, dass KI-Systeme derzeit nicht selbstlernend im Betrieb arbeiten dürfen. Dies entspricht auch den regulatorischen Anforderungen, etwa nach MDR, und wird von den Interviewten durchwegs als notwendig beurteilt.

Die empfohlene Integration ethischer Leitlinien sowie die regelmäßige Reflexion über Systemverhalten und gesellschaftliche Auswirkungen ist in der Praxis bislang nur punktuell institutionalisiert. Zwar besteht ein Bewusstsein für die Relevanz dieser Themen, deren organisatorische Umsetzung bleibt jedoch häufig auf projektbezogene Initiativen beschränkt. Querschnittsthemen wie strategische Bedeutung, organisatorische Verankerung und interdisziplinäre Zusammenarbeit werden in den Handlungsempfehlungen teilweise implizit angesprochen, in den Interviews jedoch ausdrücklich betont. Insbesondere die Unsicherheit im Zusammenspiel zwischen MDR, DSGVO und AI Act wird als großes Hindernis für eine proaktive Umsetzung genannt. Die fehlende einheitliche Interpretation der Vorschriften führt teilweise zu Zurückhaltung bei Investitionen und Innovationsprojekten. Gleichzeitig zeigt sich ein starkes Bedürfnis nach übergeordneten Koordinierungsstrukturen, etwa in Form von gemeinsamen Einstufungen, Wissensplattformen oder regulatorischer Klarstellungen. Auch der Schulungsbedarf wird als zentrales Handlungsfeld erkannt. Die Handlungsempfehlungen fordern gezielte Maßnahmen zur Kompetenzentwicklung, was mit den Aussagen der befragten Personen übereinstimmt. Dabei wird insbesondere betont, dass sich ein Markt für AI-Compliance-Beratung im Gesundheitsbereich erst im Aufbau befindet und die Qualität solcher Angebote noch schwer einzuschätzen ist.

Zusammenfassend lässt sich festhalten, dass die entwickelten Handlungsempfehlungen eine praxisorientierte Grundlage zur Umsetzung des EU AI Acts in Krankenhäusern darstellen. Die Interviews bestätigen ihre Relevanz weitgehend, zeigen jedoch zugleich, dass deren Umsetzung stark kontextabhängig ist und durch infrastrukturelle, organisatorische und rechtliche Herausforderungen limitiert wird. Es besteht ein breites Bewusstsein für die Notwendigkeit eines strukturierten, ethisch fundierten und rechtssicheren KI-Einsatzes. Der Weg dorthin erfordert aktuell jedoch noch pragmatische, flexible Lösungen, eine intensive Zusammenarbeit der betroffenen Stakeholder sowie kontinuierliche regulatorische und organisatorische Weiterentwicklungen.

8 FAZIT

Ziel dieser Arbeit war es, auf Grundlage bestehender regulatorischer Rahmenbedingungen, insbesondere des EU AI Acts, sowie praktischer Erfahrungen Handlungsempfehlungen zur Entwicklung und zum Einsatz von KI-Systemen in österreichischen Krankenhäusern zu formulieren. Die anschließende Befragung von Experten und Expertinnen ermöglichte eine fundierte Einordnung der theoretischen Empfehlungen in den konkreten Krankenhausalltag. Im Folgenden werden die zentrale Forschungsfrage sowie die drei zugehörigen Teilfragen zusammenfassend beantwortet.

Forschungsfrage: *„Wie sollen österreichische Krankenhäuser mit ihren spezifischen Gegebenheiten KI-Systeme entwickeln bzw. einsetzen, um den Vorgaben des EU AI Act gerecht zu werden?“*

Die Umsetzung des EU AI Acts in österreichischen Krankenhäusern erfordert einen organisationsspezifischen, schrittweisen und risikoorientierten Ansatz, der bestehende Strukturen integriert und zugleich auf neue regulatorische Anforderungen reagiert. Eine vollständige Eigenentwicklung von KI-Systemen durch Krankenhäuser findet in der Praxis selten statt. Weit häufiger werden bestehende Systeme adaptiert oder in Forschungskontexten getestet. Daraus ergibt sich ein hybrides Rollenverständnis zwischen Deployer und Provider, dass bei der Umsetzung regulatorischer Anforderungen wie Konformitätsbewertung oder Risikomanagement berücksichtigt werden muss. Entscheidend ist, dass Krankenhäuser vorhandene Qualitätssicherungs- und IT-Strukturen (z. B. ISO 13485, ISO 27001) nutzen und gezielt um KI-spezifische Komponenten erweitern. Zudem sollten alle eingesetzten Systeme, einschließlich Software mit eingebetteten KI-Modulen, systematisch identifiziert und bewertet werden. Gerade die Vielzahl eingesetzter Systeme sowie die dynamische Entwicklung neuer Funktionen durch Softwareupdates erfordern eine kontinuierliche Bewertung. Die Interviews zeigen, dass eine erfolgreiche Umsetzung des EU AI Acts nicht auf technischer Ebene allein erfolgen kann. Vielmehr braucht es ein Zusammenspiel aus organisatorischer Koordination, interdisziplinärer Zusammenarbeit und ethischer Reflexion.

Teilfrage 1: *„Welche spezifischen Anforderungen des EU AI Act sind für KI-Anwendungen in österreichischen Krankenhäusern relevant?“*

Für Krankenhäuser als typische Betreiber und seltener Anbieter, sind insbesondere die beschriebenen Anforderungen aus der Tabelle 6 relevant.

Die Interviews verdeutlichen, dass viele dieser Anforderungen bereits durch bestehende Regularien (z. B. MDR, DSGVO) teilweise abgedeckt sind. Der AI Act ergänzt diese Vorschriften um KI-spezifische Pflichten, etwa die strukturierte Nachvollziehbarkeit von Modellentscheidungen.

Teilfrage 2: *„Wie beeinflusst der EU AI Act die Auswahl und Entwicklung von KI-Systemen in österreichischen Krankenhäusern?“*

Die Auswahl von KI-Systemen erfolgt zunehmend unter Berücksichtigung regulatorischer Kriterien. Krankenhäuser achten auf CE-Zertifizierungen, MDR-Konformität und nun weiters auf die Konformität mit dem EU AI Act. Bei der Entwicklung eigener Systeme, etwa in der universitären Forschung, spielt vor allem die Frage der späteren klinischen Anwendung eine entscheidende Rolle.

