Architektur dezentraler Infrastrukturen zur souveränen Verwaltung und effektiven Nutzung von Gesundheitsdaten

Zur Erlangung des akademischen Grades einer

DOKTORIN DER INGENIEURWISSENSCHAFTEN
(Dr.-Ing.)

von der KIT-Fakultät für
Elektrotechnik und Informationstechnik
des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M.Sc. Christina Erler
geb. in Meißen

Tag der mündlichen Prüfung:	22.07.2025
Hauptreferent:	Prof. Dr. rer. nat Wilhelm Stork
Korreferent:	Prof. Dr. York Sure-Vetter

Zusammenfassung

Dezentrale Dateninfrastrukturen auf Basis der Blockchain-Technologie werden zunehmend als vielversprechende Lösungen für die sichere, selbstsouveräne Verwaltung, Speicherung und wissenschaftliche Nutzung von sensiblen Gesundheitsdaten in der Wissenschaft und Praxis diskutiert. Diese Infrastrukturen schaffen die Voraussetzung für den Gebrauch von Gesundheitsdaten in der personalisierten Medizin, wodurch sich eine effektivere Gesundheitsversorgung erhofft wird. Speziell die Architektur solcher Systeme ist herausfordernd, wie die schleppende Umsetzung der elektronischen Patientenakte (ePA) in Deutschland zeigt. Trotz des Potentials solcher Gesundheitsdateninfrastrukturen zur Primär- und Sekundärnutzung medizinischer Daten bestehen erhebliche technische, organisatorische und rechtliche Herausforderungen. Diese Herausforderungen erfordern eine systematische Berücksichtigung und Adressierung bei der Konzeption und Entwicklung derartiger Systeme, um deren Effektivität, Sicherheit und Akzeptanz langfristig zu gewährleisten.

Vor diesem Hintergrund widmet sich die vorliegende Arbeit der systematischen Untersuchung und methodischen Unterstützung der Architektur dezentraler Dateninfrastukturen im Gesundheitswesen, insbesondere unter Anwendung der Blockchain-Technologie. Ziel der Arbeit ist es, durch die Entwicklung eines Entscheidungsmodells den Architektur- und Entwurfsprozess zu unterstützen, indem eine strukturierte Herangehensweise zur Identifikation und Bewertung von Designentscheidungen bereitgestellt wird. Die Erstellung des Entscheidungsmodells erfolgt auf Grundlage einer iterativ erarbeiteten Taxonomie, welche die wissenschaftlich veröffentlichten Systemarchitekturen anhand ihrer technischen Designmerkmale kategorisiert und somit deren Vergleichbarkeit sowie Differenzierung ermöglicht.

Aufgrund der unzureichenden Berücksichtigung methodischer Sicherheitsbetrachtungen in der wissenschaftlichen Literatur bei der Konzeption von Gesundheitsdateninfrastrukturen auf Grundlage der Blockchain werden zwei Anwendungsfälle entwickelt: Erstens eine Systemarchitektur für eine Dateninfrastruktur zur selbstsouveränen Vernetzung der Primärversorgung und zweitens ein Datentreuhandmodell zur Sekundärnutzung medizinischer Daten für Forschungszwecke. Dabei werden der jeweilige Systemkontext, die spezifischen Anforderungen sowie die beteiligten Akteure systematisch analysiert. Darüber hinaus erfolgt eine Sicherheitsanalyse unter Anwendung der STRIDE-Methode, um potentielle Bedrohungen zu identifizieren, Sicherheitsmechanismen abzuleiten und darauf basierend eine Systemarchitektur zu entwickeln.

Im Anschluss wird überprüft, welche Designentscheidungen durch das entwickelte Entscheidungsmodell für die Anforderungen der Anwendungsfälle vorgeschlagen werden und analysiert, inwieweit diese mit den klassisch entwickelten Architekturen übereinstimmen. Die Evaluation des Entscheidungsmodells erfolgt außerdem durch eine Befragung von Expert*innen. Basierend auf den Erfahrungen und Erkenntnissen erfolgt abschließend eine methodische Einbettung des Entscheidungsmodells in den Prozess zur Konzeption von Softwarearchitekturen.

Abstract

Decentralized data infrastructures based on blockchain technology are being discussed as promising solutions for the secure, self-sovereign management, storage, and scientific utilization of sensitive health data in research and practice. These infrastructures provide the foundation for leveraging health data in personalized care, thereby aiming to enable more effective healthcare services. However, the architecture of such systems is inherently complex, as evidenced by the slow implementation of the electronic health record in Germany. Despite the potential of these infrastructures for the primary and secondary use of medical data, significant technical, organizational, and legal challenges remain. Addressing these challenges systematically is crucial to ensuring such systems' long-term effectiveness, security, and acceptance.

This dissertation focuses on the systematic investigation and methodological support for the architecture of decentralized data infrastructures in healthcare, with a particular emphasis on blockchain technology. The thesis aims to support the architectural design process by developing a decision model that provides a structured approach for identifying and evaluating design decisions. The decision model is based on a taxonomy that categorizes existing system architectures by their technical design characteristics, thereby enabling their comparison and differentiation.

To address the insufficient consideration of methodological security aspects in the scientific literature, the thesis develops two use cases: a data infrastructure for self-sovereign data sharing in primary care and a data trustee for the secondary use of medical data in research. For each use case, the system context, specific requirements, and involved actors are systematically analyzed. Threat modeling is

conducted using the STRIDE method to identify potential threats, derive security mechanisms, and design representative system architectures.

Subsequently, the thesis examines the design decisions proposed by the decision model for the use cases and analyzes their alignment with traditionally developed architectures. The decision model is further evaluated through expert interviews. Finally, based on the insights gained, the decision model is methodologically embedded into the software architecture design process.

Danksagung

Die vorliegende Dissertation entstand im Rahmen meiner Tätigkeit als wissenschaftliche Mitarbeiterin am FZI Forschungszentrum Informatik und dem Karlsruher Institut für Technologie. An erster Stelle möchte ich Prof. Dr. rer. nat. Wilhelm Stork meinen tief empfundenen Dank aussprechen. Durch seine engagierte und zugleich unkonventionelle Betreuung hat er mir nicht nur den Raum für eigenständiges Arbeiten und Denken gegeben, sondern mich auch stets daran erinnert, die praktische Anwendbarkeit und den Nutzen meiner Forschung im Blick zu behalten.

Ein herzlicher Dank gilt meinen Kolleg*innen, die mich während meiner Promotionszeit begleitet und unterstützt haben. Sie waren nicht nur wertvolle Wegbegleiter, sondern auch eine inspirierende Quelle des Wissens und der Motivation. Besonders hervorheben möchte ich Dr.-Ing. Markus Schinle, Christoph Zimmermann, Dr.-Ing. Marc Schroth, Friedrich Gauger, Dr.-Ing. Lukas Kohout, Matthias Diehl, Gergely Biri und Rodger Burmeister, deren Unterstützung und Expertise von unschätzbarem Wert für meine fachliche und persönliche Weiterentwicklung waren.

Ebenso möchte ich den Projektpartnern in den von mir betreuten Forschungsprojekten danken. Ein besonderer Dank gilt Busra Bedir, Alexa Danelski, Dr. Raphael Dressle und Dr. Bernd Feige. Ein großer Dank geht auch an meine Studierenden, die durch ihre tatkräftige Unterstützung maßgeblich zum Fortschritt meiner Arbeit beigetragen haben. Besonders hervorheben möchte ich Philip Andris, Patrick Mehl und Sophie-Charlotte Perret.

Mein tiefster Dank gilt meiner Familie und meinen Freunden, die mir über die gesamte Promotionszeit hinweg zur Seite standen und mir stets Rückhalt gegeben

haben. Sie waren meine unerschütterliche Stütze. Einen besonderen Dank möchte ich meinem Freund Patrick aussprechen, dessen unermüdliche Unterstützung, liebevolle Ermutigung und stete Rückendeckung mir Kraft und Zuversicht geschenkt haben.

Karlsruhe, im Juni 2025 *Christina Erler*

In dieser Arbeit wurden nach Möglichkeit geschlechtsneutrale Bezeichnungen verwendet. Zur Wahrung des Leseflusses und aufgrund fachsprachlicher Begriffe wird jedoch teilweise die männliche Form genutzt. Dies ist nicht als geschlechterspezifische Wertung zu verstehen und spricht alle Geschlechtsidentitäten gleichermaßen an.

Inhaltsverzeichnis

Abkürzungen

Abkürzungen

ACL	Access Control List
AES	Advanced Encryption Standard
BÄK	Bundesärztekammer
BDSG	Bundesdatenschutzgesetz
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
BfDI	Bundesbeauftragten für den Datenschutz und die Informationsfreiheit
BSI	Bundesamts für Sicherheit in der Informationstechnik
BGB	Bürgerliches Gesetzbuch
BMBF	Bundesministerium für Bildung und Forschung
BMG	Bundesministerium für Gesundheit
BMWK	Bundesministerium für Wirtschaft und Klimaschutz
BPMN	Business Process Model and Notation
BZgA	Bundeszentrale für gesundheitliche Aufklärung
CP-ABE	Ciphertext-Policy Attribute-Based Encryption
DAC	Discretionary Access Control
DaTraV	Datentransparenzverordnung

DEMIS	Deutsches Elektronisches Melde- und Informationssystem für den Infektionsschutz
DGA	Data Governance Act
DICOM	Digital Imaging and Communications in Medicine
DigiG	Digital-Gesetz
DKG	Deutsche Krankenhausgesellschaft
DiGA	Digitale Gesundheitsanwendung
DiPA	Digitale Pflegeanwendung
DIZ	Datenintegrationszentrum
DLT	Distributed-Ledger-Technologie
DoS	Denial of Service
DSGVO	Datenschutz-Grundverordnung
DSRM	Design Science Research Methode
DTI	Decentralized Trusted Identity
DVG	Digitale-Versorgung-Gesetz
DZG	Deutsche Zentren der Gesundheitsforschung
eAA	Elektronische Arztakte
eFA	Elektronische Fallakte
eGA	Elektronische Gesundheitsakte
eGK	Elektronische Gesundheitskarte
eHBA	Elektronische Heilberufsausweis
EHDS	European Health Data Space
ePA	Elektronischen Patientenakte
EHR	Electronic Health Record
e-Rezept	Elektronisches Rezept
ENHIS	Estonian National Health Information System

EU	Europäische Union
FDPG	Deutsche Forschungsdatenportal für Gesundheit
FDZ-Gesundheit	Forschungsdatenzentrum Gesundheit
FHIR	Fast Healthcare Interoperability Resources
F&E	Forschung und Entwicklung
FoPraNet-BW	Forschungspraxennetz Baden-Württemberg
G-BA	Gemeinsame Bundesausschuss
GDNG	Gesundheitsdatennutzungsgesetz
gematik	Gesellschaft für Telematikanwendungen der Gesundheitskarte mbH
GKV	Gesetzliche Krankenversicherung
GKV-SV	GKV-Spitzenverband
GMG	GKV-Modernisierungsgesetz
HIPAA	Health Insurance Portability and Accountability Act
HIMSS	Healthcare Information and Management Systems Society
HL7	Health Level Seven
IAM	Identitäts- und Zugriffsmanagementsystem
ICD	International Statistical Classification of Diseases and Related Health Problems
ID	Direkte Identifikationsmerkmale
IDS	International Data Spaces
IDSA	International Data Spaces Association
IDS-RAM	IDS Reference Architecture Model
IdM	Identitätsmanagement

IGES	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IQTIG	Institut für Qualitätssicherung und Transparenz im Gesundheitswesen
k_{pr}	Öffentliche Schlüssel
k_{pub}	Private Schlüssel
KIM	Kommunikation im Medizinwesen
KI	Künstliche Intelligenz
KIS	Krankenhausinformationssystem
KBV	Kassenärztliche Bundesvereinigung
KV	Kassenärztliche Vereinigungen
KZBV	Kassenzahnärztliche Bundesvereinigung
KZV	Kassenzahnärztliche Vereinigung
LF Decentralized Trust	Linux Foundation Decentralized Trust
MAC	Mandatory Access Control
MII	Medizininformatik-Initiative
MSIBW	Ministerium für Soziales und Integration Baden-Württemberg
NFDI	Nationale Forschungsdateninfrastruktur
NDAC	Non-Discretionary Access Control
ONC	Office of the National Coordinator for Health Information Technology
Opt-In	Einwilligungsbasierte Datenverarbeitung
Opt-Out	Widerspruchslösung
PAD	Persönliches Authentifizierungsgerät
PBFT	Practical Byzantine Fault Tolerance
PEI	Paul-Ehrlich-Institut

PHR	Personal Health Record
PIN	Persönliche Identifikationsnummer
PKI	Public Key Infrastructure
PoA	Proof of Authority
PoS	Proof of Stake
PoW	Proof of Work
PRE	Proxy Re-Encryption
RBAC	Rollenbasierte Zugriffskontrolle
RBFT	Redundant Byzantine Fault Tolerance
RKI	Robert Koch-Institut
RuBAC	Regelbasierte Zugriffskontrolle
SAML	Security Assertion Markup Language
SK	geheimer Schlüssel
SMC-B	elektronische Praxis- oder Institutsausweis
SNOMED-CT	Systematized Nomenclature of Medicine-Clinical Terms
SSH	Secure Shell
SSI	Self-Sovereign Identity
SUS	System Usability Scale
TLS	Transport Layer Security
THS	Unabhängige Treuhandstelle der Universitätsmedizin Greifswald
TSVG	Terminservice- und Versorgungsgesetz
UAC	Use & Access Committee
VC	Verifiable Credentials
VPN	Virtuelles Privates Netzwerk

WANDA	Weiteren Anwendungen für den Datenaustausch in der TI
WHO	World Health Organization
ZKPs	Zero-Knowledge-Proofs

1 Einleitung

In der Einleitung werden zunächst die Motivation und Relevanz der Arbeit sowie die daraus abgeleiteten Zielsetzungen und Forschungsfragen dargelegt. Anschließend wird das wissenschaftliche Umfeld skizziert, bevor das methodische Vorgehen und die Struktur der Arbeit erläutert werden.

1.1 Motivation

Die Speicherung, Verwaltung und Nutzung von Gesundheitsdaten über Dateninfrastrukturen ist ein zentraler Bestandteil der digitalen Transformation im Gesundheitswesen [47]. Vor allem im Kontext personalisierter Medizin spielt die systematische Nutzung von repräsentativen Gesundheitsdaten eine entscheidende Rolle, um mittels maschinellem Lernen Muster und Zusammenhänge zu erkennen, welche für die Entwicklung spezifischer Therapie- und Behandlungsansätze verwendet werden können [48]. Infolgedessen erwartet die Europäische Kommission ein Einsparpotential von 5,5 Milliarden Euro durch den verbesserten Zugang und den Austausch von Gesundheitsdaten in der Europäischen Union (EU) über einen europäischen Gesundheitsdatenraum (*engl. European Health Data Space - EHDS*) [49]. Darüber hinaus wird erwartet, dass durch die effektive Nutzung von Gesundheitsdaten für Forschung und Innovationen weitere Einsparungen in Höhe von 5,4 Milliarden Euro erzielt werden können [49].

Wegen dieser Potentiale verfolgen mehrere europäische Länder, inklusive Deutschland, seit Jahren das Ziel ihre Digitalisierung im Gesundheitswesen durch Dateninfrastrukturen und die Bereitstellung einer elektronischen Patientenakte (ePA)

voran zu treiben [50]. Deutschland zeigt jedoch im internationalen Vergleich Defizite in der digitalen Transformation des Gesundheitswesens, was durch den vorletzten Platz von 17 untersuchten Ländern in einer Studie der Bertelsmann Stiftung aus 2018 belegt wird [51]. Bereits im Jahr 2004 wurde die ePA gesetzlich im Fünften Buch Sozialgesetzbuch (SGB V, § 291a) verankert [52]. Erst 2018 wurde mit der Einführung der ePA auf freiwilliger Basis ein erster praktischer Schritt unternommen, kombiniert mit der Verpflichtung zur Bereitstellung einer ePA durch Krankenkassen in 2021 [53, 54]. Jener Zeitplan wurde durch die Implementierung der Telematikinfrastruktur (TI) durch die gematik unterstützt, welche als technologische Basis für den sicheren Austausch von Gesundheitsdaten dient [55]. Allerdings zeigte sich durch die späte Realisierung der zugrundeliegenden technischen Infrastruktur, dass die Konzepte aus den frühen 2000er Jahren nicht mehr zeitgemäß sind [56]. Um diesen Herausforderungen zu begegnen, wurde eine Weiterentwicklung der TI zur sogenannten TI 2.0 bis 2025 angestrebt [56]. Aufgrund der bislang geringen Nutzung der ePA seit 2021, die lediglich von etwa 1 % der Versicherten genutzt wurde [57], wurde vorgesehen, die Nutzung der ePA verpflichtend für alle Versicherten ab Januar 2025 einzuführen [58]. Durch die Aufdeckung von Sicherheitslücken durch den Chaos Computer Club im Dezember 2024 erfolgt eine Pilotierung dieser erst in einer abgegrenzten Modellregion zum Januar 2025 [59]. Neben den bisherigen Sicherheitsbedenken zeigt eine Befragung des IGES Institut im Auftrag der gematik aus den Jahren 2022 und 2023, dass das Vertrauen in die Datensicherheit der TI aus Sicht der medizinischen Leistungserbringenden (u.a. Praxen, Krankenhäusern und Apotheken) ausbaufähig ist [60, 61]. Im Jahr 2023 gaben 48% der Arztpraxen, 55% der Apotheken, 70% der Krankenhäuser, 35% der psychotherapeutischen Praxen und 52% der Zahnarztpraxen an, Vertrauen in die Datensicherheit der TI zu haben [61]. Im Studienbericht 2024 wurde jene Kennzahl nicht erfasst [62].

Zusammenfassend sind bei den Patient*innen im März 2025 die politisch anvisierten digitalen Innovationen auf Basis der TI bislang noch nicht flächendeckend und in vollem Umfang nutzbar [50, 58]. Die bundesweite Einführung der ePA für alle ist ab Ende April 2025 geplant, mit einer produktiven und flächendeckenden Nutzung ab Oktober 2025 [63]. Die Funktionalitäten der TI 2.0 sowie

die Implementierung strukturierter und standardisierter Datenformate, die über die bisher verwendeten PDF-Dateien zur effektiven Nutzung hinausgehen, befinden sich ebenfalls im Verzug und werden teilweise erst ab 2026 oder zu einem noch späteren Zeitpunkt verfügbar sein [64, 65]. Infolgedessen bestehen nach aktuellem Stand weiterhin fragmentierte Gesundheitsdaten und heterogene Gesundheitsdienstleistungen, die ineffizient sind, zu Unsicherheiten im Alltag der Patient*innen führen und für die Forschung nur eingeschränkt nutzbar sind [66].

Zur Bewältigung jener Herausforderungen werden dezentrale Gesundheitsdatenmanagementanwendungen und Datentreuhandsysteme auf Basis der Blockchain-Technologie für eine sichere Verwaltung und Nutzung von Gesundheitsdaten gesehen, welche die Abhängigkeit von zentralen Instanzen überwinden, Vertrauen in Datensicherheit durch Integrität schaffen sowie eine Verbesserung der Souveränität von Individuen adressieren [47, 11, 67, 68]. Trotz der in der Wissenschaft diskutierten Potenziale dieser Technologien gibt es bislang nur wenige Untersuchungen, welche die Konzeption und Architektur solcher dezentralen Infrastrukturen unter systematischer Berücksichtigung von Sicherheitsaspekten durchführen (siehe Abschnitt 3.1) [11]. Vor diesem Hintergrund setzt die vorliegende Arbeit an, indem sie untersucht, inwiefern die Architektur von dezentralen Dateninfrastrukturen im Gesundheitswesen unterstützt werden kann, um die Visionen der TI 2.0 zu adressieren, die Souveränität der Datensubjekte zu fördern sowie eine effektive Nutzung von Gesundheitsdaten zu ermöglichen.

1.2 Zielsetzung und Forschungsfragen der Arbeit

Die vorliegende Arbeit verfolgt das Ziel, die Konzeption und Architektur dezentraler Dateninfrastrukturen im Gesundheitswesen systematisch zu untersuchen und methodisch zu unterstützen. Dies soll ermöglichen, die Potenziale der Blockchain-Technologie zu nutzen, um den Zugang zu Gesundheitsdaten sowie deren sicheren Austausch und die souveräne Verwaltung durch Individuen zu stärken, während gleichzeitig eine effektive und sichere Nutzung der Daten für Forschungszwecke

gewährleistet wird. Dabei liegt ein besonderer Fokus auf der Berücksichtigung von Sicherheitsanforderungen, regulatorischen Rahmenbedingungen und technischen Konzepten, um innovative Lösungen für das Gesundheitswesen zu entwickeln, welche die Vision einer TI 2.0 und einem europäischen Gesundheitsdatenraum unterstützen. Davon ausgehend lautet die zentrale Forschungsfrage der Arbeit:

Wie kann die Konzeption und Architektur dezentraler Infrastrukturen auf Basis von Blockchain-Technologien im Gesundheitswesen zur souveränen Verwaltung und effektiven Nutzung von Gesundheitsdaten unterstützt werden?

Zur Beantwortung dieser Leitfrage wurden verschiedene Untersuchungen durchgeführt, welche in die folgenden Teilfragen unterteilt werden können:

- **TFF1:** Welche Designentscheidungen müssen für die Architektur von dezentralen Dateninfrastrukturen im Gesundheitswesen unter Berücksichtigung der Blockchain-Technologie getroffen werden?
- **TFF2:** Welche (Sicherheits-)Anforderungen, regulatorischen Vorgaben und technischen Rahmenbedingungen bestehen an dezentrale Dateninfrastrukturen im Gesundheitswesen?
- **TFF3:** Welche methodischen Ansätze eignen sich zur Unterstützung der Architektur dezentraler Dateninfrastrukturen im Gesundheitswesen und wie können dabei Sicherheitsbetrachtungen methodisch eingebettet werden?

1.3 Wissenschaftliches Umfeld

Wesentliche Teile der vorliegenden Arbeit entstanden im Rahmen von Forschungsprojekten, die durch das Bundesministerium für Bildung und Forschung (BMBF), das Bundesministerium für Wirtschaft und Klimaschutz (BMWK) und das Ministerium für Soziales und Integration Baden-Württemberg (MSIBW) gefördert

wurden. Diese Projekte haben durch die aktive Forschungsbeteiligung der Autorin wesentliche Beiträge zu verschiedenen Aspekten der digitalen Transformation im Gesundheitswesen, der Blockchain-Technologie sowie dem sicheren Umgang mit Gesundheitsdaten geleistet. Die von der Autorin im Rahmen dieser Projekte gewonnenen Erkenntnisse und erzielten Ergebnisse flossen maßgeblich in die Erstellung dieser Arbeit ein. Sie wurden dabei durch eigene wissenschaftliche Leistungen ergänzt und in unabhängigen Publikationen weiter ausgearbeitet. Eine vollständige Übersicht aller wissenschaftlichen Veröffentlichungen ist in Kapitel A.3 zu finden. Im Folgenden werden die relevanten Projekte kurz vorgestellt.

