Deutsches Forschungszentrum für Künstliche Intelligenz
German Research Center for Artificial Intelligence

Vorhaben	KardioInterakt Multimodale Interaktionstechnologien zur Patientenversorgung bei betreuungsintensiven und COVID-19-bedingten Herzerkrankungen im häuslichen Umfeld
Titel	Partnerspezifischer Abschlussbericht
Förderkennzeichen	16SV8907
Zuwendungsempfänger	Deutsches Forschungszentrum für Künstliche Intelligenz GmbH Trippstadter Straße 122, D-67663 Kaiserslautern
Ausführende Stelle	DFKI GmbH – Smart Enterprise Engineering
Projektleiter	Prof. Dr. Oliver Thomas
Bewilligungszeitraum	01.08.2022 – 31.01.2026
Autoren	Sophie Haas
Erstellungsdatum	11.05.26

Gefördert durch:

Bundesministerium für Forschung, Technologie und Raumfahrt

Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Forschung, Technologie und Raumfahrt unter dem Förderkennzeichen 16SV8907 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei dem Autor.

Inhaltsverzeichnis

1 PROJEKTVORHABEN 3

1.1 PROBLEMSTELLUNG 3

1.2 GESAMTZIEL 3

1.3 HAUPTAUFGABE DES DFKI 4

1.4 PLANUNG UND ABLAUF DES VORHABENS 5

1.5 STAND DER WISSENSCHAFT UND TECHNIK ZU PROJEKTBEGINN 5

1.6 ZUSAMMENARBEIT MIT ANDEREN STELLEN 7

2 ERGEBNISSE 7

2.1 ZENTRALE ERGEBNISSE UND ERKENNTNISSE DES DFKI 7

2.2 PROJEKTERFAHRUNGEN UND VERWERTBARKEIT 18

2.3 TECHNOLOGIEENTWICKLUNG 19

2.4 ERFOLGTE UND GEPLANTE VERÖFFENTLICHUNGEN 20

LITERATURVERZEICHNIS 21

1 Projektvorhaben

1.1 Problemstellung

Herz-Kreislauf-Erkrankungen, insbesondere chronische Herzinsuffizienz, sind nach wie vor die häufigste globale Todesursache. Weltweit leiden mehr als 64 Millionen Menschen an Herzinsuffizienz, die bei älteren Menschen die häufigste Ursache für Krankenhausaufenthalte ist (Savarese et al., 2023; World Health Organization, 2025). Herzinsuffizienz ist eine vielschichtige Erkrankung, die mit erheblicher Morbidität, hoher Mortalität und (Re-)Hospitalisierungsrate, sowie verminderter Lebensqualität und hohen Gesundheitskosten verbunden ist (Savarese et al., 2023). Darüber hinaus können Langzeitschäden bzw. Spätfolgen der Infektionskrankheit COVID-19, das sogenannte Post-COVID-Syndrom (PCS), auch zu Herzschädigungen führen. Die Pandemie hat aufgezeigt, dass im Gesundheitssystem noch zu wenig hybride Interaktionssysteme zur kontaktreduzierten Versorgung und intersektoralen Kommunikation eingesetzt werden. Das Nachsorgemanagement nach stationären Krankenhausaufenthalten ist seit Jahren Gegenstand von Gesetzgebung, Standardisierung und Forschung. Obwohl Krankenhäuser in den letzten Jahren für die internen Entlass- und Nachsorgeprozesse Managementstellen eingerichtet haben und somit Verbesserung in der Organisation und Digitalisierung erreichen konnten, gibt es immer noch starke Defizite sowohl in der intersektoralen Kommunikation, vor allem mit der Hausärztin oder dem Hausarzt als auch insbesondere mit Patientinnen und Patienten im häuslichen Umfeld. Dieser Bruch der Versorgungskontinuität wirkt sich negativ auf chronisch erkrankte Patientinnen und Patienten mit einem hohen Bedarf an Nachsorge nach stationären Krankenhausaufenthalten aus, einschließlich der Patientinnen und Patienten mit chronischer Herzinsuffizienz und der am PCS erkrankten Patientinnen und Patienten. Digitale Gesundheitstechnologien wie Telemonitoring-Systeme haben das Potenzial, die Belastung der Gesundheitssysteme durch chronische Krankheiten wie Herzinsuffizienz zu verringern, indem sie eine kontinuierliche Überwachung der Patientinnen und Patienten und rechtzeitige Interventionen im Falle einer Zustandsverschlechterung ermöglichen (Bachtiger et al., 2020; Koehler et al., 2018). Gleichzeitig bieten die zunehmende Verfügbarkeit tragbarer Sensoren und kontinuierlicher Datenströme (z. B. Vitalparameter) beispiellose Möglichkeiten für den Einsatz von künstlicher Intelligenz (KI) zur Unterstützung bei klinischen Entscheidungen und im Nachsorgemanagement (Johnson et al., 2018; Topol, 2019). Ein prominentes Beispiel hierfür ist die Apple Watch, welche Abweichungen der Herzfrequenz oder Unregelmäßigkeiten im Herzrhythmus erkennen kann. Bei der Verarbeitung sensibler personenbezogener Daten mit KI sind jedoch hohe Anforderungen in Bezug auf Transparenz, Ethik, Rechtsprechung und Robustheit zu erfüllen, da Fehlentscheidungen schwerwiegende Konsequenzen nach sich ziehen können.

1.2 Gesamtziel

Das Gesamtziel dieses Verbundprojekts war die Entwicklung, Erprobung und (Prozess-)Evaluation eines Präventions- und Nachsorgemanagements mit Hilfe von Sensorik zur kontaktlosen Erfassung von Gesundheitsdaten, eines multimodalen, telemetrischen Interaktionssystems und KI-unterstützter Predictive Analytics zur Vermeidung nachklinischer Komplikationen und Drehtür-Effekte am Beispiel von Patientinnen und Patienten mit chronischer Herzinsuffizienz und Post-Covid-Syndrom bedingten Herzerkrankungen. Mit Hilfe des hybriden, multimodalen Interaktionssystems sollte eine patientenorientierte Kommunikation von Patientinnen und Patienten, deren pflegenden Angehörigen, Hausärztin oder Hausarzt und Klinik sowie die Übertragung und Integration von Gesundheitsdaten zwischen diesen Akteursgruppen ermöglicht werden. Damit sollte eine Schnittstelle zwischen Patientinnen oder Patienten im häuslichen Umfeld bzw. im Alltag und Fachpersonal

in Praxis und Klinik ermöglicht werden. Der Einsatz von Sensorik und KI sollte dazu genutzt werden, automatisiert kombinierte Parameter zu erfassen, auszuwerten und Entgleisungen von Vitalparametern bzw. kritische Gesundheitszustände vorauszusagen bzw. dahingehend Patientinnen und Patienten und Gesundheitsakteure bei Erreichen bestimmter Schwellenwerte oder auffälliger Muster multimodal zu warnen. Es hat sich in zahlreichen Studien gezeigt, dass Telemonitoring die Mortalitätsrate bei Herzinsuffizienz senken kann (Deutsches Ärzteblatt, 2019). Darüber hinaus sollte ein kontinuierlicher Informationsfluss zwischen Klinik und Hausarztpraxis ermöglicht werden, wodurch die Hausärztin oder der Hausarzt in die weitere Therapie der Patientin oder des Patienten bereits während des Entlassmanagements integriert wird. Durch den Einsatz von Machine Learning (ML) Modellen, die in diesen Prozess integriert werden sollten, ergaben sich besondere Anforderungen. Für die Unterstützung der Prozesse im Gesundheitswesen war es notwendig, dass nachvollzogen werden kann, weshalb Modelle zu einer bestimmten Entscheidung/Empfehlung gelangt sind. Um diese Transparenz sicher zu stellen, sollten sowohl der Entwicklungsprozess des Modells als auch die eigentliche Entscheidungsfindung erklärt und dokumentiert werden. Zur Nachvollziehbarkeit des Entwicklungsprozesses sollten alle relevanten (Hyper-)Parameter des für eine Entscheidung genutzten Modells dokumentiert und die zugehörigen Daten und Modell-Ausprägungen versioniert werden. Zur Schaffung von Transparenz in der eigentlichen Entscheidungsfindung kamen interpretierbare Algorithmen, wie beispielsweise (gelernte) Entscheidungsbäume oder regelbasierte Systeme, in Frage. Alternativ konnte versucht werden, im Nachgang so viele Informationen wie möglich über die Entscheidung zu generieren. Ansätze hierfür sind Surrogate Model oder auch SHAP-Value zur Analyse des Einflusses der genutzten Inputparameter auf die Entscheidung des Modells. Durch den Einsatz von KI sollten auf diese Weise alle Akteure in ihrem Handeln unterstützt werden und auf Auffälligkeiten in Daten hingewiesen werden. Neben der Erklärbarkeit sollte auch dem Datenschutz eine hohe Bedeutung zukommen, da in bisherigen Ansätzen von ML meist zentrale Ansätze überwiegen. Dies ist in der Gesundheitsversorgung umstritten, da es sich bei den Gesundheitsdaten um sehr sensible Daten handelt. Um den Schutz der sensiblen Daten gerecht zu werden, sollten nach Möglichkeit die Trainingsdaten pseudonymisiert werden. Mittels Ansätzen des Federated Learnings sollte ermittelt werden, wie Modelle auf verschiedenen Ebenen bei der Patientin/dem Patienten, bei der behandelnden Ärztin oder dem behandelnden Arzt und global (weiter-)entwickelt werden können.

1.3 Hauptaufgabe des DFKI

Das DFKI übernahm im Projekt hauptverantwortlich die Entwicklung von KI-Algorithmen, die die Gesundheitsparameter der Patientinnen und Patienten analysieren und auswerten. Hierfür spielten die Entwicklung des sozio-technischen Konzepts und das Datenverarbeitungskonzepts eine zentrale Rolle. Im Rahmen von Arbeitspaket (AP) 2.2 arbeitete das DFKI an der Gestaltung einer robusten Datenverarbeitungslösung, die eine nahtlose Integration von Nachsorgedaten ermöglicht. Dabei wurden nicht nur technische Anforderungen berücksichtigt, sondern auch sozio-technische Aspekte einbezogen, um eine effektive Interaktion zwischen KI-System, medizinischem Personal und Patientinnen und Patienten sicherzustellen. Im AP3.1 unterstützte das DFKI die Strategion GmbH bei der iterativen Implementierung der Soll-Konzepte und lieferte insbesondere die KI-Modelle zu. Nach Fertigstellung der jeweiligen System-Komponenten evaluierte das DFKI, inwieweit die entwickelten KI-Methoden richtige Prognosen liefern und ob die bereitgestellten Erklärungen das medizinische Personal bei der Entscheidungsfindung unterstützen (AP4). Zudem wurden Fragestellungen bezüglich der ethischen, legalen und sozialen Implikationen (ELSI) für die Gewährleistung von Datensicherheit und Datenschutz behandelt und Onboarding-Material für Ärztinnen und Ärzte gestaltet.