Die Interviews zeigen, dass sich der EU AI Act indirekt auf die Auswahl durch die Erwartung einer besseren Dokumentation, eines klaren Nutzennachweises sowie einer belastbaren technischen Erklärbarkeit auswirkt. Gleichzeitig ist die Auswahl derzeit noch stark durch pragmatische Überlegungen (z. B. Verfügbarkeit, Kosten, Integration) geprägt, wodurch regulatorische Aspekte teils in den Hintergrund treten. Langfristig ist jedoch davon auszugehen, dass die Konformität mit dem EU AI Act ein zentrales Kriterium für die Marktzulassung und damit für die Auswahl durch Krankenhäuser wird.

Teilfrage 3: *„Welche Maßnahmen sollen österreichische Krankenhäuser ergreifen, um Compliance mit dem EU AI Act sicherzustellen?“*

Zur Sicherstellung der Compliance mit dem EU AI Act sollten österreichische Krankenhäuser folgende Maßnahmen ergreifen:

- Systematische Identifikation und Bewertung aller eingesetzten KI-Systeme, auch solcher mit eingebetteten KI-Funktionen
- Klare Rollenzuordnung als Deployer oder Provider – insbesondere bei Systemmodifikationen oder Forschungsprojekten
- Erweiterung bestehender QM- und Risikomanagementsysteme um KI-spezifische Anforderungen
- Dokumentation und technische Nachvollziehbarkeit sicherstellen, insbesondere bei Hochrisiko-Systemen
- Verankerung interdisziplinärer Zusammenarbeit, z. B. durch ethische Begleitgremien und gemeinsame Systembewertungen
- Schulung von medizinischem und technischem Personal, um Kompetenzlücken in Bezug auf KI, Datenschutz und regulatorische Anforderungen zu schließen
- Regelmäßiges Monitoring und Feedbackmechanismen implementieren, um Systemleistung, Bias und Datenqualität kontinuierlich zu bewerten
- Nutzung nationaler Foren und Kooperationen, um Synergien bei der Umsetzung zu schaffen und Doppelarbeit zu vermeiden

Die Interviews zeigen, dass viele dieser Maßnahmen bereits in Ansätzen existieren, jedoch oft ohne expliziten KI-Bezug oder verbindliche Verankerung. Eine erfolgreiche Umsetzung des EU AI Acts erfordert daher nicht nur technische Anpassungen, sondern auch strukturelle und kulturelle Veränderungen in den Einrichtungen.

ABBILDUNGSVERZEICHNIS

Abbildung 1: Aufbau eines künstlichen neuronalen Netzwerkes 9
Abbildung 2: Verantwortlichkeiten in einem vereinfachten Lebenszyklus eines KI-Systems 26

TABELLENVERZEICHNIS

Tabelle 1: Modelle eines KI-System-Lebenszyklus 11
Tabelle 2: Funktionen eines Krankenhauses 14
Tabelle 3: KI in der Pflege 17
Tabelle 4: Anforderungen an KI-Systeme aus Sicht des Gesundheitswesens 22
Tabelle 5: Risikoklassifizierung 27
Tabelle 6: Verpflichtungen für Anbieter und Betreiber nach Risikolevel 31
Tabelle 7: Fragenkomplexe des Hauptteils des Interviewleitfadens 43
Tabelle 8: Auswahl der Experten und Expertinnen 47
Tabelle 9: Codierschema MAXQDA 52
Tabelle 10: Statistik zu vielversprechende KI-Anwendungen 54

LITERATURVERZEICHNIS

Aboy, M., Minssen, T. & Vayena, E. (2024). Navigating the EU AI Act: implications for regulated digital medical products. *npj Digital Medicine*, *7*(1). https://doi.org/10.1038/s41746-024-01232-3

Arge ITA-AIT Parlament. (2020). *KI im Gesundheitswesen*. https://www.parlament.gv.at/dokument/fachinfos/zukunftsthemen/093_ki-diagnose.pdf

Auer, C., Hollenstein, N. & Reumann, M. (2019). Künstliche Intelligenz im Gesundheitswesen. In R. Haring (Hrsg.), *Gesundheit digital* (S. 33–46). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-57611-3_3

Babel, W. (2024). *Systemintegration in Industrie 4.0 und IoT*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-42987-4

Bajema, N., Smith, C. S., Garisto, D., Work, R. O. & Hampson, M. (2022). Real Worst-Case Scenarios: Who needs Terminators when you have precision clickbait and ultra-deepfakes? *IEEE Spectrum*, *59*(1), 8–14. https://doi.org/10.1109/MSPEC.2022.9676367

Barenkamp, M. (2024). Der AI Act: Historische Chance wegbürokratisiert? *Wirtschaftsinformatik & Management*, 1–7. https://doi.org/10.1365/s35764-024-00520-7

Barkane, I. (2022). Questioning the EU proposal for an Artificial Intelligence Act: The need for prohibitions and a stricter approach to biometric surveillance1. *Information Polity*, *27*(2), 147–162. https://doi.org/10.3233/IP-211524

Bayer, J. (2024). The place of content ranking algorithms on the AI risk spectrum. *Telecommunications Policy*, *48*(5), 102741. https://doi.org/10.1016/j.telpol.2024.102741

Behar, B. I., Eisenbeiß, K., Löscher, F. & Salfeld, R. (2022). *Modernes Krankenhausmanagement: Konzepte und Lösungen* (5. Auflage). Springer Gabler. https://doi.org/10.1007/978-3-662-65584-9

Berger-Grabner, D. (2022). Angewandte qualitative Sozialforschung. In D. Berger-Grabner (Hrsg.), *Lehrbuch. Wissenschaftliches Arbeiten in den Wirtschafts- und Sozialwissenschaften: Hilfreiche Tipps und praktische Beispiele* (4., überarbeitete und erweiterte Auflage, S. 137–168). Springer Gabler. https://doi.org/10.1007/978-3-658-39586-5_5