Das Projekt **ODin – Organisationsübergreifender Informationsaustausch mittels Distributed Ledger Technology** (BMBF, Laufzeit: 01/2019 - 06/2020) beschäftigte sich mit der Entwicklung von Distributed Ledger Technologien für den organisationsübergreifenden Informationsaustausch am Beispiel von Supply-Chain-Prozessen [7, 6]. Zusätzlich wurden Ansätze zur Modellierung von Geschäftsprozessen und zum Prozessmonitoring auf Basis von Blockchain-Technologien untersucht und ein entsprechendes Framework entwickelt [5].

Im Rahmen des Projekts **BloG^{s} – Blockchain-basiertes Gesundheitsdatenmanagement für gesamtheitliche Gesundheitsprofile** (BMBF, 03/2020 – 08/2023) wurde ein Blockchain-basierter Ansatz zum Gesundheitsdatenmanagement mit dazugehörigen Geschäftsmodell ausgearbeitet und evaluiert [14, 9, 8]. In diesem Kontext entstanden wissenschaftliche Beiträge zu Entscheidungsmodellen und Bedrohungsanalysen zur Konzeption und Architektur von Blockchain-basierten Infrastrukturen im Gesundheitswesen [11, 13].

Neben dem reinen Datenmanagement erfordert die Verwaltung von Zugriffen die Nutzung digitaler Identitäten. Dies wurde im Rahmen des Projekts **SDIKA – Sichere Digitale Identitäten Karlsruhe** (BMWK, 09/2021 – 12/2024) adressiert, indem ein Adapter zur Verwaltung und Nutzung digitaler Identitäten für öffentliche Dienste erforscht wurde. Ein besonderer Schwerpunkt im Kontext der Arbeiten der Autorin lag dabei auf der Konzeption einer SSI-basierten Lösung für das zentrale Knochenmarkspenderegister sowie die Evaluation des Systems im Gesundheitswesen [21].

Das Projekt **SouveMed – Vertrauenswürdiges Datentreuhandmodell zur souveränen Verwaltung und effektiven Nutzung von medizinischen Daten in der Schlafforschung** (BMBF, 01/2022 – 12/2023) widmete sich der Entwicklung eines Datentreuhandmodells zur selbstsouveränen Verwaltung und Sekundärnutzung von schlafmedizinischen Daten. Der wissenschaftliche Beitrag der Autorin bestand in der Konzeption der Architektur des Datentreuhandsystems sowie in der Erweiterung des zuvor entwickelten Entscheidungsmodells um spezifische Aspekte der Sekundärdatennutzung [2, 18].

Abschließend befasste sich das Projekt **ROUTINE – Reallabor zum Transfer digitaler Gesundheitsanwendungen und KI ins Gesundheitswesen** (MSIBW, 10/2022-12/2024) mit dem Aufbau und der Etablierung eines Reallabors zur Entwicklung und Verwertung von KI sowie Gesundheitsanwendungen. Im Rahmen dieses Projekts wurden Translationshürden bei der Entwicklung und dem Transfer solcher Anwendungen systematisch durch Expert*inneninterviews analysiert und wissenschaftlich publiziert [17]. Eine der häufigsten identifizierten Hürden betraf den begrenzten Zugang zu Gesundheitsdaten aufgrund technischer, rechtlicher und organisatorischer Problemstellung für die Entwicklung von KI-Anwendungen. Vor diesem Hintergrund wurde das Konzept des Datentreuhandsystems speziell für den Kontext eines KI-Reallabors weiterentwickelt und ein entsprechendes Geschäftsmodell abgeleitet.

1.4 Aufbau und Methodik

Die im Rahmen dieser Dissertation durchgeführten Forschungsarbeiten orientieren sich an der Design Science Research Methode (DSRM) nach Peffers et al. [69], welche insbesondere im Bereich der Forschung von Informationssystemen eingesetzt wird. Durch einen iterativen Prozess bestehend aus Theorie und Praxis unterstützt die DSRM bei der Entwicklung von generalisierbaren IT-Artefakten. Dabei umfasst die DSRM nach Peffers et al. sechs Schritte: (1) Identifikation und Motivation des Problems, (2) Definition der Ziele für die Lösung, (3) Entwurf und Entwicklung der Lösungsartefakte, (4) Demonstration der Lösungsartefakte,

(5) Bewertung der Effektivität und Effizienz sowie (6) Kommunikation. Die vorliegende Arbeit orientiert sich an dieser Struktur und gliedert sich demgemäß. Zunächst wird ein praxisbezogenes Problem bei der Architektur von dezentralen Gesundheitsinfrastrukturen identifiziert (vgl. Abschnitt 1.1). Daraufhin wird, unter Berücksichtigung relevanter Forschungsfragen, ein Ziel formuliert, das auf die Verbesserung der Konzeption zukünftiger Gesundheitsinfrastrukturen abzielt (vgl. Abschnitt 1.2). Im dritten Schritt werden Lösungsartefakte basierend auf dem Stand der Wissenschaft und Technik (vgl. Kapitel 3) entworfen, welche in dieser Arbeit durch ein Entscheidungsmodell (vgl. Kapitel 4) sowie Systemkonzepte in den Anwendungsfällen (vgl. Kapitel 5) repräsentiert werden. Diese werden in Übereinstimmung mit Schritt vier der DSRM entsprechend in den zuvor genannten Kapiteln demonstriert und diskutiert. Anschließend erfolgt im fünften Schritt die Evaluation der Effektivität und Effizienz der Lösungsansätze. In Kapitel 6 wird im Zuge dessen untersucht, inwieweit das entwickelte Entscheidungsmodell durch Expert*inneninterviews anwendbar und praxisgeeignet ist (vgl. Abschnitt 6.2) sowie inwieweit es zur Ableitung praxisnaher Systemkonzepte genutzt werden kann (vgl. Abschnitt 6.1). Die Anwendung des Modells wird in diesem Zusammenhang demonstriert, indem die Anforderungen und Rahmenbedingungen aus den Anwendungsfällen herangezogen werden, um zu prüfen, ob das Entscheidungsmodell zu denselben Designentscheidungen führt wie die traditionelle, sicherheitsorientierte Vorgehensweise zur Systemarchitektur in Kapitel 5. Abschließend wird durch die vorliegende Arbeit sowie begleitende wissenschaftliche Publikationen (vgl. Abschnitt 1.3) der letzte Schritt der DSRM umgesetzt, nämlich die schriftliche Aufbereitung und Kommunikation der gewonnenen Erkenntnisse.

2 Grundlagen

Das Grundlagenkapitel dient dazu, die theoretischen und konzeptionellen Prinzipien der vorliegenden Arbeit darzulegen. Es bietet eine strukturierte Übersicht über die zentralen Begriffe, Modelle und methodischen Ansätze, die im weiteren Verlauf der Arbeit herangezogen werden. Ziel ist es, das Verständnis für den thematischen Kontext auf medizinischer, juristischer und technischer Ebene zu vertiefen. Zu diesem Zweck werden zunächst das Gesundheitswesen sowie der medizinische Kontext von Gesundheitsdateninfrastrukturen dargelegt, gefolgt von einer Analyse der juristischen Rahmenbedingungen. Abschließend werden die technischen Grundlagen für die Konzeption und Entwicklung von Blockchain-basierten Gesundheitsdateninfrastrukturen sowie Datentreuhandsystemen erläutert.

2.1 Gesundheitswesen und digitale Gesundheitsdateninfrastrukturen

Der folgende Abschnitt befasst sich mit einer detaillierten Betrachtung der Strukturen des deutschen Gesundheitswesens sowie der technischen und organisatorischen Grundlagen für bestehende Dateninfrastrukturen und Datentreuhandsysteme im Kontext des Gesundheitssektors.

2.1.1 Aufbau des deutschen Gesundheitssystems

Das selbstverwaltete deutsche Gesundheitssystem beruht auf drei zentralen Prinzipien: der Versicherungspflicht, dem Solidarprinzip und dem Sachleistungsprinzip. Die Versicherungspflicht besagt, dass alle in Deutschland wohnenden Bürger*innen verpflichtet sind in einer Krankenversicherung versichert zu sein. Ab einer Beitragsbemessungsgrenze (im Jahr 2024: 5.175 € pro Monat) haben Versicherte – mit Ausnahme von Beamt*innen und Selbstständigen – die Wahl, sich entweder privat oder gesetzlich zu versichern [70]. Die gesetzliche Krankenversicherung (GKV) finanziert sich aus einkommensabhängigen Beiträgen. Im Gegensatz dazu erfolgt die Finanzierung der privaten Krankenversicherung (PKV) durch Prämien, die nach dem Alter sowie den individuellen Risiken der Versicherten gestaffelt sind. Ein Wechsel zwischen der PKV und der GKV ist ab einem gewissen Alter nicht mehr möglich. Das Solidarprinzip besagt hinsichtlich der Beiträge für die GKV, dass Personen mit höherem Einkommen entsprechend höhere Beiträge leisten, wodurch finanzstärkere Versicherte finanziell schwächere Versicherte unterstützen. Leistungen der GKV werden nach dem Sachleistungsprinzip erstattet. Dies bedeutet, dass gesetzlich Versicherte bei einer medizinischen Behandlung grundsätzlich nicht in finanzielle Vorleistung treten müssen, abgesehen von Eigenbeteiligungen und Zuzahlungen. Die medizinischen Leistungserbringenden im Gesundheitswesen, wie beispielsweise Ärzt*innen, Kliniker*innen und Apotheker*innen, rechnen die erbrachten Therapien und Arzneimittel direkt mit den Krankenkassen über die Kassenärztlichen Vereinigungen (KV) ab. [71, 70]

Das heutige Gesundheitssystem gliedert sich in drei Ebenen: (1) die Mikroebene, (2) die Mesoebene und (3) die Makroebene [70]. Diese Ebenen und die zentralen Akteure des deutschen Gesundheitssytems werden im Folgenden kurz erläutert (siehe Abbildung 2.1).

Die **Mikroebene** stellt die Ebene der direkten Patient*innenversorgung dar. Hierbei haben Versicherte die freie Wahl für einen passenden Leistungserbringenden. Es wird zwischen der ambulanten und stationären Versorgung unterschieden. In der ambulanten Versorgung sind niedergelassene Haus-, Fach-, Zahnärzt*innen,

Psychotherapeut*innen sowie Fachkräfte aus nichtärztlichen Heilberufen, wie etwa Physiotherapeut*innen, Logopäd*innen und Ergotherapeut*innen sowie Pflegekräfte angesiedelt, welche häufig eine Kassenzulassung besitzen und damit Mitglied in einer kassenärztlichen Vereinigung (KV) oder kassenzahnärztlichen Vereinigung (KZV) sind. Die stationäre Versorgung erfolgt hingegen in Krankenhäusern, in denen Versicherte – entweder mit einer ärztlichen Einweisung oder aufgrund eines Notfalls – unabhängig von ihrer Versicherungsart behandelt werden. Neben der rein medizinischen Behandlung umfasst das Gesundheitssystem auch die Pflege, die im Pflegefall pflegerische Leistungen gemäß SGB XI übernimmt, sowie die Arzneimittelversorgung, die durch (Krankenhaus-)Apotheken gewährleistet wird. [70]

Die **Mesoebene** umfasst die Administration der Gesundheitsversorgung durch Verbände und Organisationen der Leistungserbringenden sowie Kostenträgerschaft. Jene koordinieren die Art, den Umfang und die Voraussetzungen für Leistungserbringungen. Hierbei sind diese an das Gebot gebunden, dass Leistungen wirtschaftlich, ausreichend, notwendig und zweckmäßig sein müssen und das Maß des Notwendigen nicht überschreiten dürfen (§12 SGB V). In Deutschland gibt es verschiedene Verbände und Organisationen, die die Gesundheitsversorgung regeln und Interessen vertreten. Die Bundesärztekammer (BÄK) überwacht die Qualifikation der Ärzt*innen und vertritt die Interessen des Berufsstands, ähnlich wie die Bundespsychotherapeutenkammer und die Bundesapothekerkammer für Psychotherapeut*innen und Apotheker*innen. Die Kassenärztliche Bundesvereinigung (KBV) ist die Dachorganisation für alle vertragsärztlichen und -psychotherapeutischen Ärzt*innen und Psychotherapeut*innen und koordiniert die ambulante Versorgung, während auf Landesebene spezifische kassenärztliche Vereinigungen existieren. Die Deutsche Krankenhausgesellschaft (DKG) vertritt als Dachverband die Interessen der Einrichtungen der stationären Versorgung und der GKV-Spitzenverband (GKV-SV) repräsentiert die Interessen der gesetzlichen Kranken- und Pflegekassen. Neben der Interessensvertretung gehört zu den Aufgaben des GKV-SV die Finanzierung und die Gestaltung der Rahmenbedingungen für die gesundheitliche und pflegerische Versorgung sowie die Ausgestaltung

des Datenmanagements und Telematik im Gesundheitswesen [72]. Der Deutsche Pflegerat vertritt die Pflege. [71, 70]

Eine zentrale Institution der **Makroebene** ist der Gemeinsame Bundesausschuss (G-BA), der die erstattungsfähigen Leistungen festlegt und die Inhalte der Gesundheitsversorgung durch untergesetzliche Regelungen definiert. Der gesetzliche Rahmen für die Aufgaben, Tätigkeiten und Verfahren der Leistungserbringung wird in Form von Sozialgesetzen durch Beschlüsse des Bundestages und die Zustimmung des Bundesrates festgelegt. Innerhalb dieses Rahmens agieren die Bundesländer mit eigenen Kompetenzen und Gesetzgebungsbefugnissen und übernehmen die Umsetzung der Bundesgesetze auf Landesebene. Die Landesgesundheitsministerien sind für die Planung und teilweise Finanzierung der stationären Einrichtungen sowie für die Aufsicht über den regionalen öffentlichen Gesundheitsdienst und regionale Krankenkassen zuständig. Die Zuständigkeit für die GKV obliegt der Bundesregierung, während die konkrete Führungsverantwortung für gesetzliche Kranken- und Sozialversicherung beim Bundesministerium für Gesundheit (BMG) liegt. Für die federführende Verantwortlichkeit der Gesundheitspolitik und damit die Ausarbeitung von Gesetzesvorhaben, Verordnungen oder Vorschriften ist ebenfalls das BMG verantwortlich. [71, 70]

Zudem übt das BMG die Rechtsaufsicht über die Selbstverwaltungsorganisation des Gesundheitswesens auf Bundesebene aus. Zu diesen Organisationen zählen unter anderem das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM), das Paul-Ehrlich-Institut als Bundesinstitut für Impfstoffe und biomedizinische Arzneimittel (PEI), das Robert Koch-Institut (RKI), die Bundeszentrale für gesundheitliche Aufklärung (BZgA), das Bundesversicherungsamt (BVA), das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) sowie das Institut für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG). Auch über den G-BA hat das BMG die Rechtsaufsicht inne, wobei der G-BA das oberste beschlussfassende Gremium der gemeinsamen Selbstverwaltung darstellt. Zu den Mitgliedern dieses Gremiums gehören der GKV-Spitzenverband (GKV-SV), die Deutsche Krankenhausgesellschaft (DKG), die Kassenärztliche Bundesvereinigung (KBV) und die Kassenzahnärztliche Bundesvereinigung (KZBV). Dieses Gremium ist dafür zuständig, Richtlinien zur

Ausgestaltung der Gesetze zu erarbeiten, die erstattbaren Behandlungsleistungen festzulegen und Kriterien für die Struktur- und Prozessqualität in der Krankenversorgung zu definieren. Vertreter*innen von Patientenorganisationen, die maßgeblich die Interessen und Rechte von Menschen mit Behinderungen oder Erkrankungen vertreten, haben zwar die Möglichkeit, Anträge in diesem Gremium zu stellen und an den Beratungen teilzunehmen, jedoch verfügen sie nicht über ein Stimmrecht. Darüber hinaus ist der G-BA für die Verwaltung des sogenannten Innovationsfonds verantwortlich, der Projekte zur Versorgungsforschung sowie die (Weiter-)Entwicklung evidenzbasierter Leitlinien unterstützt. Erfolgreich evaluierte Projekte aus diesem Fonds können für eine Integration in die Regelversorgung empfohlen werden. [71, 70]

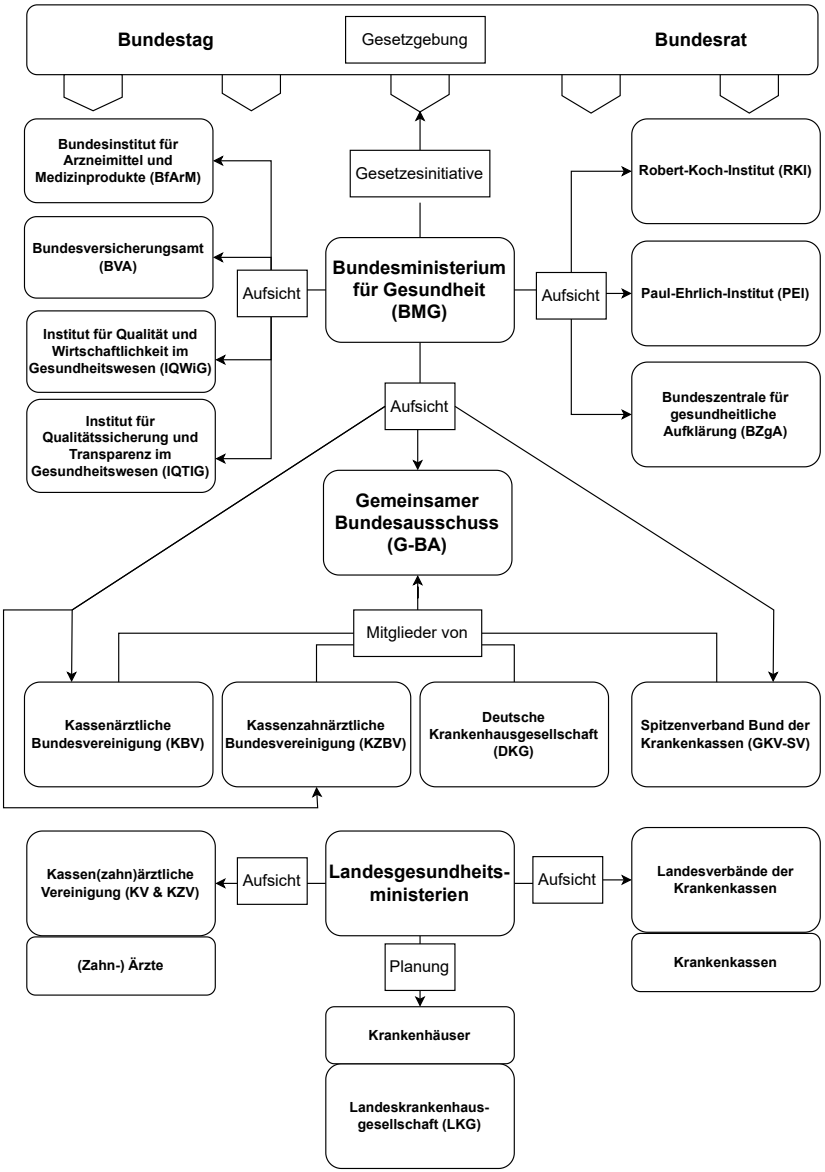

Abbildung 2.1: Aufbau und Akteure des deutschen Gesundheitssystems nach [70].

2.1.2 Die gematik: Schlüsselakteur der Digitalisierung im Gesundheitswesen

Die Gesellschaft für Telematikanwendungen der Gesundheitskarte mbH (gematik), gegründet im Jahr 2005, trägt als Schlüsselakteur die Verantwortung zum Aufbau einer digitalen Gesundheitsinfrastruktur in Deutschland [55, 56]. Sie verantwortet als Nationale Agentur für Digitale Medizin die Konzeption, die (Weiter-) Entwicklung und den Betrieb der zentralen Plattform für digitale Anwendungen im deutschen Gesundheitswesen, welche als Telematikinfrastruktur (TI) bezeichnet wird (§306, §291a und §291b SGB V). Die zentralen Aufgaben der gematik im Kontext der TI beinhaltet neben der technischen Ausgestaltung und dem Betrieb insbesondere die Definition und Durchsetzung verbindlicher Standards für Dienste, Komponenten und Anwendungen in der TI. Kernanwendungen der TI sind beispielsweise die elektronische Patientenakte (ePA), die Anwendung zur Kommunikation im Medizinwesen (KIM) und das elektronische Rezept (e-Rezept) [55, 56]. Neue Gesetze zur Digitalisierung des Gesundheitswesens, die vom Bundestag und Bundesrat verabschiedet werden, münden in der Zuweisung von Aufträgen zur Digitalisierung an die gematik. Im Falle eines Digitalisierungsauftrags erarbeitet die gematik die wesentlichen technischen und fachlichen Eckpunkte und stimmt diese in enger Abstimmung mit ihren Gesellschaftenden ab. Anschließend entwickelt die gematik die erforderlichen technischen Standards und führt Marktstudien mit potentiellen Anwendenden durch, um die Anwendungen entsprechend den Bedürfnissen der Anwendenden zu entwickeln. Nach der finalen Definition der technischen Standards durch die gematik entwickeln IT-Unternehmen die Produkte und Anwendungen, welche den festgelegten Standards gerecht werden. Die Erfüllung der Standards prüft die gematik abschließend und gibt die Produkte an die Markzulassung frei, wodurch die Produkte in der Versorgung angeschafft und verwendet werden können [73, 74].