1.4 Planung und Ablauf des Vorhabens

Die ursprünglich geplanten Meilensteine sind Tabelle 1 zu entnehmen. Im Verlauf des Verbundprojektes haben sich zu initial skizzierten Arbeits- und Zeitplan jedoch Veränderungen und Verzögerungen ergeben, die eine Verlängerung der Projektlaufzeit um sechs Monate unerlässlich gemacht haben. Die Verzögerungen bei der Entwicklung und Bereitstellung der Smart Patches (AP3.2) wirkten sich sowohl auf die vorgelagerten als auch auf die nachgelagerten APs aus. Es kam insbesondere zu einer zeitlichen Verschiebung bei der Bereitstellung von Daten, welche essenziell für die Bearbeitung des AP2.2 waren, in dem das DFKI im Rahmen des sozio-technischen Konzepts und des Datenverarbeitungskonzepts die Gesamtarchitektur entwickelte. Entsprechend beeinflussten die Verzögerungen in der Bereitstellung des Datensatzes die Entwicklung und Anpassung der KI-Modelle (AP3.1) sowie die summative Evaluation (AP4). Durch die genehmigte kostenneutrale Laufzeitverlängerung konnte ein formal angemessener Datensatz generiert werden, der als Grundlage für die konzeptionelle Ausgestaltung des KI-Demonstrators genutzt werden konnte und zusammenhängende Evaluationen ermöglichte. Dadurch konnte sichergestellt werden, dass die entwickelte Daten- und Modellpipeline etabliert ist und bei Verfügbarkeit geeigneter Telemonitoring-Daten unmittelbar weiterentwickelt und validiert werden kann.

AP	Monate ab Projektstart	Meilensteine
1	6	Prozessdokumentation und -analyse, Anforderungskatalog, User Stories
2	12	
2.1	12	Fachkonzept, SOLL-Prozesse
2.2	12	Architektur, Anwendungskonzepte, Entwicklungsplan
3	27	
3.1	27	Demonstratoren, Prozesse
3.2	27	Smart Patches, Entwicklungsvorgehen (Kosten, Effizienz, Usability)
4	33	Formative und summative Evaluationsergebnisse, Evaluations-Framework
5	36	Transferkonzept
6	36	Datenschutzkonzept, Datenmanagementplan, ethische Richtlinien, Rahmenwerke, Etablierung und Durchführung eines Arbeitskreises zu ELSI-Implikationen
7	36	Organisationsstruktur, Kommunikationskonzept, Wissenstransfer (Formate, Veröffentlichungen)

Tabelle 1: Projektmeilensteine mit Zeitplanung und zugehörigen APs.

1.5 Stand der Wissenschaft und Technik zu Projektbeginn

Einerseits konnte in der Praxis gezeigt werden, dass Patientinnen und Patienten mit einer fortgeschrittenen Herzinsuffizienz von einer telemedizinischen Betreuung stark profitieren können (Deutsches Ärzteblatt, 2019). Andererseits zeigen wissenschaftliche Beiträge auch, dass KI eine weitreichende Verbesserung in der Präzision der kardiovaskulären Medizin ermöglichen kann (Benjamins et al., 2019; Johnson et al., 2018; Krittanawong et al., 2017). Dennoch ist KI und eine nahtlose Kommunikation im Präventions- und Nachsorgemanagement in der Praxis noch kaum im Einsatz. Innovationsfonds-Projekte wie Vespeera und TRADE behandeln das Entlassungsmanagement im Rahmen der Prozessoptimierung, jedoch ohne die Veränderung

des Prozesses durch die Einbindung oder Entwicklung eines neuen IT-Instrumentariums mit dem Ziel von Kontaktreduzierungen und einer Entlastung der Gesundheitsakteure. Dabei stehen vor allem im Projekt Vespeera die Versorgungspfade zwischen Hausarztpraxen und Krankenhäusern im Vordergrund. Bereits geförderte Innovationsfonds-Projekte des aQua Instituts (Entlassungsmanagement, EMSE und USER) fokussieren mithilfe von Routinedaten und IT-Systemen die nahtlose Überleitung zu anderen Gesundheitsakteuren nach der Entlassung. Damit sollen die Patientinnen und Patienten aufeinander abgestimmte medizinische Maßnahmen erhalten. Die Abgrenzung zu dem hier beantragten Verbundprojekt liegt insbesondere darin, dass das Nachsorgemanagement und die Prävention mit Hilfe multimodaler Interaktionssysteme im Zentrum dieser beiden Projekte nicht betrachtet wurde bzw. wird, was jedoch zentraler Gegenstand des hier skizzierten Projektes ist. Die veröffentlichten Ergebnisse des 2018 abgeschlossenen Projektes EMSE, das u.a. Empfehlungen für eine Entlassungsplattform erarbeitet hat, können für die Entwicklung der multimodalen Interaktions-Plattform in diesem Projekt nutzenbringend verwendet werden. Coaching- und Monitoring-Komponenten werden vereinzelnd in Projekten mit Fokus auf Herzinsuffizienz (DMP HI PLUS, HITS sowie HerzEffektMV), dem Innovationsfonds-Projekt Telnet@NRW oder mit Fokus auf Palliativpatienten (OnCoPaTh) über eine internetbasierte eHealth-Plattform unterstützt. Darin wird weniger Fokus auf eine multimodale Interaktion sowie auf den (prädiktiven) Einsatz von KI gesetzt. In weiteren Projekten werden vereinzelnd KI-Analysen erforscht, um bspw. Risikoindikatoren für kardiopulmonale Dekompensation auf Intensivstationen durch Monitoring von Vitalparametern zu erkennen (KI-SIGS). Die Forschung steht jedoch nicht im Gesamtzusammenhang zum Prozessmanagement chronischer Krankheiten und der Übermittlung und Integration von Gesundheitsdaten und Warnungen durch multimodale hybride Interaktionssysteme in Ausnahmesituationen. Die Innovationshöhe im Bereich Telemonitoring gegenüber bisherigen Lösungen lässt sich zum einen durch das multimodale Datenerfassungssystem und zum anderen durch das innovative KI-gestützte Datenverarbeitungssystem darlegen. Die in diesem Verbundprojekt eingesetzte Sensorik sollte nicht nur Vitaldaten über gängige Hardware wie Smartwatches messen, sondern auch individualisierte Smart Patches einsetzen, deren Messumfang über herkömmliche Informationen wie Herzfrequenz oder Schritte hinausgeht. Die Smart Patches sollen perspektivisch minimal-invasiv für Herzerkrankungen relevante Parameter wie pH, Hämoglobin, Kalium und NT-proBNP messen können. Die technische Realisierung des Projektpartners addSensors GmbH soll kontinuierliche Messungen ermöglichen, ohne die Patientinnen und Patienten im Alltag zu beeinträchtigen. Diese Patches gehen über den aktuellen Stand der Technik hinaus, da sie vergleichsweise klein sein und ein geringes Gewicht haben sollen und zudem so nachhaltig gestaltet werden. Diese technische Realisierung in Verbindung mit dem gemessenen breiten Datenspektrum grenzt dieses Verbundprojekt von verwandten Innovationsfonds-Projekten wie HITS oder BMFTR-Projekte wie Fontane ab. Mittels Sensorik ermittelte Vitaldaten in Verbindung mit Informationen aus einer Patientenakte reichern die medizinische Wissensbasis über Patientinnen und Patienten in deren häuslichen Umfeld und Alltag kontinuierlich an und können mit probabilistischen KI-Modellen überwacht und ausgewertet werden. Durch die entstehenden patientenübergreifenden Datenmuster können so auch weniger offensichtliche Risikoindikatoren erkannt werden. Die daraus resultierenden Auffälligkeiten und Empfehlungen der KI können vom medizinischen Personal validiert und entsprechende Maßnahmen eingeleitet werden. Auf diese Weise wird das Datenverarbeitungssystem um eine maschinelle Ebene erweitert, bei gleichzeitiger Entlastung des medizinischen Personals. Dies unterscheidet die vorliegende Lösung von bisherigen geförderten Telemonitoring-Ansätzen, die ausschließlich auf menschliche Expertise zur Auswertung von kontinuierlich gemessenen (HITS,

conneCT CF) oder selbst eingetragenen Daten setzen (sekTOR-HF) oder andere Krankheiten als chronische Herzinsuffizienz bzw. Post-Covid-Syndrom bedingte Herzerkrankungen fokussieren (DECIDE).

Zusammenfassend liegt die Abgrenzung zu verwandten Projekten im Rahmen des Entlassmanagements und des Monitorings von chronischen Krankheiten sowie die Innovationshöhe in diesem Verbundprojekt in der nahtlosen Kommunikation von Hausärztin/-arzt und Klinik sowie der Nutzung von individuell, patientenzentriert zusammengestellter Sensorik (Smart Patches), der sektorenübergreifenden Integration von Gesundheitsdaten, der Entwicklung von KI-Algorithmen zur systematischen Erkennung von auffälligen Parametern (Vitaldaten) und dem Einleiten präventiver Maßnahmen im häuslichen Umfeld sowie der Vermittlung von Edukationsinhalten durch Heart Failure Nurses, um letztendlich die Krankenhauswiederaufnahmen bei chronisch an Herzinsuffizienz erkrankten Patientinnen und Patienten präventiv, auch durch Einsatz von KI, zu verringern und somit Kontakte zu reduzieren.

1.6 Zusammenarbeit mit anderen Stellen

Die geförderten Partner des Verbundprojekts sind Tabelle 2 zu entnehmen. Darüber hinaus wurden Ärztinnen und Ärzte aus mindestens zwölf (bei anonymer Umfrage nicht nachvollziehbar) Kliniken und Arztpraxen, sowie Heart Failure Nurses und Patientinnen und Patienten des Universitätsklinikum Münsters für die verschiedenen Evaluationszyklen einbezogen.

Institution	Ansprechpartner
Universität Osnabrück Unternehmensrechnung und Wirtschaftsinformatik (UWI)	Prof. Dr. Frank Teuteberg frank.teuteberg@uni-osnabrueck.de
Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) Smart Enterprise Engineering	Prof. Dr. Oliver Thomas oliver.thomas@dfki.de
Strategion GmbH	Dr. Florian Remark florian.remark@strategion.de
addSensors GmbH (ehemalig InnoME GmbH)	Eike Kottkamp eike.kottkamp@addsensors.com
Medizinische Fakultät, Universitätsklinikum Münster (UKM) Interdisziplinäre Sektion Herzinsuffizienz	Prof. Dr. Jürgen Sindermann juergen.sindermann@ukmuenster.de
Ärztekammer Westfalen-Lippe (ÄKWL)	Thomas Althoff thomas.althoff@aekwl.de

Tabelle 2: Übersicht der geförderten Partner des Verbundprojekts.

