

Vorhaben	<b>KI-AIM</b> KI-basierte Anonymisierung in der Medizin
Titel	<b>Abschlussbericht Teil II: Eingehende Darstellung</b>
Förderkennzeichen	16KISA115K
Zuwendungsempfänger	Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI)
Bewilligungszeitraum	15.12.2022 - 15.12.2025
Autoren	Martin Kuhn, Joscha Grüger und Ralph Bergmann
Erstellungsdatum:	15.01.2026

Gefördert durch:



Bundesministerium  
für Forschung, Technologie  
und Raumfahrt



Finanziert von der  
Europäischen Union  
NextGenerationEU

Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Forschung, Technologie und Raumfahrt sowie der Europäischen Union (EU) unter dem Förderkennzeichen 16KISA115K gefördert. Dem Muster zu Nr. 4.1 NKBF 2017 (BMFTR-Vordr. 0701/07.25) folgend, handelt es sich bei diesem Dokument um Teil II, eine eingehende Darstellung des Vorhabens. Die Teile I und III liegen als gesonderte Dokumente vor. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Verwendung der Zuwendung</b>	<b>3</b>
2.1	Projektmanagement (AP 1) . . . . .	3
2.2	Datenschutzrechtliche Aspekte (AP 2) . . . . .	3
2.3	Spezifizierung von Anwendungsfällen und des Datensatzes (AP 3) . . . . .	4
2.4	Entwicklung der Anonymisierungsplattform (AP 5) . . . . .	5
2.5	Evaluation (AP 6) . . . . .	7
2.6	Vorbereitung des wissenschaftlichen und technischen Diskurses (AP 7) . . . . .	9
<b>3</b>	<b>Wichtigste Positionen des zahlenmäßigen Verwendungsnachweises</b>	<b>11</b>
<b>4</b>	<b>Notwendigkeit und Angemessenheit der geleisteten Arbeit</b>	<b>12</b>
<b>5</b>	<b>Verwertbarkeit</b>	<b>13</b>
5.1	Wirtschaftliche Erfolgsaussichten . . . . .	13
5.2	Wissenschaftliche Erfolgsaussichten . . . . .	14
5.3	Wissenschaftliche und wirtschaftliche Anschlussfähigkeit . . . . .	14
<b>6</b>	<b>Bekannt gewordener Fortschritt</b>	<b>16</b>
<b>7</b>	<b>Erfolgte und geplante Veröffentlichungen der Ergebnisse</b>	<b>17</b>

# 1 Einleitung

Die Forschung und Entwicklung von KI-Methoden im Gesundheitswesen ist in besonderem Maße auf große Mengen qualitativ hochwertiger, realistischer Daten angewiesen, die in der klinischen Versorgung über längere Zeiträume erhoben und dokumentiert werden. Gleichzeitig ist der Zugang zu diesen Daten häufig eingeschränkt, da es sich um sensible, personenbezogene und oft stark heterogene Daten handelt, deren Nutzung strengen datenschutzrechtlichen Anforderungen unterliegt. Dies erschwert insbesondere die Sekundärnutzung für Forschung und Entwicklung von datengetriebenen Anwendungen und führt in der Praxis zu hohem organisatorischen Aufwand, langen Abstimmungsprozessen und begrenzter Verfügbarkeit geeigneter Daten.

Vor diesem Hintergrund ist das Gesamtziel des Verbundprojekts KI-basierte Anonymisierung in der Medizin (KI-AIM) die Entwicklung und Etablierung einer Anonymisierungsplattform, die die Bereitstellung großer Mengen realitätsnaher Daten ermöglicht, die keinen Personenbezug aufweisen und dadurch datenschutzrechtlich einfacher genutzt werden können. Dadurch soll der Zugang zu Daten für die Entwicklung von KI-Methoden deutlich vereinfacht und damit sowohl die Forschung als auch die Entwicklung entsprechender Anwendungen unterstützt werden. Kern des Vorhabens ist die Entwicklung der KI-AIM-Plattform Cinnamon. Die Plattform kombiniert Anonymisierungs- und Synthetisierungsverfahren über standardisierte Schnittstellen, unterstützt die Ausführung in unterschiedlichen Konfigurationen und zugleich die Bewertung der Ergebnisse anhand von Qualitäts- und Risikoindikatoren. Des Weiteren ist die KI-AIM-Plattform Cinnamon generisch ausgelegt und erweiterbar, sodass zusätzliche Verfahren, entlang einer definierten Schnittstellenspezifikation integriert werden können.

Das Teilprojekt des DFKI verfolgt innerhalb dieses Rahmens vier zentrale Ziele. Die Zugänglichkeit ausgewählter Anonymisierungs- und Synthetisierungsansätze innerhalb der gemeinsamen Plattform, die Erforschung ihrer Kombinierbarkeit unter Berücksichtigung semantischer Abhängigkeiten, die Abbildung verschiedener Kombinationsmöglichkeiten in einer für Domänenexpertinnen und Domänenexperten nutzbaren Konfiguration sowie die Entwicklung und Integration von Verfahren zur Messung der Nützlichkeit anonymisierter und synthetischer Daten. Ein starker Fokus liegt dabei auf der Datensynthese, insbesondere auf Deep-Learning-basierten Ansätzen wie Generative Adversarial

Networks (GANs), und autoregressiven Modellen (AR) sowie auf wissensbasierten Verfahren der Prozessmodellierung und Simulation.