Die Gesellschaftsstruktur der gematik setzt sich aus einer Vielzahl an zentralen Akteuren im Gesundheitswesen zusammen. Dazu gehören das Bundesministerium für Gesundheit (BMG), die Bundesärztekammer (BÄK), die Bundeszahnärztekammer (BZÄK), der Deutsche Apothekerverband (DAV), die Deutsche

Krankenhausgesellschaft (DKG), der Spitzenverband der Gesetzlichen Krankenversicherungen (GKV-SV), die Kassenärztliche Bundesvereinigung (KBV), die Kassenzahnärztliche Bundesvereinigung (KZBV) und der Verband der Privaten Krankenversicherung (PKV). Das BMG hält einen Anteil von 51% an der gematik, während 22,05% der Anteile dem GKV-SV zugeordnet sind. Die verbleibenden 24,5 % entfallen auf die restlichen Gesellschaftenden, welche die Spitzenorganisationen der Leistungserbringenden repräsentieren (vgl. Abbildung 2.2). [75]

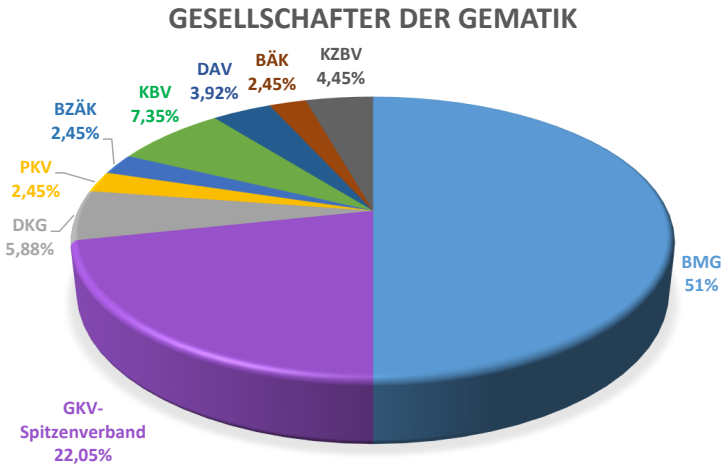

Abbildung 2.2: Prozentuale Verteilung der Gesellschafteranteile an der gematik [75].

2.1.3 Digitale Aktensysteme im Gesundheitswesen

Für Aktenkonzepte im deutschen Gesundheitswesen existieren in der Literatur diverse Bezeichnungen und Ausprägungen: Elektronische Arztakte (eAA), elektronische Gesundheitsakte (eGA), elektronische Patientenakte (ePA) und elektronische Fallakte (eFA) [76]. Die rechtlichen Grundlagen für die verschiedenen

Aktenkonzepte im deutschen Gesundheitswesen werden im Bürgerlichen Gesetzbuch (BGB) sowie im Sozialgesetzbuch (SGB) definiert (siehe Tabelle 2.1) [76]. Nachfolgend werden die für die vorliegende Arbeit relevanten Aktenkonzepte erläutert.

Grundlegend wird zwischen Patientenakten und Gesundheitsakten differenziert. Patientenakten werden von Leistungserbringenden geführt und dienen in erster Linie der Dokumentation von Behandlungen im Kontext einer spezifischen Erkrankung innerhalb einer medizinischen Einrichtung. Ursprünglich war die elektronische Patientenakte (ePA) als interne Akte konzipiert, welche die digitale Speicherung von dokumentationspflichtigen Befunden innerhalb einer einzelnen medizinischen Einrichtung ermöglichte. Dies entspricht dem Prinzip der elektronischen Arztakte (eAA), wie sie in § 630f BGB geregelt ist. Mit der zunehmenden Vernetzung von IT-Systemen und der Beteiligung verschiedener Akteure des Gesundheitswesens in einer medizinischen Behandlung wurde das Konzept zur einrichtungsübergreifenden elektronischen Patientenakte weiterentwickelt, bei der die Daten aus den Primärsystemen der jeweiligen medizinischen Einrichtung an andere relevante Einrichtungen und Akteure im Gesundheitswesen übermittelt werden. Im englischsprachigen Raum wird die ePA unter der Bezeichnung „electronic health record“ (EHR) geführt. [77]

Im Rahmen dieser Arbeit bezieht sich der Begriff ePA auf eine solche einrichtungsübergreifende Akte. Eine elektronische Gesundheitsakte (eGA) wird im Gegensatz zu einer ePA von den Patient*innen selbst verwaltet. Die Patient*innen speichern selbstständig Informationen und bestimmen, welche Personen oder Einrichtungen Zugriffsrechte auf die medizinischen Daten erhalten. Neben konkreten medizinischen Daten aus einer Behandlung, umfasst sie sogenannte Wellnessdaten, welche selbständig durch die Patient*innen erhoben werden (z.B. kontinuierlich gemessene Vitalparameter). Zudem ist für die Führung einer eGA nicht zwingend das Vorliegen einer Erkrankung erforderlich. Auf internationaler Ebene wird die eGA als „personal health record“ (PHR) bezeichnet. [77]

Akte	eAA	eGA	ePA
Rechts-grundlage	§ 630f BGB	§ 68 SGB V	§ 291a SGB
Autorität über Daten	Ärzt*innen-geführt	Patient*innen-geführt	Patient*innen-geführt
Interopera-bilität	einrichtungsintern	einrichtungsüber-greifend	einrichtungsüber-greifend
Speicher-dauer	potentiell lebens-lang	lebenslang	lebenslang

Tabelle 2.1: Übersicht über die relevanten Aktensysteme im deutschen Gesundheitswesen in Anlehnung an Kriedel [76].

2.1.4 Nationale Ansätze zu Dateninfrastrukturen und Treuhandsystemen im Gesundheitswesen

Aufgrund des rasanten technischen Fortschritts der letzten Jahre existieren verschiedene nationale Bemühungen, Projekte und Initiativen zur Erforschung und Etablierung von Dateninfrastrukturen sowie Treuhandsystemen im deutschen Gesundheitswesen. Im Folgenden werden die wesentlichen Ansätze in diesem Kontext vorgestellt.

Telematikinfrastruktur (TI)
Wie bereits in Abschnitt 2.1.2 dargelegt, wurde die gematik mit dem Aufbau der Telematikinfrastruktur (TI) in Deutschland betraut. Ziel der TI ist der Austausch von Patient*innendaten zwischen rund 73 Millionen Versicherten, 180.000 niedergelassenen Ärzt*innen und Zahnärzt*innen, 19.400 Apotheken, 1.940 Krankenhäusern sowie 109 Krankenkassen [74]. Im Folgenden wird die Architektur der TI im Detail vorgestellt.

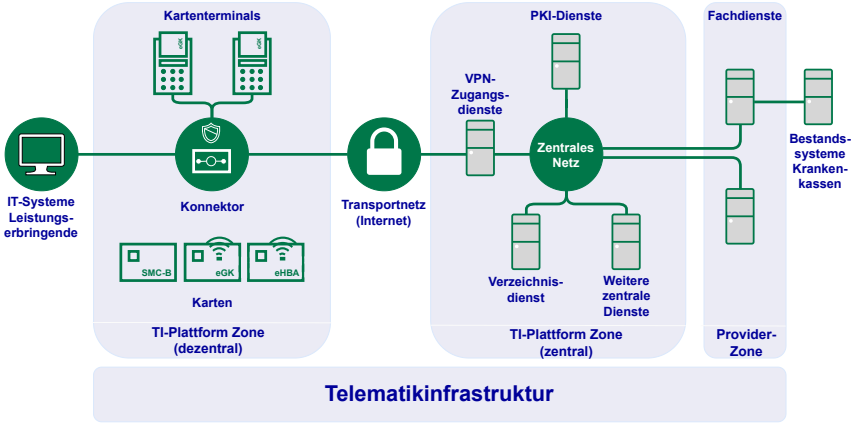

Abbildung 2.3: Übersicht über die Architektur der Telematikinfrastruktur [74].

In der Architektur der Telematikinfrastruktur (TI) wird grundsätzlich zwischen der TI-Plattform und den darauf aufbauenden Anwendungen differenziert. Die TI-Plattform stellt die zentrale Infrastruktur sowie übergreifende (Sicherheits-) Funktionen bereit, beispielsweise Authentisierungs- und Verschlüsselungsmechanismen sowie digitale Signaturen, welche von den Anwendungen genutzt werden können. Grundlegend ist die TI-Plattform in eine zentrale und eine dezentrale Zone unterteilt. Diese Zonen repräsentieren jeweils eine Teilmenge von Systemkomponenten und Diensten, welche untereinander Daten austauschen können. Ergänzend umfasst die TI neben den TI-Plattform-Zonen die sogenannte Provider-Zone [74]. Die Gesamtarchitektur mit den entsprechenden Zonen und Verbindungen zu den bestehenden IT-Systemen der Leistungserbringenden wird in Abbildung 2.3 dargestellt.

Die Teile der dezentralen TI-Plattform-Zone sind jene Komponenten und Dienste, welche dezentral in den jeweiligen medizinischen Einrichtungen (z.B. Krankenhäusern, Praxen) betrieben und genutzt werden. Dazu zählen unter anderem die Karten der an der TI Partizipierenden: (1) die elektronische Gesundheitskarte (eGK) der Versicherten, (2) der elektronische Heilberufsausweis (eHBA) sowie (3) der elektronische Praxis- oder Institutsausweis (SMC-B) der medizinischen

Leistungserbringenden. Die eGK speichert Versichertenstammdaten, optional aktivierbare Notfalldaten sowie die digitalen Schlüssel der Versicherten zur technischen Authentisierung der digitalen Identität. Im Gegensatz dazu speichert der eHBA und die SMC-B die digitalen Identitäten und damit die Berufsgruppenzugehörigkeit der Heilberufler*innen und Institutionen, welche durch Zugriffsrechte gemäß eines abgestuften Rechte- und Rollenkonzepts dazu berechtigt werden können, auf die Daten der eGK sowie der ePA zuzugreifen. Die auf den Karten gespeicherten Daten sind durch eine Persönliche Identifikationsnummer (PIN) geschützt. [74]

Darüber hinaus umfasst die dezentrale TI-Plattform-Zone die Kartenterminals und die Konnektoren mit dazugehörigen Gerätekarten, welche den Konnektoren eine eindeutige Identität zuordnen und für den Aufbau der geschützten Verbindungen innerhalb der TI genutzt werden. Der Konnektor ist die zentrale steuernde Komponente in den dezentralen Einrichtungen und bietet die Schnittstelle zu den bestehenden IT-Systemen, kontrolliert die Kommunikation mit den Kartenterminals und baut einen sicheren Kanal zur zentralen TI-Plattform über den VPN-Zugangsdienst der TI auf. Dieser Kanal ist ein Tunnel zu einem virtuellen privaten Netzwerk (VPN - *engl. Virtual Private Network*) über das Internet, bei dem die Daten auf Transportebene mittels Transport Layer Security (TLS) verschlüsselt werden und welcher im VPN-Zugangsdienst mündet. Ein VPN wird in diesem Zusammenhang als ein virtuelles privates Kommunikationsnetzwerk beschrieben, das ein bestehendes Kommunikationsnetzwerk, nämlich das Internet, verwendet, ohne klassisch eine direkte physische Verbindung zu besitzen [78]. Zusätzlich stellt der Konnektor eine Firewall sowie Basisfunktionalitäten und Sicherheitsfunktionen für die Anwendungen oder zur direkten Nutzung bereit. Auf diese Weise können Leistungserbringende direkt die Sicherheitsfunktionen verwenden, um Dokumente zu verschlüsseln oder digital zu signieren. Das Abrufen der Daten aus den bestehenden IT-Systemen ist durch die Firewall-Funktion ausschließlich mit aktiver Zustimmung des Leistungserbringenden möglich. Die Kartenterminals gewähren einen sicheren Zugriff auf die Karten und bauen eine geschützte Verbindung zum Konnektor auf, indem durch eine Transportverschlüsselung verhindert wird, dass Kartendaten von Dritten abgefangen oder manipuliert

werden. Für den ambulanten Bereich gibt es auch mobile Kartenterminals mit denen beispielsweise bei Hausbesuchen die Daten abgerufen werden können. Versicherungsstammdaten können bei den mobilen Terminals aber nicht aktualisiert werden. [74]

Die zentrale TI-Plattform-Zone umfasst die zentralen Dienste der TI-Plattform, nämlich: (1) die VPN-Zugangsdienste, (2) die Dienste der Public Key Infrastructure (PKI-Dienste), (3) den Verzeichnisdienst sowie (4) weitere zentrale Dienste. Bevor sich ein Konnektor mit dem VPN-Zugangsdienst verbinden kann, muss der Konnektor mittels Praxis- oder Institutsausweis registriert werden. Der VPN-Zugangsdienst wird über ein geschlossenes zentrales Netz mit den restlichen zentralen Diensten und Fachdiensten verbunden. Der Zugang zu diesem zentralen Netz ist ausschließlich über sichere Zugangspunkte möglich. Die Identitäten innerhalb der TI werden im PKI-Dienst verwaltet und bestehen aus einem asymmetrischen Schlüsselpaar, das einen öffentlichen Schlüssel und den entsprechenden privaten Schlüssel umfasst, sowie einem Zertifikat. Die zur Identifikation verwendeten Zertifikate werden von einer vertrauenswürdigen Stelle beglaubigt, besitzen eine begrenzte Gültigkeit und können bei Bedarf gesperrt werden. Der öffentliche Schlüssel ist gemäß dessen Definition für alle TI-Teilnehmenden bekannt, während der private Schlüssel geschützt auf der eGK, dem eHBA oder dem SMC-B gespeichert ist. Der Verzeichnisdienst enthält Email-Adressen und Verschlüsselungszertifikate von Leistungserbringenden und medizinischen Einrichtungen sowie Organisationen im Gesundheitswesen, wie beispielsweise Krankenkassen. Diese Zertifikate ermöglichen die Verschlüsselung von Informationen für die jeweiligen Akteur*innen. Zu den weiteren zentralen Diensten gehören Zeitdienste, die eine einheitliche Zeitbasis innerhalb der Telematikinfrastruktur (TI) sicherstellen, ein Konfigurationsdienst zur Aktualisierung von Software und Konfigurationen der dezentralen Komponenten sowie ein Namensdienst, der die Auffindbarkeit von Diensten ermöglicht. [74]

Die TI stellt versicherten Personen und Leistungserbringenden verschiedene Anwendungen mit mehreren fachanwendungsspezifischen Diensten, sogenannte Fachdienste, zur Verfügung, die der Provider-Zone zugeordnet sind. Diese können von Leistungserbringenden über den Konnektor und von Versicherten

mittels einer Client-Software genutzt werden. Beispiele für Anwendungen der TI sind das Versichertenstammdaten-Management, die ePA und das elektronische Rezept. Ferner gibt es Anwendungen, die ausschließlich im Konnektor genutzt werden und unabhängig von einem Fachdienst operieren. Diese werden als Fachmodule bezeichnet, wie beispielsweise das Notfalldaten-Management und der elektronische Medikationsplan. [74]

In Anbetracht dessen, dass die TI perspektivisch als die maßgebliche Infrastruktur im Gesundheitswesen verwendet werden soll, können auch Anwendungen von Drittanbietenden an die TI angebunden werden. Um als Drittanbietende an der TI teilzunehmen und die eigene Applikation über diese erreich- und nutzbar zu machen, muss diese durch die gematik als die „Weiteren Anwendungen für den Datenaustausch in der TI"(WANDA) bestätigt werden [79]. Hierfür müssen die Anbietenden nachweisen, dass ihre Anwendungen den geltenden Datenschutz- und Sicherheitsanforderungen entsprechen. Zudem sind sie verpflichtet, regelmäßig die kontinuierliche Einhaltung dieser Anforderungen zu belegen und Teil des Datenschutzmanagementsystems beziehungsweise des Managementsystems für Informationssicherheit innerhalb der TI zu sein. Ein konkretes Beispiel für eine bestätigte Anwendung im Rahmen von WANDA ist das DEMIS (Deutsches Elektronisches Melde- und Informationssystem für den Infektionsschutz) vom RKI. [74, 79]

Aufgrund der Tatsache, dass sich der Stand der Technik seit der Gründung und dem Start der Entwicklung der TI im Jahr 2005 weiterentwickelt hat, beispielsweise durch die Verbreitung von Smartphones und Cloud-Datendiensten sowie durch sich kontinuierlich ändernde IT-Sicherheitsanforderungen und erhöhte Anforderungen an Interoperabilität, Flexibilität und Skalierbarkeit, sind die zuvor beschriebenen initialen Designentscheidungen und Prämissen nicht mehr zeitgemäß. Daher wird eine iterative Weiterentwicklung der TI hin zur sogenannten TI 2.0 durch die gematik angestrebt, um den aktuellen (sicherheits-)technologischen Anforderungen gerecht zu werden (siehe Abbildung 2.4) [56].

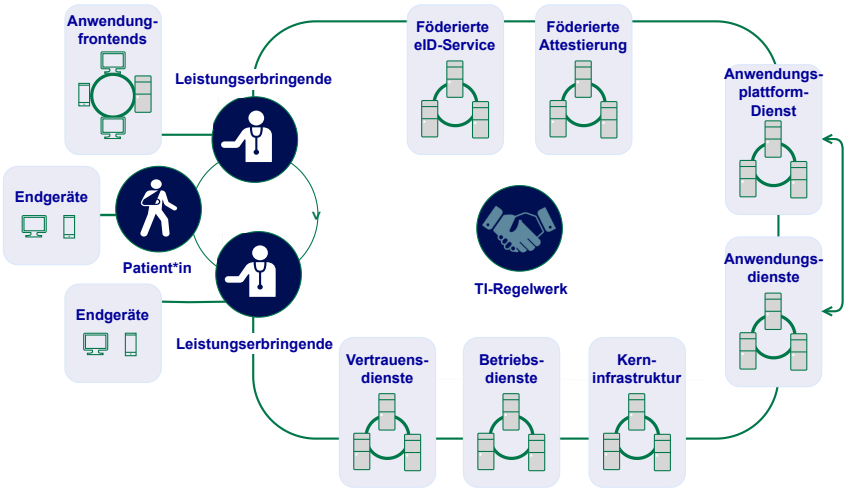

Abbildung 2.4: Übersicht über die Architektur der TI 2.0 (übersetzt aus [56]).

Zu den wesentlichen Neuerungen der TI 2.0 zählen ein konnektorunabhängiger Zugang zu den Diensten der TI, die den Einsatz physischer Konnektoren ersetzt und eine flexiblere, wartungsfreundlichere sowie skalierbarere Lösung bietet. Mit Hilfe von eigenen Endgeräten und Anwendungsfrontends sollen Versicherte und Leistungserbringende über das Internet einen direkten Zugang zu den Diensten erhalten. Zudem wird das geschlossene Netzwerk nicht mehr durch eine physikalische Netzwerkgrenze, sondern durch Kommunikationsbeziehungen auf der Anwendungsebene definiert. Dadurch stehen bei der Nutzung der TI die Anwendungsdienste (*engl. Application Services*) im Vordergrund, bei dem der Zugriff auf die TI-Dienste ausschließlich für geprüfte und authentifizierte Nutzendengruppen möglich ist, wobei der Zugang weiterhin streng kontrolliert bleibt. Teil der TI 2.0 ist hierfür die Kerninfrastruktur (*engl. Core Infrastructure*) bestehend aus wiederverwendbaren, anwendungsübergreifender Dienste in der betrieblichen Verantwortung der gematik. Die TI 2.0 setzt außerdem auf erweiterte Sicherheitsmechanismen, wie Zero-Trust-Networks, die eine Bereitstellung von Diensten durch unterschiedliche Anbietende ermöglichen, ohne dass diese sich in proprietäre Netzwerke integrieren müssen. Der Zugriff auf Dienste erfordert eine

beidseitige Authentisierung der Identitäten, die Attestation von Sicherheitseigenschaften sowie die Festlegung von Zugriffsrechten durch das Datensubjekt. Durch Betriebsdienste (*engl. Operation Services*) erfolgt beispielsweise das Monitoring der Dienste sowie der Sicherheit. Darüber hinaus wird die TI 2.0 über den Anwendungsplattform-Dienst (*engl. Application Platform Service*) mit einheitlichen Schnittstellen ausgestattet, die die Integration von Drittanbieteranwendungen erleichtern und eine verbesserte Flexibilität hinsichtlich übergreifender Anwendungsfälle bieten. Die Interoperabilität wird durch die verstärkte Nutzung internationaler einheitlicher Standards gefördert, um den Austausch von Gesundheitsdaten zwischen verschiedenen Akteuren, Institutionen und Organisationen im Gesundheitswesen zu erleichtern. Hierfür wurde der FHIR Standard als übergreifender Standard von der gematik festgelegt. Außerdem wird ein TI-Regelwerk (*engl. TI-Policy*) etabliert, dessen Einhaltung von der gematik durchgesetzt wird. Die Umsetzung dieses Regelwerks erfolgt durch die technischen Akteure unter Anwendung rechtlicher, organisatorischer und technischer Maßnahmen. Ferner wird ein föderiertes Identitätsmanagement angestrebt, um den Verwaltungsaufwand und Missbrauch von Identitäten zu minimieren. Hierdurch ist die Chipkarte nicht mehr das einzige Authentifizierungsmittel, sondern es können elektronische Identitätsnachweise verwendet werden. Um auf die Anwendungsdienste zuzugreifen, müssen sich die Nutzenden beim föderierten eID-Service (*engl. Federated eID*) des jeweiligen Sektors authentisieren, wobei Standorte und Endgeräte sowohl organisatorisch als auch technisch attestiert werden (Föderierte Attestierung - *engl. Federated Attestation*). Vertrauensdienste (*engl. Trust Services*), beispielsweise ein Dienstkatalog (*engl. Service Catalogue*), bildet die Einbindung der Akteure im angedachten föderierten System auf technischer Ebene ab. [56]

Elektronische Patientenakte (ePA)

In Deutschland haben alle gesetzlich Versicherten seit Anfang 2021 die Möglichkeit, eine elektronische Patientenakte (ePA) zu erhalten, die verpflichtend von den gesetzlichen Krankenkassen bereitgestellt wird [74]. Die ePA als Anwendung der TI wird durch den Versicherten über die ePA-App geführt. Mittels dieser können

Versicherte ihre Behandlungshistorie und -daten (z.B. Befunde, Diagnosen, Therapiemaßnahmen, Behandlungsberichte, Impfinformationen, Notfalldaten, Medikationsplan) durch die Einstellung von Leistungserbringenden einsehen und diese für weitere Leistungserbringende freigeben. Dabei entscheiden die Versicherten selbst, welche Daten eingestellt, gelöscht und eingesehen werden können. Ferner können Versicherte eigene Daten in die ePA einstellen (z.B. Blutdruckmessergebnisse). Die Daten der ePA werden in dem ePA-Fachdienst der TI, betrieben durch die Krankenkassen, gespeichert. Leistungserbringende greifen über den Konnektor auf die ePA zu. Zugriffsrechte werden durch den autorisierten Versicherten mittels der eGK oder eines alternativen Zugriffsverfahren an Leistungserbringende erteilt. Jene Berechtigungen werden im Fachdienst verwaltet, welcher die Zugriffsrechte der Leistungserbringenden bei einem Zugriff prüft. Versicherte können sich durch einen Zugehörigen vertreten lassen, falls dieser eine eGK besitzt. [74]

Nach Angaben des zum Beginn des Jahres 2025 amtierenden Bundesgesundheitsministers Karl Lauterbach nutzen Anfang 2024 nur weniger als ein Prozent der gesetzlich Versicherten in Deutschland die ePA, weshalb ab Januar 2025 die Verpflichtung zum Anlegen der neuen ePA eingeführt wurde, es sei denn der Versicherte widerspricht explizit (Opt-Out-Verfahren) [80, 81]. Zur Testung und weiteren Skalierung hat am 15. Januar die Pilotphase der neuen ePA in den Modellregionen der gematik (Hamburg, Hamburger Umfeld sowie Franken) und den KV-Regionen Westfalen-Lippe und Nordrhein gestartet [81]. Die Einführung der ePA für alle soll ab Ende April 2025 bundesweit erfolgen. Eine flächendeckende und produktive Nutzung wird ab Oktober 2025 angestrebt [63]. Die neue ePA bietet auch die Möglichkeit, Gesundheitsdaten im öffentlichen Interesse für Forschungszwecke bereitzustellen. Versicherte, die einer solchen Forschungsdatennutzung nicht zustimmen, haben die Option, die Verwendung ihrer Daten durch einen Widerspruch abzulehnen. Die Gesundheitsdaten werden im Aktensystem der ePA pseudonymisiert und dem Forschungsdatenzentrum Gesundheit (FDZ-Gesundheit) zur Verfügung gestellt (§ 363 SGB V). Das RKI als Vertrauensstelle sorgt dafür, dass keine Rückschlüsse auf die Identität der einzelnen Versicherten möglich sind. Zur Nutzung der Daten für Forschungszwecke ist ein Antrag beim FDZ-Gesundheit erforderlich. Das FDZ-Gesundheit überprüft die eingereichten

Anträge hinsichtlich der vorgesehenen Nutzungszwecke und entscheidet auf Basis gesetzlich definierter Kriterien über die Bereitstellung der Daten. Dabei werden ausschließlich aggregierte und anonymisierte Datensätze an die Antragstellenden weitergegeben [81]. Die Bereitstellung pseudonymisierter Abrechnungsdaten aller gesetzlich Krankenversicherten in Deutschland erfolgt aus der ePA bereits über das FDZ-Gesundheit [82]. Die Übermittlung weiterer Gesundheitsdaten an das FDZ-Gesundheit soll Mitte 2025 schrittweise beginnen [82].