2 Ergebnisse

2.1 Zentrale Ergebnisse und Erkenntnisse des DFKI

AP1: Erfassung der Problemstellung und Definition der Anforderungen

Der Fokus des AP1 lag auf der systematischen Analyse der Ausgangssituation sowie der strukturierten Erhebung und Konsolidierung der Anforderungen aller relevanten Akteursgruppen. Das Ziel bestand darin, eine

belastbare Grundlage für die Konzeption eines KI-gestützten, vitalparameterbasierten Nachsorgemanagements für Patientinnen und Patienten mit chronischer Herzinsuffizienz zu schaffen. Zu diesem Zweck wurden die Anforderungen der beteiligten Gesundheitsakteure, insbesondere Ärztinnen und Ärzte, erhoben und analysiert. Die Ergebnisse wurden in enger Zusammenarbeit mit der Strategion GmbH in Form von Personas und User Stories aufbereitet, um Nutzungskontexte, Informationsbedarfe und Interaktionsanforderungen nachvollziehbar abzubilden. Die Anforderungserhebung umfasste sowohl technische Aspekte (etwa hinsichtlich der eingesetzten Plattformen und Architekturen) als auch organisatorische und datenschutzbezogene Fragestellungen zur Datenhaltung, -verarbeitung und -kommunikation.

Parallel dazu erfolgte eine umfangreiche Recherche zum Stand der Technik im Bereich maschineller Lernverfahren zur Analyse von Gesundheitsdaten im Telemonitoring. Der Schwerpunkt lag dabei auf der Verarbeitung tabellarischer Zeitreihendaten, existierenden ML-Lösungen mit vergleichbarer Problemstellung sowie etablierten Datenaufbereitungs- und Evaluationsansätzen. Ein weiterer zentraler Aspekt war die Identifikation geeigneter Methoden zur Erklärbarkeit von KI-Modellen, um den Anforderungen an Transparenz, Nachvollziehbarkeit und Akzeptanz im medizinischen Kontext gerecht zu werden.

Auf Basis der nationalen Versorgungsleitlinie für chronische Herzinsuffizienz (Bundesärztekammer (BÄK) et al., 2023) wurden die für das Vorhaben relevanten medizinischen Parameter identifiziert. Zur Erhebung der erforderlichen Vitalparameter wurde eine Marktanalyse geeigneter Wearables durchgeführt. Initial wurde die Huawei Watch D2 ausgewählt, da sie zum Zeitpunkt der Recherche als einziges Gerät alle benötigten Parameter, insbesondere eine eigenständige Blutdruckmessung, medizinproduktkonform erfassen und plattformunabhängig bereitstellen konnte. Aufgrund des geringen Marktanteils und Bedenken hinsichtlich des Datenschutzes wurde jedoch im Nachgang entschieden, dass die Nutzung von Smartwatches von Samsung oder Apple in Kombination mit einem separaten Blutdruckgerät vorzuziehen ist. Die Auswahl der Parameter und die zugehörige Erhebungsmethoden sind in Tabelle 3 gelistet. Parameter mit explizitem COVID-19-Bezug wie die Häufigkeit und der Zeitpunkt vergangener Infektionen oder das Vorliegen eines Post-COVID-Syndroms wurden im Rahmen der Anforderungsanalyse inhaltlich berücksichtigt, aufgrund des hohen Durchseuchungsgrads und der daraus resultierenden eingeschränkten Kausalität für das vorliegende Anwendungsszenario jedoch bewusst nicht in den initialen Parameterumfang für den Proof-of-Concept aufgenommen.

Personenbezogene Daten (einmalig bei Registrierung)	**Körperliche Untersuchung** (täglich über Tagebuchfunktion)	**Vitalzeichen** (täglich, mit Smartwatch und Blutdruckmessgerät)	**Laborwerte** (täglich mit Smart Patch)
Geschlecht	Körpergewicht	Blutdruck	pH
Alter	Dyspnoe (Atemnot bei geringer Belastung)	Herzfrequenz	Laktat
NYHA-Klasse	Orthopnoe (Kurzatmigkeit im Liegen)	Atemfrequenz	Hämoglobin
	Periphere Ödeme (Ansammlung von Flüssigkeit in den Beinen)	Sauerstoffsättigung	Glukose (Blutzucker)
	Schwindel	Temperatur	Kalium
			Natrium
			Serumkreatinin
			Korrigiertes NT-proBNP

Tabelle 3: Betrachtete Parameter für die Vorhersage einer bevorstehenden Dekompensation.

Anschließend wurde die existierende Literatur hinsichtlich realistischer Wertebereiche, zeitlicher Dynamiken und krankheitsspezifischer Unterschiede analysiert. Basierend auf den vorliegenden Daten wurden potenzielle Zielvariablen (z. B. Dekompensationsstatus, Diagnosen, Operationen und medizinische Maßnahmen) systematisch verglichen und bewertet. Da keine öffentlich verfügbaren Telemonitoring-Datensätze mit ausreichender Granularität existierten, wurden zudem Möglichkeiten der synthetischen Datengenerierung sowie alternative Datenquellen geprüft.

AP2: Soll-Konzepterarbeitung – Fachkonzept und Datenverarbeitungskonzept

AP2.1: Fachkonzeptionierung des nahtlosen, vitalparameterbasierten Nachsorgemanagements

In AP2.1 wurde das fachliche Soll-Konzept für ein durchgängiges, vitalparameterbasiertes Nachsorgemanagement entwickelt. Zunächst wurde definiert, an welchen Stellen im Versorgungsprozess relevante Gesundheitsdaten erhoben werden und wie diese strukturiert in das Gesamtsystem einfließen. Entsprechend des in Tabelle 3 dargestellten Datenkonzepts hat das DFKI bei der Entwicklung der Architektur des multimodalen Interaktionssystems (siehe Abbildung 1) sowie des zugehörigen Prozessdiagramms unterstützt.

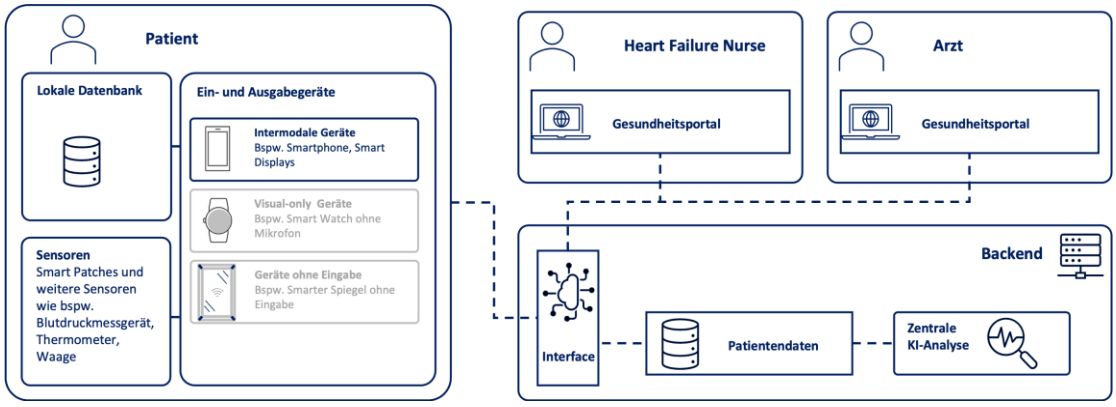


Abbildung 1: Architektur für das multimodale Interaktionssystem.

Ein wesentlicher Bestandteil der Fachkonzeption war die Entwicklung geeigneter Erklär- und Informationsformate. Für Ärztinnen und Ärzte wurden begleitende Onboarding-Formate konzipiert, die die Auswahl der zugrunde liegenden Daten für das Training des KI-Modells bzw. für das Gesamtsystem einordnen. Für Patientinnen und Patienten wurden verständliche Erklärtexte zu den einzelnen Parametern und deren Bedeutung innerhalb der App erstellt. Darüber hinaus wurden erste Überlegungen zur Ausgestaltung eines Human in the Loop Konzepts angestellt, das die Rollen der Akteure bei der Datenselektion, dem Modelltraining und der Bewertung der KI-Ergebnisse berücksichtigt. Ein manuelles Annotieren (Labeln) der Trainingsdaten durch medizinisches Fachpersonal ist dabei grundsätzlich nicht vorgesehen, da der Zeitpunkt einer Dekompensation den zugrunde liegenden klinischen Daten entnommen werden und somit eine retrospektive Zuordnung automatisiert erfolgen soll. Ergänzend ist innerhalb der Anwendung eine Feedbackfunktion vorgesehen, über die behandelnde Ärztinnen und Ärzte die Plausibilität von Warnmeldungen oder Prognosen rückmelden können. Dadurch soll eine kontinuierliche Qualitätssicherung des Modells ermöglicht werden. Die initiale Datenselektion erfolgte in enger Abstimmung mit erfahrenen Kardiologinnen und Kardiologen sowie unter Berücksichtigung etablierter Versorgungsleitlinien. Erweiterungen der Parameterauswahl und Anpassungen an neue wissenschaftliche Erkenntnisse sind im Systemdesign ausdrücklich vorgesehen.

AP2.2: Sozio-technische Konzeption / Datenverarbeitungskonzept

Aufbauend auf dem Fachkonzept aus AP2.1 wurde in AP2.2 das sozio-technische Gesamtkonzept sowie das zugehörige Datenverarbeitungskonzept ausgearbeitet. Ziel war es, die fachlichen Anforderungen in eine tragfähige technische und organisatorische Architektur zu überführen. Hierzu wirkte das DFKI maßgeblich bei der Gestaltung des Anwendungs- und Interaktionsdesigns sowie bei der inhaltlichen Überarbeitung der Mockups entsprechend des Datenkonzepts mit. Dies erfolgte in enger Abstimmung mit dem Konsortium sowie in Form von Interviews mit sieben erfahrenen Kardiologinnen und Kardiologen. Das Online-Portal für Ärztinnen und Ärzte wurde als zentrale Komponente zur Teleüberwachung bei chronischer Herzinsuffizienz konzipiert. Es umfasst KI-basierte Warnmeldungen, die das Risiko einer bevorstehenden Dekompensation frühzeitig anzeigen (siehe Abbildung 2).