Als konkreter Anwendungsfall wird die KI-Forschung zur Entwicklung klinischer Entscheidungsunterstützungssysteme in der dermatologischen Onkologie adressiert. Gerade onkologische Behandlungsdaten sind aufgrund ihrer Komplexität, Heterogenität und Diversität schwer zu anonymisieren und zu synthetisieren, zugleich besteht ein hoher Bedarf an datenschutzkonform nutzbaren Daten für Forschung und Entwicklung. Im Rahmen der Evaluation wird daher überprüft, inwieweit anonymisierte und synthetisierte Daten im Vergleich zu Originaldaten für ausgewählte medizinische Use Cases nutzbar sind und welche Risiken dabei in Kauf genommen werden müssen. Ergänzend wird die Gebrauchstauglichkeit der Plattform in einer Usability-Evaluation untersucht. Die Plattform wurde unter der Apache-2.0 (Open-Source-Lizenz) veröffentlicht, um Nachnutzung, Weiterentwicklung und Transfer insbesondere auch in Richtung Industrie zu ermöglichen.

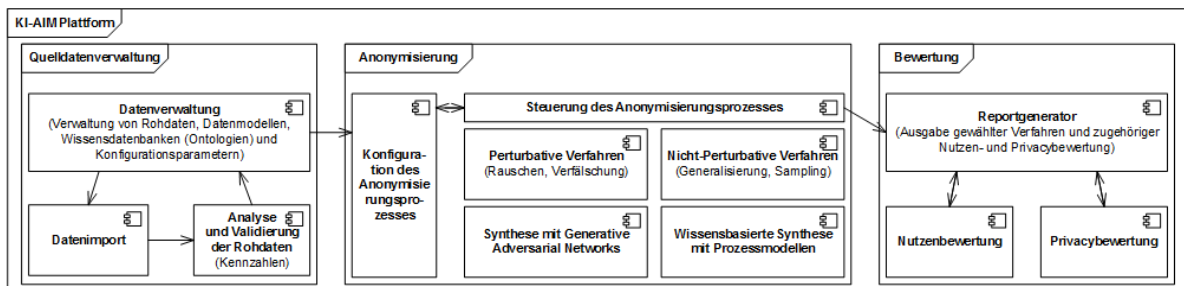


Abbildung 1.1 Plattformarchitektur der KI-AIM Plattform Cinnamon.

## 2 Verwendung der Zuwendung

Im Folgenden werden die Ziele des Vorhabens und die erzielten Ergebnisse im Einzelnen anhand der Arbeitspakete mit Beteiligung des DFKI dargestellt.

### 2.1 Projektmanagement (AP 1)

In diesem Arbeitspaket lag der Schwerpunkt auf der Koordination des KI-AIM-Projektes. Die Zusammenarbeit der Partner (UniMS/HTZ, UniMS/IMI, BIH, DATATREE und DFKI) wurde organisatorisch gebündelt, der Informationsfluss gesichert und die Zielerreichung fortlaufend nachverfolgt. Die Projektkoordination übernahm die UniMS, das DFKI unterstützte durch Mitarbeit in Abstimmungen, Zuarbeiten zu Berichten und eine strukturierte interne Nachverfolgung der eigenen Ergebnisse. Im Rahmen des Projektmanagements wurden regelmäßige Projekt- und Teilprojekttreffen zur Abstimmung des Arbeitsstands und zur gemeinsamen Ableitung weiterer Schritte genutzt. Das DFKI übernahm dabei die Koordination von Terminen im eigenen Teilprojekt, bereitete Status- und Ergebnisberichte auf und stellte diese der Projektkoordination (UniMS) zur Verfügung. Ebenfalls wurden Beiträge zur Projektdokumentation, zu Sachstandsberichten sowie zu inhaltlichem Material für Öffentlichkeitsarbeit und Dissemination geliefert. Ebenfalls brachte sich das DFKI durch die frühzeitige Benennung technischer Risiken wie Datenverfügbarkeit, Schnittstellen, Umsetzungsaufwände sowie durch Vorschläge zu geeigneten Gegenmaßnahmen und Priorisierungen ein. Des Weiteren wurde vom DFKI in Trier ein Quartalsmeeting organisiert, das einen zentralen Austausch zwischen den Projektpartnern ermöglichte.

### 2.2 Datenschutzrechtliche Aspekte (AP 2)

In AP 2 wurden Aspekte des Personenbezugs mit den technischen Arbeitspaketen zur Plattformentwicklung und Evaluation verzahnt. Ziel war es, ein nachvollziehbares Vorgehen bereitzustellen, mit dem das verbleibende Risiko strukturiert eingeschätzt werden kann, und dieses Vorgehen anschließend auf die im Projekt eingesetzten Verfahren anzuwenden.

Im Arbeitsschritt UAP2.1 (Definition eines Bewertungsverfahrens für den Grad der Anonymisierung) wurde auf Grundlage von Best-Practice-Ansätzen sowie regulatorischen Leitlinien (u. a. EU AI Act, DSGVO) ein Bewertungsverfahren konzipiert. Im Vordergrund stand dabei die Anforderung, die eingesetzten Methoden so zu beschreiben, dass Annahmen, Parametrisierung und Grenzen der Verfahren nachvollziehbar und konsistent dokumentierbar sind. Das DFKI stellte in diesem Schritt den Bezug zu den in der Plattform verwendeten Synthetisierungsverfahren her und prüfte die Anwendbarkeit des Bewertungsverfahrens unter den Bedingungen praktischer KI-Forschung, insbesondere mit Blick auf typische Deep-Learning-gestützte Generatoren und deren Konfigurationsparameter. Dieser Arbeitsschritt konnte aufgrund von Personalengpässen bei dem Projektpartner DATATREE nicht im ursprünglich geplanten Zeitrahmen abgeschlossen werden. Da inhaltliche Synergien mit der Algorithmenbewertung in UAP2.2 bestehen, wurde UAP2.1 auf Dezember 2025 verschoben und auch erfüllt. Durch die Verschiebung des UAP2.1 auf Dezember 2025 haben sich keine Verzögerungen im Projekt ergeben und die Ziele des Projektes konnten erreicht werden.