Blase, N., Diehl, A. & Wasem, J. (2023). Digitalisierung im Krankenhaus. In J. Klauber, J. Wasem, A. Beivers & C. Mostert (Hrsg.), *Krankenhaus-Report 2023* (S. 165–178). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-66881-8_11

Bogner, A., Littig, B. & Menz, W. (2014). *Interviews mit Experten: Eine praxisorientierte Einführung*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-19416-5

Botero Arcila, B. (2024). AI liability in Europe: How does it complement risk regulation and deal with the problem of human oversight? *Computer Law & Security Review*, *54*, 106012. https://doi.org/10.1016/j.clsr.2024.106012

Bundesministerium für Arbeit, Soziales, Gesundheit und Konsumentenschutz. (2019). *Das österreichische Gesundheitssystem: Zahlen - Daten - Fakten* (3. vollständig überarbeitete Auflage). https://broschuerenservice.sozialministerium.at/Home/Download?publicationId=636

Bundesministerium für Arbeit, Soziales, Gesundheit, Pflege und Konsumentenschutz. (2024). *Klassifikation der österreichischen Krankenanstalten*. https://www.sozialministerium.gv.at/dam/jcr:62b2d9d5-e6d3-4519-bb50-8509fd40f83b/klassifikation_krankenanstalten_20201008.pdf

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: Der Mensch steht im Mittelpunkt. *Innere Medizin (Heidelberg, Germany)*, *64*(11), 1025–1032. https://doi.org/10.1007/s00108-023-01597-9

Busse, R., Schreyögg, J., Tiemann, O., Gibis, B., Milstein, R., Bäuml, M., Müller, J., Hildebrandt, H., Gröne, O., Schulte, T., Steimle, T. & Verheyen, F. (2022). Leistungsmanagement. In R. Busse, J. Schreyögg & T. Stargardt (Hrsg.), *Management im Gesundheitswesen: Das Lehrbuch für Studium und Praxis* (5. Auflage, S. 11–217). Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-64176-7_2

Capparelli, F., Finocchiaro, G. & Pini, S. (2024). Towards ethical AI: risk management and privacy in the age of innovation. *IET Conference Proceedings*(7), 36–41. https://doi.org/10.1049/icp.2024.2525

Chen, M. & Decary, M. (2020). Artificial intelligence in healthcare: An essential guide for health leaders. *Healthcare management forum*, *33*(1), 10–18. https://doi.org/10.1177/0840470419873123

Clough, R. A. J., Sparkes, W. A., Clough, O. T., Sykes, J. T., Steventon, A. T. & King, K. (2024). Transforming healthcare documentation: harnessing the potential of AI to generate discharge summaries. *BJGP Open*, *8*(1), BJGPO.2023.0116. https://doi.org/10.3399/BJGPO.2023.0116

Czvetkó, T., Kummer, A., Ruppert, T. & Abonyi, J. (2022). Data-driven business process management-based development of Industry 4.0 solutions. *CIRP Journal of Manufacturing Science and Technology*, *36*, 117–132. https://doi.org/10.1016/j.cirpj.2021.12.002

Domhoff, D., Seibert, K., Bergmann, L., Theune, S., Bießmann, F., Fürstenau, D., Schulte-Althoff, M. & Wolf-Ostermann, K. (2024). Forschungs- und Entwicklungsprojekte in der Pflege erfolgreich planen und umsetzen – Voraussetzungen und Gelingensbedingungen am Beispiel des Einsatzes von künstlicher Intelligenz. In W. Swoboda & N. Seifert (Hrsg.), *Digitale Innovationen in der Pflege* (S. 379–399). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-658-39302-1_21

Döring, N. (2023). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-64762-2

Dos Santos, F. C., Johnson, L. G., Madandola, O. O., Priola, K. J. B., Yao, Y., Macieira, T. G. R. & Keenan, G. M. (2024). An example of leveraging AI for documentation: ChatGPT-generated nursing care plan for an older adult with lung cancer. *Journal of the American Medical Informatics Association : JAMIA*, *31*(9), 2089–2096. https://doi.org/10.1093/jamia/ocae116

Ertel, W. (2021). *Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung* (5. Auflage). *Computational intelligence*. Springer Vieweg. https://doi.org/10.1007/978-3-658-32075-1

Europäische Kommission, Generaldirektion Kommunikationsnetze, Inhalte und Technologien. (2019). *Ethik-leitlinien für eine vertrauenswürdige KI*. Publications Office. https://doi.org/10.2759/22710

VERORDNUNG (EU) 2024 DES EUROPÄISCHEN PARLAMENTS UND DES RATES zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz) (2024). https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138-FNL-COR01_DE.pdf

Europäisches Parlament. (2024, 13. März). *Gesetz über künstliche Intelligenz: Parlament verabschiedet wegweisende Regeln* [Pressemitteilung].

EY. (2023). *KI wird menschlicher, aber werden Menschen ihr auch vertrauen? Ein Überblick über die EU KI-Verordnung und deren Auswirkungen für Unternehmen*. https://assets.ey.com/content/dam/ey-sites/ey-com/en_ch/noindex/ey-de-eu-ai-act-2023.pdf?mkt_tok=NTIwLVJYUC0wMDMAAAGXPq0M5t0Re6MEhD43N3_PhbN61Hj0wLZkez1GH8gS-XRk54JCZp7hN32h79Je-yxA7IlWZ6ZFeEP-FvHEpajyLx4Opna8UvR48WdbWdywEZn08cvvCw

EY. (2024). *The European Union Artificial Intelligence Act*. https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/ai/ey-eu-ai-act-political-agreement-overview-february-2024.pdf

Fraunhofer-Institut für Kognitive Systeme IKS. (2024a). *Künstliche Intelligenz (KI) und maschinelles Lernen*. https://www.iks.fraunhofer.de/de/themen/kuenstliche-intelligenz.html

Fraunhofer-Institut für Kognitive Systeme IKS. (2024b). *Künstliche Intelligenz in der Medizin*. https://www.iks.fraunhofer.de/de/themen/kuenstliche-intelligenz/kuenstliche-intelligenz-medizin.html