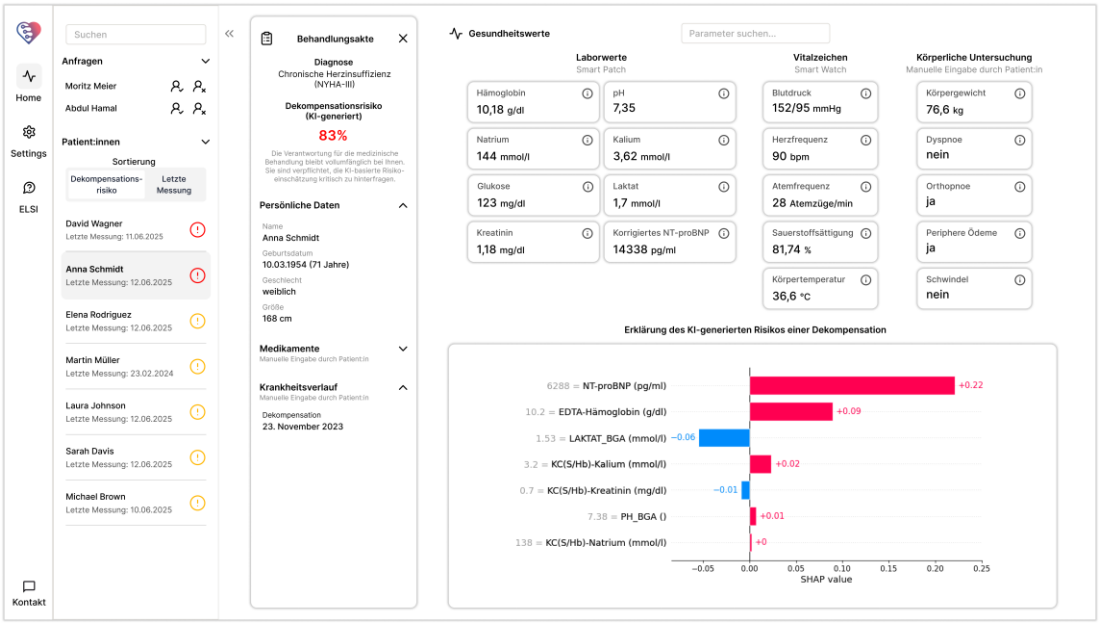


Abbildung 2: Mockup des Online-Portals für Ärztinnen und Ärzte.

Die korrespondierende App für Patientinnen und Patienten dient der Erfassung der mittels Sensoren gemessenen Gesundheitsparameter sowie zusätzlicher körperlicher Symptome, die über eine Tagebuchfunktion dokumentiert werden können (siehe Abbildung 3).

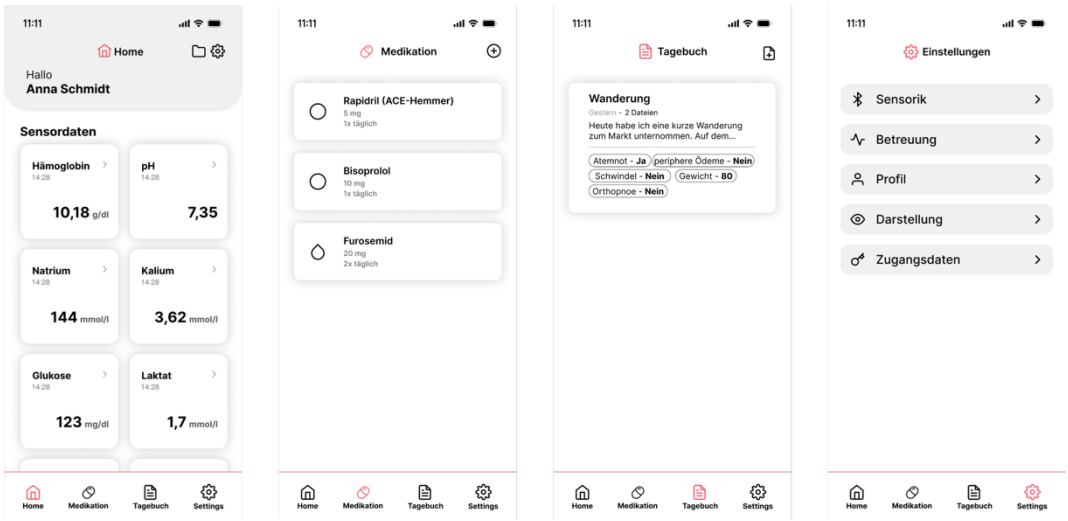


Abbildung 3: Mockup der App für Patientinnen und Patienten.

Im Rahmen des Datenverarbeitungskonzepts wurden die zu erhebenden Daten, deren Herkunft sowie deren zeitliche Verfügbarkeit detailliert beschrieben (entsprechend Tabelle 3). Aufgrund von Verzögerungen bei der Entwicklung der Smart Patches musste jedoch von der ursprünglich geplanten Datenerhebung abgewichen werden. Stattdessen wurden in einer neuen Studie des UKM Daten erfasst, die als Grundlage für die Entwicklung der KI-Ansätze dienen sollten. In diesem Zusammenhang wurden auch Active-Learning-Ansätze zur Bewältigung fehlender Daten bzw. Labels untersucht. Aufgrund des hohen Ressourcenbedarfs für die Entwicklung einer entsprechenden technischen Infrastruktur (einschließlich der Konzeption und Implementierung einer dedizierten Plattform zur Annotation der medizinischen Daten, der Integration geeigneter Vorselektions- und Priorisierungsalgorithmen sowie der Sicherstellung regulatorischer und datenschutzrechtlicher Anforderungen) und des erheblichen zeitlichen Aufwands für qualifiziertes medizinisches Fachpersonal zur manuellen Annotation der Daten, wurde jedoch von einer Umsetzung abgesehen.

Darüber hinaus wurden verschiedene Strategien zur Datenversionierung sowie unterschiedliche Ansätze zur datenschutzkonformen Modellierung analysiert, darunter zentrale Modellerstellung und Federated Learning. Beim Federated Learning wird das KI-Modell dezentral auf mehreren Endgeräten mit jeweils lokalen Datensätzen trainiert, ohne dass die zugrunde liegenden Rohdaten das jeweilige Gerät verlassen. Stattdessen werden in regelmäßigen Abständen ausschließlich die aktualisierten Modellparameter an einen zentralen Server übermittelt, dort aggregiert und zu einem globalen Modell zusammengeführt, welches anschließend wieder an die lokalen Instanzen zurückgespielt wird. Für die initiale Umsetzung wurde bewusst auf ein zentrales Modelltraining gesetzt, da dieses eine schnellere Anpassung an neue Daten und Anforderungen ermöglichte und angesichts der noch begrenzten und nicht ausreichend stabilen Datengrundlage zweckmäßiger war. Da die verfügbaren Datensätze weder stark heterogen noch institutionell verteilt waren, bestand keine Notwendigkeit für ein dezentrales Trainingsverfahren. Ein Federated Learning Ansatz, insbesondere unter Einbindung mehrerer Arztpraxen oder Kliniken als eigenständige Instanzen, wird jedoch perspektivisch im Rahmen einer Skalierung als sinnvoll erachtet.

AP3: Implementierung: Interaktionsplattform und Smart Patches

AP3.1: Iterative sozio-technische Implementierung

In AP3.1 erfolgte die iterative Umsetzung der zuvor entwickelten Konzepte und die schrittweise Überprüfung der Anforderungen. Hierzu wurde der vom UKM bereitgestellte Datensatz analysiert, aufbereitet und als Grundlage für den Einsatz vom ML-Modellen nutzbar gemacht. Dieser umfasste routinemäßige Blutuntersuchungsergebnisse (z. B. Sauerstoffsättigung), die Klassifizierung der Herzinsuffizienz (z. B. NYHA-Klassifikation), Befunde aus diagnostischen Verfahren (z. B. Spiroergometrie), klinische Diagnosen (z. B. Klassifikation der Kardiomyopathie), Komplikationen (z. B. Dekompensation), medizinische Interventionen (z. B. Implantation eines linksventrikulären Herzunterstützungssystems), Krankenhausaufenthalte (z. B. Aufnahmedatum) und personenbezogene Merkmale (z. B. Alter) von 118 Patientinnen und Patienten. Im Verlauf der Analyse zeigte sich, dass die Daten insbesondere hinsichtlich der Vollständigkeit und Granularität nur eingeschränkt für ein Telemonitoring-Szenario geeignet waren. Während das Telemonitoring die tägliche Erfassung mehrerer Messwerte erfordert, um subtile Trends zu erkennen, lieferten die klinischen Kontrolluntersuchungen in dieser Kohorte in der Regel nur einen einzigen Messwert alle drei bis sechs Monate. Diese unzureichende Messfrequenz beeinträchtigte die für unseren speziellen Anwendungsfall erforderliche Repräsentativität. Zudem wurde die Messung hochrelevanter Parameter wie NT-proBNP nur durchgeführt, wenn eine klinische und finanzielle Rechtfertigung dafür vorlag, was zu einem erheblichen Anteil fehlender Werte

führte. Andere beinhaltete Parameter, wie beispielsweise Spiroergometrie-Ergebnisse, lassen sich wiederum nicht auf ein Fernüberwachungssystem mit tragbaren Geräten übertragen. Ein weiteres wichtiges Problem betraf die Zielvariable (d. h. den Dekompensationsstatus), die lediglich pro medizinischem Fall verfügbar war und nicht einem genauen Zeitstempel oder einer bestimmten Messung zugeordnet werden konnte. Wir prüften Medikamentendaten als möglichen Ersatz für diese fehlenden Zeitstempel, da die Verabreichung eines bestimmten Medikaments (Furosemid) auf eine aktive Dekompensation hindeutet. Die Zeitstempel der Medikamentengabe stimmten jedoch selten mit denen der Messwerte aus den Laboruntersuchungen überein, und die Aufzeichnungen waren nur für wenige Patientinnen und Patienten verfügbar. Zur Erweiterung und Strukturierung der verfügbaren Informationen wurde ergänzend der Einsatz eines Large Language Models (LLM) geprüft und umgesetzt, um Inhalte aus unstrukturierten Arztbriefen automatisiert zu extrahieren. Arztbriefe enthalten wichtige Gesundheitsinformationen (z. B. Diagnosen, Anamnese, Untersuchungsbefunde, Medikation), liegen jedoch in heterogen formatierten Freitexten mit variierender Struktur und Terminologie vor. Die technische Herausforderung bestand insbesondere in der zuverlässigen Kontextinterpretation sowie der standardisierten Überführung der extrahierten Inhalte in eine tabellarische, maschinenlesbare Form. Hierzu wurde ein vortrainiertes Sprachmodell (Llama 3.1) integriert, das relevante Informationen kontextsensitiv identifiziert und vordefinierten Parametern (z. B. Medikationsgruppen) zuordnet. Die extrahierten Daten wurden anschließend automatisiert kodiert (z. B. binäre oder numerische Repräsentation kategorialer Merkmale), um eine konsistente Eingliederung in die bestehende Datenstruktur zu gewährleisten. Im bereitgestellten Datensatz waren die Arztbriefe jedoch nur für eine geringe Anzahl von Patientinnen und Patienten vorhanden, so dass die bestehenden Lücken dadurch nicht systematisch gefüllt werden konnten.