Im Arbeitsschritt UAP2.2 (Bewertung der Algorithmen) wurde das in UAP2.1 definierte Vorgehen zur Bewertung der im Projekt eingesetzten Verfahren angewandt. Das DFKI fokussierte dabei Synthetisierungsansätze aus dem Bereich des Deep-Learnings und führte die Bewertung praktisch anhand frei verfügbarer medizinischer Datensätze sowie im Projektkontext genutzter Forschungsdaten durch. Dabei wurden sowohl die Nutzbarkeit der synthetisierten Daten als Grundlage für nachgelagerte Analysen betrachtet als auch Verfahren zur Einschätzung verbleibender Risiken einbezogen. Dazu gehörten unter anderem Holdout-Analysen sowie Attribut-Inferenz, Dataset-Linkage und Singling-Out-Analysen, die im Projektkontext durch das BIH für die Risikoanalyse implementiert wurden. Die Evaluation zeigte, dass die implementierten Synthetisierungsverfahren die Privatheit der Daten grundsätzlich gewährleisten. Diese Ergebnisse wurden genutzt, um Optimierungspotenziale in Verfahren und Parametrisierung zu identifizieren und als Rückkopplung für die Weiterentwicklung der Synthetisierungskomponenten in der Plattform bereitzustellen.

## **2.3 Spezifizierung von Anwendungsfällen und des Datensatzes (AP 3)**

In diesem Arbeitspaket lag der Schwerpunkt auf der Spezifizierung von Anwendungsfällen sowie der daraus abgeleiteten Definition eines geeigneten Datensatzes als Grundlage für die nachfolgenden technischen Arbeitspakete und die Evaluation. Der Beitrag des DFKI konzentrierte sich dabei auf die Prüfung der Umsetzbarkeit der Anwendungsfälle mit verfügbaren KI-Methoden sowie auf die Bewertung der KI-Readiness der vorgesehenen klinischen Daten im Hinblick auf Nutzbarkeit für Modelltraining und -evaluation.

Im Arbeitsschritt UAP3.1 (Spezifizierung von Anwendungsfällen) wurden im Rahmen einer Workshopserie mit klinischen Expertinnen und Experten aus dem Netzwerk Westdeutsches Tumorzentrum potenzielle Anwendungsfälle für KI-Verfahren diskutiert und abgestimmt. Das DFKI beteiligte sich an der Planung des Workshops und übernahm insbesondere die technische Machbarkeitsprüfung der vorgeschlagenen Use Cases anhand der potenziell verfügbaren Daten. Dazu gehörte die Einschätzung, welche Datenarten und -qualitäten für die jeweiligen Modellierungsansätze erforderlich sind und ob diese im klinischen Kontext realistisch bereitgestellt werden können. Auf dieser Grundlage wurden drei Anwendungsfälle als umsetzbar und fachlich sinnvoll klassifiziert: (1) ein Clinical Decision Support System, das auf Basis vorhandener Patientinnendaten Behandlungsempfehlungen ableitet, (2) die Vorhersage des Behandlungsverlaufs, also welche diagnostischen oder therapeutischen Schritte im Krankheitsverlauf als nächstes zu erwarten sind, sowie (3) die Vorhersage der Metastasierungswahrscheinlichkeit anhand klinischer Parameter zur Identifikation von Risikopatientinnen und -patienten. Für die nachgelagerten Arbeiten wurde der Anwendungsfall Clinical Decision Support System als Schwerpunkt gesetzt und im weiteren Projektverlauf detaillierter ausgearbeitet.

Im Arbeitsschritt UAP3.2 (Spezifizierung des Datensatzes) wurden die verfügbaren klinischen Daten durch die UniMS/IMI gesichtet und durch das DFKI systematisch bewertet. Das DFKI übernahm hierbei insbesondere die Bewertung der KI-Readiness der Daten sowie eine Potenzialanalyse, ob die Datensatzdefinition die Anforderungen der vorgesehenen KI-Methoden zur Adressierung der ausgewählten Anwendungsfälle abdeckt. Konkret prüfte das DFKI die Nutzbarkeit der ermittelten Daten als Input für Modelltraining und -evaluation (unter anderem hinsichtlich Strukturierungsgrad, Vollständigkeit und Konsistenz). Dadurch wurde sichergestellt, dass die Daten die jeweiligen Fragestellungen ausreichend adressieren und den klinischen Verlauf sowie relevante Einflussfaktoren angemessen repräsentieren. Dabei zeigte sich, dass direkte Datenexporte (ohne Vorverarbeitung) aus dem Krankenhausinformationssystem der Hautklinik Münster für den unmittelbaren Einsatz für KI-Methoden nur eingeschränkt geeignet sind. Das liegt daran, dass das verwendete Format (oBDS) für die Berichterstattung optimiert wurde. Gleichzeitig konnte im Projektkontext nachgewiesen werden, dass eine Nutzung möglich ist, sofern geeignete Vorverarbeitungsschritte zur Harmonisierung, Bereinigung, Strukturierung und Merkmalsableitung durchgeführt werden.

## 2.4 Entwicklung der Anonymisierungsplattform (AP 5)

In AP 5 wurde die KI-AIM-Plattform Cinnamon in einem iterativen, modularen Entwicklungsprozess implementiert. Ziel war die Entwicklung einer generischen und erweiterbaren Plattform, die die Verwaltung von Quelldaten und Konfigurationen, die Ausführung von Anonymisierungs- und

Synthetisierungsverfahren sowie die Risikobewertung der Ergebnisse unterstützt. Der Schwerpunkt der Beiträge des DFKI lag dabei auf der Entwicklung der Komponenten zur Synthese der Daten.

