

Das Projekt DATA CARE

Künstliche Intelligenz für klinische Studien

Das Projekt **DATA CARE**

Künstliche Intelligenz für klinische Studien

Autorin und Autoren

Sina Mackay
Dario Antweiler
Dr. Stefan Rüping

Fraunhofer IAIS

Oktober 2024

Das Fraunhofer IAIS

Als Teil einer der führenden Organisationen für anwendungsorientierte Forschung ist das Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS mit Sitz in Sankt Augustin/Bonn und einem Standort in Dresden eines der führenden Wissenschaftsinstitute auf den Gebieten Künstliche Intelligenz (KI), Maschinelles Lernen und Big Data in Deutschland und Europa. Rund 370 Mitarbeitende unterstützen Unternehmen bei der Optimierung von Produkten, Dienstleistungen und Prozessen sowie bei der Entwicklung neuer digitaler Geschäftsmodelle. Das Fraunhofer IAIS gestaltet die digitale Transformation unserer Arbeits- und Lebenswelt: mit innovativen KI-Anwendungen für Industrie, Gesundheit und Nachhaltigkeit, mit zukunftsweisenden Technologien wie großen KI-Sprachmodellen oder Quantum Machine Learning, mit Angeboten für die Aus- und Weiterbildung oder für die Prüfung von KI-Anwendungen auf Sicherheit und Vertrauenswürdigkeit.

Das Fraunhofer IAIS steht im Zentrum eines starken Forschungs- und Transfernetzwerks: Als einer von vier führenden Partnern betreibt das Fraunhofer IAIS KI-Spitzenforschung im »Lamarr-Institut für Maschinelles Lernen und Künstliche Intelligenz«, das als Teil der KI-Strategie der Bundesregierung dauerhaft vom Bund und Land Nordrhein-Westfalen gefördert wird. Den direkten Transfer der Forschungsergebnisse, insbesondere in mittelständische Unternehmen, sichert unter anderem die »Kompetenzplattform KI.NRW«, die das Fraunhofer IAIS als zentrale Vernetzungsinitiative des Landes Nordrhein-Westfalen zum Thema Künstliche Intelligenz leitet.

Weiterhin koordiniert das Fraunhofer IAIS als geschäftsführendes Institut die »Fraunhofer-Allianz Big Data und Künstliche Intelligenz«, welche die branchenübergreifende Expertise von über 30 Fraunhofer-Instituten bündelt und Fachkräfte aus Unternehmen mit einem etablierten und umfassenden Schulungsprogramm in Data Science und KI weiterbildet. Daneben bestehen langjährige enge Kooperationen in Forschung und Lehre mit der Universität Bonn und dem »Bonn-Aachen International Center for Information Technology« (b-it), das von der Fraunhofer-Gesellschaft, der Universität Bonn, der RWTH Aachen und der Hochschule Bonn-Rhein-Sieg getragen wird.

Das Fraunhofer IAIS engagiert sich zudem federführend in wichtigen Projekten und Initiativen zur Förderung der digitalen Souveränität Europas und der vertrauenswürdigen und nachhaltigen Anwendung von Künstlicher Intelligenz. Dazu zählen »OpenGPT-X« zur Entwicklung eines großen KI-Sprachmodells für Europa, »ZERTIFIZIERTE KI« zur Etablierung von Prüf- und Zertifizierungsverfahren für KI-Technologien oder die erfolgreiche Bildungsinitiative »Roberta«, die junge Menschen bereits ab dem Schulalter spielerisch an die Informatik heranführt.

www.iais.fraunhofer.de

Inhalt

Einleitung & Glossar	6
1. Der Einsatz von Künstlicher Intelligenz im Rahmen klinischer Studien	7
1.1. Auswertung von Dokumenten	7
1.2. Der Prozess klinischer Studien	7
2. Das DATACARE-Konzept	8
2.1. Datensouveränität und informierte Einwilligungsprozesse für klinische Studien	8
2.2. Herausforderungen des Patient-Trial-Matchings	8
2.3. Automatisierung des Patient-Trial-Matchings mit Künstlicher Intelligenz	9
2.4. Die Rolle von Foundation Models für das Patient-Trial-Matching	9
2.5. Die DATACARE App als Prototyp	10
Hintergrund & Projektpartner	11
3. Umfrageergebnisse zur Digitalisierung und Datensouveränität	12
4. Diskussion und Ausblick	13
Literaturverzeichnis	14
Impressum	15

Einleitung

Mit der digitalen Transformation wird die Nutzung von Gesundheitsdaten in Europa zunehmend relevant. Der von der EU-Kommission geplante Europäische Gesundheitsdatenraum (European Health Dataspace EHDS) soll eine grundlegende Voraussetzung dafür sein, die Kontrolle über persönliche Gesundheitsdaten zu stärken und diese für Forschung und Innovation zu nutzen. Der EHDS zielt darauf ab, eine sichere und effiziente Verwendung von digitalen Gesundheitsdaten zu gewährleisten, die sowohl den Schutz der Privatsphäre der Bürgerinnen und Bürger als auch den Fortschritt im Gesundheitswesen fördert.