Aufgrund dieser Faktoren waren die klinischen Kontrolluntersuchungen sowohl für das Modelltraining als auch als Grundlage für die Generierung synthetischer Telemonitoring-Daten, für die repräsentative, vielfältige und kontinuierliche Datenströme erforderlich sind (Benaim et al., 2020), unzureichend. Um hinsichtlich der KI-Entwicklung dennoch einen Proof-of-Concept realisieren zu können, wurde ein formal angemessener Telemonitoring-Datensatz generiert, der sich an den relevanten Informationen aus den klinischen Kontrolluntersuchungen, medizinischen Leitlinien (Bundesärztekammer (BÄK) et al., 2023) und Erkenntnissen aus der Literatur orientierte, jedoch keinen Anspruch auf klinische Validität erhob. Dabei wurde die experimentelle Anpassung der Messfrequenz pro Tag, des Gesamtbeobachtungszeitraums sowie der Anzahl der simulierten Patientinnen und Patienten ermöglicht. Die generierten Messwerte wurden in realistischen Wertebereichen modelliert und zur Abbildung realistischer Krankheitsdynamiken wurden Baseline-Zustände modelliert, welchen kontrollierte Zustandsverbesserungen oder -verschlechterungen folgten, einschließlich mehrfach täglicher Vitalzeichenmessungen mit tageszeitabhängigen Schwankungen. Darüber hinaus erfolgte eine strukturierte Differenzierung nach klinisch relevanten Gruppenmerkmalen, insbesondere hinsichtlich Dekompensation gegenüber stabilem Verlauf, Geschlecht sowie NYHA-Klassifikation. Zusätzlich wurden Sonderfälle (z. B. tödliche Verläufe) berücksichtigt. Der generierte Datensatz wurde abschließend von den medizinischen Experten des Konsortiums auf Plausibilität und interne Konsistenz überprüft. Auf dieser Grundlage konnte die entwickelte Vorverarbeitungs- und Modellpipeline im Sinne eines methodischen Proof-of-Concept umgesetzt und erprobt werden. Bei Verfügbarkeit geeigneter Telemonitoring-Daten ist eine Weiterentwicklung und klinische Validierung unmittelbar möglich. Die Einschränkungen der Datenverfügbarkeit betreffen somit vorrangig die Datengrundlage, nicht jedoch die konzeptionelle Ausgestaltung des KI-Demonstrators.

Im nächsten Schritt wurden umfangreiche Vorverarbeitungsschritte implementiert, darunter Verfahren zum Umgang mit fehlenden Werten (Imputation), zur Skalierung von Merkmalen (Standardisierung), zur Kompensation von Klassenungleichgewichten (Over-/Undersampling) sowie zur Merkmalsselektion. Auf dieser Basis wurden verschiedene ML-Modelle (z. B. Long Short-Term Memory, Recurrent Neural Network) für die Vorhersage des Dekompensationsrisikos implementiert und anhand aussagekräftiger Metriken (z. B. Loss, Accuracy, F1-Score) miteinander verglichen.

Ergänzend wurde unter Berücksichtigung der Anforderungen der involvierten Akteure und der betrachteten Parameter eine Auswahl geeigneter (Darstellungs-)Methoden getroffen, um den Einfluss der betrachteten Parameter auf die KI-Vorhersage zu visualisieren: SHapley Additive exPlanations (SHAP), Counterfactual Explanations (CFEs) und Anchors. Ziel war es, durch die Modelltransparenz die Akzeptanz und das Vertrauen des medizinischen Fachpersonals zu erhöhen. Um sicherzustellen, dass der Einsatz von ML und die genaue Funktionsweise für die Nutzenden nachvollziehbar sind, wurden für die Ärztinnen und Ärzte sowie für die Patientinnen und Patienten jeweils aufklärende Onboarding-Formate zum Gesamtsystem vorbereitet.

AP3.2: Entwicklung individueller Smart Patches

Das AP3.2 zur Entwicklung individueller Smart Patches konnte aufgrund externer Abhängigkeiten und Verzögerungen nicht wie ursprünglich geplant umgesetzt werden. Stattdessen erfolgte eine enge Zusammenarbeit mit dem UKM zur Analyse und Nutzung von Daten aus klinischen Kontrolluntersuchungen. Die Qualität und Struktur der verfügbaren Daten wurden kritisch geprüft, zu formal angemessenen Telemonitoring-Daten erweitert und durch geeignete Vorverarbeitungsschritte aufbereitet. Dabei wurde die KI-Pipeline, einschließlich Vorverarbeitung, Modelltraining und Evaluation, modular und flexibel konzipiert, sodass eine nahtlose Übertragung auf künftig verfügbare, real erhobene Telemonitoring-Daten gewährleistet ist. Die ursprünglich angestrebte frühzeitige Mitgestaltung der Smart-Patches und zugehörige Analysen hinsichtlich Kosten, Effizienz und Usability konnten entsprechend jedoch nicht realisiert werden.

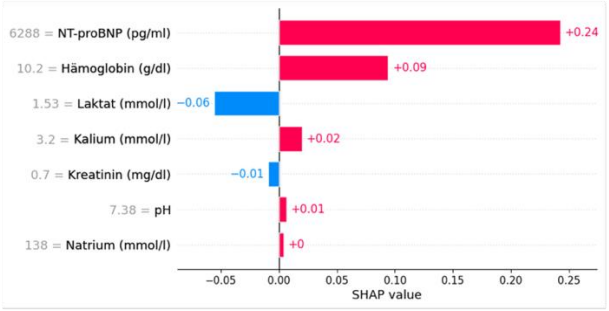
AP4: Formative und summative Evaluation

In AP4 wurden die entwickelten ML-Modelle sowie die zugehörigen Erklärungen evaluiert. Zur quantitativen Bewertung der Prognosegüte wurden verschiedene Modelle verglichen. Das Recurrent Neural Network erzielte die besten Ergebnisse, getestet auf simulierten Telemonitoring-Daten mit 300 Patientinnen und Patienten über einen Zeitraum von 30 Tagen, mit einer Messung pro Tag. Die Daten wurden in einem 70/15/15-Split in Trainings-, Validierungs- und Testdaten aufgeteilt und mittels Min-Max-Scaling normalisiert. Es wurden 14 Tage als Input- und 14 Tage als Output-Fenster verwendet. Das bedeutet, dass das Modell auf Basis der Gesundheitsdaten der jeweils vorangegangenen zwei Wochen eine Prognose darüber erstellt, ob innerhalb der darauffolgenden zwei Wochen ein erhöhtes Risiko für eine Dekompensation besteht. Durch dieses Zeitfenster können potenzielle Zustandsverschlechterungen mit ausreichend Vorlauf für die Ergreifung entsprechender medizinischer Maßnahmen erkannt werden. Das Training erfolgte über 45 Epochen. Auf den Testdaten erzielte das Modell einen Loss von 0,2, sowie 94% Accuracy, 90% Precision, 97% Recall, 93% F1-Score und 98% Area Under the Receiver Operating Characteristic Curve Score. Die Ergebnisse demonstrieren die technische Machbarkeit und eine hohe Trennschärfe im Proof-of-Concept-Szenario. Aufgrund der simulierten Datengrundlage ist jedoch von einem gewissen Overfitting auszugehen, d.h., das Modell passt

sich zu stark an die spezifischen Muster der Trainingsdaten an und verliert dadurch an Generalisierungsfähigkeit gegenüber neuen, ungesehenen Daten. Eine klinische Validierung auf Basis realer Telemonitoring-Daten ist perspektivisch erforderlich.

In Bezug auf die erforderliche Transparenz der Modellentscheidung stellte sich die Herausforderung, dass es sich um ein „Black Box"-Modell handelt, dessen Entscheidungslogik ohne zusätzliche Maßnahmen nicht nachvollziehbar ist. Durch den Einsatz von Methoden der Explainable AI (XAI) lässt sich der Einfluss einzelner Parameter im Nachgang jedoch aufschlüsseln, um Nachvollziehbarkeit, Akzeptanz und Vertrauen bei Ärztinnen und Ärzten zu stärken. Zur Evaluation der Erklärungen wurden 15 Interviews mit Kardiologinnen, Kardiologen und Heart Failure Nurses durchgeführt. Anhand eines echten Fallbeispiels wurden drei XAI-Methoden mithilfe der Explanation Satisfaction Scale (Hoffman et al., 2018) verglichen: SHAP, CFEs und Anchors (siehe Abbildung 4). Der Einsatz von XAI-Methoden wurde insgesamt positiv bewertet. Es zeigte sich jedoch, dass keine einzelne Methode für alle Zwecke geeignet ist, sondern ein mehrschichtiges Konzept erforderlich ist. Aufgrund der intuitiven grafischen Darstellung wurde SHAP als besonders geeignet für den hektischen klinischen Alltag eingeschätzt. CFEs wurden als nützliche, handlungsorientierte Ergänzung zur Ableitung therapeutischer Schritte oder für didaktische Zwecke gewertet, jedoch als zu komplex für die Routineanwendung. Anchors wurden als eher oberflächlich und trivial wahrgenommen. Der Nutzen der Erklärungen wurde deutlich hervorgehoben. So wurde beispielsweise geäußert: „Ja, die [Erklärung] ist von größter Bedeutung, weil wir uns in der Verlaufskontrolle dieser Patienten ja exakt nach den Parametern richten, die hier aufgeführt sind." Zudem zeigte sich eine hohe Übereinstimmung zwischen den erklärten Modellentscheidungen und der klinischen Erfahrung, etwa: „Es deckt sich halt wieder mit dem, was man im klinischen Alltag sieht. Vor allem, dass das NT-proBNP eigentlich so der relevante ausschlaggebende Faktor ist für eine Dekompensation."

1



3

WENN NT-proBNP (pg/ml) > 1174.99
UND Hämoglobin (g/dl) <= 10.20
DANN SAGE VORHER: hohe Dekompensationswahrscheinlichkeit (67-100 %)

2

Hämoglobin (g/dl)	NT-proBNP (pg/ml)	Laktat (mmol/l)	Kreatinin (mg/dl)	Natrium (mmol/l)	pH	Kalium (mmol/l)	Wahrscheinlichkeit Dekompensation
14.8 (+4.6)	6288.0	1.53	0.7	155.3 (+17.3)	7.38	3.2	50 % (-33 %)
10.2	2643.85 (-3644.15)	1.53	0.7	138.0	7.38	3.2	17 % (-66 %)
10.2	844.45 (-5443.55)	1.53	0.7	138.0	7.38	3.2	17 % (-66 %)

Abbildung 4: Vergleich verschiedener Erklärungsmethoden für die Vorhersage des KI-Modells.