Medizininformatik-Initiative (MII)

Die Medizininformatik-Initiative ist ein Förderprojekt des Bundesministeriums für Bildung und Forschung (BMBF), welches von 2016 bis 2027 gefördert wird und darüber hinaus betrieben werden soll [83]. Die MII zielt darauf ab, eine Infrastruktur für die standardisierte Integration klinischer Daten aus der Patient*innenversorgung und der medizinischen Forschung aufzubauen. Diese soll als Grundlage für die datenbasierte Medizin in Deutschland dienen. Zu diesem Zweck wurde in Zusammenarbeit mit der deutschen Universitätsmedizin eine dezentral föderierte Forschungsdateninfrastruktur aufgebaut (siehe Abbildung 2.5). Den Kern dieser Infrastruktur bildet die Errichtung von Datenintegrationszentren (DIZ) an den verschiedenen Universitätskliniken und medizinischen Standorten der derzeit vier Konsortien (DIFUTURE, HIGHmed, MIRACUM und SMITH), die Patient*innen behandeln und Gesundheitsdaten generieren. Diese DIZ fördern den einheitlichen Zugang zu den Gesundheitsdaten und Primärsystemen, indem sie standardisierte Formate und einheitliche Antragsverfahren im Sinne der MII unterstützen sowie für Datenqualität und Datensicherheit sorgen. [83]

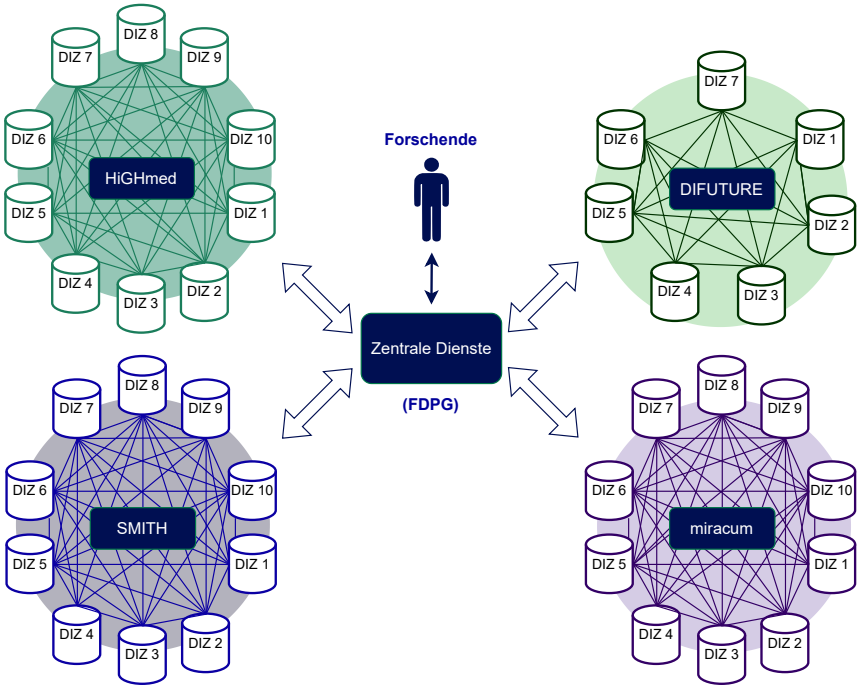

Abbildung 2.5: Aufbau der Forschungsdateninfrastruktur der MII [83].

Die zentrale Anlaufstelle für Forschende zur Erfassung, Beantragung und Bereitstellung von Daten aus den DIZ der MII ist das Deutsche Forschungsdatenportal für Gesundheit (FDPG) [84]. Der vom FDPG vorgegebene Prozess zur Beantragung und Nutzung von Daten durch Forschende [85] wird nachfolgend anhand von Abbildung 2.6 detailliert beschrieben. Forschende mit wissenschaftlicher Tätigkeit und konkreter Forschungsfrage können sich zunächst im FDPG registrieren, indem sie beispielsweise ihre Zugehörigkeit zu einer Forschungsgruppe nachweisen [86]. Nach der Freischaltung im FDPG haben Forschende die Möglichkeit, mittels einer Machbarkeitsanfrage zu prüfen, ob eine ausreichende Datenmenge zur Machbarkeit der Forschungsfrage vorhanden ist. Dazu gibt der Forschende in der Machbarkeitsanfrage spezifische Ein- und Ausschlusskriterien an und

erhalten als Rückmeldung die Anzahl der in den DIZ vorhandenen Daten, die den angegebenen Kriterien entsprechen [87]. Im Anschluss kann der Forschende einen elektronischen Datennutzungsantrag stellen, welcher von den betroffenen Standorten und deren Use & Access Committee (UAC) geprüft wird. Bei positiver Bewertung wird ein Nutzungsvertrag abgeschlossen, welcher zur Datennutzung berechtigt [88]. Die Datennutzung ist über zentrale oder verteilte Analysen möglich [89]. Bei zentralen Analysen erhalten die Forschenden pseudonymisierte Patient*innendaten direkt über FDPG, vorausgesetzt, die betroffenen Patient*innen haben zuvor ihre Einwilligung über einen Broad Consent erteilt. Die vollständige Implementierung erfolgt schrittweise bis 2034 und umfasst zentrale Meilensteine, darunter das Inkrafttreten der Regeln für die Primär- und Sekundärnutzung vorrangiger Kategorien von Gesundheitsdaten wie Patient*innenkurzakten und elektronische Verschreibungen/Verabreichungen (2029), die sukzessive Erweiterung der unterstützten Datentypen durch beispielsweise medizinische Bildgebung (2031) sowie die Möglichkeit für Drittstaaten und internationale Organisationen, sich dem System anzuschließen (2034). erteilt. Diese Daten werden von den DIZ extrahiert, an eine zentrale Instanz übermittelt und dort zu einem Gesamtdatensatz zusammengeführt. Neben der Qualitätssicherung und Archivierung ermöglicht die zentrale Instanz den Download der pseudonymisierten Daten, wodurch Forschende mit großer Flexibilität und Geschwindigkeit arbeiten können. Allerdings ist die verfügbare Fallzahl aufgrund der Notwendigkeit einer Einwilligung häufig geringer als bei verteilten Analysen [89]. Bei verteilten Analysen verbleiben die Daten an den jeweiligen Standorten, wo sie anhand eines von den Forschenden bereitgestellten Analyseskripts verarbeitet werden. Die Ergebnisse sind vollständig anonymisiert und werden nach Prüfung des Personenbezugs zentral zusammengeführt. Dieses Verfahren basiert auf alternativen Rechtsgrundlagen, wodurch größere Fallzahlen einbezogen werden können. Perspektivisch sollen Technologien wie DataSHIELD genutzt werden, um die Sicherheit und Effizienz verteilter Analysen weiter zu optimieren [89].

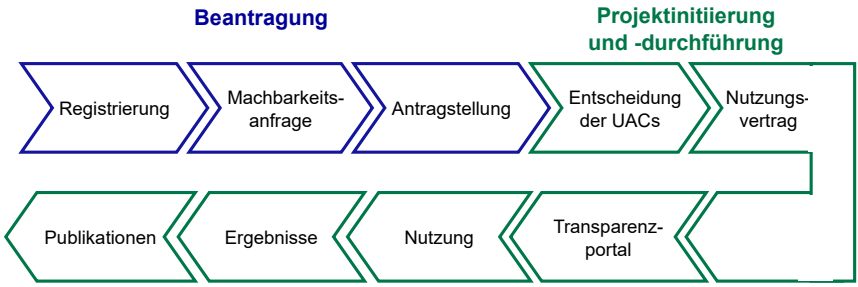

Abbildung 2.6: Abläufe zur Beantragung und Nutzung von Daten über das Forschungsdatenportal in Anlehnung an [85].

Aktuell befindet sich die MII in der Ausbau- und Erweiterungsphase (2023-2026), in der die erfolgreich demonstrierten Lösungen aus der Aufbau- und Vernetzungsphase (2018-2022) auf weitere Universitätskliniken und Forschungseinrichtungen übertragen werden [83]. Seit 2021 sollen hierzu neben den universitären Einrichtungen auch weitere Einrichtungen der ambulanten Versorgung und private Kliniken durch sechs sogenannte digitale FortschrittsHubs involviert werden [90]. Unter anderem die Muster-Einverständniserklärungen auf Basis einer breiten Einwilligung für Patient*innen [91] und der Kerndatensatz basierend auf internationalen IT- und Terminologie Standards (HL7 FHIR, SNOMED CT) [92] legen den Grundstein für zukünftige Forschungsdateninfrastrukturen.

Nationale Forschungsdateninfrastruktur (NFDI)

Die Nationale Forschungsdateninfrastruktur (NFDI) soll die aktuell dezentralen Datenbestände von Wissenschaft und Forschung systematisch erschließen, nachhaltig sichern, zugänglich machen und hierdurch eine (inter-) nationale Vernetzung ermöglichen [93]. Zu diesem Zweck wurde der Verein Nationale Forschungsdateninfrastruktur (NFDI) e.V. 2020 durch den Bund und die Bundesländer gegründet, welcher zum Ziel hat eine zukünftige Forschungsdateninfrastruktur in Deutschland zu gestalten. Unter dem Dach des Vereins sollen sich 30 Konsortien in unterschiedlichen Anwendungsfällen etablieren und konsortienübergreifend vernetzten, welche die Breite der Wissenschaftslandschaft in Deutschland repräsentieren. Durch den konsortienübergreifenden Austausch zu Querschnittthemen,

welche alle Konsortien betreffen, sollen beispielsweise gemeinsame Standards entwickelt werden. [94]

Im Rahmen der Förderung zur Bildung von Konsortien gibt es drei Initiativen in der medizinischen Domäne: NFDI4MED, welche die Deutschen Zentren der Gesundheitsforschung (DZG) mit den DIZ der MII verbindet, NFDI4Health mit epidemiologischem Schwerpunkt sowie GHGA mit der Erschließung und Zurverfügungstellung molekularer (Omics-) Daten [95].

Unabhängige Treuhandstelle der Universitätsmedizin Greifswald (THS)
Zur Trennung von Forschungsdaten und identifizierenden Daten haben sich Modelle etabliert, die auf einer klaren Verteilung der juristischen Rollen basieren. Auf dieser Grundlage wurde im Jahr 2024 die Unabhängige Treuhandstelle der Universitätsmedizin Greifswald (THS) gegründet, welche als Datentreuhand sowohl von der datengenerierenden Einrichtung als auch von nachgeordneten Forschungseinrichtungen unabhängig ist [96]. Die THS unterstützt die datenschutzkonforme Verwaltung von personenbezogenen Daten, Einwilligungserklärungen und der Pseudonymisierung, indem sie spezialisierte Softwarelösungen (MOSAIC Tools: E-PIX, gPAS und gICS) zum Schutz jener Daten bereitstellt. Zu ihren Aufgaben gehören die Führung von Patient*innenlisten mit eindeutiger Identifizierung, Dopplungsausschluss und Record Linkage, die Verwahrung und Prüfung von Einverständniserklärungen von Patient*innen und Proband*innen sowie die Pseudonymisierung und Anonymisierung von personenbezogenen Daten, etwa für die Speicherung im Forschungskontext oder die Weitergabe an Dritte. Außerdem erfolgt der Abgleich von Personendaten mit externen Datenquellen, wie etwa Melderegistern. Die THS unterstützt mit ihren Lösungen deutschlandweit und international weitere Forschungsprojekte und Studien, wie etwa die NAKO Gesundheitsstudie. [97] Die NAKO Gesundheitsstudie ist eine Langzeitbevölkerungsstudie durchgeführt von einem Netzwerk deutscher Forschungseinrichtungen deren Ziel die Identifizierung von Ursachen bei der Entstehung von Volkskrankheiten wie Diabetes oder Krebs ist [98].

CenTrust
CenTrust ist die Datentreuhand-Plattform der Bundesdruckerei GmbH, die als

unabhängige Vertrauensinstanz, die Daten zwischen Datengebenden und Datennutzenden sicher und gesetzeskonform vermittelt [99]. CenTrust positioniert sich als vertrauenswürdiger Akteur, der als Sicherheitsunternehmen des Bundes und integraler Bestandteil der Bundesdruckerei GmbH höchsten Sicherheits- und Vertrauensstandards entspricht [99]. Die zentrale Anwendung von CenTrust ist der sogenannte Vertrauensstellendienst, welcher on-demand Daten aus unterschiedlichen Quellen verknüpft und pseudonymisiert. Der Vertrauensstellendienst übermittelt die pseudonymisierten Daten anschließend an die berechtigte Stelle, die diese gemäß den vereinbarten Vorgaben nutzen oder weitergeben darf [99]. Durch dieses Berechtigungs- und Zugriffsmanagement erhalten die Einrichtungen ausschließlich die für ihren Zweck erforderlichen Informationen [100]. Eine Zuordnung zwischen der Identität des Datengebenden und dem Pseudonym kann nur der Vertrauensstellendienst herstellen. Die Bundesdruckerei unterstützt mit ihrem Vertrauensstellendienst verschiedene medizinische Forschungsprojekte unter anderem das Robert-Koch-Institut (RKI) für die Übermittlung medizinischer Daten von HIV-Patient*innen aus verschiedenen Kliniken zu Forschungszwecken. Zusätzlich unterstützen sie das Multiple Sklerose Register mit ihrem Dienst und das RKI beim Digitalen Impfquoten-Monitoring zur COVID-19-Impfung [99].

Forschungspraxennetz Baden-Württemberg (FoPraNet-BW)
Im Zeitraum von Februar 2020 bis Januar 2025 wurde der Aufbau des Forschungspraxennetz Baden-Württemberg (FoPraNet-BW) durch das BMBF gefördert, welches eine Forschungsinfrastruktur in Baden-Württemberg mit qualifizierten hausärztlichen Forschungspraxen und allgemeinmedizinischen Einrichtungen zum Ziel hatte [101]. Im Rahmen der studienbasierten Infrastruktur erfolgt zunächst ein „Case Finding“ bei dem potentiell relevante Betroffene identifiziert werden und anschließend eine Kontaktaufnahme zur Studienteilnahme erfolgt, bei dem eine informierte Einwilligung zur pseudonymen Datennutzung erteilt werden kann. Nach erfolgter Einwilligung und anschließenden Pseudonymisierung der Daten in der Praxis werden die pseudonymen Daten an einen zentralen Speicher übermittelt, wo diese durch die Studienmitarbeitenden analysiert werden können. [102]

2.1.5 Internationale Initiativen zu Gesundheitsdateninfrastrukturen und Treuhandsystemen

Neben Deutschland bemühen sich diverse Länder weltweit um die Digitalisierung ihres Gesundheitswesens, unter anderem durch die Einführung von ePA-Systemen und den Aufbau von Datentreuhandsystemen. Die Europäische Union (EU) strebt dabei insbesondere die Schaffung eines gemeinsamen Gesundheitsdatenraums an, der einen sicheren und rechtlich konformen grenzüberschreitenden Datenaustausch innerhalb Europas ermöglichen soll [103]. Gemäß einer Studie der Bertelsmann Stiftung aus dem Jahr 2018 zur digitalen Transformation nationaler Gesundheitssysteme schneidet Deutschland im internationalen Vergleich schlecht ab und landet auf Rang 16 von 17 untersuchten OECD-Ländern [51]. Die Studie identifiziert Estland, Kanada, Dänemark, Israel und Spanien als Spitzenreiter in der digitalen Gesundheitsversorgung [51]. Nachstehend werden ausgewählte internationale Initiativen und Ansätze vorgestellt, die in diesem Kontext von Bedeutung sind. Ein besonderes Augenmerk wird bei den ePA-Systemen auf die estnische Lösung gelegt, die im internationalen Vergleich entsprechend der Studie der Bertelsmann Stiftung herausragt und zusätzlich die Blockchain-Technologie anwendet.

European Health Data Space (EHDS)

Der European Health Data Space (EHDS) ist eine europäische Initiative für die gemeinsame Nutzung von Gesundheitsdaten zwischen nationalen Gesundheitssystemen in der europäischen Union sowohl für die Primärnutzung im Versorgungskontext als auch die Sekundärnutzung für Forschungszwecke im öffentlichen Interesse. Auf diese Weise soll im Allgemeinen die Versorgungen der einzelnen Gesundheitssysteme der EU verbessert und gestärkt werden. Patient*innen profitieren von einem verbesserten Zugang zu ihren Gesundheitsdaten, starken Datenschutzmaßnahmen und mehr Kontrolle über deren Nutzung. Gesundheitsfachkräfte erhalten effizienteren Zugriff auf Patient*innenakten, während Forschende und Unternehmen Zugang zu hochwertigen, pseudonymisierten Daten für Innovationen und wissenschaftliche Zwecke erhalten. Die Sekundärnutzung

soll nach dem Opt-Out-Prinzip erfolgen. Regulierungsbehörden und politische Entscheidungsträger*innen profitieren von besseren Daten für öffentliche Gesundheitsüberwachung und politische Maßnahmen. Gleichzeitig gewährleisten strikte Datenschutz- und Sicherheitsvorkehrungen, basierend auf bestehenden EU-Regelungen, die sichere Verarbeitung sensibler Gesundheitsdaten. Um dieses Ziel zu erreichen, wird ein Governance-Rahmen etabliert, der Regeln, Normen, Praktiken und Infrastrukturen für das europäische Datenökosystem definiert. Beispielsweise soll der Datenaustausch basierend auf einem europäischen Standard, dem European Electronic Health Record exchange Format (EEHRxF), erfolgen [104]. Am 5. März 2025 wurde die Verordnung über den EHDS durch die EU veröffentlicht, welcher am 26. März 2025 in Kraft getreten ist, womit die Übergangsphase zur Umsetzung der Durchführungsrechtakte beginnt [105]. Die vollständige Implementierung erfolgt schrittweise bis 2034 und umfasst zentrale Meilensteine, darunter das Inkrafttreten der Regeln für die Primär- und Sekundärnutzung vorrangiger Kategorien von Gesundheitsdaten wie Patient*innenkurzakten und elektronische Verschreibungen/Verabreichungen (2029), die sukzessive Erweiterung der unterstützten Datentypen durch beispielsweise medizinische Bildgebung (2031) sowie die Möglichkeit für Drittstaaten und internationale Organisationen, sich dem System anzuschließen (2034). [105]

GAIA-X

Das 2019 gestartete Projekt GAIA-X hat das Ziel, eine nachhaltige und innovative Datenwirtschaft in Europa zu fördern und digitale Souveränität zu gewährleisten [106]. GAIA-X strebt dazu die Schaffung eines vertrauenswürdigen und föderierten Datenökosystems in Europa an, in dem Daten sicher und selbstsouverän durch die Dateninhabenden mit Datennutzenden geteilt werden können [107]. Die Mission von GAIA-X besteht darin, Standards für das angedachte föderierte Datenökosystem zu etablieren. Zu diesem Zweck erfolgt im Rahmen des Projekts die Entwicklung von gesamteuropäischen Spezifikationen, Richtlinien, Regeln und einem Verifizierungsrahmen, welche die Grundlage für eine interoperable und souveräne Datenverarbeitung schaffen [106]. Im Jahr 2022 begannen elf Konsortien in Deutschland im Rahmen eines Förderwettbewerbs des BMWKs mit der

konkreten Umsetzung verschiedener Anwendungsfälle (u.a. Gesundheit, Finanzen, Mobilität und Landwirtschaft), sowie der ersten Inbetriebnahme von Diensten und Datenräumen, wie beispielsweise dem Mobility Data Space. Zwei der Projekte im Gesundheitswesen sind die Projekte Health-X dataLoft und TEAM-X [107].

International Data Spaces (IDS)
Die International Data Spaces Association (IDSA) ist eine Non-Profit-Organisation und steht hinter der Entwicklung des International Data Spaces (IDS), welcher ein zentrales Element der GAIA-X Architektur darstellt [108, 109]. Der IDS ist ein virtueller Datenraum, der bestehende Standards und Technologien nutzt, um einen sicheren und standardisierten Datenaustausch in einem vertrauenswürdigen Daten-Ökosystem zu erreichen [108].