Die Integration von Künstlicher Intelligenz (KI) im Gesundheitswesen eröffnet darüber hinaus neue Möglichkeiten, insbesondere in Bereichen, in denen Texte und Dokumente eine große Rolle spielen. KI-Technologien wie Natural Language Processing (NLP) und Large Language Models (LLMs) ermöglichen es dabei, große Mengen an Dokumenten zu analysieren und wertvolle Informationen aus unstrukturierten Daten zu extrahieren. Beispiele hierfür sind die automatisierte Erstellung von Arztbriefen oder die Zuordnung von Abrechnungs-codes, die den administrativen Aufwand für klinisches Personal erheblich reduzieren können.

Das Projekt DATACARE widmet sich speziell dem Bereich »Klinische Studien«, denn hier spielt textbasierte Dokumentation eine wichtige Rolle. Es zielt darauf ab, das Potenzial von Gesundheitsdaten zu erschließen und gleichzeitig die Datensouveränität der Patientinnen und Patienten zu schützen.

Ein zentraler Bestandteil des im DATACARE-Projekt entwickelten Konzepts ist ein optimiertes Patient-Trial-Matching (PTM): Hier geht es darum, geeignete Patientinnen und Patienten für klinische Studien zu finden und den Einwilligungsprozess zu vereinfachen. Dabei wird sowohl auf die Bedürfnisse des medizinischen Fachpersonals sowie auch der Patientinnen und Patienten eingegangen, um eine verständliche Kommunikation und höchste Datenschutz-Standards zu gewährleisten.

Im folgenden Whitepaper werfen wir zunächst einen Blick auf die technischen Möglichkeiten und die Funktionsweisen von KI und deren Potenzial zur Erleichterung von Aufgaben in klinischen Studien. Dann fokussieren wir auf den Ansatz des DATACARE-Projekts und dessen Ergebnisse. Abschließend geben wir einen Ausblick auf die praktische Anwendung der entwickelten Lösungen.

Glossar

Künstliche Intelligenz (KI):

Systeme, die komplexe Probleme intelligent lösen.

Foundation Models:

Große, vortrainierte Modelle für verschiedene KI-Aufgaben, die spezifisch angepasst werden können.

Large Language Models (LLMs):

Unterkategorie der Foundation Models für die Verarbeitung und Erzeugung natürlicher Sprache.

Natural Language Processing (NLP):

Ein Bereich der KI, der sich mit der Interaktion zwischen Computern und menschlicher Sprache befasst.

Generative Pre-trained Transformer 4 (GPT-4):

Ein LLM von OpenAI, das menschenähnlichen Text generieren kann und auf der Transformer-Architektur basiert.

1. Der Einsatz von Künstlicher Intelligenz im Rahmen klinischer Studien

1.1. Auswertung von Dokumenten

Textdokumente sind ein wesentlicher Bestandteil klinischer Studien, da sie wichtige Informationen über Studienprotokolle und Patientenberichte enthalten. KI hat das Potenzial, enorme Mengen solcher Dokumente automatisiert zu analysieren und wird so zu einer Schlüsseltechnologie für die Durchführung von klinischen Studien.

Insbesondere im Bereich NLP spielt die Informationsextraktion eine zentrale Rolle, um unstrukturierte klinische Texte zu analysieren und sie in strukturierte Formate zu überführen. Dies geschieht durch die Identifikation und Kategorisierung wichtiger Entitäten wie Diagnosen, Medikamente und Symptome in den Textdokumenten.

Dieses Vorgehen ermöglicht anschließend die Erstellung umfangreicher klinischer Datenbanken, die neben manuell erstellten Datenbanken [1] für Forschungszwecke genutzt werden können, etwa zur Erkennung von Trends in Krankheitsverläufen und zur Unterstützung der Therapieentwicklung. Diese Details entscheiden darüber, ob ein Patient oder eine Patientin für eine bestimmte Studie in Frage kommt.

Eine besonders leistungsfähige Technologie in diesem Zusammenhang sind die sogenannten Foundation Models, zu denen auch LLMs wie GPT-4 gehören.

Diese Modelle haben gezeigt, dass sie effektiv Informationen aus Fließtexten extrahieren können. Eine aktuelle Studie belegt, dass GPT-4 erfolgreich als Gutachter in systematischen Literaturübersichten eingesetzt wird und dabei eine nahezu perfekte Übereinstimmung mit menschlichen Gutachtern erzielt [2]. Die Erkenntnisse und Erfolge aus der Literaturrecherche können somit als wertvolle Grundlage dienen, um ähnliche Fortschritte im Bereich des PTM zu erzielen und die Effizienz der Patientenauswahl in klinischen Studien zu verbessern.

1.2. Der Prozess klinischer Studien

Klinische Studien sind ein wesentlicher Bestandteil des Zulassungsprozesses für Medikamente. Jeder Schritt ist mit Herausforderungen verbunden, die häufig zu Verzögerungen oder Abbrüchen führen (Abb. 1). Beispielsweise kann in der Prä-Klinik oder in der Forschungsphase eine unzureichende Wirkung der Therapie festgestellt werden, die einen Abbruch der Entwicklung nach sich zieht. Nur ein Bruchteil der Medikamente, die Gegenstand der Forschungen sind, erreichen am Ende das Ziel der Zulassung und Produktion. Durch den Einsatz automatisierter Textverarbeitungstechnologien kann ein Teil der Herausforderungen im Prozess jedoch erheblich gemildert werden: Denn KI ermöglicht es, große Mengen an textbasierten Daten effizient zu analysieren und wertvolle Informationen zu extrahieren, was die Effizienz und Genauigkeit in klinischen Studien deutlich steigern kann. Das Projekt DATA CARE fokussiert sich insbesondere auf den Prozess des PTM.