Ergänzend wurde das Design des Online-Portals und der App jeweils mit Ärztinnen und Ärzten und Patientinnen und Patienten evaluiert. Eine Erhebung, inwieweit die Interaktion der Nutzenden den Versorgungsprozess effektiv unterstützt, konnte hingegen nicht erfolgen, da im Projektkontext kein vollständig einsatzfähiger Demonstrator unter Realbedingungen implementiert werden konnte. Eine wirtschaftliche Bewertung konnte

aufgrund noch nicht abschließend bestimmbarer Prozess- und Hardwarekosten ebenfalls nicht durchgeführt werden. Insgesamt wird das System jedoch als perspektivisch wertvolle Entwicklung zur Unterstützung medizinischer Versorgungsprozesse eingeschätzt, vorbehaltlich einer klinischen Validierung und Integration in bestehende Strukturen.

Um relevante ELSI-Fragestellungen zu adressieren, wurde im Rahmen der Onboarding-Materialien ein besonderer Fokus auf Datensicherheit, Datenschutz sowie die Rollen- und Verantwortungsabgrenzung zwischen dem KI-System und dem medizinischem Fachpersonal gelegt. Hierzu wurden drei komplementäre Formate entwickelt und mit Ärztinnen und Ärzten evaluiert: ein interaktiver Rundgang zur funktionalen Einführung (Link zum Figma Prototyp), ein analoger Beipackzettel zur Vermittlung von Hintergrundwissen, sowie ein Self-Assessment zur Reflexion und Verständnissicherung. Es wird darüber aufgeklärt, welche Daten zu welchem Zweck verarbeitet werden und dass diese ausschließlich zur medizinischen Betreuung genutzt, verschlüsselt gespeichert und nur autorisiertem Fachpersonal zugänglich gemacht werden. Zudem wird explizit klargestellt, dass das KI-Modell keine automatischen Entscheidungen trifft und die ärztliche Verantwortung für Interpretation und Therapieentscheidungen vollständig bestehen bleibt. Darüber hinaus werden der bewusste Umgang mit möglichen Fehlbewertungen des Systems, die aktive Rückmeldung von Diskrepanzen an die Entwicklerinnen und Entwickler sowie organisatorische Schutzmaßnahmen (z. B. kontrollierte Zugriffe, Abmeldung, Bildschirmsperre) thematisiert. Auf diese Weise werden Datenschutzanforderungen, Verantwortungsbewusstsein und ein realistisches Verständnis der Systemgrenzen integraler Bestandteil der Anwendungseinführung. Die Evaluation mit 24 Ärztinnen und Ärzten ergab, dass die drei Onboarding-Formate hinsichtlich Darstellungsart, Verständlichkeit und Praxisrelevanz überwiegend positiv bewertet wurde. In dem Ranking der Formate nach Nützlichkeit wurde die Kombination der Formate, ihre logische Reihenfolge und ihre gegenseitige Ergänzung als besonders sinnvoll hervorgehoben, da sie unterschiedliche Informationsbedarfe bedienen. Gleichzeitig wurde insbesondere der technische Hintergrundteil im Beipackzettel, insbesondere die Beschreibung des KI-Modells und seiner Funktionsweise, als nicht durchgängig leicht verständlich bewertet, was weiteren Schulungs- und Vereinfachungsbedarf nahelegt.

AP5: Transfer und Verwertung durch Projektpartner

In AP5 wurden die entwickelten KI-basierten Analyseansätze sowie die zugrunde liegende Datenverarbeitungskonzeption hinsichtlich ihrer Übertragbarkeit und Skalierbarkeit bewertet. Ziel war es, die konzeptionelle Anschlussfähigkeit der entwickelten Lösung über das konkrete Projektsetting hinaus sicherzustellen.

Die entwickelte KI-Pipeline wurde bewusst modular und reproduzierbar aufgebaut, sodass Vorverarbeitung, Modelltraining und Evaluation strukturiert dokumentiert und grundsätzlich auf andere kontinuierliche Gesundheitsdatensätze übertragbar sind. Aufgrund der fehlenden realitätsgetreuen Telemonitoring-Datengrundlage konnte eine produktionsreife Modellbereitstellung im Projektkontext nicht erfolgen. Gleichwohl wurden die technischen und architektonischen Grundlagen so gestaltet, dass eine spätere Integration real erhobener Daten ohne grundlegende Anpassungen möglich ist. Ein weiterer zentraler Aspekt des Transferkonzepts ist die Übertragbarkeit in andere Regionen und Versorgungskontexte. Hierbei wurden insbesondere Unterschiede in Datenverfügbarkeit, Versorgungsstrukturen und infrastrukturellen Rahmenbedingungen berücksichtigt. Die Lösung ist prinzipiell skalierbar, sowohl im Hinblick auf verschiedene Krankheitsbilder (z. B. Diabetes oder Leberinsuffizienz) und eine steigende Anzahl von Patientinnen und Patienten als auch auf die Einbindung zusätzlicher Datenquellen und Versorgungseinrichtungen. Mit zunehmender regionaler Verteilung der Datenbestände könnte perspektivisch die Implementierung eines Federated Learning Ansatzes

sinnvoll werden, um institutionsübergreifende Modellverbesserungen datenschutzkonform zu ermöglichen. Darüber hinaus sind die entwickelten Onboarding-Formate und das methodische Vorgehen zur Evaluation erklärbarer KI-Methoden ebenfalls übertragbar und können als strukturierte Grundlage für vergleichbare KI-basierte Anwendungen im Gesundheitswesen dienen. Die weitergehende wissenschaftliche und wirtschaftliche Verwertung wird in Kapitel 2.2 detailliert dargestellt.

AP6: Ethische, soziale und rechtliche Begleitung

Im Rahmen von AP6 wurde der Einsatz von KI über die gesamte Projektlaufzeit hinweg systematisch und kritisch unter ethischen, sozialen und rechtlichen Gesichtspunkten begleitet. Ziel war es, regulatorische Anforderungen frühzeitig zu berücksichtigen, Verantwortlichkeiten klar zu definieren und potenzielle Auswirkungen des KI-Systems auf alle beteiligten Akteure transparent zu reflektieren.

Zu Beginn erfolgte eine strukturierte Analyse der für das Vorhaben relevanten Regulatorik. Hierzu zählten insbesondere Datenschutzanforderungen im Kontext Digitaler Gesundheitsanwendungen (DiGA) und Digitaler Pflegeanwendungen (DiPA) sowie die Vereinbarung von Qualitätssicherungsmaßnahmen gemäß § 135 Abs. 2 SGB V zum Telemonitoring bei Herzinsuffizienz. Ergänzend erfolgte eine vertiefte Auseinandersetzung mit der europäischen KI-Verordnung (AI-Act) und dem Medizinproduktegesetz. Dabei wurde insbesondere untersucht, ob und unter welchen Voraussetzungen das KI-System als Hochrisiko-KI-System im Sinne der europäischen KI-Verordnung einzuordnen ist und welche daraus resultierenden Anforderungen zu berücksichtigen sind. Die entsprechenden Anforderungen (z. B. in Bezug auf Transparenz, Dokumentation, Risikomanagement und menschliche Aufsicht) wurden systematisch dokumentiert. Darüber hinaus wurden relevante Fragestellungen zusammengetragen, die eine strukturierte Abgrenzung zwischen Software und potenziell als Medizinprodukt einzustufenden Systemen ermöglichen.

Neben der in AP4 beschriebenen Aufarbeitung der ELSI-Fragestellungen, lag ein weiterer Schwerpunkt lag auf der Analyse der Auswirkungen des KI-Einsatzes auf die verschiedenen Akteure. Dabei wurde insbesondere die Frage der Deutungshoheit adressiert. Die KI-basierte Risikoeinschätzung wurde konzeptionell klar als unterstützendes Instrument definiert, während die fachliche Bewertung, Entscheidungsfindung und Verantwortung uneingeschränkt bei der behandelnden Ärztin oder dem behandelnden Arzt verbleiben. Gleichzeitig wurden datenschutzrechtliche Aspekte wie verschlüsselte Datenübertragung, Zugriffsbeschränkungen und Zweckbindung der Datennutzung konzeptionell berücksichtigt. Darüber hinaus wurde darauf geachtet, den Eingriff in den Alltag der Patientinnen und Patienten gering zu halten. Die Datenerhebung erfolgt überwiegend automatisiert über tragbare Sensoren, ergänzende Angaben sind freiwillig. Um Verunsicherung zu vermeiden, erhalten Patientinnen und Patienten keine direkte Einsicht in die vom KI-Modell vorhergesagte Dekompensationswahrscheinlichkeit. Um Benachteiligungen zu vermeiden, soll allen geeigneten Patientinnen und Patienten die Nutzung des Systems aktiv angeboten und transparent über alternative Behandlungsoptionen aufgeklärt werden.

AP7: Projektmanagement und Öffentlichkeitsarbeit

Für die Kommunikation der Projektergebnisse ist das DFKI an drei Publikationen beteiligt. Im Folgenden werden die Vorhaben skizziert. Eine Übersicht über die vollständigen Titel, Autorenschaft, Zielkonferenz bzw. das Zieljournal sowie den Einreichungsstatus ist in Kapitel 2.4 zu finden. Nach Veröffentlichung werden die Publikationen auf der Projektwebseite ergänzt: https://kardiointerakt.de.

„Prescribing a Dose of Transparency“: Existierende Literatur verdeutlicht, dass die intransparente Black-Box-Natur von KI das Vertrauen von Ärztinnen und Ärzten untergraben kann, wodurch die Bedeutung von erklärbarer KI (XAI) unterstrichen wird. Diese Studie evaluiert XAI innerhalb eines telemedizinischen Frühwarnsystems am Beispiel einer drohenden Herzinsuffizienz-Dekompensation. Drei XAI-Methoden werden anhand der Explanation Satisfaction Scale (Hoffman et al., 2018) verglichen: SHAP, CFEs und Anchors (siehe AP4). In Interviews mit fünfzehn deutschen medizinischen Fachkräften wurde SHAP aufgrund seiner intuitiven grafischen Darstellung weitgehend bevorzugt. CFEs wurden als handlungsorientierte Ergänzung zur Ableitung therapeutischer Schritte geschätzt, während Anchors als zu oberflächlich angesehen wurden. Unsere Ergebnisse legen nahe, SHAP für transparente Einblicke und CFEs für umsetzbare Leitlinien zu kombinieren, um das Vertrauen zu stärken und die Akzeptanz von KI zu fördern.