Die Akteure des IDS umfassen Dateninhabende (*engl. Data Owner*), Datenbereitstellende (*engl. Data Provider*), Datenkonsumierende (*engl. Data Consumer*) und Datennutzende (*engl. Data User*) [110]. Der Begriff der Dateninhabenden wird durch Otto et al. nicht im rechtlichen Sinne verstanden, sondern aus einer Managementperspektive. Dateninhabende sind laut Otto et al. juristische oder natürliche Personen, die Daten erstellen und/oder über sie Kontrolle ausüben, wodurch sie in der Lage sind, Nutzungsrichtlinien sowie Zahlungsmodelle zu definieren und den Zugang zu den Daten festzulegen. Gemäß Otto et al. übernehmen Datenbereitstellende die Aufgabe, Daten für den Austausch zwischen Dateninhabenden und Datenkonsumierenden bereitzustellen. Typischerweise übernimmt laut Otto et al. der Dateninhabende auch die Rolle des Datenbereitstellenden, wobei es in bestimmten Fällen, wie etwa bei der Nutzung externer Cloud-Dienste oder Rechenzentren, auch möglich ist, dass Datenanbietende und Dateninhabende nicht identisch sind. Ein ähnliches Bild zeigt sich bei den Rollen der Datenkonsumierenden und Datennutzenden, welche häufig durch die gleiche Entität repräsentiert werden, jedoch nicht zwangsläufig identisch sein müssen. Beispielsweise kann der Betreibende einer ePA-Anwendung als Datenkonsumierende fungieren, während Patient*innen, welche die ePA-Anwendung zur Verwaltung ihrer Gesundheitsdaten nutzen, als Datennutzende gelten. [110, 109]

Die grundlegende Architektur und die zugrundeliegenden Konzepte des IDS werden im IDS Reference Architecture Model (IDS-RAM) von Otto et al. beschrieben (siehe Abbildung 2.7) [110, 109] und im Nachfolgenden genauer erläutert. Der IDS ermöglicht den Austausch von Daten zwischen Akteur*innen, ohne eine dazwischengeschaltete zentrale Infrastruktur, durch den Einsatz von Peer-to-Peer Konnektoren (*engl. Connector*). Diese Konnektoren gewährleisten einen direkten Datenaustausch zwischen den Datenquellen der Datenbereitstellenden und den Datensenken der Datenkonsumierenden, ohne dass eine externe Speicherung oder Verarbeitung der Daten durch Dritte erfolgt. Um Daten Dritten bereitstellen zu können, stellt der Datenbereitstellende Metadaten über den sogenannten Broker bereit. Diese Metadaten enthalten beispielsweise Informationen darüber, über welchen Konnektor die Daten verfügbar sind, sowie die vom Dateninhabenden definierten Nutzungsrichtlinien. Anschließend können Datenkonsumierende über ihren Konnektor eine Anfrage an den Konnektor des Datenbereitstellenden stellen, um direkt Zugriff auf die Daten zu erhalten. Die Übertragung der Daten erfolgt dabei nur, sofern die vom Dateninhabenden festgelegten Nutzungsrichtlinien mit den Nutzungsabsichten der Datenkonsumierenden übereinstimmen. In diesem Fall protokolliert der Konnektor die Datentransaktion und sendet den Datensatz an das sogenannte Clearing House. Für das Clearing House kann laut Otto et al. die Nutzung der Blockchain-Technologie in Betracht gezogen werden [109]. Der App-Store stellt Data-Apps bereit, die den Konnektoren hinzugefügt werden können und die ausgetauschten Daten weiterverarbeiten. Zur Annotation und Beschreibung von Datensätzen werden spezifische Vokabulare vom sogenannten Vocabulary Provider bereitgestellt. Eine weitere Schlüsselkomponente ist der Identity Provider, der die Identitätsinformationen der Akteur*innen im IDS erstellt, verwaltet, überprüft und validiert. [110, 109]

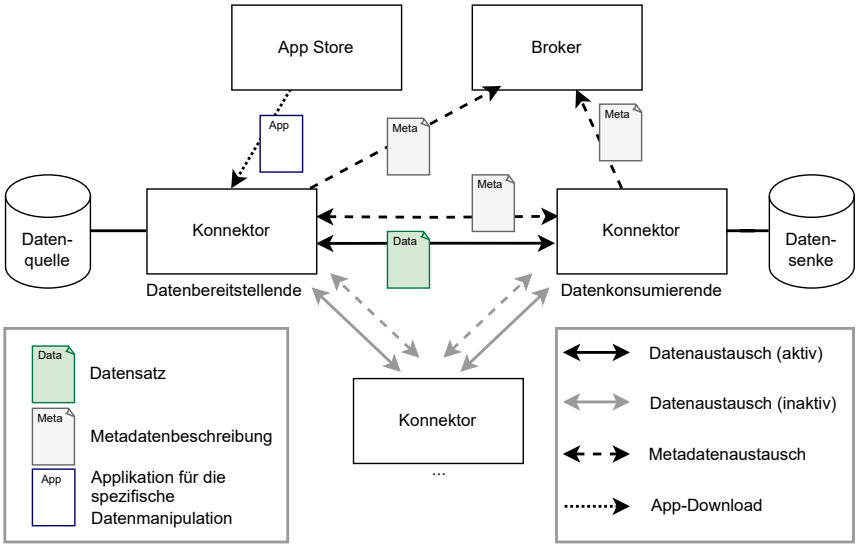

Abbildung 2.7: Übersicht der technischen Komponenten des IDS-RAM und deren Interaktionsweise (übersetzt aus [110]).

Estland

X-Road ist eine in Estland entwickelte verteilte Informationsaustauschplattform, die seit 2001 den sicheren und effizienten Datenaustausch zwischen verschiedenen Informationssystemen staatlicher Organisationen ermöglicht. Sie gewährleistet die Interoperabilität dieser Systeme, schützt die Daten während der Übertragung vor unbefugtem Zugriff und Manipulation und stellt sicher, dass nur autorisierte Personen Zugriff auf die Daten erhalten [111]. Eines an die X-Road-Infrastruktur angebundenes System ist das Estonian National Health Information System (ENHIS), eine seit 2008 betriebene zentrale Plattform, welche die estnischen Gesundheitsdienstleistenden vernetzt und über 40 Millionen medizinische Dokumente speichert [112]. Ergänzt wird dieses durch die Integration genomischer Daten von über 200.000 Personen aus dem Estonian Genome Center [112]. Datenschutz und

Datenhoheit bilden zentrale Prinzipien der estnischen e-Health-Lösung. Gesundheitsdaten gehören den betroffenen Individuen, während Gesundheitsdienstleistende verpflichtet sind, diese Daten in das ENHIS einzupflegen. Der Zugang zum System ist auf autorisierte medizinische Fachkräfte beschränkt. Patient*innen besitzen das Recht, ihre Gesundheitsdaten über ein Patient*innenportal einzusehen, den Zugriff auf diese Daten mittels eines Opt-Out-Verfahrens einzuschränken und detailliert nachzuverfolgen, welche Personen ihre Informationen eingesehen haben. In Estland verfügen 100% der Patient*innen über eine landesweit digitale Gesundheitsakte [112]. Das ENHIS nutzt die KSI-Blockchain, welche die Datenintegrität und Sicherheit der sensiblen Daten vor internen Angriffen durch ein Logging schützt [113].

FINDATA

Die FINDATA-Behörde ist die zentrale nationale Behörde für die Verwaltung und Freigabe von Gesundheits- und Sozialdaten der finnischen Bevölkerung. Aktuell informiert sie Forschende über die Datenverfügbarkeit, erteilt Genehmigungen zur Sekundärnutzung und ermöglicht die pseudonymisierte Datenverarbeitung in den von ihr bereitgestellten sicheren Umgebungen [114]. Der Zugriff auf die Gesundheits- und Sozialdaten der nationalen eHealth-Infrastruktur Kanta ist seit Januar 2021 möglich [115].

MyHealthRecord

MyHealthRecord ist das nationale Gesundheitsaktensystem Australiens, in dem Patient*innendaten gespeichert und den Bürger*innen zur Verfügung gestellt werden [116]. Es ermöglicht Patient*innen, Gesundheitsinformationen mit Ärzt*innen, Krankenhäusern und anderen Gesundheitsdienstleistenden zu teilen. Die Nutzung und Teilnahme an dem System ist standardmäßig für jede Australier*in aktiviert und es muss aktiv widersprochen werden [117]. Die Daten von MyHealthRecord stehen noch nicht für Forschungszwecke zur Verfügung, dies ist aber geplant. Hierzu werden aktuell Regelungen für die Forschung und das öffentliche Gesundheitswesen festgelegt [118].

2.1.6 Interoperabilitätsstandards im Gesundheitswesen

Durch die Standardisierung von Datenformaten, Kommunikationsprotokollen und Codierungssystemen wird Interoperabilität (siehe Abschnitt 2.3.13) zwischen Systemen und Geräten gewährleistet [119]. Durch die zunehmende Digitalisierung und Vernetzung von Systemen im Gesundheitswesen wurden verschiedene relevante Standards entwickelt und etabliert, die im Folgenden kurz erläutert werden.

Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT): SNOMED-CT ist ein Terminologiestandard für die Dokumentation von medizinischem Wissen in ePAs der von der Standardisierungsorganisation Systematized Nomenclature of Medicine (SNOMED) International herausgegeben wird [119]. Der Standard besteht aus Konzepten, Beschreibungen und den Beziehungen zwischen den Konzepten. Konzepte sind dabei klinisch relevante Begriffe für die Dokumentation von Prozeduren und Befunden. Die Daten werden in dem Standard strukturiert, indem ihnen weltweit gültige Codes zugeordnet werden. [120]

Fast Healthcare Interoperability Resources (FHIR): FHIR ist ein internationaler Standard der Organisation Health Level Seven (HL7) zum Datenaustausch zwischen unterschiedlichen Softwaresystemen im Gesundheitswesen. FHIR beschreibt Datenformate und Elemente als Ressourcen und bietet Standardisierung für Programmierschnittstellen (APIs) für die Übertragung der Ressourcen an. Zusätzlich kann semantische Interoperabilität durch die Verwendung von Kodierungssystemen innerhalb der FHIR Ressourcen erreicht werden. [119, 120]

International Statistical Classification of Diseases and Related Health Problems (ICD): ICD ist ein internationaler Standard für die Meldung von Krankheiten, welcher von der World Health Organization (WHO) herausgegeben wird. Darin werden Erkrankungen, Verletzungen und andere gesundheitsrelevante Zustände definiert und in einer hierarchischen Struktur angeordnet. Der ICD-11 ist seit 2022 die aktuelle Fassung des Standards. [119]

Digital Imaging and Communications in Medicine (DICOM): DICOM ist ein international verbreiteter Standard zur Kommunikation und Verwaltung von medizinischen Bilddaten und zugehöriger Metainformationen. Der DICOM-Standard dient dem interoperablen Austausch medizinischer Bilder zwischen verschiedenen Informationssystemen und unterstützt deren Speicherung und Integration in Bildarchivierungssystemen [119].

2.2 Juristische Grundlagen und Rahmenbedingungen

Die Konzeption und Implementierung von Dateninfrastrukturen sowie Datentreuhandsystemen zur Nutzung von Gesundheitsdaten unterliegen der Einhaltung spezifischer rechtlicher Vorgaben. Diese ergeben sich aus nationalen und europäischen Gesetzen sowie Verordnungen. Ziel dieses Abschnittes ist es, die zentralen gesetzlichen und rechtlichen Anforderungen, Terminologien und Gesetze zu beleuchten, welche für die Entwicklung und den Betrieb solcher Systeme maßgeblich sind und zum Verständnis der nachfolgenden Konzepte beitragen. Besondere Beachtung finden die gesetzlichen Regelungen und Änderungen, die während der Bearbeitungszeit dieser Arbeit verabschiedet und in Kraft getreten sind.

2.2.1 Regulatorische Rahmenbedingungen für Infrastrukturen im Gesundheitswesen

Der Grundstein zur Schaffung der TI und ePA in Deutschland wurde 2004 mit dem **GKV-Modernisierungsgesetz (GMG)** gelegt [52]. Durch die Ergänzung der Paragraphen § 291 und § 291a im Fünften Sozialgesetzbuch (SGB V) wurde die Einführung der eGK und der TI geregelt. In § 291a des SGB V wird insbesondere der Auftrag des Aufbaus der TI beschrieben, welche 2005 die Grundlage zur Gründung der gematik gebildet hat [121].

Im Jahr 2015 konkretisierte das Gesetz für sichere digitale Kommunikation und Anwendungen im Gesundheitswesen (**E-Health-Gesetz**) den Fahrplan zur Einführung der TI und spezifizierte Sanktionsmöglichkeiten bei Nichteinhaltung sowie die vorgesehenen Anwendungen innerhalb dieser Infrastruktur, einschließlich der ePA [53]. Dieser Fahrplan sah die Anbindung von Krankenhäusern und Praxen an die TI bis Mitte 2018, die Einführung eines Medikationsplans in Papier- und elektronischer Form bis 2018 sowie die Schaffung der technischen Rahmenbedingungen für die ePA bis Ende 2018 vor [122]. Das Ziel bestand darin, ab 2019 allen gesetzlich Versicherten die Möglichkeit zur freiwilligen Nutzung der ePA anbieten zu können, um Arztbriefe, Notfalldaten, den Medikationsplan sowie weitere Gesundheitsdaten eigenständig verwalten zu können. Zusätzlich wurden finanzielle Anreize geschaffen, um die Nutzung der Funktionen der eGK zu fördern. Darüber hinaus öffnete das Gesetz die TI für weitere Berufsgruppen und ermöglichte die Einführung telemedizinischer Leistungen, wie beispielsweise Online-Sprechstunden. [122]

Zur Verbesserung der Leistungen der Krankenkassen und der medizinischen Versorgung sowie effizienteren Arztterminvergabe trat im Mai 2019 das **Terminservice- und Versorgungsgesetz (TSVG)** in Kraft [54]. Im Kontext der Digitalisierung der Gesundheitsversorgung und der Förderung ihrer breiteren Anwendung im Alltag sah das Gesetz vor, dass die Krankenkassen bis spätestens 2021 verpflichtet sind, ihren Versicherten eine ePA zur Verfügung zu stellen. Zusätzlich sollte der Zugriff auf die Daten der ePA entkoppelt von der eGK ermöglicht werden. Damit Entscheidungsprozesse innerhalb der gematik optimiert und die Digitalisierung sowie den Aufbau technischer Infrastrukturen im Gesundheitswesen beschleunigt werden können, übernahm das Bundesministerium für Gesundheit (BMG) 51 % der Geschäftsanteile der gematik. [54]

Ebenfalls 2019 trat das **Digitale-Versorgung-Gesetz (DVG)** in Kraft, welches umfassende Maßnahmen zur Förderung der Digitalisierung im Gesundheitswesen vorsah [123]. Im Rahmen des Gesetzes wurden Apotheken bis Ende September 2020 und Krankenhäuser bis zum 1. Januar 2021 verpflichtet, sich an die TI anzuschließen. Zudem wurde die Möglichkeit geschaffen, dass Hebammen,

Physiotherapeut*innen sowie Pflege- und Rehabilitationseinrichtungen sich freiwillig an die TI anbinden können, wobei die Kosten für eine solche freiwillige Anbindung erstattet werden. Zur Durchsetzung der TI-Anbindung wurden auch Sanktionen verschärft: Ärzt*innen, die sich der Anbindung an die TI weiterhin verweigern, unterliegen seit dem 1. März 2020 einem erhöhten Honorarabzug von 2,5 %, nachdem dieser zuvor bei 1 % lag. Ziel dieser Regelungen war es, die flächendeckende Nutzung der TI voranzutreiben. Im Rahmen des DVG wurde in Deutschland erstmals die Nutzung von Sozialdaten für die Verbesserung der medizinische Forschung vorgesehen. Hierzu sollte ein Forschungsdatenzentrum (siehe Abschnitt 2.1.4 für weitere Informationen zum FDZ-Gesundheit) geschaffen werden, welches die pseudonymen Abrechnungsdaten der Krankenkassen aggregiert. Auf Antrag können der Forschung aus diesen Daten anonymisierte Auswertungen zur Verfügung gestellt werden. Mit dem DVG wurden zudem die Grundlagen für die Entwicklung offener und standardisierter Schnittstellen geschaffen, wodurch die Interoperabilität der verschiedenen Systeme im Gesundheitswesen gefördert wird. [123]

Die Neufassung der **Datentransparenzverordnung (DaTraV)** aus 2020 konkretisiert das DVG hinsichtlich Aufgaben und Verfahren der Datentransparenz im vorgesehenen Forschungsdatenzentrum [124]. Sie benennt das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) als verantwortliche Stelle für die Aufgaben des Forschungsdatenzentrums sowie das Robert Koch-Institut (RKI) für die Wahrnehmung der Aufgaben der Vertrauensstelle. Die DaTraV umfasst außerdem Regelungen zu den Datenübermittlungsfristen, zur Datenverarbeitung durch den GKV-SV, Verfahren zur Pseudonymisierung, zur Kostenerstattung, Evaluation und Weiterentwicklung des Forschungsdatenzentrums. [124]

Das ebenfalls 2020 in Kraft getretene **Patientendaten-Schutz-Gesetz (PDSG)** konkretisiert hingegen die DVG aus datenschutzrechtlicher Sicht [125]. Hierzu legt das Gesetz klare Regelungen hinsichtlich Datenschutz und Datensicherheit in der TI fest sowie Regelungen für den Umgang mit Störungen und Sicherheitsmängeln bei Diensten und Komponenten der TI und entsprechende Bußgelder bei nicht ordnungsgemäßer Meldung an die gematik. Zusätzlich sieht es einen Anspruch der Versicherten vor, dass die ePA von Ärzt*innen befüllt wird. Diese

erhalten für das erste Befüllen und Verwalten der ePA einen monetären Ausgleich. Das PDSG sieht ebenfalls vor, dass Patient*innen ab 2022 per Smartphone oder Tablet einzelfallbasiert entscheiden können, wer Zugriff auf welche Daten erhält. Außerdem sieht das PDSG vor, dass ab 2023 Versicherte die Möglichkeit besitzen ihre ePA-Daten über eine Datenspende der Forschung zur Verfügung zu stellen. Abseits von Befunden, Arztbriefen und Röntgenaufnahmen sollen sich ab 2022 zusätzlich noch folgende Datenkategorien in der ePA abspeichern lassen: (1) Impfausweis, (2) Mutterpass, (3) U-Heft für Kinder und (4) das Zahn-Bonusheft. [125]

Das **Digital-Gesetz (DigiG)** sieht vor, dass die ePA ab Anfang 2025 für alle gesetzlich Versicherten zur Verfügung gestellt wird [126]. Versicherte haben die Möglichkeit, der Nutzung der ePA zu widersprechen (Opt-Out). Auch private Krankenversicherungen (PKV) sollen eine widerspruchsbasierte ePA anbieten können. Mit der ePA erhalten die Versicherten eine automatisch erstellte digitale Medikationsübersicht, die in enger Verbindung mit dem E-Rezept steht. Diese Verknüpfung soll helfen, unerwünschte Wechselwirkungen zwischen Arzneimitteln zu vermeiden. Darüber hinaus ist die Einrichtung eines Digitalbeirats bei der gematik vorgesehen, bestehend aus Vertreter*innen des Bundesbeauftragten für den Datenschutz und die Informationsfreiheit (BfDI), des Bundesamts für Sicherheit in der Informationstechnik (BSI) sowie Expert*innen aus den Bereichen Ethik und Medizin. Er soll die gematik kontinuierlich beraten und Empfehlungen zu Datenschutz, Datensicherheit, Datennutzung und Anwendendenfreundlichkeit abgeben. [126]

Mit dem **Gesundheitsdatennutzungsgesetz (GDNG)** sollen Gesundheitsdaten verstärkt für die gemeinwohlorientierte Forschung zugänglich gemacht werden [127]. Hierzu soll eine zentrale Datenzugangs- und Koordinierungsstelle eingerichtet werden, um bürokratische Hürden abzubauen und durch eine zentrale Anlaufstelle den Zugang zu Forschungsdaten erleichtern. Das GDNG sieht dabei vor, dass die Gesundheitsdaten dezentral gespeichert bleiben und nur für spezifische Forschungsanträge in einer sicheren Umgebung zugänglich gemacht werden. Zudem wird zukünftig die Datenschutzaufsicht für länderübergreifende Forschungsvorhaben durch eine*n Landesdatenschutzbeauftragte*n koordiniert und

ein Forschungsgeheimnis eingeführt. Dieses Forschungsgeheimnis bedeutet, dass Forschende Gesundheitsdaten ausschließlich im Rahmen der gesetzlich zulässigen Bestimmungen nutzen und weitergeben dürfen. Sie sind zudem verpflichtet, die Daten vertraulich zu behandeln. Bei Verstößen gegen dieses Forschungsgeheimnisses wird künftig eine Strafnorm wirksam. Das GDNG sieht zusätzlich eine Weiterentwicklung des FDZ-Gesundheit beim BfArM vor, um pseudonymisierte Daten mit medizinischen Registern zu verknüpfen. Ferner ist für die Antragsberechtigung nicht mehr ausschlaggebend, wer einen Forschungsantrag stellt, sondern der gemeinwohlorientierte Zweck der Forschung. Darüber hinaus wird für die pseudonymisierte Freigabe von Daten aus der ePA ein Opt-Out-Verfahren eingeführt, um Behandlungsdaten für die Forschung besser nutzen zu können. Hierzu soll eine digitale Verwaltung von Widersprüchen den Patient*innen zur Verfügung gestellt werden. Über Ombudsstellen der Krankenkassen soll ein nicht digitaler Widerspruch ermöglicht werden. Kranken- und Pflegekassen erhalten zudem die Möglichkeit, personalisierte Gesundheitswarnungen auf Basis von Abrechnungsdaten an Versicherte zu geben mit dem Ziel den Schutz der Individuen zu gewährleisten, z.B. bei Arzneimitteltherapiesicherheit oder der Identifikation von Krebserkrankungen. Hierbei regelt das GDNG Transparenzpflichten der Kranken- und Pflegekassen sowie Ordnungswidrigkeiten bei Verstößen. [127]

Zusätzlich zu den oben beschriebenen nationalen Gesetzen ist das europäische **Daten-Governance-Gesetz** (*engl. Data Governance Act* - DGA) von Bedeutung [128]. Der DGA zielt darauf ab, das Vertrauen in den freiwilligen Datenaustausch von Unternehmen und Bürger*innen zu stärken, da beispielsweise Unternehmen den Verlust ihres Wettbewerbsvorteils befürchten. Daher legt der DGA spezifische Regelungen für Anbietende von Datenvermittlungsdiensten fest, um sicherzustellen, dass diese als neutrale und vertrauenswürdige Instanzen den Datenaustausch im gemeinsamen europäischen Datenraum fördern. Dabei wird das Ziel verfolgt ein alternatives Modell zu den Datenverarbeitungspraktiken der Big-Tech-Plattformen zu schaffen, welche durch den Besitz großer Datenmengen einen hohen Marktanteil erlangen. Gleichzeitig unterstreicht der DGA die Bedeutung von Transparenz und Neutralität bei Datenmittlern, während gleichzeitig Einzelpersonen und Unternehmen die Kontrolle über ihre Daten behalten. Dies

ermöglicht es den Dateninhabenden, ihre Daten auf freiwilliger Basis und ohne finanzielle Gegenleistung im Sinne des Datenaltruismus bereitzustellen. Der DGA regelt hierbei die Bedingungen zur Weiterverwendung geschützter Daten des öffentlichen Sektors, wobei zentrale Anforderungen und Prozesse definiert werden. Zu den wesentlichen Regelungen gehören technische Anforderungen, die sicherstellen, dass die Privatsphäre und Vertraulichkeit der betroffenen Daten gewahrt bleiben. Hierzu zählen Maßnahmen wie Anonymisierung, Pseudonymisierung und die Nutzung sicherer Verarbeitungsumgebungen. Darüber hinaus sieht der DGA eine Unterstützung von Forschenden durch öffentliche Stellen vor, insbesondere beim Zugang zu relevanten Datenbeständen. Diese Unterstützung kann etwa die Hilfe bei der Einholung notwendiger Einwilligungen oder Genehmigungen umfassen. Ein weiterer zentraler Punkt betrifft die Erhebung angemessener Gebühren für die Genehmigung von Anträgen zur Weiterverarbeitung von Daten. Diese Gebühren dürfen die tatsächlich anfallenden Kosten nicht übersteigen, wodurch die finanzielle Hürde für Forschende und andere Akteure minimiert wird. Zudem wird im DGA eine maximale Frist von zwei Monaten für die Bearbeitung und Entscheidung über solche Anträge festgelegt. Die Mitgliedstaaten sind verpflichtet, geeignete zuständige Stellen für die Genehmigung zu benennen, um eine effiziente Bearbeitung sicherzustellen. Zur Förderung der Transparenz sieht der DGA außerdem die Einrichtung einer zentralen Informationsstelle sowie eines zugehörigen Registers vor. Diese dienen dazu, eine Übersicht darüber bereitzustellen, welche Daten bei welchen Behörden gespeichert sind. [128]

2.2.2 Datenschutzrelevante Rahmenbedingungen

Personenbezogene Daten sind alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen und dadurch ein direkter oder indirekter Bezug zu einer Identität ermöglichen (vgl. Art. 4 Abs. 1 DSGVO) [129]. Personenbezogene Gesundheitsdaten unterliegen aufgrund ihrer Sensibilität und Schutzbedürftigkeit einer umfassenden Regulierung auf nationaler und europäischer Ebene, da sie zu den besonderen Kategorien personenbezogener Daten nach § 22 Abs. 1 Bundesdatenschutzgesetz (BDSG) gehören [130, 131]. Im

Folgenden werden die wesentlichen datenschutzrechtlichen Regularien auf den verschiedenen Ebenen vorgestellt.