Im Arbeitsschritt UAP5.1 (Quelldatenverwaltung und Konfiguration) entwickelte das DFKI die Quelldatenverwaltung als zentrale Grundlage für die weiteren Plattformkomponenten. Hierfür wurden vom DFKI Funktionen implementiert, um Datensätze, Konfigurationsparameter und Modelle sowie Ontologien und Taxonomien zu verwalten. Zusätzlich wurden Mechanismen entwickelt, um Subdatensätze zu erzeugen, Daten zu filtern und Daten samt Konfigurationen so vorzubereiten, dass sie von nachgelagerten Komponenten genutzt werden können. Ergänzend wurden Vorschaufunktionen sowie statistische Auswertungen und Visualisierungen (Histogramme, Dichteverteilungen und Heatmaps) implementiert, um frühe Plausibilitätsprüfungen und eine Exploration der Datensätze zu unterstützen. Weiterhin erstellte das DFKI Konfigurationsansätze für die Parametrisierung von Synthetisierungsverfahren, um die Reproduzierbarkeit und Vergleichbarkeit verschiedener Durchläufe zu ermöglichen.

Im Arbeitsschritt UAP5.3 (GANs-basierte Synthese) entwickelte das DFKI Komponenten zur Datensynthese auf Basis von Deep-learning Methoden und integrierte diese in die Plattform. Zunächst erfolgte die Konzeptionalisierung von Klassen, Methoden und Konfigurationsdateien, die über einheitliche Interfaces die problemlose Integration neuer Synthetisierungsalgorithmen ermöglichen. Im nächsten Schritt wurden geeignete Synthesearchitekturen identifiziert, die in die Plattform integriert werden können. Bei der Auswahl wurde besonderer Wert auf wissenschaftlich evaluierte Verfahren gelegt und darauf geachtet, unterschiedliche Modellfamilien abzudecken. Neben GAN-basierten Generatoren wurden mehrere State-of-the-art Architekturen implementiert, darunter das Conditional Tabular Generative Adversarial Network (CTGAN), der Tabular Variational Autoencoder (TVAE), Adversarial Random Forests (ARF), Bayesian Networks (BN), Denoising Diffusion Probabilistic Models (DDPM), Probabilistic Autoregressive Regression (PAR) sowie Generation of Realistic Tabular Data (GReaT), das einen autoregressiven Ansatz auf Basis generativer Sprachmodelle nutzt. Die Verfahren wurden so eingebunden, dass sie über eine einheitliche, reproduzierbare Konfiguration nutzbar sind und auch Szenarien unterstützen, in denen bereits anonymisierte Daten im Anschluss synthetisiert werden. Dazu wurden generische Schnittstellen zur Frontend-Backend-Kommunikation implementiert und eine Konfigurationsoberfläche vorgesehen, die Parameterwahl und Ausführung über standardisierte Eingabelemente vereinheitlicht.

Im Arbeitsschritt UAP5.4 (wissensbasierte Synthese mit Prozessmodellen) wurden Verfahren zur Generierung synthetischer Daten mittels Simulation von Prozessmodellen entwickelt. Dafür wurden Process-Mining-Ansätze ausgewählt und angepasst, um aus Rohdaten Prozessmodelle über medizinische Abläufe abzuleiten und diese für eine zufallsgesteuerte, kontrollierte Simulation nutzbar zu machen. Das DFKI entwickelte hierfür einen Algorithmus, der auf Basis von Prozessdaten eine

interne Repräsentation eines Prozessmodells erzeugt und diese sowohl für Simulation als auch für die Synthetisierung nutzbar macht. Die Implementierung folgt einem autoregressiven Prinzip und nutzt Deep-Learning-Methoden, um insbesondere temporale Aspekte longitudinaler Daten und Prozessdaten adäquat zu erfassen und prozedurale Abhängigkeiten über Sequenzen hinweg abzubilden. Der Ansatz unterstützt unterschiedliche Modellarchitekturen, darunter Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) und Gated Recurrent Units (GRU), sowie optional bidirektionale Varianten, wodurch eine flexible Anpassung an verschiedene Anwendungsszenarien möglich ist. Durch die Nutzung der erlernten Prozessstrukturen und ergänzender fachlicher Modellannahmen wird die Generierung synthetischer, prozedural konsistenter Verlaufsdaten unterstützt. Die methodischen Ergebnisse wurden zudem vom DFKI wissenschaftlich aufgearbeitet und in einem Springer-Journal publiziert <sup>1</sup>.

Im Arbeitsschritt UAP5.6 (Integration von GANs und Prozessmodellsimulation) wurden die Ergebnisse aus UAP5.3 und UAP5.4 zusammengeführt. Ziel war es, synthetische, Prozessdaten zu erzeugen, die sowohl statistisch plausibel als auch prozedural konsistent sind. Dazu wurde der in UAP5.4 entwickelte Ansatz durch das DFKI erweitert, um semantische Informationen und prozedurale Randbedingungen explizit in die Generierung einzubringen. Hierzu wurden Methoden integriert, die bestehende Prozessmodelle wie Directly-Follow-Graphs (DFG) und Petri-Netze (PN) als Bedingungen für die Datengenerierung nutzen, wodurch die erzeugten Sequenzen semantisch durch Experten angereichert werden können. Für die praktische Umsetzung wurden Schnittstellen bereitgestellt, über die unterschiedliche Prozessmodelltypen in den Generierungsprozess eingebunden werden können. Ergänzend wurden Funktionen implementiert, die die im Prozessmodell definierten Bedingungen auf die Simulation und Sequenzgenerierung mit Deep-Learning-Verfahren übertragen, insbesondere für LSTM, RNN und GRU. Dadurch lässt sich die Generierung so steuern, dass über längere Zeiträume konsistente Sequenzen erzeugt werden können, wobei Expertenwissen gezielt berücksichtigt wird. Die in diesem Rahmen erarbeiteten Ergebnisse wurden zusätzlich wissenschaftlich aufgearbeitet und werden als Konferenzbeitrag („Synthetic Log Generation under Control-Flow Conditions using Autoregressive Models“) auf der CAiSE-Konferenz 2026 eingereicht.