„Symptoms, Causes, and Mitigation of Obstacles in AI-Based Telemonitoring”: Trotz vielversprechender technischer Ergebnisse scheitern viele KI-Systeme an der Integration in den klinischen Alltag. Zu den größten Herausforderungen zählen Einschränkungen hinsichtlich der Datenqualität, begrenzte Transparenz, komplexe Vorschriften und Schwierigkeiten bei der Integration von KI-Ergebnissen in etablierte Arbeitsabläufe. Darüber hinaus sind die zugrunde liegenden End-to-End-Entwicklungsprozesse oft unzureichend dokumentiert und weisen einen positiven Publikationsbias auf, wodurch die Möglichkeit, aus früheren Bemühungen zu lernen, eingeschränkt wird. Diese Studie untersucht die Hindernisse, die während des gesamten Entwicklungsprozesses von KI-basierten Telemonitoring-Systemen unter Verwendung des Cross Industry Standard Process for Data Mining (CRISP-DM) auftreten. Aufbauend auf verschiedenen CRISP-DM-Varianten analysieren wir deren konzeptionelle Eignung für regulierte Gesundheitskontexte. Empirische Erkenntnisse werden aus einer qualitativen, prozessorientierten Fallstudie des KardioInterakt Projekts gewonnen. Durch die systematische Rekonstruktion von Entscheidungen, Hindernissen und Bewältigungsstrategien, synthetisiert mit früheren Studien, leiten wir übertragbare, umsetzbare Handlungsempfehlungen für die Entwicklung von KI-Systemen im Gesundheitswesen ab und ordnen diese den kanonischen CRISP-DM-Phasen zu. Dazu gehören Umfangs- und Machbarkkeitsbewertung, Einbeziehung von Interessengruppen, Datenstrategie, Modellierung, Erklärbarkeit, Evaluation, Schulung der Nutzenden, Integration in Arbeitsabläufe und regulatorische Verpflichtungen. Die Ergebnisse bieten somit einen praxisnahen Leitfaden für die Entwicklung robuster, tragfähiger, vertrauenswürdiger und regulierungskonformer KI-basierter Telemonitoring-Lösungen.

„Richtig dosierte Information zu KI-Systemen im Gesundheitswesen“: Die Integration von KI-Systemen in die klinische Praxis scheitert oft an mangelnder Akzeptanz und Unsicherheiten des medizinischen Personals hinsichtlich der Funktionsweise und Entscheidungshoheit. Vor dem Hintergrund untersucht diese Studie, welche Informationsbedarfe Ärztinnen und Ärzten in der Einarbeitungsphase mit einem KI-basierten Entscheidungsunterstützungssystem haben und welche Onboarding-Formate ein fundiertes Systemverständnis fördern. Am Beispiel der KardioInterakt-Anwendung wurden drei komplementäre Formate entwickelt: ein interaktiver Rundgang zur funktionalen Einführung, ein analoger Beipackzettel zur Vermittlung von Hintergrundwissen (z. B. technische, ethische, legale und soziale Implikationen) sowie ein Self-Assessment zur Reflexion und Verständnissicherung (siehe AP4). Die Wirksamkeit hinsichtlich der Darstellungsart, Verständlichkeit und Relevanz wurde in einer Umfrage mit 24 Ärztinnen und Ärzten evaluiert. Erste Ergebnisse zeigen, dass die strukturierte Kombination der Formate als besonders wirkungsvoll eingeschätzt wurde, da sie unterschiedliche Informationsbedarfe adressiert. Zugleich wurde hinsichtlich der Verständlichkeit technischer Detailerläu-

terungen weiterer Schulungs- und Vereinfachungsbedarf festgestellt. Die Arbeit leistet damit einen praxisnahen Beitrag zur Gestaltung von Mensch-KI-Schnittstellen im klinischen Alltag. Die abgeleiteten Handlungsempfehlungen bieten wertvolle Implikationen für Entwicklerinnen und Entwickler, sowie Entscheidungstragende im Gesundheitswesen, um KI-Systeme durch gezielte Informationsvermittlung erfolgreich und sicher in die Versorgungspraxis zu integrieren.

2.2 Projekterfahrungen und Verwertbarkeit

Die wirtschaftliche Verwertung im Verbundprojekt wird insbesondere durch die assoziierten und antragstellenden Unternehmen geleistet. Von Seiten der Unternehmen wird bspw. das Feld der Produktion und Bereitstellung der geeigneten Hard- und Software (Sensorik, KI-Algorithmen, Cloud-basierte Plattform, multimodale User Interfaces) aufgespannt. Somit hat das dargelegte Projekt eine hohe wirtschaftliche Bedeutung in mehrfacher Hinsicht. Das DFKI setzte an den Schnittstellen der beteiligten Unternehmen an. So wurden u.a. vertrauenswürdige KI-Algorithmen und Onboarding-Formate entwickelt. Der entwickelte KI-Service bietet vielversprechende wirtschaftliche Erfolgsaussichten, insbesondere durch seine Fähigkeit, personalisierte Risikoeinschätzungen auf Basis von Zeitreihendaten effizient und skalierbar zu liefern. Er übertrifft bestehende Lösungen durch seine modular aufgebaute Pipeline, die leicht für unterschiedliche chronische Erkrankungen (z. B. Diabetes, Leberinsuffizienz) oder für Zeitreihenklassifikation in anderen Industrien (z. B. zur Verkaufsprognose im Einzelhandel oder zur Maschinenüberwachung in der Produktion) adaptiert werden kann. Diese breite Einsetzbarkeit bietet ein hohes Marktpotenzial für Anbieter im Gesundheitswesen, insbesondere in Deutschland mit seiner stark wachsenden eHealth-Infrastruktur. Die Algorithmen können bei zukünftigen Aufträgen als Blaupause dienen und damit die (Weiter-)Entwicklung neuer KI-Algorithmen signifikant beschleunigen. Zudem baute das DFKI weitere Kompetenzen und Expertise bei der Gestaltung multimodaler Serviceplattformen und Interaktionssystemen auf, die in der Wirtschaft immer stärker nachgefragt werden. Neben der Erweiterung des Produktportfolios konnten somit auch relevante sozio-technische Erkenntnisse gewonnen und praxisnahe Ansätze für eine Verbesserung der medizinischen Versorgung der Zielgruppe in der langfristigen Perspektive entwickelt werden.

Das wissenschaftliche Verwertungspotenzial ist als sehr hoch einzustufen, da erste Erfahrungen mit der telemedizinischen Überwachung chronischer Erkrankungen am Beispiel eines weit verbreiteten Krankheitsbilds (chronische Herzinsuffizienz) eine wertvolle Grundlage für die positive Beeinflussung des Krankheitsverlaufs darstellen. Nach jahrelangen Diskussionen hat der Gemeinsame Bundesausschuss im Dezember 2020 die telemedizinische Versorgung bei schwerer Herzinsuffizienz zum Teil der ambulanten ärztlichen Versorgung gemacht (Gemeinsamer Bundesausschuss, 2020). Dabei war noch völlig offen, mit welcher Struktur und welchen sozio-technischen Informationssystemen diese telemedizinische Betreuung konkret ausgestaltet werden soll. Somit ist das hier dargelegte Projekt von unmittelbarer Relevanz für die aktuelle medizinische Versorgungsforschung im Bereich chronischer (Herz-)Erkrankungen. In der Etablierung der geeigneten Form der telemedizinischen Versorgung liegt aus gesundheitspolitischer und wirtschaftlicher Sicht das Potential die (Re-)Hospitalisierungsrate bei Herzinsuffizienz und damit die teuren stationären Leistungen sowie physische Kontakte zu reduzieren. In Deutschland werden die durch Herzinsuffizienz bedingten Kosten für das Gesundheitssystem jährlich auf ca. 2,7 Milliarden Euro geschätzt, von denen bis zu 50% durch frühzeitige Behandlung und Aufklärung eingespart werden könnten (Zugck et al., 2005). Die in diesem Verbundprojekt etablierten Grundlagen für die telemedizinische Versorgung kann zu diesen Kosteneinsparungen beitragen. Darüber hinaus sind die Projektergebnisse für weitere wissenschaftliche Forschung von Nutzen, z. B.

zur Verbesserung von Vorhersagegüte, Interpretierbarkeit oder Fairness von KI-Modellen im Gesundheitswesen. Zudem eröffnen sich Möglichkeiten zur Integration in größere, datengetriebene Versorgungssysteme. Die enge Zusammenarbeit mit medizinischen Einrichtungen und Technologiepartnern im Projektverlauf hat die Praxisnähe der Lösung erhöht und bildet die Grundlage für eine direkte Verwertung durch diese Akteure. Eine solche telemedizinische Versorgung lässt sich auch auf andere chronische, betreuungsintensive Krankheiten, die Nachsorgemanagement benötigen, sowie strukturschwache Regionen übertragen, in denen die adäquate hausärztliche Versorgung zunehmend knapper wird.

2.3 Technologieentwicklung

Im Verlauf der Projektlaufzeit wurden in der Forschung und Anwendung einige technologische Ansätze zur Unterstützung der Diagnostik, Überwachung und Versorgung von Patientinnen und Patienten mit Herzinsuffizienz entwickelt.

Ein wesentlicher Schwerpunkt lag auf der Weiterentwicklung maschineller Lernverfahren zur Vorhersage von Erkrankungen und klinischen Ereignissen. Doki et al. (2022) adressieren die Problematik begrenzter Datenmengen und stellen ein simples Diagnosesystem zur Vorhersage von Herzerkrankungen vor. Durch den Einsatz von Extreme Gradient Boosting (XGBoost), Feature-Selektion und Hyperparameter-Optimierung konnten auch bei einem vergleichbar kleinen Datensatz von 303 Patientinnen und Patienten mit jeweils einer einzelnen Messung eine hohe Vorhersagegüte erzielt werden.

Mehrere Arbeiten fokussieren sich auf die Vorhersage von Rehospitalisierung bei Herzinsuffizienz. Zhang et al. (2023) entwickelten ein ML-Modell zur Prognose ungeplanter Wiederaufnahme innerhalb von 30 Tagen bei 14843 älteren Patientinnen und Patienten. Nach dem Vergleich mehrerer Algorithmen erwies sich ein CatBoost-Modell als leistungsfähigstes Verfahren. Neben der reinen Prädiktion wurde auch die XAI-Methode SHAP eingesetzt, um einflussreiche klinische und labordiagnostische Merkmale zu identifizieren. Diesen Ansatz greifen auch Luo et al., (2024) auf und erweitern ihn um eine explizite Mensch-Maschine-Kollaboration. In der Studie wurde ein Datensatz von 8396 hospitalisierten älteren Patientinnen und Patienten mit Herzinsuffizienz analysiert. Dabei wurden automatisch durch ML ausgewählte Merkmale mit von Kardiologinnen und Kardiologen festgelegten Variablen kombiniert. Das resultierende hybride Modell, ebenfalls auf Basis von CatBoost, zeigte eine bessere Vorhersageleistung als rein datengetriebene oder ausschließlich Experten-basierte Modelle. Mittels SHAP-Analyse wurden die fünf wichtigsten Prädiktoren identifiziert.