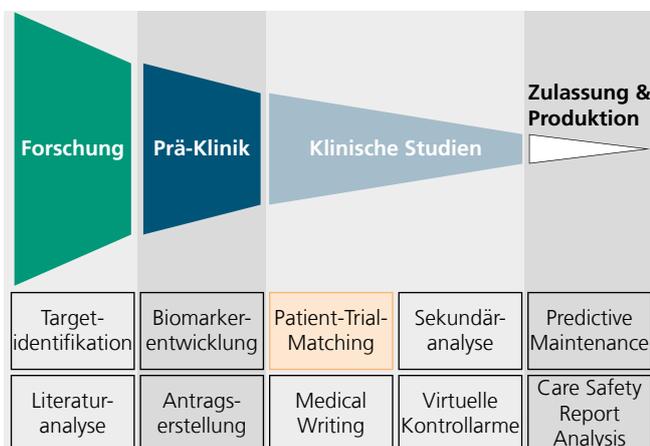


Abbildung 1: Zulassungsprozess für Medikamente: Die oberen Segmente repräsentieren die Hauptschritte: Forschung, Prä-Klinik, Klinische Studien und schließlich Zulassung und Produktion. Die nach rechts schmal werdenden Segmente repräsentieren die abnehmende Anzahl an Medikamenten, die weiterhin für die Zulassung in Betracht gezogen werden. Unter jedem Schritt sind relevante Teilprozesse aufgeführt, die die Herausforderungen und Aktivitäten innerhalb des jeweiligen Segments verdeutlichen. Die erste Zeile fokussiert sich auf Forschungs- und Entwicklungsprozesse, während die zweite Zeile administrative und regulatorische Aspekte behandelt.

2. Das DATACARE-Konzept

2.1. Datensouveränität und informierte Einwilligungsprozesse für klinische Studien

Im Fokus des DATACARE-Projekts stand die Optimierung der Rekrutierung von Studienteilnehmenden, wobei das PTM eine bedeutende Rolle spielt. Die zentrale Frage war, wie mit Hilfe von KI die Informations- und Einwilligungsprozesse für Patientinnen und Patienten sowie für das Fachpersonal erleichtert werden können, während gleichzeitig die Hoheit über die Gesundheitsdaten gewährleistet bleibt.

Ziel des DATACARE-Projekts war die Entwicklung eines Konzepts für Datensouveränität und informierte Einwilligungsprozesse, die es Patientinnen und Patienten ermöglichen, die Kontrolle über ihre persönlichen Gesundheitsdaten zu behalten und aktiv an der klinischen Forschung teilzunehmen. Dafür wurde ein interdisziplinäres Forschungsteam gebildet: Ein besonderer Fokus lag hierbei auf der Integration rechtlicher, klinischer, ökonomischer sowie technischer Forschungsexpertise. Darüber hinaus war eine Interessenvertretung der Patientinnen und Patienten aktiver Teil des Konsortiums, so dass die Perspektive der Betroffenen mit einfließen konnte.

Das entwickelte Konzept fördert Nutzerorientierung und Partizipation und nimmt Rücksicht auf die unterschiedlichen Perspektiven von Patientinnen und Patienten, Selbsthilfeorganisationen, Klinikpersonal und Pharmaindustrie. Durch das PTM wird das Angebot potenzieller klinischer Studien breiter und zugänglicher gestaltet, indem automatisierte Prozesse es ermöglichen, die Bedürfnisse der Patientinnen und Patienten gezielt zu erfassen und eine passgenaue Zuordnung zu geeigneten Studien zu gewährleisten.

Die Ergebnisse wurden in einem Prototyp als Smartphone-App umgesetzt, um die Anwendung des Konzepts praxisnah testen zu können.

2.2. Herausforderungen des Patient-Trial-Matchings

Im Rahmen des DATACARE-Projekts wurden mehrere zentrale Herausforderungen im Patient-Trial-Prozess identifiziert:

- Um den Matching-Prozess mit Hilfe einer App zu verbessern und zu erleichtern, muss die App trotz einer Vielzahl an Funktionen **benutzerfreundlich und übersichtlich** bleiben. Ein ausgewogener Mittelweg zwischen Detailtiefe und Klarheit ist daher unerlässlich.
- Zusätzlich müssen **rechtliche Fragen** geklärt werden, wie zum Beispiel die Dauer des Datenzugriffs und die Verwaltung der Daten durch die Patientinnen und Patienten selbst.
- Auch die **ethische Frage, ob Anreize zur Teilnahme an Studien** geschaffen werden sollten, indem das Freigeben von personenbezogenen und vertraulichen Daten oder die Teilnahme selbst monetär belohnt werden, muss betrachtet werden.
- Die Gewährleistung des **Datenschutzes** bei der Einführung neuer Tools, wie sie etwa Smartphone-Apps darstellen, erfordert strenge Datenschutzmaßnahmen sowie kontinuierliche Anpassungen der Sicherheitsprotokolle.
- Ein weiterer Aspekt ist die Klärung des **Zugriffs auf die elektronische Patientenakte (ePA)**. Die App muss mit einem KI-Modell kommunizieren, das aufgrund der erforderlichen Rechenleistung und Datenverarbeitungskapazität nicht lokal auf dem Smartphone gehostet werden kann, sondern an einem externen, sicheren Standort untergebracht sein muss. Eine geeignete Lösung für das Hosting der KI-Modelle, die im Rahmen einer App-Anwendung Daten aus der ePA verarbeiten, muss gefunden werden. Aus Datenschutzgründen sollten diese Modelle lokal bei einer vertrauenswürdigen Institution innerhalb Deutschlands gehostet werden. Cloud-Lösungen könnten in diesem Kontext weniger geeignet sein, da es oft schwieriger ist, so die erforderlichen Datenschutzstandards zu garantieren.
- Eine weitere Hürde besteht darin, medizinische und wissenschaftliche Inhalte in einer für Patientinnen und Patienten **zugänglichen Sprache** zu vermitteln. Dies ist entscheidend, um ihnen fundierte Entscheidungen zu ermöglichen und die eigenständige Verwaltung ihrer Gesundheitsdaten zu unterstützen.