Auf EU-Ebene bildet seit 2016 die **Datenschutz-Grundverordnung (DSGVO)** den rechtlichen Rahmen für den Schutz personenbezogener Daten, einschließlich der Anforderungen an die Verarbeitung jener Daten, wie sie in Art. 9 DSGVO festgelegt sind [129]. Die DSGVO gilt seit 2018 unmittelbar in allen Mitgliedstaaten der Europäischen Union, wobei sie durch Öffnungsklauseln Raum für nationale Regelungen lässt. Zentrales Prinzip der DSGVO ist das Verbotsprinzip, welches die Verarbeitung personenbezogener Daten grundsätzlich untersagt, wenn diese nicht den Anforderungen der Verordnung entsprechen. Der Begriff der „Verarbeitung" umfasst dabei alle manuellen oder automatisierten Vorgänge im Umgang mit personenbezogenen Daten, wie etwa deren Erhebung, Speicherung, Veränderung, Abfrage, Verwendung, Löschung, Verknüpfung und Offenlegung (Art. 4 Abs. 2 DSGVO). Bei der Verarbeitung personenbezogener Daten besteht die Verpflichtung zur informierten und freiwilligen Einwilligung durch die betroffene Person, welche laut DSGVO der Verarbeitung zustimmen muss (Art. 4 Abs. 11 DSGVO). Bezüglich Art. 7 DSGVO benötigt eine Einwilligung keine bestimmte Form, jedoch hat die betroffene Person das Recht, ihre Einwilligung jederzeit zu widerrufen. Nach Erwägungsgrund 33 der DSGVO wird erlaubt, dass Einwilligungen zur wissenschaftlichen Forschung auch für bestimmte Bereiche erteilt werden können, da der Zweck zum Zeitpunkt der Erhebung unklar sein kann. Das Recht auf Vergessenwerden ist in Art. 17 der DSGVO verankert und gewährt der betroffenen Person das Recht, die unverzügliche Löschung ihrer personenbezogenen Daten zu verlangen, sofern die Verarbeitung dieser Daten nicht aus anderen, rechtlich begründeten Gründen erforderlich ist. In Art. 17 Abs. 3 DSGVO werden Ausnahmen genannt, die dieses Recht einschränken. Das Löschungsrecht kann nicht geltend gemacht werden, wenn die Daten zur Erfüllung einer rechtlichen Verpflichtung, zur Wahrung öffentlicher Interessen – wie im Bereich der Gesundheitsversorgung – oder zu wissenschaftlichen Zwecken erforderlich sind. Gesundheitsdaten werden gemäß Art. 4 Abs. 15 DSGVO als personenbezogene Daten definiert, „die sich auf die körperliche oder geistige Gesundheit einer natürlichen

Person, einschließlich der Erbringung von Gesundheitsdienstleistungen, beziehen und aus denen Informationen über deren Gesundheitszustand hervorgehen“ [129]. In Art. 5 DSGVO werden die folgenden Grundsätze für die Verarbeitung von personenbezogenen Daten beschrieben:

- **Rechtmäßigkeit, Verarbeitung nach Treu und Glauben, Transparenz:** Die Daten müssen auf rechtmäßige Weise, nach Treu und Glauben und auf eine nachvollziehbare Weise verarbeitet werden.
- **Zweckbindung:** Die Erhebung der Daten darf nur zu festgelegten, eindeutigen und legitimen Zwecken erfolgen. Die Weiterverarbeitung im öffentlichen Interesse für beispielsweise die wissenschaftliche Forschung gilt gemäß Art. 89 Abs. 1 DSGVO nicht als unvereinbar mit dem ursprünglichen Zweck.
- **Datenminimierung:** Die Menge der zu verarbeitenden Daten soll auf das notwendige Maß zur Erfüllung des Zwecks beschränkt sein.
- **Richtigkeit:** Die Daten müssen sachlich richtig und auf dem neusten Stand sein.
- **Speicherbegrenzung:** Daten dürfen nur so lange gespeichert werden, wie es für die Erfüllung des Verarbeitungszwecks notwendig ist.
- **Integrität und Vertraulichkeit:** Die Daten müssen durch geeignete technische und organisatorische Maßnahmen vor unbefugtem Zugriff und Missbrauch geschützt werden sowie in einer Weise verarbeitet werden, dass der Schutz und die Sicherheit der Daten gewährleistet wird.

Auf nationaler Ebene ergänzt und konkretisiert das **Bundesdatenschutzgesetz (BDSG)** die Vorgaben der DSGVO für Deutschland [131]. Das BDSG enthält zusätzliche Vorschriften zur Verarbeitung personenbezogener Daten, einschließlich Regelungen zur Verarbeitung von Daten durch öffentliche und nichtöffentliche Stellen. Gemäß § 22 Abs. 2 ist bei der Verarbeitung von Gesundheitsdaten als besonderer Kategorie an personenbezogenen Daten insbesondere die Wahrung der Interessen der betroffenen Person durch geeignete Maßnahmen vorzusehen.

Darüber hinaus existieren die **Datenschutzgesetze der Länder und Kirchen**, die den Umgang mit personenbezogenen Daten auf regionaler Ebene bzw. innerhalb spezifischer Anwendungsbereiche weiter konkretisieren [132].

2.2.3 Arten von Einwilligungen

Grundlegend wird bei der Verarbeitung von personenbezogenen Daten zwischen der einwilligungsbasierten Datenverarbeitung (Opt-In) und der Widerspruchslösung (Opt-Out) unterschieden. Bei der einwilligungsbasierten Datenverarbeitung ist entsprechend die aktive Zustimmung zur Verarbeitung durch die betroffene Person mittels einer Willensbekundung erteilt worden (Art. 4 Abs. 11 DSGVO) [129]. Im Gegensatz dazu wird bei der Widerspruchslösung die Datenverarbeitung geduldet, es sei denn, es liegt ein Widerspruch vor. [95]

Für die einwilligungsbasiere Datenverarbeitung gibt es diverse Varianten von Einwilligungen:

- **Blanket Consent:** Blanket Consent beschreibt eine pauschale Einwilligung, bei der Datensubjekte ihre Zustimmung zur uneingeschränkten Nutzung ihrer Daten für jegliche zukünftige Forschungszwecke geben, ohne dass ihnen Informationen über die zukünftige Verwendung bereitgestellt werden oder diese weitere Kontrollmöglichkeiten haben. Jener Ansatz erleichtert die Durchführung von Forschungsprojekten, da keine wiederholten Einwilligungen eingeholt werden müssen, wenn sich neue Forschungsfragen oder Anwendungsfälle ergeben. [133]
- **Open Consent:** Der Open Consent basiert als offene Einwilligung auf einem Transparenzansatz, bei dem Datensubjekte einer uneingeschränkten Offenlegung ihrer Daten zustimmen. Ein typisches Beispiel hierfür ist die Veröffentlichung von Daten in einer Open-Access-Datenbank. [133]

- **Broad Consent:** Der Broad Consent erlaubt eine abgestufte oder dynamische Zustimmung, bei der Datensubjekte verschiedene Restriktionen bezüglich der Verwendung ihrer Daten festlegen können. Diese Art der Einwilligung ermöglicht es Datensubjekten, einer Vielzahl von Forschungsprojekten zuzustimmen, auch wenn der spezifische Zweck der Datennutzung zum Zeitpunkt der Einwilligung noch nicht festgelegt ist. [133, 95] In Erwägungsgrund 33 der DSGVO wird dieser Ansatz beispielsweise durch die Angabe des allgemeinen Zwecks „medizinische Forschung“ anerkannt [129].
- **Dynamic Consent:** Mittels des Dynamic Consent hat das Datensubjekt die Wahl zur Datennutzung über einen anhaltenden Kommunikationskanal zwischen Datengebenden und Datennutzenden. Über digitale Plattformen und Anwendungen können hierdurch Datensubjekte ihre Einwilligung erteilen und diese im weiteren Verlauf anpassen. Dieser Ansatz fördert die aktive Beteiligung der Datensubjekte. [133]

2.3 Technische Grundlagen

Die nachfolgend dargestellten Definitionen und Erläuterungen zentraler Begriffe dienen als Fundament für die technische Analyse sowie die thematische Kontextualisierung der vorliegenden Arbeit.

2.3.1 Datentreuhandmodelle

Die Grundlage für die datenbasierte Forschung sowie Entwicklung datenbasierter Innovationen und Dienstleistungen bildet die verantwortungsvolle Bereitstellung und Nutzung von Daten. Das BMBF betrachtet Datentreuhandmodelle als einen vielversprechenden Lösungsansatz zur Bewältigung dieser Herausforderungen [134]. Dabei definiert das BMBF Datentreuhandmodelle als *"[...] neutrale Intermediäre, die einen fairen Ausgleich der Interessen der beteiligten Akteure*

und einen vertrauensvollen Austausch von Daten inklusive des dafür notwendigen technischen und organisatorischen Zugangs ermöglichen" [134]. Lindner und Straub [135] sowie Feth et al. [136] betonen, dass bislang keine einheitliche und allgemein anerkannte Definition des Begriffs „Datentreuhänder“ vorliegt. Lindner und Straub heben jedoch hervor, dass der Begriff in zwei unterschiedlichen Konzepten verstanden werden kann: Erstens als treuhänderische Verwaltung und neutrale Vermittlung von Daten und zweitens als Instanz, die Vertrauen zwischen Datengebenden und Datennutzenden schafft und im Falle von Konflikten als Schlichtungsstelle agiert [135]. Beide Ansätze finden laut Lindner und Straub in der Praxis Anwendung [135]. Feth et al. [136] führen die fehlende einheitliche Definition von Datentreuhändern auf deren vielfältige Funktionen, unterschiedlichen Zielsetzungen sowie die Diversität der spezifischen Anwendungsbranchen und -sektoren zurück. Auf Grundlage der Gemeinsamkeiten und des kleinsten gemeinsamen Nenners definieren Feth et al. einen Datentreuhänder als *"[...] eine Vertrauensinstanz, die schützenswerte Daten zwischen Datengebern und Datennutzern unter Wahrung der Interessen beider Seiten digital vermittelt"* [136]. Gemäß Blankertz et al. [137] kann ein Treuhänder verschiedene Funktionen übernehmen, darunter: (1) die Verwaltung von Zugriffsrechten auf Daten, (2) die Speicherung der Daten, (3) die Aufbereitung von Daten, wie beispielsweise Pseudonymisierung, Verschlüsselung und Qualitätssicherung, (4) die Datenveredelung sowie die Durchführung von Datenauswertungen, (5) der Bereitstellung und gegebenenfalls Weitergabe von Daten auf einer neutralen Plattform, (6) die Verhandlung über Datenzugriffsrechte mit den Datennutzenden, (7) den Betrieb einer treuhänderischen Infrastruktur sowie (8) die Schaffung von Transparenz bezüglich der Datenzugriffe.

Auf Grundlage der allgemeinen Definitionen wird im Rahmen dieser Arbeit ein Datentreuhandmodell als eine technische, neutrale Instanz definiert, die schützenswerte Daten zwischen Datengebenden und Datennutzenden unter Wahrung der Interessen beider Seiten digital vermittelt und dabei Mehrwertdienste zur Bereitstellung und Nutzung der Daten zur Verfügung stellt. Diese Definition umfasst alle von Blankertz et al. [137] genannten Funktionen. Durch diese umfassende Sichtweise grenzt sich die verwendete Definition des Datentreuhandmodells von

anderen, spezialisierten Begriffen wie „Treuhandstelle“ oder „Vertrauensstelle“ ab, die oft nur einzelne Funktionen übernehmen [96, 99]. Ein Beispiel hierfür ist die unabhängige Treuhandstelle Greifswald, die sich auf das Einverständnis- und Pseudonym-Management sowie den Abgleich mit externen Datenquellen wie Melderegistern konzentriert [96]. Die Begriffe „Datentreuhandmodell“, „Datentreuhand“, „Treuhänder“ und „Datentreuhandsystem“ werden im Kontext dieser Arbeit als Synonyme verwendet, um die zentralen Konzepte einheitlich und auf Basis der zuvor genannten Begriffsdefinition zu beschreiben.

2.3.2 Softwarearchitektur

Die Softwarearchitektur ist ein zentrales Konzept in der Softwareentwicklung und bildet das Fundament für den Entwurf und die Implementierung von Softwaresystemen. In der Fachliteratur existiert eine Vielzahl unterschiedlicher Definitionen des Begriffs „Softwarearchitektur“ [138]. Im folgenden Abschnitt werden ausgewählte, häufig zitierte Definitionen näher vorgestellt.

Softwarearchitektur ist nach Fowler *"[...] a notion of the core elements of the system, the pieces that are difficult to change. A foundation on which the rest must be built"* [139].

Die ISO/IEC/IEEE 42010:2022 definiert Architektur als *"[...] fundamental concepts or properties of a system in its environment embodied in its elements, relationships, and in the principles of its design and evolution"* [140].

Nach Booch et al. [141] ist Softwarearchitektur die Summe verschiedener wichtiger Entscheidungen über die Organisation eines Softwaresystems, die Auswahl von Strukturelementen und deren Schnittstellen, aus denen das System zusammengesetzt ist, das Verhalten und Zusammenspiel dieser Elemente, den hierarchischen Aufbau von Subsystemen und den zugrundeliegenden Architekturstil.

Eoin Woods definiert Softwarearchitektur als *"[...] the set of design decisions which, if made incorrectly, may cause your project to be canceled"* [142].

Diese Definitionen betonen die Bedeutung von zentralen Designentscheidungen als entscheidenden Einflussfaktor für den Erfolg eines Softwaresystems bzw. Projekts. Des Weiteren hebt die Definition von Woods [142] hervor, dass sich Softwarearchitektur durch schwer änderbare Entscheidungen auszeichnet, deren nachträgliche Anpassung Auswirkungen auf die zentralen Projektparameter wie Budget, Zeit und Qualität haben kann. Entscheidungen, die leicht anpassbar sind, gelten demnach nicht als architekturrelevant und erfordern keine Architekturaufwände, da ihre Veränderbarkeit keine nachhaltigen Auswirkungen auf das Gesamtsystem oder die Projektplanung hat [143]. Toth definiert „schwere Änderbarkeit“ als Änderungen, welche teuer, aufwendig oder qualitätsgefährdend sind [143].

Unter Zugrundelegung dieser allgemeinen Definitionen, wird im Rahmen dieser Arbeit Softwarearchitektur als die Gesamtheit wichtiger Entscheidungen über die Organisation und Struktur eines Softwaresystems verstanden, die aufgrund ihrer schwer änderbaren Natur eine nachhaltige und entscheidende Rolle für den Erfolg des Systems spielen.

Ähnlich wie in der Medizin, wo unterschiedliche Sichten auf die Struktur des menschlichen Körpers verwendet werden, ist auch bei Softwaresystemen eine differenzierte Betrachtung der Architektur erforderlich. Dabei wird in der Softwarearchitektur zwischen Strukturen und Sichten unterschieden [144]. Die vorliegende Arbeit verwendet die Definition von Strukturen und Schichten von Bass, Clements und Kazman [144]:

- **Strukturen** bezeichnen die Menge von Architekturelementen, die in der Software oder Hardware physisch existieren.
- **Sichten** sind Darstellungen von Strukturen und dienen als abstrahierte Darstellungen, die sowohl die Architekturelemente als auch die Beziehungen zwischen diesen darstellen.

Eine grafische Modellierungsmethode zur Darstellung und hierarchischen Abstraktion der Architektur von Softwaresystemen ist das C4-Modell von Simon Brown [145]. Es beschreibt die Softwarearchitektur durch die Verwendung von

vier unterschiedlichen Sichten: Systemkontext-, Container-, Komponenten und Quelltext-Diagramme. Diese Sichten sind hierarchisch in die folgenden Ebenen gegliedert und repräsentieren verschiedene Abstraktionsebenen [145]:

- Ebene 1 - Systemkontext-Diagramm: Ein Systemkontext-Diagramm stellt den Ausgangspunkt dar und veranschaulicht, wie das zu entwickelnde Softwaresystem in die umgebende Welt eingebettet ist, indem es dessen Interaktionen mit Benutzenden und externen Systemen darstellt.
- Ebene 2 - Container-Diagramm: In einem Containerdiagramm wird ein Softwaresystem durch die Darstellung von Containern beschrieben, die miteinander interagieren. Ein Container kann dabei eine Anwendung oder einen Datenspeicher darstellen, die jeweils spezifische Verantwortlichkeiten innerhalb des Systems übernehmen.
- Ebene 3 - Komponenten-Diagramm: Ein Komponenten-Diagramm untergliedert Container in zusammenhängende Komponenten und stellt die Interaktionen zwischen diesen Komponenten sowie zu anderen Containern und Systemen dar.
- Ebene 4 - Quelltext-Diagramm: Das Quelltext-Diagramm verdeutlicht die konkrete Implementierung von Komponenten durch beispielsweise UML-Klassendiagramme oder Entity-Relationship-Diagramme.

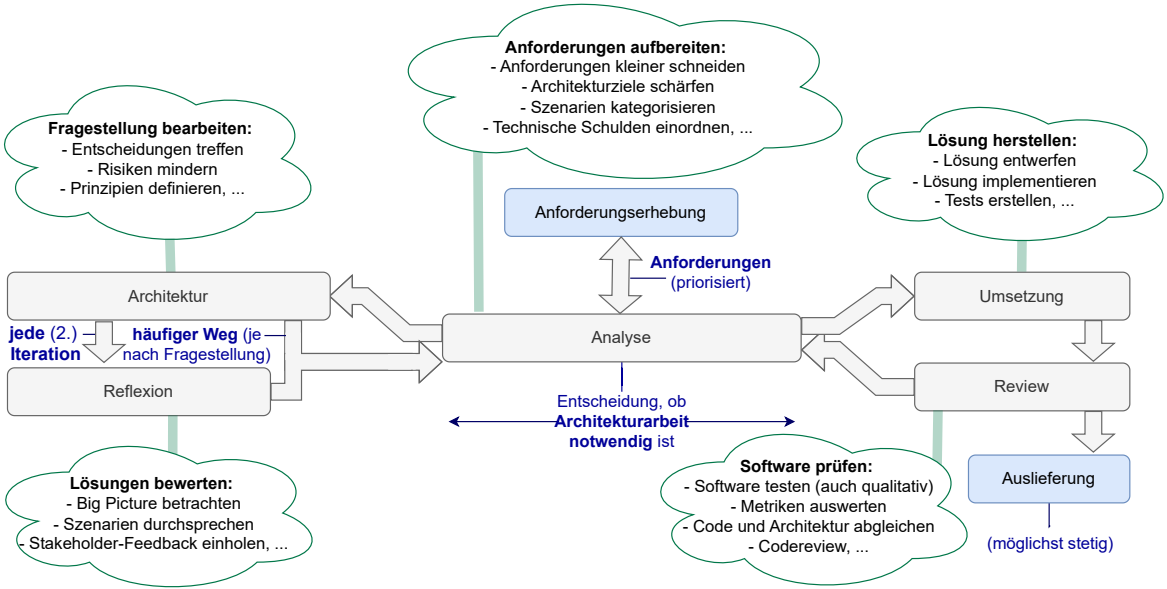

Abbildung 2.8: Softwarearchitektur-Prozess nach Toth [143].

2.3.3 Softwarearchitektur-Prozess

Toth [143] beschreibt den Softwarearchitektur-Prozess als einen iterativen Ansatz (siehe Abbildung 2.8). Im Folgenden wird der von Toth [143] beschriebene Softwarearchitektur-Prozess im Detail erläutert. Die Basis für den durch Anforderungen getriebenen Prozess bilden priorisierte Anforderungen, welche die Bedarfe, Wünsche und Anforderungen der Stakeholder repräsentieren. Diese Anforderungen sollen neben den rein funktionellen Anforderungen ebenfalls die nicht-funktionalen Qualitätsanforderungen enthalten, welche beschreiben, wie eine Funktionalität bereitgestellt werden soll und Aspekte wie Performance, Sicherheit, Zuverlässigkeit sowie Benutzendenfreundlichkeit adressieren. Eine Kategorisierung dieser Qualitätsmerkmale wird durch die ISO/IEC 25010 [146] beschrieben und wird in Abbildung 2.9 veranschaulicht.