## 2.5 Evaluation (AP 6)

In AP 6 wurde die in AP 5 entwickelte KI-AIM-Plattform Cinnamon evaluiert. Die Koordination lag beim DFKI. Ziel war es, ein belastbares Evaluationskonzept zu erstellen, die Plattform und ihre Ergebnisse in zwei Iterationen funktional zu bewerten sowie ergänzend die Gebrauchstauglichkeit der Plattform in einer Usability-Evaluation zu untersuchen. Die Evaluation kombinierte quantitative

<sup>1</sup><https://link.springer.com/article/10.1007/s44311-025-00033-5>

Metriken zur Vergleichbarkeit von Real-, anonymisierten und synthetischen Daten mit der Einbindung klinischer Domänenexpertinnen und Domänenexperten zur fachlichen Einordnung.

Im Arbeitsschritt UAP6.1 (Evaluationskonzept und Metrikauswahl/-definition) wurde vom DFKI das Evaluationskonzept für die KI-AIM-Plattform erarbeitet und in drei Phasen strukturiert. In der ersten Phase steht die Plattform selbst im Fokus, wobei Metriken zur Analyse der Nutzbarkeit der erzeugten Daten implementiert und auf die Plattformläufe angewendet werden. Die zweite Phase adressiert die Prüfung der in AP 3 spezifizierten Anwendungsfälle mit zusätzlichen KI-Werkzeugen, die nicht Teil der Plattform sind, um die Nutzbarkeit der anonymisierten und synthetischen Daten in konkreten Anwendungsszenarien nachvollziehbar zu testen. In der dritten Phase wird ergänzend die Benutzerfreundlichkeit untersucht, indem Nutzende nach der Durchführung definierter Aufgaben strukturiert zur Usability befragt werden. Zur Umsetzung dieser drei Phasen wurden, ausgehend von den in AP 3 definierten Anwendungsfällen, geeignete KI-Methoden und Werkzeuge ausgewählt, für die Nutzung auf den Testdaten vorbereitet und bei Bedarf angepasst. Dazu gehörten die Spezifikations- und Implementierungsarbeiten für Metriken, die die Nutzbarkeit der erzeugten Daten für KI-Methoden abbilden, sowie die erforderliche Datenaufbereitung und Wissensmodellierung für reproduzierbare Testläufe. Ein besonderer Schwerpunkt lag auf der Berücksichtigung medizinischer Taxonomien, da die klinischen Routinedaten Diagnosen und Prozeduren über standardisierte Codes (ICD- und OPS-Codes) verfügen. Um diese hierarchischen Strukturen in der Evaluation und im Use-Case sinnvoll zu nutzen, wurden am DFKI Methoden zur Abbildung von Ähnlichkeiten innerhalb der Taxonomien weiterentwickelt und für die Vorhersageaufgabe optimiert. Die Ergebnisse wurden in der Publikation „Leveraging Taxonomy Similarity for Next Activity Prediction in Patient Treatment“ und im Jahr 2024 im EURO Journal eingereicht und sind aktuell in Revision.

Im Arbeitsschritt UAP6.2 (Durchführung der Algorithmenevaluation) wurde basierend auf dem erstellten Evaluationskonzept eine zweistufige Bewertung umgesetzt. In der ersten Iteration führte das DFKI eine metrikbasierte Evaluation durch. Diese umfasste Resemblance als Maß für statistische Ähnlichkeit, ML-Utility als Maß für die Nutzbarkeit der Daten für Machine-Learning-Methoden sowie die praxisorientierte Eignung der Daten für den in AP 3 definierten medizinischen Use Case. Zur Quantifizierung der ML-Utility implementierte und nutzte das DFKI mehrere Referenzverfahren, darunter Support Vector Machines und Random Forests. Dadurch konnten Nutzbarkeitsunterschiede zwischen Real-, anonymisierten und synthetischen Daten nachvollziehbar und vergleichbar gemessen werden. In der zweiten Iteration entwickelte das DFKI ein Clinical Decision Support System auf Basis von Case-Based Reasoning, um den Use Case aus AP 3 gezielt zu adressieren. Der Fokus lag dabei auf der Next Activity Prediction im Behandlungskontext. Die Fallrepräsentation sowie die Ähnlichkeitsbestimmung wurden dabei in Workshops gemeinsam mit klinischen Domänenexpertinnen und -experten erarbeitet und anschließend technisch umgesetzt. Darauf aufbauend implementierte das DFKI die CBR-Methodik inklusive Ergebnisableitung und ergänzte diese um spezifische Skripte sowie

eine passende Auswertelogik. Die Anwendung der entwickelten Werkzeuge auf die Realdaten sowie auf anonymisierte und synthetische Daten erfolgte anschließend durch das Hauttumorzentrum in Münster. Die technische Bewertung wurde konsistent mit den in M6.1 definierten Metriken durchgeführt, sodass die Nutzbarkeit der unterschiedlichen Datensatzvarianten objektiv gegenübergestellt werden konnte. Ergänzend organisierte das DFKI Workshops mit Domänenexpertinnen und -experten. In diesen wurden ausgewählte Fälle aus dem Datensatz zusammen mit den Ergebnissen des CBR-Ansatzes strukturiert diskutiert und fachlich bewertet. Sämtliche Ergebnisse beider Iterationen wurden systematisch in der Evaluationsdokumentation festgehalten. Die wissenschaftlich relevanten Erkenntnisse wurden zusätzlich aufbereitet und im Rahmen der Publikation "Clinical Decision Support for Skin Tumor Treatment: A Case-Based Reasoning Approach" bei der International Conference on Case-Based Reasoning (ICCBR 2025) eingebracht und vor Fachpublikum diskutiert.