Gama, F., Tyskbo, D., Nygren, J., Barlow, J., Reed, J. & Svedberg, P. (2022). Implementation Frameworks for Artificial Intelligence Translation Into Health Care Practice: Scoping Review. *Journal of medical Internet research*, *24*(1), e32215. https://doi.org/10.2196/32215

García-Gómez, J. M., Blanes-Selva, V. & Doñate-Martínez, A. (2024). Proposing an AI Passport as a Mitigating Action of Risk Associated to Artificial Intelligence in Healthcare. *Studies in health technology and informatics*, *316*, 547–551. https://doi.org/10.3233/SHTI240472

Gesundheitsportal. (2024). *Gesundheitssystem - Krankenversorgung*. https://www.gesundheit.gv.at/gesundheitsleistungen/gesundheitswesen/gesundheitssystem.html

Gilbert, S. (2024). The EU passes the AI Act and its implications for digital medicine are unclear. *npj Digital Medicine*, *7*(1), 135. https://doi.org/10.1038/s41746-024-01116-6

Gläser, J. & Laudel, G. (2010). *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen* (4. Auflage). VS Verlag. http://d-nb.info/1002141753/04

Granigg, W. & Lichtenegger, K. (2024). Data Science und Künstliche Intelligenz. In A. Klein, S. Dennerlein & H. Ritschl (Hrsg.), *Health Care und Künstliche Intelligenz: Ethische Aspekte verstehen – Entwicklungen gestalten* (S. 51–68). utb GmbH.

Häder, M. (2015). *Empirische Sozialforschung: Eine Einführung* (3. Auflage). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-19675-6

Halwachs-Baumann, G. (2024). *Gesundheitsökonomie für Mediziner* (1. Auflage). Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-68911-0

Heinrichs, B. (2023). *Künstliche Intelligenz - zwischen Hype und Hysterie.* Jülich Forschungszentrum. https://www.fz-juelich.de/de/aktuelles/news/meldungen/2023/ki-chancen-risiken

Hensen, P. (2022). *Qualitätsmanagement im Gesundheitswesen: Grundlagen für Studium und Praxis* (3. Auflage). Springer Gabler Wiesbaden. https://doi.org/10.1007/978-3-658-38299-5

Hevner, March, Park & Ram (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75. https://doi.org/10.2307/25148625

Iivari, J. (2015). Distinguishing and contrasting two strategies for design science research. *European Journal of Information Systems*(24), 107–115. https://doi.org/10.1057/ejis.2013.35

Kaiser, R. (2021). *Qualitative Experteninterviews: Konzeptionelle Grundlagen und praktische Durchführung* (2. Auflage). Springer VS Wiesbaden. https://link-1springer-1com-1b1168j9w092c.perm.fh-joanneum.at/chapter/10.1007/978-3-658-30255-9_2 https://doi.org/10.1007/978-3-658-30255-9_2

Keskinbora, K. H. (2019). Medical ethics considerations on artificial intelligence. *Journal of clinical neuroscience : official journal of the Neurosurgical Society of Australasia*, *64*, 277–282. https://doi.org/10.1016/j.jocn.2019.03.001

Kilian, R., Jäck, L. & Ebel, D. (2025). European AI Standards - Technical Standardization and Implementation Challenges under the EU AI Act, 1–32. https://doi.org/10.2139/ssrn.5155591

Kirste, M. & Schürholz, M. (2019). Einleitung: Entwicklungswege zur KI. In V. Wittpahl (Hrsg.), *Künstliche Intelligenz* (S. 21–35). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-58042-4_1

Klein, A., Dennerlein, S. & Ritschl, H. (2024). Einleitung. In A. Klein, S. Dennerlein & H. Ritschl (Hrsg.), *Health Care und Künstliche Intelligenz: Ethische Aspekte verstehen – Entwicklungen gestalten* (S. 19–48). utb GmbH.

KPMG. (2024). *Decoding the EU AI Act Understanding the AI Act's impact and how you can respond*. https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2024/02/decoding-the-eu-artificial-intelligence-act.pdf.

Küster, D. & Schultz, T. (2023). Künstliche Intelligenz und Ethik im Gesundheitswesen – Spagat oder Symbiose? [Artificial intelligence and ethics in healthcare-balancing act or symbiosis?]. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, *66*(2), 176–183. https://doi.org/10.1007/s00103-022-03653-5

Lakhani, P. & Sundaram, B. (2017). Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology*, *284*(2), 574–582. https://doi.org/10.1148/radiol.2017162326

Lang, V. (2023). *Digitale Kompetenz: Grundlagen der Künstlichen Intelligenz, Blockchain-Technologie, Quanten-Computing und deren Anwendungen für die Digitale Transformation*. Springer Berlin. https://doi.org/10.1007/978-3-662-66285-4

Li, X., Pang, T., Liu, W. & Wang, T. (2017). Fall detection for elderly person care using convolutional neural networks. *International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 1–6. https://doi.org/10.1109/CISP-BMEI.2017.8302004

Liebl, A. & Klein, T. (2023). *Empfehlungen für innovationsfreundliche Klassifizierungsregeln in der KI-Verordnung*. https://www.appliedai.de/assets/files/AI-Act_WhitePaper_final_CMYK_GER.pdf

Liyanage, H., Liaw, S.-T., Jonnagaddala, J., Schreiber, R., Kuziemsky, C., Terry, A. L. & Lusignan, S. de (2019). Artificial Intelligence in Primary Health Care: Perceptions, Issues, and Challenges. *Yearbook of medical informatics*, *28*(1), 41–46. https://doi.org/10.1055/s-0039-1677901

Madigan, M. *Disaster Response Practices: Guide to Mass Casualty/Mass Fatality Management Planning*. *Management for Professionals*. Springer Cham. https://doi.org/10.1007/978-3-031-42147-1_6

Martín-Noguerol, T., López-Úbeda, P. & Luna, A. (2024). AI in radiology: Legal responsibilities and the car paradox. *European Journal of Radiology*, *175*, 111462. https://doi.org/10.1016/j.ejrad.2024.111462

Mayring, P. (2020). Qualitative Inhaltsanalyse. In G. Mey & K. Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie* (S. 495–511). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-26887-9_52