Diese Herausforderungen verdeutlichen die Komplexität des Projekts und unterstreichen die Notwendigkeit eines interdisziplinären Ansatzes. Die Beteiligung von Patientinnen und Patienten an dem Entwicklungsprozess ist essenziell, um nachhaltige und vertrauenswürdige Lösungen zu entwickeln.

2.3. Automatisierung des Patient-Trial-Matchings mit Künstlicher Intelligenz

Die Vorplanung einer klinischen Studie kann mehrere Monate dauern. Nach der Festlegung klinischer Endpunkte muss die Fallzahl festgelegt werden, die eine statistisch ausreichende Datengrundlage erwarten lässt. Anschließend besteht die Herausforderung darin, diese Fallzahl durch Rekrutierung einer ausreichend großen Anzahl an Patientinnen und Patienten zu erreichen. Tatsächlich scheitert ein großer Anteil klinischer Studien an unzureichender Rekrutierung (Abb. 1). Je nach Region und Indikation sind es einer Metaanalyse zufolge mindestens 20 Prozent [3].

PTM verwendet KI-Algorithmen, um die Eignung von Patientinnen und Patienten für klinische Studien zu bewerten. Dafür werden mit Hilfe von KI unstrukturierte Dokumente aus zwei Quellen ausgewertet: patientenbezogene und studienbeschreibende Dokumente (Abb. 2).

Je nach Ansatz können beide Quellen getrennt ausgewertet und extrahierte Informationen in eine strukturierte Form gebracht werden. Auf dieser Grundlage kann die Übereinstimmung zwischen den extrahierten Angaben berechnet werden [4], [5]. Alternativ können beide Quellen in einen gemeinsamen Vektorraum eingebettet werden. Das Matching-Problem kann als Klassifikationsmodell umgesetzt werden, das darauf abzielt, Patientinnen und Patientenakten und Studien als passend, unpassend oder neutral einzustufen. Ein entsprechender

Ansatz namens »DeepEnroll« verfügt über eine spezielle Komponente zur Handhabung numerischer Daten wie Alter, Laborergebnisse und Dosierungen. Diese numerischen Daten werden ebenfalls in den Vektorraum eingebettet. Ein ergänzendes Entailment-Modul gleicht numerische Informationen in den Ein- und Ausschlusskriterien und den elektronischen Patientenakten explizit ab [6].

Mit dem Automatisieren dieser Prozesse kann die Effizienz klinischer Studien entscheidend erhöht werden und die Zeit, die für die Entwicklung neuer Therapien benötigt wird, verkürzt sich deutlich. Der Einsatz von Foundation Models ist hierbei zentral.

2.4. Die Rolle von Foundation Models für das Patient-Trial-Matching

Foundation Models sind leistungsstarke LLMs, die auf umfangreichen Textdaten trainiert werden. Sie haben die Fähigkeit, große Textmengen zu verstehen und komplexe Muster und Beziehungen innerhalb dieser Texte zu erkennen. Sie eignen sich daher in besonderer Weise dafür, beim PTM eingesetzt zu werden:

Die wichtigsten Funktionen:

- **Textanalyse:** Foundation Models extrahieren relevante Informationen aus klinischen Studiendokumenten und Patientenakten.
- **Sprachverständnis:** Sie interpretieren die Bedeutung von Texten, um die Ein- und Ausschlusskriterien von Studien mit Patienteninformationen abzugleichen.
- **Lernfähigkeit:** Sie verbessern sich kontinuierlich durch Training auf neuen Daten, was zu präziseren Matching-Ergebnissen führt.