Im Arbeitsschritt UAP6.3 (Durchführung der Usabilityevaluation) wurde die Gebrauchstauglichkeit der Plattform gemeinsam mit der Universität Münster vorbereitet. Das Vorgehen wurde als Kombination aus szenariobasierten Tests und einer anschließenden Befragung mit der System Usability Scale konzipiert. Dafür wurde ein Testszenario entwickelt, das drei Aspekte abbildet. Es umfasst erstens ein Einstiegsszenario zur Prüfung zentraler Grundfunktionen, zweitens einen realitätsnahen Anwendungsfall mit vordefinierten Anonymisierungs- und Synthetisierungskonfigurationen und drittens Aufgaben zur Anpassung von Konfigurationen durch die Nutzenden. Das DFKI trug insbesondere zur Auswahl und Erstellung geeigneter Synthetisierungskonfigurationen bei. Weiterhin unterstützte das DFKI die Identifikation und Rekrutierung der Teilnehmenden und führte Interviews sowie User-Studies mit Expertinnen und Experten durch. Die Rückmeldungen wurden anschließend genutzt, um konkrete Verbesserungen der Bedienbarkeit und der Arbeitsabläufe in der KI-AIM Plattform Cinnamon abzuleiten.

## **2.6 Vorbereitung des wissenschaftlichen und technischen Diskurses (AP 7)**

In AP 7 lag der Schwerpunkt darauf, die Nachnutzung der Projektergebnisse vorzubereiten und den Transfer in Forschung und Praxis anzustoßen. Hierzu gehörte die fortlaufende Aufbereitung wissenschaftlicher Resultate sowie die Bereitstellung der in AP 5 und AP 6 entwickelten KI-AIM-Plattform Cinnamon als Open-Source-Software unter der Apache-2.0 Lizenz <sup>2</sup>. Ergänzend wurden Austauschformate und Workshops und Konferenzen genutzt, um Rückmeldungen aus der Community

---

<sup>2</sup><https://github.com/KI-AIM/Cinnamon>

systematisch aufzunehmen und den Aufbau einer nachhaltigen Anwender- und Entwicklercommunity zu unterstützen.

Im Arbeitsschritt UAP7.1 (Bereitstellung der Projektergebnisse) zur wissenschaftlichen Verwertung beteiligte sich das DFKI an der Erstellung projektbegleitender Dokumente und wissenschaftlicher Beiträge gemeinsam mit den Projektpartnern. Ein Schwerpunkt lag auf der aktiven Vorstellung und Diskussion der Plattform und der zugrunde liegenden Methoden in geeigneten Fachformaten. Besonders hervorzuheben ist ein gemeinsam mit der UniMS/IMI durchgeführter Workshop auf der GMDS, in dem die KI-AIM Plattform Cinnamon vorgestellt und durch Konferenzteilnehmende in einer Hands-on-Session praktisch erprobt wurde. Das DFKI stellte das KI-AIM-Projekt auch auf anderen Konferenzen vor und bereitete Publikationen vor. Eine Auflistung ist in Kapitel 7 zu finden.

Im Arbeitsschritt UAP7.2 (Planung der wirtschaftlichen Verwertung) unterstützte das DFKI die Veröffentlichung als Open-Source-Projekt durch Beiträge zur technischen und organisatorischen Aufbereitung. Dazu gehörten insbesondere Dokumentation und Nutzungshinweise, um Einstiegshürden zu reduzieren und eine verlässliche Community-Arbeit zu ermöglichen. Darüber hinaus beteiligte sich das DFKI an konsortialen Abstimmungen zu Verwertungsaspekten, soweit diese für einen nachhaltigen Betrieb und die langfristige Weiterentwicklung des Open-Source-Projekts relevant sind. Insgesamt wurden damit wesentliche Voraussetzungen geschaffen, um die Projektergebnisse über die Laufzeit hinaus öffentlich zugänglich zu machen, die Wiederverwendung in Forschung und Entwicklung zu erleichtern und eine aktive Community rund um die KI-AIM Plattform Cinnamon zu fördern.

### **3 Wichtigste Positionen des zahlenmäßigen Verwendungsnachweises**

Im Verbundprojekt ist der wesentliche Teil der Kosten durch Personalausgaben entstanden. An zweiter Stelle folgten Reisekosten. Die Reisekosten wurden geringfügig mit 10% überschritten, dies liegt jedoch im Rahmen der Austauschbarkeit der Mittel. Insgesamt wurden die Positionen des zahlenmäßigen Nachweises im Rahmen des Verwendungsnachweises und das Gesamtbudget um weniger als 1 % überschritten.

## 4 Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die Notwendigkeit und Angemessenheit der öffentlichen Förderung für die im Vorhaben KI-AIM geleisteten Arbeiten sind einerseits in der Charakteristik des Forschungsprojektes und andererseits in der hohen Allgemeinrelevanz der bearbeiteten Fragestellungen begründet. Die im Projekt geleisteten Arbeiten waren dem Arbeitsplan angemessen. Ebenfalls waren die durchgeführten Forschungsarbeiten im Teilprojekt des Verbundprojektes sowie die dafür aufgewandten Ressourcen notwendig und angemessen, da sie der im Projektantrag formulierten Planung entsprachen und alle wesentlichen im Arbeitsplan formulierten Aufgaben erfolgreich bearbeitet wurden. Darüber hinaus waren keine zusätzlichen Ressourcen für das Projekt notwendig.

## 5 Verwertbarkeit

In dem Teilvorhaben wurden keine Erfindungen getätigt bzw. Schutzrechtsanmeldungen vorgenommen.