Mitchell, A. D., Let, D. & Tang, L. (2023). AI Regulation and the Protection of Source Code. *International Journal of Law and Information Technology*, *31*(4), 283–301. https://doi.org/10.1093/ijlit/eaad026

Mockenhaupt, A. (2021). *Digitalisierung und Künstliche Intelligenz in der Produktion: Grundlagen und Anwendung*. *Lehrbuch*. Springer Vieweg. https://doi.org/10.1007/978-3-658-32773-6

Morales Santos, Á., Lojo Lendoiro, S., Rovira Cañellas, M. & Valdés Solís, P. (2024). The legal regulation of artificial intelligence in the European Union: A practical guide for radiologists. *Radiologia*, *66*(5), 431–446. https://doi.org/10.1016/j.rxeng.2023.11.011

Neuwirth, R. J. (2023). Prohibited artificial intelligence practices in the proposed EU artificial intelligence act (AIA). *Computer Law & Security Review*, *48*, 105798. https://doi.org/10.1016/j.clsr.2023.105798

Offermann, P., Levina, O., Schönherr, M. & Bub, U. (2009). Outline of a design science research process. *DESRIST '09: Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*, Artikel 7, 1–11. https://doi.org/10.1145/1555619.1555629

Palmer, A. & Schwan, D. (2024). More Process, Less Principles: The Ethics of Deploying AI and Robotics in Medicine. *Cambridge quarterly of healthcare ethics : CQ : the international journal of healthcare ethics committees*, *33*(1), 121–134. https://doi.org/10.1017/S0963180123000087

Pesapane, F., Volonté, C., Codari, M. & Sardanelli, F. (2018). Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States. *Insights into imaging*, *9*(5), 745–753. https://doi.org/10.1007/s13244-018-0645-y

Priyanka, N. (2016). *Applying Deep Learning to ICD-9 Multi-label Classification from Medical Records.* Department of Computer Science Stanford University. https://cs224d.stanford.edu/reports/priyanka.pdf

Puppe, F., Krug, M. & Kempf, S. (2024). Neuronale Netze und Deep Learning. In W. Swoboda & N. Seifert (Hrsg.), *Digitale Innovationen in der Pflege* (S. 325–350). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-67914-2_13

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*. http://arxiv.org/pdf/1711.05225

Ranjbar, A. & Ravn, J. (2024). Monitoring of Artificial Intelligence in Hospitals. *Studies in health technology and informatics*, *316*, 683–684. https://doi.org/10.3233/SHTI240506

Raposo, V. L. (2022). Ex machina: preliminary critical assessment of the European Draft Act on artificial intelligence. *International Journal of Law and Information Technology*, *30*(1), 88–109. https://doi.org/10.1093/ijlit/eaac007

Schreyögg, J., Moos, G., Brüggemann, F., Fleßa, S., Weber, W., Sohn, S., Gudd, K., Voss, H., Schöffski, O., Mühlbacher, A., Pimperl, A., Dangel, P., Schlösser, R. & Everding, J. (2022). Informationsmanagement und Controlling. In R. Busse, J. Schreyögg & T. Stargardt (Hrsg.), *Management im Gesundheitswesen: Das Lehrbuch für Studium und Praxis* (5. Auflage, S. 529–630). Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-64176-7_6

Semigran, H. L., Linder, J. A., Gidengil, C. & Mehrotra, A. (2015). Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ*, *351*, h3480. https://doi.org/10.1136/bmj.h3480

Solaiman, B. & Malik, A. (2025). Regulating algorithmic care in the European Union: evolving doctor-patient models through the Artificial Intelligence Act (AI-Act) and the liability directives. *Medical law review*, *33*(1). https://doi.org/10.1093/medlaw/fwae033

Steiermärkische Krankenanstaltengesellschaft m.b.H. (2025a, 25. Januar). *Über uns*. https://www.uniklinikumgraz.at/ueber-uns#c4411

Steiermärkische Krankenanstaltengesellschaft m.b.H. (2025b, 25. Januar). *Universitätsklinik für Anästhesiologie und Intensivmedizin*. https://www.uniklinikumgraz.at/anaesthesie

Studer, S., Bui, T. B., Drescher, C., Hanuschkin, A., Winkler, L., Peters, S. & Müller, K.-R. (2021). Towards CRISP-ML(Q): A Machine Learning Process Model with Quality

Assurance Methodology. *Machine Learning and Knowledge Extraction*, *3*(2), 392–413. https://doi.org/10.3390/make3020020

Stuurman, K. & Lachaud, E. (2022). Regulating AI. A label to complete the proposed Act on Artificial Intelligence. *Computer Law & Security Review*, *44*, 105657. https://doi.org/10.1016/j.clsr.2022.105657

Taulli, T. (2023). *Grundlagen der Kunstlichen Intelligenz: Eine nichttechnische Einführung*. Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-66283-0

van Kolfschooten, H. & van Oirschot, J. (2024). The EU Artificial Intelligence Act (2024): Implications for healthcare. *Health Policy*, *149*, 105152. https://doi.org/10.1016/j.healthpol.2024.105152

Visengeriyeva, L. (2025). *The AI Engineer's Guide to Surviving the EU AI Act*. https://www.oreilly.com/library/view/the-ai-engineers/9781098172480/

Wagner, M., Borg, M. & Runeson, P. (2024). Navigating the Upcoming European Union AI Act. *IEEE Software*, *41*(1), 19–24. https://doi.org/10.1109/MS.2023.3322913

Wahl, F. & Wilhelm, S. (2024). Sensorik und künstliche Intelligenz in der Pflege. In W. Swoboda & N. Seifert (Hrsg.), *Digitale Innovationen in der Pflege* (S. 307–324). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-67914-2_12

Wendland, M. (2024). Künstliche Intelligenz in der Medizin: Rechtsrahmen, Haftung und ethische Grundlagen. In A. Klein, S. Dennerlein & H. Ritschl (Hrsg.), *Health Care und Künstliche Intelligenz: Ethische Aspekte verstehen – Entwicklungen gestalten* (S. 330–341). utb GmbH.