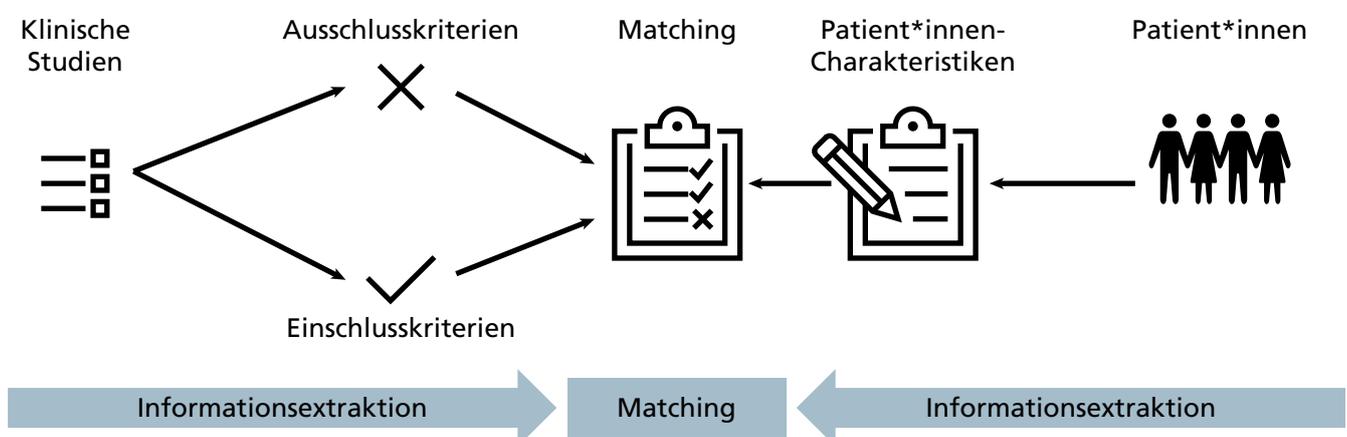


Abbildung 2: Übersicht der Funktionsweise eines automatisierten Systems für Patient-Trial-Matching

Die wichtigsten Vorteile:

- **Effizienzsteigerung:** Automatisierung des Matching-Prozesses spart Zeit und Ressourcen.
- **Genauigkeit:** Hochentwickelte Algorithmen erhöhen die Wahrscheinlichkeit eines korrekten Matchings.
- **Skalierbarkeit:** KI-Modelle können mit steigender Datenmenge wachsen und sich anpassen.

Herausforderungen und Risiken:

- **Datenqualität:** Die Genauigkeit des Matchings hängt von der Qualität der Patientendaten ab.
- **Bias und Fairness:** Verzerrungen in den Trainingsdaten können zu unfairen oder diskriminierenden Ergebnissen führen.
- **Transparenz und Vertrauen:** Es muss sichergestellt werden, dass die Nutzerinnen und Nutzer verstehen, wie die Empfehlungen zustande kommen, und dass sie den Ergebnissen vertrauen können.

Insgesamt ist das Potenzial von Foundation Models im Bereich des PTM enorm, doch es ist wichtig, dass die Entwicklung und Anwendung dieser Technologie verantwortungsvoll und mit Blick auf ethische Aspekte erfolgt.

2.5. Die DATACARE App als Prototyp

Im Rahmen des DATACARE-Projekts wurde ein Prototyp der bereits erwähnten App entwickelt, mit deren Hilfe die Rekrutierung von Teilnehmerinnen und Teilnehmern für klinische Studien erleichtert werden soll. Die Automatisierung des Rekrutierungs- und Kommunikationsprozesses im Rahmen klinischer Studien soll sowohl den Studienzentren als auch den Patientinnen und Patienten zugutekommen. Die App scannt dabei ein breites Angebot an klinischen Studien und stellt eine Übersicht dar. Nutzerinnen und Nutzer sind dann in der Lage, selbständig Anfragen zur Teilnahme zu stellen.

Dabei soll die App den Zugriff auf die persönliche ePA ermöglichen, die ab 2025 für alle gesetzlich Versicherten bereitgestellt werden wird. Patientinnen und Patienten sollen innerhalb der

App so in der Lage sein, ihre Gesundheitsdaten selbst zu verwalten und unter Umständen sogar zu monetarisieren.

Die im Rahmen des Forschungsprojekts prototypisch entwickelte App greift dabei auf synthetische Datensätze zu und simuliert das Bereitstellen von Daten. Der Prototyp arbeitet mit strukturierten synthetischen Daten im FHIR-Format (Fast Healthcare Interoperability Resources). FHIR ist ein standardisiertes Format zur elektronischen Übertragung von Gesundheitsinformationen, das Interoperabilität zwischen verschiedenen Systemen fördert. Es ermöglicht die einfache Integration und den Austausch von Gesundheitsdaten zwischen Anwendungen und Systemen.

Es ist jedoch zu erwarten, dass die ePA auch unstrukturierte Daten wie beispielsweise Fließtexte in pdf-Dokumenten beinhaltet. Um auf dem Prototyp aufbauend eine Anwendung zu entwickeln, die in der Praxis funktioniert, müsste die App solche unstrukturierten Daten verarbeiten können. Dabei kann KI helfen, denn sie kann aus unstrukturierten Daten wie Texten Informationen herausfiltern und in ein strukturiertes Format wie etwa eine Tabelle überführen. Dieser Zwischenschritt erleichtert automatisches Filtern und Gruppieren von unstrukturierten Dokumenten. So wird Patientinnen und Patienten die Handhabung und Auswahl von relevanten Dokumenten für Facharzttermine erleichtert.

Ein Schlüsselmerkmal des Prototyps ist die Zugänglichkeit der Inhalte und Funktionen, um auch weniger technikaffine Menschen abzuholen und einzubeziehen. Nur so können Nutzerinnen und Nutzer ermächtigt werden, selbstständig Entscheidungen über die Teilnahme an klinischen Studien zu treffen.

Das entwickelte Konzept und der auf dieser Basis entstandene Prototyp zeigen insbesondere Ansätze zur Verbesserung der klinischen Forschung auf. Um einen realen Einsatz in der Praxis in Zukunft umsetzen zu können, ist es entscheidend, dass der Prototyp weiter optimiert und in realen klinischen Umgebungen getestet wird. Dabei müssen kontinuierlich Rückmeldungen von den Nutzerinnen und Nutzern eingeholt und die Systeme entsprechend angepasst werden.