### 5.1 Wirtschaftliche Erfolgsaussichten

Das Projekt KI-AIM hat gezeigt, dass die kombinierte Nutzung von Anonymisierung und KI-basierter Datensynthese in einer integrierten Plattform technisch umsetzbar ist und einen klaren Mehrwert für Forschung und Entwicklung im Gesundheitswesen bietet. Mit der KI-AIM Plattform Cinnamon wurden wiederverwendbare Komponenten, Schnittstellen und Konfigurationsmöglichkeiten geschaffen, mit denen Aufbereitung, Anonymisierung, Synthese und Bewertung sensibler Gesundheitsdaten strukturiert und nachvollziehbar durchgeführt werden können. Dadurch wird die Durchführung von KI- und Datenanalyseprojekten mit medizinischen Routinedaten erleichtert und der Aufwand für Datenbereitstellung, Abstimmung und Risikoabwägung perspektivisch reduziert, insbesondere durch schnellere Bereitstellung geeigneter Trainings- und Testdaten.

Wirtschaftliche Potenziale ergeben sich insbesondere dort, wo datenschutzkonforme Datenräume, Trainingsdaten für KI und nachvollziehbare Qualitätsnachweise benötigt werden, etwa in Kliniken, Forschungsnetzwerken, bei Anbietern medizinischer Software und datengetriebenen Dienstleistern. Die Open-Source-Veröffentlichung erhöht die Verwertbarkeit, da Unternehmen die KI-AIM Plattform Cinnamon als Basis für Integrationen, domänenspezifische Erweiterungen oder serviceorientierte Angebote nutzen können. Insgesamt ergeben sich daraus gute wirtschaftliche Erfolgsaussichten, da die Projektergebnisse als Grundlage für Anschlussprojekte, den Ausbau der Plattform sowie Kooperationen mit klinischen und industriellen Partnern dienen können.

## 5.2 Wissenschaftliche Erfolgsaussichten

Die wissenschaftlichen Ergebnisse des DFKI umfassen die Entwicklung und Integration von Syntheseverfahren in die KI-AIM-Plattform Cinnamon, insbesondere für tabellarische, longitudinale und Prozessdaten. Hierzu zählen sowohl Deep-Learning-basierte Synthetisierungsansätze als auch wissensbasierte Verfahren, die Prozesswissen und Eventdaten zur Generierung konsistenter Verlaufsdaten nutzen. Ergänzend wurden im Projekt evaluative Methoden und Werkzeuge erarbeitet, um die Nutzbarkeit synthetischer und anonymisierter Daten in konkreten Anwendungsszenarien zu bewerten, einschließlich eines fallbasierten Clinical-Decision-Support-Ansatzes als Demonstrator für die Use-Case-orientierte Evaluation.

Das Projekt hat die Expertise des DFKI in der datenschutzkonformen Nutzung medizinischer Daten, in der Synthese sensibler Daten sowie in der Überführung von KI-Methoden in nutzbare Softwareartefakte weiter gestärkt. Dieses Wissen kann in thematisch verwandten Vorhaben und möglichen Anschlussprojekten gezielt wiederverwendet und ausgebaut werden. Es wird erwartet, dass die erzielten Ergebnisse und die darauf aufbauenden wissenschaftlichen Veröffentlichungen auf Konferenzen und in Journals die wissenschaftlich-technische Kompetenz und Reputation weiter erhöhen. Mittel- und langfristig kann dies zu weiteren Forschungs- und Entwicklungsprojekten in Verbänden sowie gegebenenfalls auch zu Industriekooperationen führen. Bestandteil des wissenschaftlichen Verwertungskonzepts ist zudem, die Projektergebnisse in Forschung und Lehre einzubringen und als Anwendungsfeld in KI-orientierten Studiengängen zu verankern.

## 5.3 Wissenschaftliche und wirtschaftliche Anschlussfähigkeit

Das DFKI strebt als gemeinnützige Forschungseinrichtung grundsätzlich keine eigenständige wirtschaftliche Verwertung seiner Forschungsergebnisse an. Ein Transfer in die Anwendung erfolgt typischerweise in Kooperation mit dedizierten Wirtschaftspartnern in Folgevorhaben, bilateralen Projekten oder im Einzelfall über Ausgründungen. Im Sinne der wissenschaftlichen Anschlussfähigkeit sollen aufbauend auf den Projektergebnissen weitere Vorhaben initiiert werden, die gemeinsam mit wissenschaftlichen Partnern, Forschungseinrichtungen sowie Akteuren aus der Daten- und Gesundheitsbranche die Weiterentwicklung, Validierung und Nutzbarmachung der Ergebnisse vorantreiben. Der Transfer wird zudem durch wissenschaftliche Publikationen, Workshops und Weiterbildungsangebote unterstützt.

Auf wissenschaftlicher Ebene ermöglicht KI-AIM eine nachhaltige Erweiterung der Forschungsaktivitäten in den Themenfeldern Datensynthese, Privacy-Assessment sowie Generierung longitudinaler

Daten. Durch die Open-Source-Veröffentlichung von der KI-AIM Plattform Cinnamon besteht darüber hinaus eine klare technische Anschlussfähigkeit. Die Plattform kann als Integrations- und Experimentierumgebung genutzt und um weitere Verfahren, Domänenadaptionen und Bewertungsmethoden erweitert werden. Insgesamt rechnet das DFKI aus KI-AIM mit Ergebnissen, die in nachfolgenden Grundlagen- und Anwendungsprojekten gemeinsam mit Partnern aus Forschung, Klinik und Industrie umgesetzt werden können, sowohl im bestehenden Partnernetzwerk als auch darüber hinaus. Die zentralen F&E-Themen des Projekts besitzen das Potenzial, in zukünftigen Verbundprojekten vertieft zu werden.