Wiemschulte, J. (2024). *Gesundheitsökonomie, Gesundheitssystem, Gesundheitswesen: Kompendium zum QB 3 des Studiums Humanmedizin*. Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-68457-3

Wolf-Ostermann, K., Fürstenau, D., Theune, S., Bergmann, L., Biessmann, F., Domhoff, D., Schulte-Althoff, M. & Seibert, K. (2021). *Konzept zur Einbettung von KI-Systemen in der Pflege: Sondierungsprojekt zu KI in der Pflege (SoKIP).* https://doi.org/10.26092/ELIB/496

ANHANG

Anhang 1: Interviewleitfaden

Phase	U-Phase	Ziel/ Ablauf	Information oder Fragestellungen
Einleitung	Begrüßung & Vorstellung	**Ziel:** Schaffung einer angenehmen Gesprächsatmosphäre, kurzer Überblick über Interviewverlauf **Ablauf:** Vorstellung; Dank für die Teilnahme; Kurze Erklärung der Struktur (Phasen und Dauer)	**I:** Herzlich willkommen und vielen Dank, dass Sie sich Zeit für dieses Interview nehmen. Ich heiße Georg Strahlhofer und schreibe meine Masterarbeit zum Thema *„Künstliche Intelligenz im Krankenhaus - Gesetzliche Rahmenbedingungen und ethische Aspekte“* Im Folgenden gebe ich Ihnen einen kurzen Überblick über den Ablauf: Wir klären zuerst ein paar organisatorische Punkte, dann stelle ich kurz meine Arbeit vor und gehe auf das Forschungsdesign ein. Anschließend würde ich gerne Ihre Einschätzungen und Erfahrungen zu verschiedenen Aspekten hören. Zum Abschluss fasse ich die wichtigsten Punkte zusammen und bedanke mich für Ihre Zeit.
	Administrative Punkte	**Ziel**: Klärung der Datenverwendung, Vertraulichkeit und Einverständnis zur Aufnahme **Ablauf:** Fragen zur Neutralisierung oder Nennung von Namen/Unternehmen; Einverständniserklärung zur Aufzeichnung	**F:** Wären Sie damit einverstanden, dass Ihre Daten (Name, Position und Unternehmen) in der Arbeit erwähnt werden oder sollen sie bis auf Ihre Rolle neutralisiert werden? **F:** Wäre es in Ordnung, wenn ich das Interview zur Dokumentation aufzeichne?

Phase	U-Phase	Ziel/ Ablauf	Information oder Fragestellungen
	Vorstellung Masterarbeit	**Ziel:** Kurzvorstellung des Themas, der Forschungsfragen und Ziele **Ablauf:** Darstellung der Relevanz von KI im Gesundheitswesen; Hinweis auf den EU AI Act und dessen Bedeutung für Hochrisikobereiche; Forschungsfrage(n) und Zielsetzung der Arbeit; Methodik (Design Science Research)	**I:** **Thematischer Überblick:** „KI ist im Alltag überall präsent und ihr Potenzial im Gesundheitswesen wird als hoch eingeschätzt. Gleichzeitig gibt es Risiken (z. B. Bias, Privatsphäre). Der EU AI Act (verabschiedet April 2024) legt strenge Regeln für KI in Hochrisikobereichen wie dem Gesundheitswesen fest. **Problemstellung:** „Österreichische Krankenhäuser müssen diesen Anforderungen gerecht werden. Meine zentrale Forschungsfrage lautet: *Wie sollen österreichische Krankenhäuser KI-Systeme entwickeln und einsetzen, um den Vorgaben des EU AI Act zu entsprechen?“* **Zielsetzung:** „Die Arbeit entwickelt Handlungsempfehlungen, die auf Metaebene dabei unterstützen, die neuen gesetzlichen Vorgaben einzuhalten. **Methodik:** „Ich folge dem Design Science Research (DSR). Dabei identifiziere ich das Problem, entwickle erste Lösungen (Artefakte), evaluiere sie (hier im Interview) und finalisiere sie.“
	Vorstellung Hauptteil	**Zie**l: Kernteil des Interviews, in dem die konkreten Erfahrungen, Einschätzungen und Validierungen erfolgen. **Ablauf:** Allgemeine Erfahrungen mit KI und Bekanntheit des EU AI Act; Validierung der Handlungsempfehlungen entlang CRISP-ML(Q); Zukunftsaussichten	**I:** CRISP-ML(Q) steht für „Cross-Industry Standard Process for Machine Learning with Quality“ und ist eine Erweiterung des bekannten CRISP-DM-Standards, der ursprünglich für Data-Mining-Projekte entwickelt wurde. Im Unterschied zu CRISP-DM liegt bei CRISP-ML(Q) der Fokus stärker auf den spezifischen Anforderungen und Qualitätsaspekten von Machine-Learning-Projekten (z. B. Risikobewertung, Modellvalidierung, Monitoring). Dadurch wird der gesamte Lebenszyklus eines KI-Systems – von der Problemdefinition über die Datenaufbereitung und Modellierung bis hin zur kontinuierlichen Überwachung im Betrieb – strukturierter und qualitätsgesicherter abgedeckt.