Hintergrund

Forschungsergebnisse aus dem Projekt PARIS liefern eine zentrale Komponente für die DATACARE App

Die Herausforderung, gezielt Informationen in Fließtexten zu identifizieren, war bereits Gegenstand des Projekts PARIS (Patient Recruitment Information System). In diesem Kontext wurde ein innovativer Ansatz zur systematischen Literaturrecherche entwickelt, der auf der Nutzung von Sprachmodellen basiert [2]. Das Projekt PARIS hat Fortschritte in der Entwicklung von Technologien zur Patientenversuchszuordnung und zur systematischen Literaturübersicht erzielt. Ein wesentlicher Erfolg war die Bereitstellung und Nutzung der Annotations- und Extraktionssoftware *AnEx*. Dafür wurde ein NLP-Modell zur Informationsextraktion trainiert und angepasst.

Durch umfangreiche Markt- und Literaturrecherchen wurden spezifische Anforderungen und Anwendungsfälle definiert. Ein Demonstrator mit Datenlade- und Visualisierungsfunktionen wurde entwickelt und steht für die Modellintegration bereit. Zudem wurden Technologien zur Automatisierung der systematischen Literaturübersicht mithilfe von LLMs entwickelt.

Die DATACARE Projektpartner

Das Projekt wird vom Bundesministerium für Bildung und Forschung (BMBF) gefördert und von dem Deutschen Zentrum für Luft- und Raumfahrt Projektträger (DLR-PT) organisiert.

Die Projektpartner sind die Fraunhofer-Institute ITMP, IMW und IAIS, das Universitätsklinikum Frankfurt, die Goethe-Universität Frankfurt und das Karlsruher Institut für Technologie (KIT).

Das Projektteam sammelte klinische Studien, entwickelte Annotationsleitfäden und automatisierte die Datenextraktion aus wissenschaftlichen Publikationen. Ein Prototyp für die Patientenversuchszuordnung wurde erstellt, der die Übereinstimmung zwischen klinischen Narrativen und Einschlusskriterien bewertet. Die Ansätze wurden auf verschiedenen Veranstaltungen und Konferenzen präsentiert, um die Ergebnisse zu teilen und Kooperationen mit Kliniken und Institutionen zu fördern.

Die Resultate aus PARIS tragen direkt zur Zielsetzung des DATACARE-Projekts bei, in dem ein Konzept für Datensouveränität und informierte Einwilligung in der klinischen Forschung entwickelt wurde. Insbesondere die Technologien zur Patientenversuchszuordnung verbessern die Nutzerorientierung und Partizipation, indem sie Patienten befähigen, ihre eigenen Daten für die Teilnahme an klinischen Studien zu verwalten. Dies unterstützt die übergeordneten Ziele von DATACARE, die Datensouveränität der Patientinnen und Patienten zu stärken und die Effizienz der klinischen Forschung zu erhöhen.

3. Umfrageergebnisse zur Digitalisierung und Datensouveränität

Über die Entwicklung eines Konzepts und App-Prototyps für das PTM hinaus wurde im Rahmen des DATA-CARE-Projekts auch eine Umfrage zum Thema Digitalisierung und Datensouveränität durchgeführt. Dabei wurden Meinungen und Erwartungen von Patientinnen und Patienten sowie ärztlichem Fachpersonal abgefragt. Die Ergebnisse bieten wertvolle Einblicke, welches Potenzial der Einsatz digitaler Technologien und KI im Gesundheitswesen allgemein bietet:

Patientinnen und Patienten nutzen digitale Technologien und sehen Vorteile der KI

- Die Mehrheit der Patientinnen und Patienten nutzt regelmäßig digitale Technologien und viele sind bereit, persönliche medizinische Daten digital an ihre Ärztinnen und Ärzte zu übermitteln.
- Bezüglich der Nutzung von KI im Gesundheitswesen fühlen sich die Patientinnen und Patienten unterschiedlich gut informiert.
- Vorteile der KI wurden als diagnostische Entscheidungsunterstützung, Terminmanagement, kosteneffiziente Behandlungen und Früherkennung von Krankheiten identifiziert.

Ärztinnen und Ärzte sehen das Potenzial digitaler Anwendungen und betonen die Bedeutung von Zuverlässigkeit und Sicherheit

- Auch die Ärztinnen und Ärzte sehen großes Potenzial in digitalen Anwendungen im Gesundheitswesen.
- Benutzerfreundlichkeit, Zeitersparnis und Datensicherheit wurden als wichtige Merkmale zukünftiger digitaler Plattformen genannt.
- Bezüglich der KI im Gesundheitswesen betonten die Ärztinnen und Ärzte die Wichtigkeit von Zuverlässigkeit und Sicherheit, um das Vertrauen in die Systeme zu stärken und sicherzustellen, dass sie gegen Ausfälle und Manipulationen geschützt sind.

Diese Ergebnisse unterstützen das Projekt-Vorhaben, patientenfreundliche und praxisorientierte Lösungen zu entwickeln. Der DATA-CARE-Prototyp stellt dabei einen wichtigen ersten Schritt dar für eine technische Lösung, welche Patientinnen und Patienten die Verwaltung und Nutzung von Gesundheitsdaten erleichtert und zudem die Teilnahme an klinischen Studien fördert.