Neben klassischen klinischen und labordiagnostischen Daten werden zunehmend neuartige Datenquellen erschlossen. Murton et al. (2023) untersuchten Sprach- und Stimmmerkmale von 52 Patientinnen und Patienten als potenzielle Biomarker für akut dekompensierte Herzinsuffizienz. Die Analyse von Sprachaufnahmen zeigte Veränderungen der Phonation, Sprechgeschwindigkeit und Stimmqualität im Verlauf der Behandlung. Ein weiterer Entwicklungspfad betrifft die kontinuierliche Überwachung von Patientinnen und Patienten im häuslichen Umfeld. Chausiaux et al. (2022) berichten über den Einsatz eines automatisierten Point-of-Care-Geräts zur Erkennung peripherer Ödeme bei einer Patientin mit Herzinsuffizienz. Das System nutzt 3D-Aufnahmen und berechnet die Volumenzunahme in der Fuß- und Unterbeinregion und generiert frühzeitige Warnungen, um klinische Entscheidungen zu unterstützen.

Auch KI-gestützte Vorhersagemodelle auf Basis von Telemonitoring-Daten wurden weiterentwickelt. Kerexeta et al. (2023) nutzten Vitalparameter und Fragebogendaten von 488 Patientinnen und Patienten, um kardiale Dekompensationsereignisse bei chronischer Herzinsuffizienz vorherzusagen. Das verwendete

XGBoost-Modell identifizierte sowohl physiologische Veränderungen (Gewichtszunahme, Sauerstoffsättigung in den letzten Tagen, Herzrate) als auch subjektive Symptome (allgemeines Wohlbefinden, Orthopnoe und Knöchelbeschwerden) als relevante Prädiktoren und verdeutlicht den Mehrwert multimodaler Datensätze. Ciccarelli et al. (2025) untersuchten in der SMART-CARE Studie den prospektiven Einsatz drei verschiedener CE-zertifizierter, multiparametrischer Wearables in Kombination mit KI-Algorithmen zur Früherkennung klinischer Verschlechterungen anhand eines Datensatzes von 300 Patientinnen und Patienten mit chronischer Herzinsuffizienz. Über einen Zeitraum von sechs Monaten wurden kontinuierlich physiologische Parameter wie Sauerstoffsättigung, Herzratenvariabilität, Atemfrequenz, Hauttemperatur und Schlafmetriken erfasst und in Echtzeit analysiert. Ziel war es, Hospitalisierung zu reduzieren und Lebensqualität zu verbessern. Zudem entwickelte die Charité Berlin ein Telemonitoring System, bestehend aus einer Waage, einem Blutdruckmessgerät, einem Pulsoximeter, einem mobilen EKG-Gerät und einem Tablet (Haserück & Kurz, 2023). Die erfassten Daten werden täglich von Ärztinnen und Ärzten und Pflegepersonal überprüft und im Interventionsfall wird die Patientin oder der Patient kontaktiert. Durch die frühzeitige Identifikation kritischer Veränderungen ermöglicht das System frühzeitige therapeutische Maßnahmen und markiert einen wichtigen Schritt von projektbasierter Telemedizin hin zu einer etablierten, digital unterstützten Versorgungsform.

2.4 Erfolgte und geplante Veröffentlichungen

Die aus dem Projekt hervorgegangenen und geplanten Publikationen, an denen das DFKI beteiligt ist, sind Tabelle 4 zu entnehmen. Tiefergehende Informationen zu den jeweiligen Veröffentlichungen finden sich in Kapitel 2.1 unter AP7.

Publikationstitel	Autorenschaft	Journal / Konferenz	Status
Prescribing a Dose of Transparency: A Qualitative Evaluation of Explainable AI Methods with Healthcare Professionals	Sophie Haas[1], Malte Högemann[1,2], Oliver Thomas[1,2]	Frontiers in Artificial Intelligence	Einreichung voraussichtlich März 2026
Symptoms, Causes, and Mitigation of Obstacles in AI-Based Telemonitoring: Strategies for System Development Through a CRISP-DM Lens	Sophie Haas[1,2], Jonas Hammer[3], Fabia Marie Hettler[3], Patricia Kajüter Rodrigues[3], Frank Teuteberg[3], Oliver Thomas[1,2]	Frontiers in Artificial Intelligence	Einreichung voraussichtlich März 2026
Richtig dosierte Information zu KI-Systemen im Gesundheitswesen: Gestaltung und Evaluation von Onboarding-Formaten für medizinisches Personal	Sophie Haas[1,2], Malte Högemann[1,2], Mara Rehmer[1], Jonas Hammer[3], Frank Teuteberg[3], Oliver Thomas[1,2]	HMD – Praxis in der Wirtschaftsinformatik	Einreichung voraussichtlich März 2026

Tabelle 4: Aus dem Projekt hervorgegangene und geplante Publikationen mit Beteiligung des DFKI.

[1] Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), Smart Enterprise Engineering

[2] Universität Osnabrück, Informationsmanagement und Wirtschaftsinformatik (IMWI)

[3] Universität Osnabrück, Unternehmensrechnung und Wirtschaftsinformatik (UWI)

Literaturverzeichnis

Bachtiger, P., Plymen, C. M., Pabari, P. A., Howard, J. P., Whinnett, Z. I., Opoku, F., Janering, S., Faisal, A. A., Francis, D. P., & Peters, N. S. (2020). Artificial intelligence, data sensors and interconnectivity: Future opportunities for heart failure. *Cardiac Failure Review*, *6*, e11.

Benaim, A. R., Almog, R., Gorelik, Y., Hochberg, I., Nassar, L., Mashiach, T., Khamaisi, M., Lurie, Y., Azzam, Z. S., Khoury, J., & others. (2020). Analyzing medical research results based on synthetic data and their relation to real data results: Systematic comparison from five observational studies. *JMIR medical informatics*, *8*(2), e16492.

Benjamins, J. W., Hendriks, T., Knuuti, J., Juarez-Orozco, L. E., & van der Harst, P. (2019). A primer in artificial intelligence in cardiovascular medicine. *Netherlands Heart Journal*, *27*(9), 392–402.

Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), & Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). (2023). *Nationale VersorgungsLeitlinie Chronische Herzinsuffizienz – Langfassung* (Version 4.0) [Application/pdf]. https://doi.org/10.6101/AZQ/000510

Chausiaux, O. E., Keyser, M., Williams, G. P., Nieznański, M., Downer, P. J., Garnett, R. E., Berry, R., & Husheer, S. L. G. (2022). Heart failure decompensation alerts in a patient's home using an automated, AI-driven, point-of-care device. *BMJ Case Reports CP*, *15*(4), e248682.

Ciccarelli, M., Bramanti, A., Carrizzo, A., Garofano, M., Visco, V., Izzo, C., Rusciano, M. R., Galasso, G., Loria, F., Bruno, G., & others. (2025). Artificial intelligence-based remote monitoring for chronic heart failure: Design and rationale of the SMART-CARE study. *Frontiers in Digital Health*, *7*, 1719562.

Deutsches Ärzteblatt. (2019, Oktober 29). *Telemonitoring kann Mortalität bei Herzinsuffizienz senken*. https://www.aerzteblatt.de/news/telemonitoring-kann-mortalitaet-bei-herzinsuffizienz-senken-ea56ef80-7037-4ce7-a123-2ff7e3e0d812

Doki, S., Devella, S., Tallam, S., Gangannagari, S. S. R., Reddy, P. S., & Reddy, G. P. (2022). Heart disease prediction using xgboost. *2022 third international conference on intelligent computing instrumentation and control technologies (ICICICT)*, 1317–1320.

Gemeinsamer Bundesausschuss. (2020, Dezember 17). *Richtlinie Methoden vertragsärztliche Versorgung: Telemonitoring bei Herzinsuffizienz*. https://www.g-ba.de/beschluesse/4648/

Haserück, A., & Kurz, C. (2023). Telemonitoring: Digitaler Lebensretter. *Deutsches Ärzteblatt*, *120*(49), 7927-
.

Hoffman, R. R., Mueller, S. T., Klein, G., & Litman, J. (2018). Metrics for explainable AI: Challenges and prospects. *arXiv preprint arXiv:1812.04608*.

Johnson, K. W., Torres Soto, J., Glicksberg, B. S., Shameer, K., Miotto, R., Ali, M., Ashley, E., & Dudley, J. T. (2018). Artificial intelligence in cardiology. *Journal of the American College of Cardiology*, *71*(23), 2668–2679.

Kerexeta, J., Larburu, N., Escolar, V., Lozano-Bahamonde, A., Macia, I., Beristain Iraola, A., & Grana, M. (2023). Prediction and analysis of heart failure decompensation events based on telemonitored data and artificial intelligence methods. *Journal of Cardiovascular Development and Disease*, *10*(2), 48.

Koehler, F., Koehler, K., Deckwart, O., Prescher, S., Wegscheider, K., Kirwan, B.-A., Winkler, S., Vettorazzi, E., Bruch, L., Oeff, M., & others. (2018). Efficacy of telemedical interventional management in patients with heart failure (TIM-HF2): A randomised, controlled, parallel-group, unmasked trial. *The Lancet*, *392*(10152), 1047–1057.

Krittanawong, C., Zhang, H., Wang, Z., Aydar, M., & Kitai, T. (2017). Artificial intelligence in precision cardiovascular medicine. *Journal of the American College of Cardiology*, *69*(21), 2657–2664.

Luo, H., Xiang, C., Zeng, L., Li, S., Mei, X., Xiong, L., Liu, Y., Wen, C., Cui, Y., Du, L., & others. (2024). SHAP based predictive modeling for 1 year all-cause readmission risk in elderly heart failure patients: Feature selection and model interpretation. *Scientific reports*, *14*(1), 17728.

Murton, O. M., Dec, G. W., Hillman, R. E., Majmudar, M. D., Steiner, J., Guttag, J. V., & Mehta, D. D. (2023). Acoustic voice and speech biomarkers of treatment status during hospitalization for acute decompensated heart failure. *Applied sciences*, *13*(3), 1827.

Savarese, G., Becher, P. M., Lund, L. H., Seferovic, P., Rosano, G. M. C., & Coats, A. J. S. (2023). Global burden of heart failure: A comprehensive and updated review of epidemiology. *Cardiovascular Research*, *118*(17), 3272–3287. https://doi.org/10.1093/cvr/cvac013

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature medicine*, *25*(1), 44–56.

World Health Organization. (2025, Juli 31). *Cardiovascular diseases (CVDs)*. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Zhang, Y., Wang, H., Yin, C., Shu, T., Yu, J., Jian, J., Jian, C., Duan, M., Kadier, K., Xu, Q., & others. (2023). Development of a prediction model for the risk of 30-day unplanned readmission in older patients

with heart failure: A multicenter retrospective study. *Nutrition, Metabolism and Cardiovascular Diseases*, *33*(10), 1878–1887.

Zugck, C., Nelles, M., Frankenstein, L., Schultz, C., Helms, T., Korb, H., Katus, H., & Remppis, A. (2005). Telemedizinisches Monitoring bei herzinsuffizienten Patienten. *Herzschrittmachertherapie & Elektrophysiologie*, *16*(3), 176–182.