Phase	U-Phase	Ziel/ Ablauf	Information oder Fragestellungen
Hauptteil - Allgemeine Erfahrungen mit KI im Gesundheitswesen und dem EU AI Act		**Ziel:** Einstieg ins Thema, Verständnis für vorhandene KI-Anwendungen und Wissen über EU AI Act **Ablauf:** Fragen zu bisherigen KI-Projekten, Erfolgen und Herausforderungen; Einschätzung der Akzeptanz bei Personal und Patient:innen; Kenntnisstand EU AI Act	**F:** Welche Anwendungen von KI sind besonders vielversprechend in Pflege, Diagnostik oder Verwaltung? **F:** Gibt es Beispiele für den erfolgreichen Einsatz von KI in Ihrer Einrichtung? **F:** Welche Herausforderungen sehen Sie in der Implementierung von KI-Systemen in Krankenhäusern? **F:** Wie bewerten Sie die Akzeptanz von KI-Systemen bei medizinischem Fachpersonal und bei Patient:innen?“ **F:** Wie vertraut sind Sie mit dem EU AI Act?
Hauptteil - Validierung der Handlungsempfehlungen anhand des CRISP-ML(Q)-Modells	Phase 1: Geschäfts- und Datenverständnis	**Ziel**: Abprüfen, wie Teilnehmende KI-Systeme identifizieren, Verantwortlichkeiten klären und Risiken managen **Ablauf:** Fragen zum Umgang mit KI-Systemen, Rolle(n) im Lebenszyklus, Risikomanagement und rechtlichen Anforderungen- Erfassung praktischer Erfahrungen und möglicher Lücken	Identifikation von KI-Systemen: **F:** Wie stellen Sie sicher, dass alle KI-Systeme in Ihrer Organisation erfasst und geprüft werden? Rolle im KI-Lebenszyklus: **F:** Welche Verantwortlichkeiten nimmt ein Krankenhaus in der Regel basierend auf den Vorgaben des EU AI Acts ein? (Provider bzw. Deployer) **F:** Ist diese Trennung in der Praxis immer klar oder werden hier Grauzonen in diesem spezifischen Kontext erwartet? Risikomanagementsystem: **F:** Gibt es ein Risikomanagementsystem (z. B. IEEE AI Ethics oder NIST AI RMF), das den gesamten Lebenszyklus abdeckt? **F:** Wie werden Risken in Hinblick auf Gesundheit, Sicherheit und Grundrechte in diesem Risikomanagementsystem berücksichtigt? Registrierungsanforderungen und Überwachung:

Phase	U-Phase	Ziel/ Ablauf	Information oder Fragestellungen
			F: Falls Sie eine öffentliche Behörde sind, wie setzen Sie Registrierungsanforderungen bzw. Grundrechts-Folgenabschätzungen um? Überwachungssystem für Provider: **F:** Wie kontrollieren oder überwachen Sie, ob Provider die technischen Dokumentationen und Standards einhalten?
	Phase 2: Datenvorbereitung	**Ziel:** Prüfen von Datenqualität, Bias-Vermeidung, Relevanz der Daten für den Einsatzzweck **Ablauf:** Fragen zur Data Governance und konkreten Prozessen; Bewertung, wie systematisch Daten aufbereitet werden	Datenqualität und Datenverwaltung: **F:** Wie stellen Sie sicher, dass Datensätze vollständig und konsistent sind? **F:** Nutzen Sie bestimmte Tools oder Prozesse? Relevanz der Eingabedaten (Deployer): **F:** Wie gewährleisten Sie, dass die von Ihnen verwendeten Daten relevant für Ihren konkreten Anwendungsfall sind? Bias-Erkennung und Vermeidung: **F:** Wie stellen Sie sicher, dass Bias erkannt und vermieden wird? Automatisierte Datenvalidierung: **F:** Wie überprüfen Sie kontinuierlich die Datenqualität?
	Phase 3: Modellierung	**Ziel:** Einschätzung zu Modellrückverfolgbarkeit, Erklärbarkeit, Fairness und Risikoprüfung **Ablauf:** Fragen zu Mechanismen der Fairness-Prüfung, Modellrisikoanalyse und Testverfahren	Rückverfolgbarkeit und Erklärbarkeit: **F:** Wie stellen Sie sicher, dass Ihre Modelle in einem klinischen Umfeld nachvollziehbar und auditierbar sind? Verzerrungen und Diskriminierung: **F:** Welche Mechanismen haben Sie, um diskriminierende Ergebnisse zu erkennen und dagegen vorzugehen? Modellrisikoanalyse: **F:** Nutzen Sie eine formale Risikoanalyse (z. B. Over-/Underfitting, Unexpected Outcomes)?

Phase	U-Phase	Ziel/ Ablauf	Information oder Fragestellungen
			F: Wie dokumentieren Sie diese Risiken? Fairness-Metriken: **F:** Wenden Sie definierte Fairness-Kennzahlen an? **F:** Wie interpretieren Sie die Resultate?
	Phase 4: Evaluation	**Ziel:** Feststellen, wie umfangreich die Evaluations- und Testverfahren sind (z. B. technische Dokumentation, Robustheit, Konformität) **Ablauf:** Fragen zu Pre-Market-Tests, Dokumentation, Registrierung und Robustheitsprüfungen	Vor-Markteinführungstests: **F:** Wie prüfen Sie vor der Einführung, ob das System für den vorgesehenen Zweck geeignet ist? Technische Dokumentation (Provider): **F:** Erstellen Sie eine vollständige Dokumentation, die Zweck, Version und Benutzeranweisungen enthält? Registrierung und Konformitätsbewertung (Provider): **F:** Für Hochrisiko-KI-Systeme: Wie gehen Sie bei der Registrierung in einer EU-Datenbank oder bei Konformitätsbewertungen vor? Testarten und -methoden: **F:** Führen Sie strukturierte Tests (z. B. Unit-, Integrations-, System- oder A/B-Tests) durch? (z. B. gegen adversarielle Angriffe)
	Phase 5: Bereitstellung	**Ziel:** Verständnis für den Produktions-/Livegang der KI und Einhaltung von Transparenz, Sicherheitsvorkehrungen, Meldewesen **Ablauf:** Fragen zum Risikokommunikationsprozess,	Risikokommunikation: **F:** Wie informieren Sie medizinisches Personal und Patient:innen über verbleibende Risiken? Kennzeichnung von KI-Systemen: **F:** Ist klar ersichtlich, dass Patient:innen mit einem KI-System interagieren (wo relevant)? Überwachungssysteme und menschliche Kontrolle: **F:** Gibt es ein Not-Aus oder andere Interventionsmöglichkeiten?