Gleichzeitig ist damit das volle Potenzial digitaler Anwendungen in der Gesundheitsversorgung noch lange nicht ausgeschöpft: Krankenhäuser und Pflegeeinrichtungen setzen bereits digitale Technologien wie maschinelles Lernen, intelligente Sensoren und Big-Data-Analysen ein, um die Qualität der Versorgung und die betriebliche Effizienz zu verbessern. Die Anwendung solcher Technologien zeigt weitreichendes Potenzial, von der einfachen Innovation operativer Prozesse bis hin zu den anspruchsvollsten Behandlungen von Notfall-Patientinnen und -Patienten.

Bei all diesen technologischen Fortschritten kann die Bedeutung einer ordnungsgemäßen Einwilligung, Datensicherheit und eines verantwortungsvollen Umgangs mit diesen Technologien nicht genug betont werden – entsprechende Richtlinien und ethische Leitlinien für Gesundheitsdienste, die kontinuierlich an den technologischen Fortschritt angepasst werden, müssen entwickelt werden. Für eine breite Akzeptanz und zielgerichtete Implementierung zukünftiger Systeme ist es essenziell, dass die Akteure der medizinischen Daten in einem transparenten Prozess so weit wie möglich einbezogen werden.

In der projektinternen DATA-CARE-Umfrage wurden diesbezüglich Gemeinsamkeiten zwischen Patientinnen und Patienten sowie Ärztinnen und Ärzten festgestellt. Obwohl es spezifische Abweichungen in der Bewertung bestimmter Schwerpunkte gab, legten beide Gruppen Wert auf wesentliche Aspekte bei der Implementierung zukünftiger Datenmanagementsysteme. Sowohl Patientinnen und Patienten als auch Ärztinnen und Ärzte erachteten die folgenden Punkte in diesem Kontext als besonders wichtig:

- **Verbesserte Effizienz:** Die Befragten erwarten durch die Implementierung von KI eine Verbesserung der Lebensqualität für Patientinnen und Patienten, verbesserte Behandlungsqualität durch Ärztinnen und Ärzte und verbessertes Ressourcenmanagement im Allgemeinen.
- **Klare Kommunikation über (KI-bezogene) Prozesse:** Die Befragten wünschen sich eine klare Kommunikation darüber, was technisch möglich ist und wie es implementiert wird.
- **Datenmanagement:** Die Befragten legen großen Wert auf die Gewährleistung von Datensicherheit und Vertraulichkeit.

- **Digitalisierung von Dokumenten:** Die Befragten erwarten die Digitalisierung von ärztlichen Bewertungen, Bildgebung oder Einwilligungsformularen als Grundlage für eine unkomplizierte Übermittlung an Gesundheitsdienstleister und eine vereinfachte Bereitstellung und Freigabe durch Patientinnen und Patienten.

Zusammenfassend lässt sich sagen, dass die DATACARE-Umfrage wichtige Einblicke in die Bedürfnisse und Erwartungen von Patientinnen und Patienten und Ärztinnen und Ärzten im Kontext der Digitalisierung des Gesundheitswesens geliefert hat. Gleichzeitig wird durch die Umfrage deutlich, dass das Potenzial

noch nicht ausgeschöpft ist und noch viele Möglichkeiten im Bereich der Patientenversorgung und -kommunikation liegen. Besonders die Integration von KI und NLP-Technologien und hier speziell der Einsatz von Foundation Models und Large Language Models ist vielversprechend.

Ein weiterer wichtiger Schritt wird die Entwicklung von Richtlinien und Standards sein, die den sicheren und ethischen Einsatz dieser Technologien gewährleisten. Hierbei spielt die Zusammenarbeit mit regulatorischen Behörden, Datenschutzexperten und Ethikkommissionen eine entscheidende Rolle.

4. Diskussion und Ausblick

Die Ergebnisse des DATACARE-Projekts verdeutlichen die Relevanz von KI und NLP in der klinischen Forschung, insbesondere im Kontext des PTM. Die Fähigkeit, geeignete Patientinnen und Patienten für klinische Studien effizient zu identifizieren, stellt einen entscheidenden Fortschritt dar. Durch den Einsatz von KI-Technologien kann der Matching-Prozess automatisiert und optimiert werden, was nicht nur die Effizienz der Rekrutierung erhöht, sondern auch die Wahrscheinlichkeit, dass Patientinnen und Patienten in die für sie passenden Studien aufgenommen werden.

Sollte die prototypisierte App tatsächlich für den Einsatz in der Praxis weiterentwickelt werden, könnte dies zu einer signifikanten Verbesserung des Zugangs zu klinischen Studien führen. Patientinnen und Patienten hätten die Möglichkeit, einfacher auf relevante Studienangebote zuzugreifen und sich aktiv an der Auswahl zu beteiligen. Dies würde nicht nur die Rekrutierung beschleunigen, sondern auch die Patientenzufriedenheit erhöhen, da sie die Kontrolle über ihre Teilnahme an klinischen Studien hätten.

Langfristig könnte die erfolgreiche Umsetzung dieser Technologien im PTM nicht nur die klinische Forschung revolutionieren, sondern auch die Patientenversorgung nachhaltig verbessern. Dies geschieht durch bessere Studienangebote, schnellere Medikamentenzulassungen und individuellere Therapieansätze. Es ist entscheidend, dass zukünftige Entwicklungen verantwortungsvoll erfolgen und ethische Standards eingehalten werden, um das Vertrauen in digitale Gesundheitslösungen zu sichern.