## 6 Bekannt gewordener Fortschritt

Im Projektzeitraum wurden keine projektspezifischen Ergebnisse Dritter identifiziert, die die Durchführung des Vorhabens wesentlich verändert hätten. Hierzu wurden begleitend in regelmäßigen Abständen Literaturrecherchen durchgeführt.

## 7 Erfolgte und geplante Veröffentlichungen der Ergebnisse

Im Oktober 2023 wurde das Verbundprojekt im Rahmen der International Conference on Process Mining (ICPM 2023) im Workshop Process-oriented Data Science for Healthcare (PODS4H) mit einem Posterbeitrag vorgestellt.<sup>1</sup> Neben der Sichtbarkeit des Vorhabens im internationalen Umfeld diente der Austausch dazu, Rückmeldungen zur weiteren Ausgestaltung der Evaluation einzuholen.

Im Jahr 2024 wurde die KI-AIM-Plattform Cinnamon als Open-Source-Software auf GitHub veröffentlicht.<sup>2</sup> Die Veröffentlichung wurde durch eine begleitende Dokumentation sowie mehrere virtuelle Austauschformate ergänzt, um erste Rückmeldungen einzuholen und eine Community zur Nutzung und Weiterentwicklung aufzubauen.

Im Jahr 2025 wurden auf Basis der gesammelten Erkenntnisse und des Community-Feedbacks methodische Arbeiten zur Evaluation weiterentwickelt und als Preprint veröffentlicht: „Leveraging Taxonomy Similarity for Next Activity Prediction in Patient Treatment“.<sup>3</sup> Das zugehörige Manuskript wurde Ende 2024 beim EURO Journal on Decision Processes (EJDP) eingereicht und befindet sich zum Zeitpunkt der Berichtserstellung in Revision.

Im Jahr 2025 wurde auf der International Conference on Case-Based Reasoning (ICCBR 2025) der Beitrag „Clinical Decision Support for Skin Tumor Treatment: A Case-Based Reasoning Approach“ vorgestellt<sup>4</sup>. Diese Ergebnisse wurden im Projekt genutzt, um die Nutzbarkeit der Daten in dem klinischen Anwendungsfall systematisch zu bewerten. Des Weiteren wurde die Konferenz genutzt, um die Sichtbarkeit des Vorhabens im internationalen Umfeld zu stärken.

Im Jahr 2025 wurden mehrere Projektergebnisse im Rahmen der Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS 2025) kommuniziert. Dazu zählt die Publikation „Towards Fairness in Synthetic Healthcare Data: A Framework for

---

<sup>1</sup><https://www.pods4h.com/>

<sup>2</sup><https://github.com/KI-AIM/Cinnamon>

<sup>3</sup><https://arxiv.org/abs/2503.07638>

<sup>4</sup>[https://link.springer.com/chapter/10.1007/978-3-031-96559-3\\_27](https://link.springer.com/chapter/10.1007/978-3-031-96559-3_27)

the Evaluation of Synthetization Algorithms“ die gemeinsam mit der UniMS/IMI veröffentlicht wurde<sup>5</sup>. Ebenfalls wurde ein gemeinsamer Workshop durchgeführt, der durch das DFKI und die UniMS/IMI geleitet wurde. Dort wurde die KI-AIM Plattform Cinnamon zur Kombination von Anonymisierung, Synthese und Risikobewertung vorgestellt und in einer Hands-On-Session praktisch erprobt. Die Diskussion diente der Einordnung des Bedarfs sowie der Identifikation und Priorisierung von Verbesserungen.

Im Dezember 2025 wurde die Publikation „PALSYN: a method for synthetic multi-perspective event log generation with differential private guarantees“ im Process Science Journal von Springer veröffentlicht.<sup>6</sup> Die Arbeit schließt inhaltlich an die im Projekt entwickelten Ansätze zur Synthese prozeduraler und longitudinaler Daten an.

Für 2026 ist eine Einreichung bei der International Conference on Advanced Information Systems Engineering (CAiSE 2026) vorgesehen.<sup>7</sup> Der Beitrag mit dem Titel „Synthetic Log Generation under Control-Flow Conditions using Autoregressive Models“ führt die im Projekt erarbeiteten Ansätze zur Generierung von Prozessdaten unter Berücksichtigung von Expertenwissen weiter. Die Ergebnisse liegen bereits vor.

Ebenfalls ist für 2026 eine Einreichung zur Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS 2026)<sup>8</sup> zusammen mit der UniMS/IMI geplant, die aus Arbeiten zum Hyperparameter-Tuning hervorgegangen ist. Ziel ist es, Tuning-Strategien zur Optimierung Deep-Learning-basierter Syntheseverfahren bereitzustellen und damit die praktische Nutzbarkeit der Methoden in der KI-AIM-Plattform Cinnamon weiter zu erhöhen.

Als zusammenfassender Disseminationsschritt ist darüber hinaus für Anfang 2026 ein konsortialer Beitrag vorgesehen, der die KI-AIM-Plattform Cinnamon sowie zentrale methodische und evaluative Ergebnisse des Gesamtvorhabens konsolidiert beschreibt. Der vorläufige Titel lautet „Cinnamon: A Clinical Data Sharing Platform Unifying Anonymization, Synthetization, and Risk Assessment“. Eine Einreichung ist beim Journal Scientific Data geplant.<sup>9</sup>

---

<sup>5</sup><https://pubmed.ncbi.nlm.nih.gov/40899524/>

<sup>6</sup><https://link.springer.com/article/10.1007/s44311-025-00033-5>

<sup>7</sup><https://caise26.polimi.it>

<sup>8</sup><https://iscbgmgs2026.info>

<sup>9</sup><https://www.nature.com/sdata/>