Phase	U-Phase	Ziel/ Ablauf	Information oder Fragestellungen
		Kennzeichnungspflicht und menschlicher Kontrolle	Protokollierung und Meldung von Vorfällen: **F:** Wie werden kritische Vorfälle dokumentiert und an Behörden gemeldet? Ethik- und Governancekommission: **F:** Gibt es ein Gremium, das ethische Fragestellungen laufend prüft? Interdisziplinäre Zusammenarbeit und Schulung: **F:** Wie organisieren Sie die Zusammenarbeit zwischen KI-Entwicklern, medizinischem Personal und Patientengremien? **F:** Welche Schulungen finden statt?
	Phase 6: Überwachung und Wartung	**Ziel:** Einsicht in langfristige Qualitätssicherung, Umgang mit Daten-/Modell-Drift, Ethikleitlinien und Notfallplanung **Ablauf:** Fragen zu QM-System, Vertragsanpassungen, Monitoring des Systems im Betrieb; Einblicke in Kontinuität und Weiterentwicklung	Qualitätsmanagementsystem (Provider): **F:** Nutzen Sie ein etabliertes QMS (z. B. ISO/IEC 13485)? **F:** Wie unterstützt dies den langfristigen Betrieb? Daten- und Modell-Drift: **F:** Monitoren Sie das System laufend, um Abweichungen (Drift) früh zu erkennen? **F:** Wie reagieren Sie? Chaos-Engineering und Notfallplanung: **F:** Simulieren Sie technische Ausfälle oder Cyberangriffe? **F:** Existiert ein Notfallplan? Ethikleitlinien und Stakeholder-Feedback: **F:** Wie oft aktualisieren Sie Ihre Ethikleitlinien, und wer entscheidet darüber? **F:** Binden Sie externe Stakeholder ein?

Phase	U-Phase	Ziel/ Ablauf	Information oder Fragestellungen
	Querschnittsfragen	**Ziel:** Ergänzende Fragen, die mehrere Phasen betreffen (organisatorische Verankerung, strategische Bedeutung, Herausforderungen, Schulungsbedarf) **Ablauf:** Fragen ggf. am Ende oder verteilt stellen- Sammeln zusätzlicher Infos zu Hürden und Rahmenbedingungen	Organisatorische Verankerung: **F:** Welche Abteilungen sind für die Einhaltung der KI-Regulierung verantwortlich? **F:** Wie ist die Zusammenarbeit organisiert? Strategische Bedeutung: **F:** Welchen Stellenwert hat der EU AI Act auf Führungsebene? **F:** Sind die nötigen Ressourcen verfügbar? Herausforderungen und Verbesserungsmöglichkeiten: **F:** Wo liegen die größten Hindernisse (z. B. Budget, Fachkräfte, technische Barrieren)? **F:** Welche Unterstützung könnte helfen? Wissensstand und Schulungsbedarf: **F:** Wo sehen Sie Bedarf an zusätzlichen Schulungen oder Informationen, um den EU AI Act besser umzusetzen?
Hauptteil - .Zukunftsaussicht		**Ziel:** Blick in die Zukunft, Rolle von KI in Krankenhäusern, mögliche Kooperationen mit Gesetzgebern und Technologieentwicklern **Ablauf:** Fragen zu künftigen KI-Entwicklungen und strategischer Relevanz- Ideen zur Verbesserung der Zusammenarbeit zwischen Stakeholdern	**F:** Welche Entwicklungen in der KI-Technologie sehen Sie als besonders relevant für Krankenhäuser? **F:** Welche Rolle sollte KI in der strategischen Ausrichtung von Krankenhäusern spielen? **F:** Wie könnte die Zusammenarbeit zwischen Gesetzgebern, Technologieentwicklern und Krankenhäusern verbessert werden?

Phase	U-Phase	Ziel/ Ablauf	Information oder Fragestellungen
Schluss – Zusammenfassung, Ausblick und Danksagung		**Ziel:** Zusammenfassung der wichtigsten Punkte, Raum für offene Fragen, wertschätzender Abschluss **Ablauf:** Kurzer Rückblick auf das Gespräch; Möglichkeit für Ergänzungen; Dank für die Teilnahme	**Zusammenfassung**: Vielen Dank für Ihre wertvollen Einblicke. Ich fasse kurz zusammen, was wir heute besprochen haben… **Ausblick**: Möchten Sie noch etwas ergänzen oder anregen, dass Ihrer Ansicht nach bisher zu kurz kam? **Danksagung**: Ich bedanke mich nochmals herzlich für Ihre Zeit und Ihre Offenheit. Sollten Sie im Nachhinein Fragen haben oder etwas ergänzen wollen, können Sie sich gerne bei mir melden.

Anhang 2: E-Mail Vorlage für die Kontaktaufnahme

Von: Strahlhofer Georg David
Gesendet: [redacted]
Bis: [redacted]
Betreff: Einladung zum Experteninterview: KI & EU AI Act im Gesundheitswesen

Sehr geehrter [redacted]

mein Name ist **Georg Strahlhofer**, und ich befinde mich aktuell in der Endphase meiner Masterarbeit im Masterstudium IT & Wirtschaftsinformatik an der FH CAMPUS 02 . Ihr Kontakt wurde mir von meinem Betreuer, **Ing. Dipl.-Ing. Dipl.-Ing. (FH) Dr.techn. Michael Georg Grasser, MBA MPA**, weitergeleitet, da er Sie als einen geeigneten Experten für mein Forschungsvorhaben im Bereich **Künstliche Intelligenz im Gesundheitswesen** empfiehlt.

In meiner Masterarbeit untersuche ich, wie österreichische Krankenhäuser **KI-Systeme entwickeln und einsetzen können, um die Vorgaben des EU AI Act zu erfüllen**. Ziel ist es, fundierte **Handlungsempfehlungen** zu erarbeiten, die eine praxisnahe Umsetzung erleichtern.

Dafür führe ich **Experteninterviews** durch, um wertvolle Einblicke aus der Praxis zu gewinnen. Das Interview dauert circa eine Stunde und wird online über Microsoft Teams stattfinden, um Ihren zeitlichen Aufwand möglichst gering zu halten.

Falls Sie Interesse haben, sende ich Ihnen gerne im nächsten Schritt den Interviewleitfaden zur weiteren Information zu. Bei Rückfragen stehe ich Ihnen jederzeit telefonisch unter +43 664 8274661 oder per E-Mail unter 51810664@edu.campus02.at zur Verfügung.

Ich würde mich über Ihre positive Rückmeldung freuen und danke Ihnen im Voraus für Ihre Zeit und Unterstützung!

Mit freundlichen Grüßen
Georg Strahlhofer

Anhang 3: Die Datei Masterarbeit_Georg_Strahlhofer_Experteninterviews.mx24 und die Transkriptionen sind gesondert der Abgabe beigelegt.