Die Herausforderungen, die bei der Vermittlung medizinischer Inhalte in verständlicher Sprache an Patientinnen und Patienten bestehen, sind dabei nicht zu unterschätzen. Eine klare und transparente Kommunikation ist essenziell, um das Vertrauen der Patientinnen und Patienten zu gewinnen und ihre aktive Teilnahme an klinischen Studien zu fördern. Außerdem müssen ethische Bedenken und Datenschutzfragen stets im Blick behalten werden, insbesondere wenn es um die Verwendung ihrer Daten im PTM geht.

Die Beteiligung von Patientinnen und Patienten im Entwicklungsprozess ist unerlässlich, um nachhaltige und vertrauenswürdige Lösungen zu schaffen. Nur so können die entwickelten Technologien effizient und verantwortungsvoll in die Praxis umgesetzt werden.

Das DATACARE-Projekt hat ein umfassendes Konzept entwickelt, das eine komplexe Thematik aus verschiedenen Blickwinkeln beleuchtet. Eine der wichtigsten Komponenten dieses Konzepts ist die Nutzung von KI zur Verarbeitung von Textdokumenten. Die entwickelten Prototypen und Konzepte zeigen, dass diese Technologien erhebliche Vorteile bei der Identifizierung geeigneter Patientinnen und Patienten für klinische Studien bieten. Durch die Automatisierung des Matching-Prozesses wird nicht nur die Effizienz erhöht, sondern auch die Datensouveränität der Patientinnen und Patienten gestärkt. Insgesamt bietet das Projekt großes Potenzial, um die klinische Forschung nachhaltig effizienter zu gestalten und die Patientenversorgung zu verbessern.

Literaturverzeichnis

- [1] N. J. Dobbins, T. Mullen, Ö. Uzuner, and M. Yetisgen, 'The Leaf Clinical Trials Corpus: a new resource for query generation from clinical trial eligibility criteria', *Sci. Data*, vol. 9, no. 1, p. 490, Aug. 2022, doi: 10.1038/s41597-022-01521-0.
- [2] A. Landschaft *et al.*, 'Implementation and evaluation of an additional GPT-4-based reviewer in PRISMA-based medical systematic literature reviews', *Int. J. Med. Inf.*, vol. 189, p. 105531, Sep. 2024, doi: 10.1016/j.ijmedinf.2024.105531.
- [3] D. B. Fogel, 'Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review', *Contemp. Clin. Trials Commun.*, vol. 11, pp. 156–164, Aug. 2018, doi: 10.1016/j.conctc.2018.08.001.
- [4] S. Datta *et al.*, 'AutoCriteria: a generalizable clinical trial eligibility criteria extraction system powered by large language models', *J. Am. Med. Inform. Assoc.*, vol. 31, no. 2, pp. 375–385, Jan. 2024, doi: 10.1093/jamia/ocad218.
- [5] Y. Ni *et al.*, 'Increasing the efficiency of trial-patient matching: automated clinical trial eligibility Pre-screening for pediatric oncology patients', *BMC Med. Inform. Decis. Mak.*, vol. 15, no. 1, p. 28, Apr. 2015, doi: 10.1186/s12911-015-0149-3.
- [6] X. Zhang, C. Xiao, L. M. Glass, and J. Sun, 'DeepEnroll: Patient-Trial Matching with Deep Embedding and Entailment Prediction', in *Proceedings of The Web Conference 2020*, Taipei Taiwan: ACM, Apr. 2020, pp. 1029–1037. doi: 10.1145/3366423.3380181.

Impressum

Förderkennzeichen

BMBF 01GP2112

Herausgeber

Fraunhofer-Institut für Intelligente Analyse-
und Informationssysteme IAIS
Schloss Birlinghoven 1
53757 Sankt Augustin

Redaktion

Silke Loh
Eléna Zay-Vanvoorden

Grafik und Layout

Achim Kapusta
Asra-Soraya Neumeister

Bildnachweise

Cover: © Pisan - stock.adobe.com
Abbildung 1: Fraunhofer IAIS
Abbildung 2: Fraunhofer IAIS

Autorin und Autoren

Sina Mackay ist Data Scientist am Fraunhofer IAIS und arbeitet im Bereich Healthcare Analytics. Sie leitet Projekte, die sich mit der Anwendung von Künstlicher Intelligenz in medizinischen Kontexten befassen. Ihr Arbeitsfokus liegt auf der Analyse und Nutzung medizinischer Daten zur Unterstützung von Forschungs- und Entwicklungsprojekten im Gesundheitswesen.

Dario Antweiler ist Teamleiter des Geschäftsfelds Healthcare Analytics am Fraunhofer IAIS und betreut Projekte in den Themenbereichen Digitalisierung im Krankenhaus sowie KI in der Pharmakologie. Sein Forschungsfeld ist Machine Learning und Visual Analytics im Gesundheitswesen.

Dr. Stefan Rüping ist Leiter der Abteilung Knowledge Discovery am Fraunhofer IAIS. Er forscht in den Bereichen Explainable AI und Natural Language Processing.

Stand

Oktober 2024

© Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS, Sankt Augustin, Oktober 2024

Kontakt

Sina Mackay
Data Scientist
Abteilung Knowledge Discovery
sina.mackay@iais.fraunhofer.de

Fraunhofer-Institut für Intelligente
Analyse- und Informationssysteme IAIS
Schloss Birlinghoven 1
53757 Sankt Augustin

www.iais.fraunhofer.de