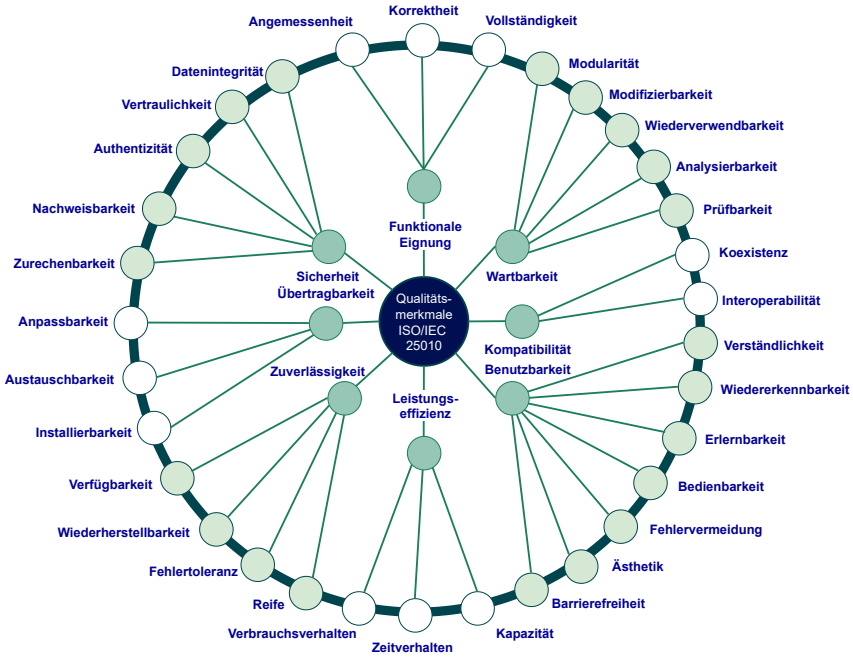

Abbildung 2.9: Kategorisierung der Qualitätsanforderungen nach ISO/IEC 25010 [146].

Qualitätsanforderungen sind laut Toth häufig mit hohen Risiken verbunden, wodurch die Rücknahme entsprechender Architekturentscheidungen mit erheblichem Aufwand oder hohen Kosten verbunden ist [143]. Aus diesem Grund haben sie eine besondere Bedeutung für die Softwarearchitektur und sind somit architekturrelevante Größen. Nach Toth bietet sich ein initialer Anforderungs-Workshop für die Erhebung des zentralen Anforderungskatalogs und der Erarbeitung einer groben Produktvision an [143]. Darauf aufbauend sollen Anforderungspflege-Workshops durchgeführt werden, um die Anforderungen kontinuierlich zu überprüfen und anzupassen. Im Rahmen einer Analyse als ersten Schritt des Softwarearchitektur-Prozesses bilden diese Anforderungen die Grundlage. Zunächst werden die Anforderungen für die nächste Iteration ausgewählt und entschieden, ob die Anforderungen direkt umgesetzt werden können oder Architekturarbeit erforderlich ist.

Je nach Ergebnis der Analyse teilt sich der Prozess in zwei Zyklen auf: den Architekturzyklus (links) und den Implementierungszyklus (rechts). Grundsätzlich erfolgt im Implementierungszyklus die Implementierung, Testung, Integration und Auslieferung der Software. Nach Toth [143] soll in diesem Zyklus der Großteil der Zeit aufgewendet werden, da dort der produktive Teil der Softwareentwicklung erfolgt und auslieferbare Software durch das kontinuierliche prüfen der Software (Review) entsteht. Die direkte Umsetzung von Anforderungen birgt jedoch Risiken, insbesondere, wenn diese unzureichend verstanden oder schwer in die bestehende Applikation integrierbar sind. Dies kann zu suboptimalen Lösungen führen, deren nachträgliche Anpassung schwer änderbar ist und daher mit hohen Kosten und erheblichem Aufwand verbunden sein kann. Werden auf Grundlage der Anforderungen weitreichende und schwer änderbare Entscheidungen in der Umsetzung erforderlich, wird der Architekturzyklus durchlaufen. Das Ziel dieses Zyklus ist es, die Anforderungen und die damit verbundene Problemstellung umfassend zu verstehen, Unsicherheiten zu beseitigen und potenzielle Risiken zu minimieren (z.B. durch die Entwicklung von Prototypen). Als Ergebnis der Architektur werden Entscheidungen auf architekturrelevante Fragestellungen vorbereitet.

Zur Erkennung von architekturrelevanten Fragestellungen empfiehlt Toth [143], sich von den folgenden Fragen leiten zu lassen:

- Ist die Entscheidung später nur schwer zu ändern?
- Ist die Umsetzung der Entscheidung eher teuer?
- Werden sehr hohe, qualitative Anforderungen gestellt? (Hochsicherheit, Hochverfügbarkeit, Hochperformanz etc.)
- Lassen sich Anforderungen nur schwer in Bestehendes abbilden?
- Ist die eigene Erfahrung im Lösungsspektrum schwach?

Werden diese Fragen mehrheitlich mit „Ja“ beantwortet, liegt mit hoher Wahrscheinlichkeit eine architektonische Fragestellung vor. Bei dem Vorliegen solcher Fragestellungen und Entscheidungen empfiehlt Toth [143] architektonische Mittel, um das Risiko potenzieller Fehlentscheidungen zu minimieren. Unter architektonischen Mitteln versteht Toth [143]:

- Genaue Analyse nicht-funktionaler Anforderungen
- Konzeption und ggf. Modellierung möglicher Lösungen
- Definition von Prinzipien
- Breite Vermittlung von Architekturentscheidungen
- Verarbeitung von Erkenntnissen aus der Umsetzung
- Verankerung von qualitativen Tests zum Erhalt von Rückschlüssen zur Architekturidee.

Die Entscheidungen sollen im Anschluss dokumentiert werden. Diese werden daraufhin entweder als Architekturidee direkt in einem Implementierungszyklus umgesetzt oder in einer Reflexionsphase mit weiteren Stakeholdern validiert. Falls weitere Risiken, Kompromisse, offenen Punkte oder Probleme auftreten, kann ein weiterer Architekturzyklus erforderlich werden. Die Häufigkeit des Durchlaufens

des Architekturzyklus hängt von der Komplexität und den Anforderungen des jeweiligen Projekts ab. Während bei komplexen, technologisch anspruchsvollen oder verteilten Systemen häufiger Architekturarbeit notwendig ist, sind kleinere und weniger komplexe Vorhaben weniger stark auf diesen Zyklus angewiesen. Auch das verwendete Vorgehensmodell beeinflusst die Integration des Architekturzyklus. In klassischen Softwareentwicklungsprojekten, wie sie beispielsweise im Wasserfallmodell angewendet werden, wird dieser typischerweise einmal zu Beginn des Projekts durchlaufen, wobei dieser mehrere Monate in Anspruch nehmen kann. Im Gegensatz dazu ermöglichen es agile Vorgehensmodelle wie Scrum, den Architekturzyklus mehrfach innerhalb einer Iteration zu durchlaufen. Jede architektonisch relevante Anforderung, da potentiell schwer änderbar, wird in einem Architekturzyklus betrachtet und kommuniziert, bevor dieser umgesetzt wird. Die Reflexion erfolgt dann nicht bei jeder Anforderung, sondern abhängig von der Anzahl der getroffenen Architekturentscheidungen als Workshop in jeder oder jeder zweiten Iteration.

2.3.4 Architekturmuster und -taktiken

Für spezifische Problemstellungen in unterschiedlichen Domänen wurden bestimmte Zusammensetzungen von architektonischen Elementen identifiziert und dokumentiert, die als wiederverwendbare Lösungen für Probleme in der Softwarearchitektur dienen. Diese systematischen Kompositionen architektonischer Elemente und Bündel an Architekturentscheidungen als bewährte Lösungen werden nach Bass, Clements und Kazman [144] als Architekturmuster (*engl. Architectural Patterns*) bezeichnet. Beispiele für Architekturmuster sind das Schichten-Architekturmuster oder das Client-Server-Architekturmuster. Architekturmuster lassen sich von Entwurfsmustern (*engl. Design Patterns*) abgrenzen, welche auf einer niedrigeren Abstraktionsebene ansetzen. Während Architekturmuster die Entwicklung der grundlegenden Organisation eines Softwaresystems unterstützen, konzentrieren sich Entwurfsmuster auf Probleme in der Implementierung und des Design von Software (z.B. Singleton-Muster, Factory-Muster)[144, 147].

Architekturtaktiken (*engl. Architectural Tactics*) hingegen sind nach Bass, Clements und Kazman [144] Entwurfsentscheidungen, welche das Erreichen von Qualitätsanforderungen zum Ziel haben. Sie zeichnen sich dadurch aus, dass sie sich in der Regel auf eine einzelne Struktur oder einen spezifischen Mechanismus konzentrieren, um gezielt eine bestimmte architektonische Anforderung zu adressieren, ohne dabei Trade-Offs zu berücksichtigen. Im Gegensatz dazu beinhalten Architekturmuster bereits integrierte Trade-Offs, die durch das Muster selbst gelöst werden. Im Vergleich zu Architekturmustern sind Architekturtaktiken einfacher. Sie dienen als grundlegende Bausteine des Designs, aus denen Architekturmuster entwickelt werden können. Ein Beispiel für eine Architekturtaktik wäre die Ausnahmebehandlung (*engl. exception handling*, bei der Mechanismen zur Erfassung und den Umgang mit Fehlern implementiert werden, ohne das System zum Stillstand zu bringen [144]. Im Weiteren sollen die zwei Architekturmuster Client-Server sowie Peer-to-Peer vorgestellt werden (vgl. Abbildung 2.10).

Client-Server: Das Client-Server-Architekturmuster basiert auf der Kommunikation zwischen mindestens zwei unabhängigen Prozessen: (1) einem sogenannten Client, der Dienste anfragt, und (2) einem Server, der diese bereitstellt. Die Interaktion erfolgt synchron, wobei der Server die angeforderten Aufgaben ausführt und die Ergebnisse über das Netzwerk zurücksendet. Dieses Architekturmuster zeichnet sich durch hohe Flexibilität und eine klare Trennung der Verantwortlichkeiten aus, bringt jedoch Herausforderungen in Bezug auf Sicherheit, Entwicklungsaufwand und Fehleranfälligkeit mit sich. [147]

Peer-to-Peer (P2P): P2P-Architekturen basieren auf gleichberechtigten und vernetzten Komponenten (Peers), die sowohl Funktionen eines Clients als auch eines Servers übernehmen und Ressourcen wie Speicherplatz, Rechenleistung oder Dateien gemeinsam teilen. Diese Architekturen zeichnen sich durch eine hohe Ausfallsicherheit aus, da sie keinen zentralen Schwachpunkt (*engl. Single Point of Failure*) darstellen. Sie finden primär in großflächigen Netzwerken wie dem Internet Anwendung, beispielsweise für ausfallsichere Datenverteilung, digitale Telefonie oder Instant Messaging. Vorteile sind die hohe Ausfallsicherheit durch Redundanz und die dezentrale Organisation, die keine zentrale Administration erfordert. Allerdings ergeben sich auch Herausforderungen, wie die Erkennung

und Verhinderung von Datenmanipulationen durch kompromittierte Peers sowie die aufwendige Suche von Fehlerquellen in der verteilten Kommunikation. [147]

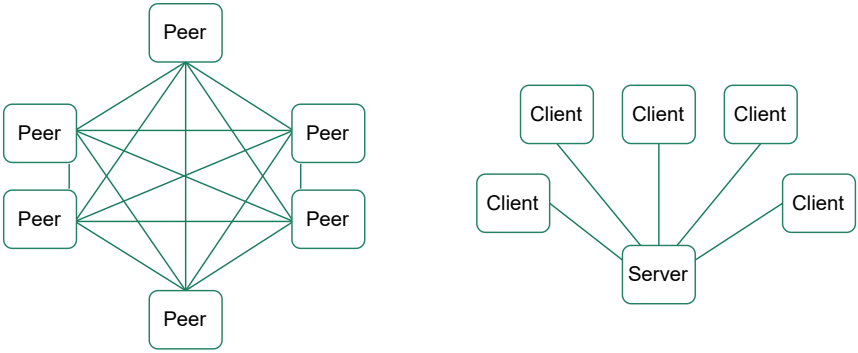

Abbildung 2.10: Darstellung der Architekturmuster Client-Server und Peer-to-Peer [147].

2.3.5 Vorgehensmodelle in der Softwareentwicklung

Softwareentwicklungsprojekte nutzen verschiedene Vorgehensmodelle (vgl. [148]), die sich grundsätzlich in lineare und iterative Ansätze unterteilen lassen. Klassische lineare Modelle, wie beispielsweise das Wasserfallmodell, folgen einem sequenziellen Ablauf. Bei diesem Vorgehensmodell wird die nächste Projektphase erst begonnen, wenn die vorherige vollständig abgeschlossen ist. Der Übergang und der Abschluss zwischen den Phasen wird durch definierte Meilensteine kontrolliert, die den Projektfortschritt überprüfen. Diese Modelle bieten bei gleichbleibenden Anforderungen eine hohe Planungssicherheit und ermöglichen eine präzise Aufwands- sowie Zeitschätzung. Allerdings sind diese weniger dynamisch, da sie eine geringe Flexibilität gegenüber Änderungen aufweisen. Dies kann dazu führen, dass Fehler zu einem späten Zeitpunkt im Entwicklungsprozess erkannt werden. [148]

Softwarelösungen als Ergebnis der Softwareentwicklung dienen der Bewältigung von Problemen, die in ihrer Komplexität variieren können. Mit zunehmender

Komplexität eines Softwaresystems steigt häufig auch die Fehleranfälligkeit, wodurch eine erhöhte Flexibilität in der Entwicklung und Anpassung erforderlich wird [143]. Hier setzen iterative Modelle an. Iterative Modelle basieren auf kurzen Entwicklungszyklen (Iterationen). Agile Methoden, insbesondere das Scrum-Framework, fokussieren die Flexibilität, Anpassungsfähigkeit und kontinuierliche Verbesserung im Rahmen dieser Iterationen. In selbstorganisierten, interdisziplinären Teams werden sogenannte Sprints durchgeführt, die gemeinsam in zeitlich begrenzten Abschnitten lauffähige Produktinkremente entwickeln. Durch eine regelmäßige Evaluation dieser Inkremente mit Stakeholdern im Rahmen von sogenannten Sprint Reviews erfolgt eine kontinuierliche Bewertung und eine schrittweise Annäherung an das Projektergebnis. Hierdurch wird eine ständige Anpassung an sich verändernde Anforderungen ermöglicht. Hingegen sind agile Ansätze abhängig von einer engen Zusammenarbeit der Projektbeteiligten und erfordern ein hohes Maß an Vertrauen sowie Aufwände zur Umsetzung und Anpassung bestehender Prozesse. [148]

Im Kontext dieser Arbeit wird Softwarearchitektur vor dem Hintergrund agiler Vorgehensweisen untersucht, da in der Praxis in der Softwareentwicklung agile Ansätze weit verbreitet sind [148].

2.3.6 Entscheidungsmodelle

Ein Entscheidungsmodell ist eine systematische Orientierungshilfe zur Unterstützung von Entscheidungsprozessen, indem es die Auswahl bestimmter Technologien, Lösungen oder Muster erleichtert, die beispielsweise bei der Konzeption von Softwarearchitekturen von Relevanz sind [149, 150].

Abbildung 2.11 zeigt die im Rahmen dieser Arbeit verwendete Notation für das Entscheidungsmodell. Diese lehnt sich an der Notation der Business Process Model and Notation (BPMN) an [151]. Ein Kreis symbolisiert den Beginn des Entscheidungsprozesses, während ein Kreis mit einer dicken Umrandung das Ende des Entscheidungsprozesses markiert. Ein paralleles Gateway repräsentiert eine Gabelung des Entscheidungsprozesses, bei der alle durch gerichtete Pfeile dargestellten

Pfade simultan durchlaufen werden. Im Gegensatz dazu überprüft ein exklusives Gateway spezifische Bedingungen und teilt den Entscheidungsprozess abhängig von deren Erfüllung in einen der alternativen, sich gegenseitig ausschließenden Pfade. Abgerundete Rechtecke stellen die Designentscheidungen dar. Teilprozesse werden durch abgerundete blaue Rechtecke mit einem Plus-Symbol dargestellt. Diese Teilprozesse werden durch ein gesondertes Teilmodell beschrieben, das über einen eigenen Start- und Endpunkt verfügt.

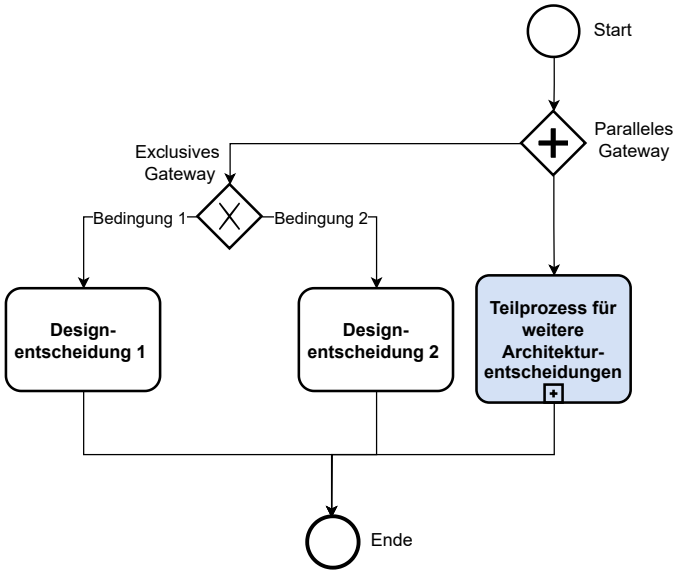

Abbildung 2.11: Verwendete Notation für Entscheidungsmodelle, angelehnt an die BPMN.

2.3.7 FAIR-Prinzipien

Im Jahr 2016 wurden durch Wilkinson et al. die „FAIR Guiding Principles for scientific data management and stewardship“ veröffentlicht [152]. Das Ziel der Prinzipien von Wilkinson et al. besteht darin, Leitlinien bereitzustellen, um die Auffindbarkeit (***F**indability*), Zugänglichkeit (***A**ccessibility*), Interoperabilität

(***I**nteroperability*) und Wiederverwendbarkeit (***R**eusability*) von digitalen Ressourcen in Datenmanagementsystemen und -infrastrukturen zu verbessern [152]. Die FAIR-Prinzipien bieten damit einen strukturierten und umfassenden Rahmen zur Förderung einer nachhaltigen, offenen und effektiven Nutzung von (wissenschaftlichen) Daten [152].

2.3.8 Blockchain-Technologie

Distributed-Ledger-Technologie (DLT) ist ein Oberbegriff für vertrauenswürdige dezentrale Dateninfrastrukturen, welche manipulationssicher Transaktionen mehrerer Netzwerkteilnehmenden protokollieren [153]. Wie der Begriff DLT bereits andeutet, bildet ein Hauptbuch (*engl. Ledger*), dessen Kopien dezentral über alle Netzwerkknoten (*engl. Nodes*) eines P2P-Netzwerks verteilt verwaltet werden, den Kern dieser Technologie [153]. Die bekannteste Repräsentation einer DLT ist die 2008 unter dem Pseudonym Satoshi Nakamoto veröffentlichte Blockchain, welche als Schlüsseltechnologie hinter der Kryptowährung Bitcoin allgemein bekannt ist [154]. Im Fall der Blockchain stellt der Ledger eine verkettete Liste von Blöcken dar, wobei jeder Block eine Reihe von Transaktionen speichert und durch einen kryptographischen Hashwert mit dem vorhergehenden Block verknüpft ist (siehe Abbildung 2.12) [155]. Sollte ein Knoten mit betrügerischer Absicht eine Transaktion in einem früheren Block verändern, ändern sich die Hashwerte aller nachfolgenden Blöcke. Diese Veränderung wird als ungültig erkannt, wodurch die Integrität des Systems sichergestellt wird und die Blockchain vor Manipulation geschützt wird [153]. Der erste Block der Blockchain wird als Genesis-Block bezeichnet und wird bei der Initialisierung des Netzwerks erzeugt. Er kann beispielsweise Konfigurationen enthalten oder, im Fall von Bitcoin, anfängliche Kontostände. Jeder Knoten der Blockchain besitzt eine Kopie des Ledgers und kann dessen Transaktionen verifizieren. Die Identifikation der teilnehmenden Knoten in einer Blockchain erfolgt durch eine Blockchain-Adresse kombiniert mit Public-Key-Kryptographie, wobei diese mit ihrem privaten Schlüssel Transaktionen signieren können. Nach der Signierung wird die Transaktion an einen Knoten im Netzwerk übermittelt, der sie validiert und bei Gültigkeit an einen verarbeitenden

Knoten (*engl. Miner*) weiterleitet. Jene Miner aggregieren valide Transaktionen in Blöcke und hängen diese der Blockchain an. Neue Blöcke werden an das gesamte Netzwerk gesendet. Um an dieser Stelle eine Einigung zu finden, welcher Block als nächstes an die Blockchain angehangen wird, existieren sogenannte Konsensverfahren. [155] Deren Funktionsweise und weitere Blockchain-relevante Konzepte werden nachfolgenden im Detail beleuchtet.

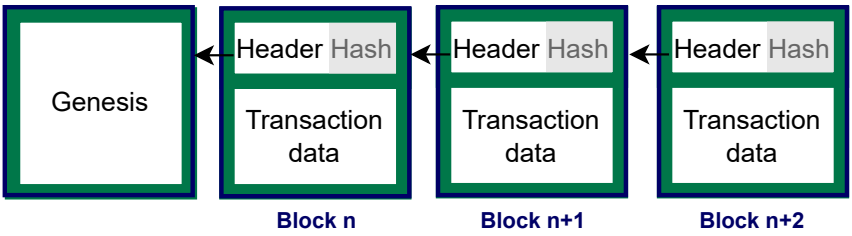

Abbildung 2.12: Strukturelle Aufbau des Ledgers einer Blockchain in Anlehnung an Xu et al. [155].

Konsensverfahren

Mittels des gewählten Konsensverfahrens erfolgt innerhalb des Netzwerks eine Abstimmung darüber, ob eine Transaktion oder ein Block zur Blockchain hinzugefügt wird. Aufgrund der dezentralen Struktur des Netzwerks kann es vorkommen, dass mehrere Miner gleichzeitig Blöcke generieren, die an der nächsten Position der Blockchain eingefügt werden sollen. Dies geschieht, da Miner möglicherweise nicht rechtzeitig darüber informiert wurden, dass andere bereits ebenfalls einen Block generiert haben. [155]

Ein übergeordnetes Verfahren zur Konsensfindung ist der Nakamoto-Konsens, welcher mit der Bitcoin-Blockchain eingeführt wurde und demgemäß häufig in öffentlichen Blockchains Anwendung findet [155]. Werden beispielsweise zwei Blöcke, Block 1 und Block 2, parallel erstellt und an den aktuellen Block 0 angehängt, resultiert daraus eine Aufspaltung der Blockchain (*engl. Fork*) in zwei Ketten gleicher Länge. Unter diesen Umständen entscheidet der nachfolgende Block, ob die Kette mit Block 1 oder Block 2 fortgeführt wird. Auch hier kann es vorkommen, dass beide Ketten parallel einen nachfolgenden Block anhängen. In

diesem Fall trifft der darauffolgende Block die endgültige Entscheidung darüber, welche Kette fortgeführt wird. Das Bitcoin-Protokoll begrenzt derartige parallele Verzweigungen typischerweise auf maximal ein bis zwei Blöcke, wodurch die Hauptkette (*engl. main chain* frühzeitig konsolidiert wird. [155]

Der Nakamoto-Konsens findet Anwendung im Proof of Work (PoW) Konsensalgorithmus, bei dem ein kryptographisches Puzzle zum Hinzufügen eines Blocks gelöst werden muss. In Kombination mit PoW entspricht die längste Kette derjenigen, die im Durchschnitt die meiste Rechenleistung erhalten hat. Das Puzzle vom PoW ist so gewählt, dass es schwierig zu lösen, aber einfach zu verifizieren ist, wodurch eine zufällige Zeitdauer für dessen Lösung erforderlich wird. Die Miner treten in einen Wettbewerb, um das kryptographische Puzzle für jeden Block zu lösen, indem sie ihre Rechenleistung steigern. Dadurch erhöhen sie ihre Wahrscheinlichkeit, den nächsten Block erfolgreich an die Blockchain anzuhängen. Im Erfolgsfall werden sie mit einem monetären Anreiz in Form eines bestimmten Betrags der jeweiligen Kryptowährung der Blockchain belohnt. Es existieren verschiedene Proof-of-Work-Mechanismen, wie beispielsweise Ethash, das von Ethereum verwendet wird, und Hashcash, welches im Bitcoin-Netzwerk zum Einsatz kommt. [155]

Proof of Stake (PoS) verwendet ebenfalls den Nakamoto-Konsens. Im Rahmen von PoS wird der nächste hinzuzufügende Block in Abhängigkeit davon gewählt, welche Menge native Kryptowährung des Blockchain-Netzwerkes der Miner besitzt. Im Vergleich zum PoW ist PoS kosteneffizienter, da weniger Rechenleistung für das Mining notwendig ist und die Latenzzeit kürzer ausfällt. Ein Nachteil ist jedoch, dass dadurch das passive Halten von Kryptowährungen erschwert wird. [155]

Ein Konsensverfahren, das den Nakamoto-Konsens nicht anwendet, ist die Practical Byzantine Fault Tolerance (PBFT) [155]. PBFT wird verwendet, um Konsens in permissioned Blockchains zu finden und die Integrität des Systems auch bei fehlerhaften oder böswilligen Knoten aufrechtzuerhalten. Mit PBFT können bis zu einem Drittel der Knoten ($f = \frac{n-1}{3}$) als fehlerhaft oder betrügerisch toleriert werden. Der Prozess beginnt, indem ein Client eine Anfrage an den sogenannten

Primary Node sendet. Dieser leitet die Anfrage an die restlichen Knoten, die als Backup Nodes bezeichnet werden, weiter. Die Backup Nodes führen die Anfrage aus und senden die Ergebnisse an den Client zurück. Der Client akzeptiert die Antwort nur, wenn er $f + 1$ übereinstimmende Antworten von unterschiedlichen Backup Nodes erhält, wodurch die Antwort als gültig bestätigt wird. Die Antwort der Anfrage muss jedoch deterministisch sein. Vorteil des PBFT ist die starke Konsistenzgarantie und geringe Latenz. Jedoch führt der Austausch vieler Nachrichten zu einer beschränkten Skalierbarkeit. [155, 156]

Kryptowährungen

Mit Hilfe der Blockchain können digitale Assets erzeugt und verwaltet werden, beispielsweise Token oder Währungen. Eine Kryptowährung stellt demgemäß ein digitales Zahlungsmittel dar, welches auf der Blockchain verwaltet wird. Mit Hilfe von Transaktionen können Kryptowährungen zwischen den Teilnehmenden des Blockchain-Netzwerkes transferiert werden. Damit diese Transaktionen in der Blockchain aufgenommen werden, müssen die Teilnehmenden Transaktionsgebühren zahlen, welche abhängig von der Größe der Transaktion und nicht deren Wert sind. Beispiele für Kryptowährungen sind Bitcoin sowie Ether. Ether ist die native Währung der Etherium Blockchain. [155]

Smart Contract

Ein Smart Contract ist ein deterministisches in der Blockchain gespeichertes Programm, über das Transaktionen ausgeführt werden können. Ein Smart Contract kann eine Vielzahl von Funktionen ausführen und ist nicht notwendigerweise ein rechtlicher Vertrag. Allerdings kann er zur Überprüfung und Automatisierung bestimmter Aspekte rechtsverbindlicher Verträge eingesetzt werden. Smart Contracts ermöglichen die Automatisierung von Prozessen, indem sie Trigger, Bedingungen und Geschäftslogik implementieren, die komplexe, programmierbare Transaktionen ermöglichen. Typischerweise werden Smart Contracts zu Verwaltung von digitalen Assets verwendet.[155]

Orakel

Ein Orakel (*engl. Oracle*) wird verwendet, um externe Informationen in das Blockchain-System zu integrieren. Dies kann entweder durch die Ausführung

einer regulären Transaktion eines Smart Contracts oder durch die Nutzung speziell für Orakel definierter Smart Contracts erfolgen. Da Smart Contracts auf den durch Orakel bereitgestellten Daten basieren, muss ein Orakel eine vertrauenswürdige Quelle darstellen, um die Integrität der ausgeführten Transaktionen zu gewährleisten. [155]

Wallet
Ein Wallet ist eine digitale Brieftasche, die es den Teilnehmenden eines Blockchain-Netzwerks ermöglicht, ihre Sammlung privater Schlüssel zu verwalten, die zur Erstellung und Signatur von Transaktionen ausgehend von ihrem Blockchain-Konto verwendet werden. Hierbei wird zwischen Software- und Hardware-Wallets unterschieden. Hardware-Wallets sind physische Geräte, welche zur Speicherung von privaten Schlüsseln verwendet werden. [155]

Die Blockchain-Frameworks Hyperledger Indy und Aries
Hyperledger ist eine Sammlung von Open-Source-Projekten, die von der Hyperledger Foundation ins Leben gerufen wurden. Diese Projekte wurden 2024 in die Linux Foundation Decentralized Trust (LF Decentralized Trust) übernommen, ebenfalls einer Stiftung der Linux Foundation, die sich breiter auf die Entwicklung und Bereitstellung von Open-Source-Lösungen für dezentrale Ökosysteme konzentriert [157]. Hyperledger Aries ermöglicht eine vertrauenswürdige P2P-Kommunikation, die auf dezentralen selbstbestimmten Identitäten (SSI - *engl. Self-Sovereign Identity*) und verifizierbaren Nachweisen (VC - *engl. Verifiable Credentials*) basieren [158]. Eine entsprechende Erläuterung von SSI und VC erfolgt in Abschnitt 3.3. Initial wurden die Werkzeuge und Protokolle zum Aufbau von Aries Agenten (P2P-Knoten) und Netzwerken im Rahmen von Hyperledger Indy entwickelt. Damit diese jedoch Blockchain-agnostisch, also unabhängig von gewissen Blockchain-Frameworks sind, wurden diese in Hyperledger Aries migriert [158]. Hyperledger Indy stellt ein Blockchain-Framework bereit, welches speziell für die Verwaltung und Speicherung von Identitäten entwickelt wurde und als Blockchain-Lösung in Aries verwendet werden kann [158]. Hyperledger Indy ist eine öffentliche Blockchain und verwendet als Konsensverfahren Plenum eine Implementierung des Redundant Byzantine Fault Tolerance (RBFT) Konsensalgorithmus [159]. RBFT stellt eine Erweiterung des klassischen PBFT

dar, indem es zusätzliche redundante Validierungs- und Bestätigungsschritte für Transaktionen implementiert. Bei RBFT werden mehrere Instanzen des gleichen BFT-Protokolls parallel ausgeführt, wobei jede Instanz auf einer eigenen Maschine läuft. Alle Instanzen überprüfen und ordnen die eingehenden Anfragen, aber nur die Anfragen, die von der Master-Instanz genehmigt wurden, werden tatsächlich ausgeführt [160].

2.3.9 Zugriffs- und Identiätsmanagement

Ein Identitäts- und Zugriffsmanagementsystem (IAM - *engl. Identity and Access Management System*) bezeichnet die Verfahren und Strukturen zur Verwaltung und Steuerung von Identitäten in cyber- und cyber-physischen Systemen [161]. Gemäß Bouras et al. [161] umfasst ein IAM-System die Funktionen der Authentifizierung von Entitäten, Methoden zur Autorisierung und Gewährung von Zugriffsrechten sowie den gesamten Lebenszyklus digitaler Identitäten. Jener Lebenszyklus beinhaltet die Erstellung, Aktualisierung und Wartung einer Identität gemäß vom Administrator des IAMs vordefinierten Regeln und Bedingungen [161]. In diesem Zusammenhang besteht das Identitätsmanagement aus zwei Hauptbereichen: (1) die Ausstellung von Anmeldeinformationen (*engl. Credentials*) und eindeutigen Identifikatoren an Benutzende bei der initialen Registrierung sowie (2) die Authentifizierung der Benutzenden und die Kontrolle ihres Zugriffs auf Dienste und Ressourcen auf der Grundlage ihrer Identifikatoren und Anmeldeinformationen während dem Betrieb des Dienstes [162]. Die Kernentitäten eines IAMs sind Benutzende (*engl. User*), Identitätsanbietende (*engl. Identity Provider*) und Dienstanbietende (*engl. Service Provider*) [161].

Identitätsmanagementansätze lassen sich in zentrale, föderierte, nutzerzentrierte und dezentrale Identitätsmanagementsysteme (IdM) einteilen, Dezentrale IdMs werden in Self-Sovereign Identity (SSI) und Decentralized Trusted Identity (DTI) unterteilt. [161] Zugriffsmanagementansätze unterscheiden sich gemäß ihrer Zugriffskontrollstrategie in Discretionary Access Control (DAC) und Non-Discretionary Access Control (NDAC) [163]. Eine genauere Erläuterung und

Differenzierung der Identitäts- und Zugriffsmanagementansätze erfolgt im Kapitel 3. Nachfolgend wird die Definition des Begriffs „Identität“ dargelegt.

Identität
Identitäten repräsentieren eine Entität innerhalb einer spezifischen Anwendungsdomäne. Zum Beispiel stellen die in einer ePA gespeicherten persönlichen Gesundheitsdaten einer Patientin, wie Name, Geburtsdatum und Krankengeschichte, sowie deren physische Merkmale, wie sie vom medizinischem Personal wahrgenommen werden, die Identität dieser Patientinnen-Entität dar. Identitäten stehen in der Regel in Bezug zu realen Entitäten. Typische reale Entitäten sind Personen oder Organisationen. [162]

Eine Identität wird durch eine Sammlung von Merkmalen definiert, die als Identifikatoren (*engl. Identifiers*) bezeichnet werden und der Identifikation dienen. Sie können unterschiedliche Merkmale besitzen, wie etwa vorübergehende oder permanente Gültigkeit, selbst gewählt oder von einer Autorität zugewiesen, sowie die Fähigkeit zur Interpretation durch Menschen oder ausschließlich durch digitale Systeme. Die Merkmale einer Identität variieren je nach Art der realen Entität, die identifiziert wird. Beispielsweise gelten bestimmte Identifikatoren wie das Geburtsdatum für Individuen, jedoch nicht für Organisationen, während andere, wie eine nationale Unternehmensregistrierungsnummer, auf Unternehmen, jedoch nicht auf Einzelpersonen anwendbar sind. Ein Identitätsdomäne ist ein Bereich, in dem jede Identität einzigartig ist. [162]

2.3.10 De-Identifikation

Jede Person verfügt über direkte Identifikationsmerkmale (ID) sowie indirekte Identifikationsmerkmale, die als Quasi-Identifier (QID) bezeichnet werden. Direkte Identifikationsmerkmale umfassen alle Daten, die eine unmittelbare Identifikation einer Person ermöglichen, wie etwa der Name oder eindeutige Nummern wie die Personalausweis- oder Reisepassnummer. Im Gegensatz dazu sind Quasi-Identifier Merkmale, die in Verbindung mit externen Informationen oder durch

die Kombination mehrerer QIDs in einer QID-Gruppe eine Identifikation ermöglichen. Beispiele für Quasi-Identifier sind Geburtsdatum, Wohnort, Beruf oder andere persönliche Attribute. Eine Person kann somit durch ihre ID und die zugehörige QID-Gruppe charakterisiert werden. Je spezifischer die QID-Gruppe ist, desto weniger Menschen erfüllen die entsprechenden Merkmale, wodurch letztlich eine eindeutige Identifikation möglich wird. [164]

Die De-Identifikation bezeichnet den Prozess der Reduktion bzw. Entfernung der personenbezogenen Daten, wodurch der direkte Bezug zu einer individuellen Person eingeschränkt oder vollständig entfernt wird. Abhängig vom Grad der Entfernung des Personenbezugs wird hierbei zwischen Pseudonymisierung und Anonymisierung unterschieden [164]. Nachfolgend werden diese Begriffe näher erläutert.

Pseudonymisierung: Bei der Pseudonymisierung wird der Personenbezug nicht irreversibel entfernt, sondern die Daten werden von der individuellen Person getrennt, sodass eine Identifikation nur unter Hinzunahme zusätzlicher Informationen möglich ist (Art.4 Nr.5 DSGVO). Dadurch wird der ungehinderte Zugang durch Dritte verhindert oder eingeschränkt. Ein typisches Beispiel für Pseudonymisierung ist das Ersetzen eines Klarnamens durch eine Nummer. Die Zuordnung zwischen Nummer und Klarname wird in einem separaten Dokument hinterlegt, welches den weiteren Verarbeitenden nicht zugänglich ist. Dadurch bleibt die Möglichkeit der Re-Identifikation, also der Wiederherstellung des Personenbezugs, erhalten. [164]

Anonymisierung: Das Zehnte Buch des Sozialgesetzbuchs (§3 Abs. 6 BDSG und §67 Abs. 8 SGB X) definiert Anonymisierung als "das Verändern personenbezogener Daten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können". Praktisch versteht man darunter die Veränderung oder Löschung von Daten, mit dem Ziel, den Personenbezug aus den Daten zu entfernen. Dabei wird zwischen drei Formen der Anonymisierung unterschieden: der formalen, faktischen und absoluten Anonymisierung. [164, 130]

Die absolute Anonymisierung beschreibt die Tatsache, dass der Bezug zur betreffenden Person in keiner Weise hergestellt werden kann. Dies wird durch die Entfernung aller direkten und indirekten Identifizierungsmerkmale erreicht. Die faktische Anonymisierung hingegen zielt darauf ab, dass eine Person mit unverhältnismäßig großem Aufwand an Zeit, Kosten und Arbeitskraft re-identifiziert werden kann. Hierzu werden die QIDs entfernt oder verändert (z.B. durch Generalisierung, Permutation oder Unterdrückung). Die formale Anonymisierung ist das Entfernen der direkten Identifikationsmerkmale, welche die Re-Identifikation anhand der QID nicht ausschließt. [164, 130]

2.3.11 Informationssicherheit und Kryptographie

Nach Wegener et al. bezeichnet IT-Sicherheit den Schutz von IT-Systemen, wobei der Fokus primär auf technischen Maßnahmen und insbesondere den Schutz elektronisch gespeicherter Informationen sowie deren Verarbeitung innerhalb von IT-Systemen umfasst. Im Gegensatz zur IT-Sicherheit beschränkt sich der Begriff der Informationssicherheit nicht allein auf den Schutz elektronisch gespeicherter Informationen. Stattdessen umfasst er laut Wegener et al. alle Informationen in sämtlichen Phasen der Geschäftsprozesse. Die klassischen Grundwerte der Informationssicherheit werden dabei häufig durch die drei zentralen Begriffe Vertraulichkeit, Integrität und Verfügbarkeit charakterisiert. [165] Nachstehend erfolgt eine Erläuterung der drei zentralen Ziele der Informationssicherheit:

- Vertraulichkeit (*engl. Confidentiality*) beschreibt die Eigenschaft, dass Informationen ausschließlich berechtigten Personen oder Entitäten zugänglich sind. Zugriffskontrolle als Schutzmaßnahme dient beispielsweise dazu, unbefugten Zugriff auf Informationen zu verhindern [165].
- Integrität (*engl. Integrity*) bezieht sich auf die Unversehrtheit und Korrektheit von Informationen. Dabei geht darum durch geeignete Maßnahmen sicherzustellen, dass Manipulationen erkannt und nachverfolgt werden können [165].

- Verfügbarkeit (*engl. Availability*) bezeichnet die Eigenschaft, dass Informationen Berechtigten innerhalb eines angemessenen Zeitrahmens zur Verfügung stehen [165].

Die benannten Ziele der Informationssicherheit werden durch kryptographische Verfahren sichergestellt [165]. Kryptographie stellt einen Teilbereich der Kryptologie dar, die zusätzlich auch die Kryptanalyse umfasst, welche als die Kunst des Brechens kryptographischer Systeme verstanden wird. Infolgedessen beschäftigt sich die Kryptographie mit der Wissenschaft des geheimen Schreibens, deren Ziel das Verschleiern von Nachrichten oder Informationen ist [166]. Im Folgenden erfolgt eine detaillierte Darstellung der relevanten kryptographischen Verfahren.

Symmetrische Verschlüsselung
Die symmetrische Verschlüsselung basiert darauf, dass sowohl Sendende als auch Empfangende denselben Schlüssel für die Ver- und Entschlüsselung von Nachrichten verwenden [166]. Anhand eines Beispiels verschlüsselt im Kontext der symmetrischen Verschlüsselung eine Sendende Alice ihre Nachricht x mithilfe eines symmetrischen Schlüssels k, wodurch das Chiffrat y entsteht. Bob empfängt dieses Chiffrat und entschlüsselt es ebenfalls mit dem Schlüssel k, um die ursprüngliche Nachricht wiederherzustellen. Die Entschlüsselung stellt somit den inversen Prozess der Verschlüsselung dar. Für die Verteilung des Schlüssels k zwischen Alice und Bob benötigt das System einen sicheren Kommunikationskanal [166].

Der Advanced Encryption Standard (AES) ist die weltweit am häufigsten verwendete symmetrische Verschlüsselungsmethode. AES wird in zahlreichen Internetsicherheitsstandard wie IPsec, Transport Layer Security (TLS) und Secure Shell (SSH) eingesetzt. [166]

Asymmetrische Verschlüsselung
Im Gegensatz zur symmetrischen Verschlüsselung basiert die asymmetrische Verschlüsselung, auch als Public-Key-Verschlüsselung bezeichnet, auf der Verwendung eines Schlüsselpaares. Jenes Verfahren wurde von Diffie, Hellman und

Merkle basierend auf der Idee entwickelt, dass die Person, die die Nachricht verschlüsselt keinen geheimen Schlüssel verwenden muss und damit kein sicherer Kommunikationskanal notwendig wird. Ausschließlich der Empfangende besitzt einen geheimen Schlüssel zur Entschlüsselung, welcher ebenfalls nicht einmal dem Sendenden bekannt sein muss. Hierzu veröffentlicht Bob einen öffentlichen Verschlüsselungsschlüssel k_{pub}, welcher allgemein und damit Alice bekannt ist. Alice verwendet diesen Schlüssel k_{pub} zur Verschlüsselung und erhält das Chiffrat y. Bob empfängt das Chiffrat und entschlüsselt es mit seinem privaten Schlüssel k_{pr}. Die Sicherheit dieser Verfahren beruht auf der Nutzung von Einwegfunktionen, die zwar in polynomieller Zeit berechnet werden können, jedoch deren Umkehrfunktionen mit der aktuellen Hardware als praktisch unmöglich (zeitlich auf mehr als 10.000 Jahre geschätzt) zu berechnen sind. [166]

Zusätzlich können durch asymmetrische Verfahren Nachrichten signiert und deren Integrität sichergestellt werden, da Empfangende durch den privaten Schlüssel nachweisen können, dass sie im Besitz des Geheimnisses sind und damit ihre Identität authentifizieren. Ein Nachteil von asymmetrischen Verfahren ist jedoch ein erhöhter Rechenaufwand, sowie die Anfälligkeit für zukünftige Angriffe durch Quantencomputer bei nicht Post-Quantum-sicheren Algorithmen. [166]

Hybride Verschlüsselung

Um die Vorteile von symmetrischer und asymmetrischer Verschlüsselung zu verwenden, gibt es auch hybride Protokolle, beispielsweise TLS. Hier wird asymmetrische Verschlüsselung zunächst verwendet, um den Schlüssel für die symmetrische Verschlüsselung auszutauschen, und dann wird der symmetrische Schlüssel für die eigentliche Kommunikation genutzt, um die Effizienz zu maximieren. [166]

Hashfunktionen

Hashfunktionen sind ein wichtiges Werkzeug der Kryptographie und werden in verschiedenen Protokollen sowie bei der Blockchain und Digitalen Signaturen angewandt. Mit Hilfe von Hashfunktionen lassen sich beliebig lange Nachrichten in kurze Bitstrings mit fester Länge überführen. Hierbei werden diese Bitstrings als Hashwert bezeichnet, welche unabhängig von der Länge der Eingabe sind.

Praktische Hashfunktionen haben eine Ausgabelänge von 128 bis 512 Bits. [166] Laut Paar und Pelz [166] müssen Hashfunktionen die folgenden drei Eigenschaften erfüllen, um als sicher zu gelten:

- **Urbildresistenz:** Eine Hashfunktion muss eine Einwegfunktion sein. Das bedeutet, dass es rechnerisch unmöglich sein soll, die ursprüngliche Nachricht x aus dem Hashwert z der Hashfunktion $h(x) = z$ zu berechnen [166].

- **Schwache Kollisionsresistenz:** Für digitale Signaturen mit Hashfunktionen ist es von entscheidender Bedeutung, dass zwei verschiedene Nachrichten ($x_1 \neq x_2$) nicht denselben Hashwert ($z_1 = h(x_1) = h(x_2) = z_2$) erzeugen. Aufgrund der Tatsache, dass Hashfunktionen eine unendlich große Anzahl von möglichen Eingaben haben, während die Länge der Ausgaben begrenzt ist, wird es immer mehrere Eingaben mit dem gleichen Hashwert als Ausgabe geben. Durch die Existenz solcher Kollisionen, werden Hashfunktionen in der Praxis so entworfen, dass es rechnerisch unmöglich ist, die Kollisionen zu finden. Durch das Durchprobieren verschiedener Eingaben, kann jedoch eine Kollision gefunden werden. Durch eine Wahl einer ausreichend großen Länge des Hashwerts (80 Bits) können jene Angriffe verhindert werden [166].

- **Starke Kollisionsresistenz:** Hashfunktionen werden als stark kollisionsresistent bezeichnet, wenn es rechnerisch praktisch unmöglich ist, zwei verschiedene Eingaben ($x_1 \neq x_2$) zu finden, für die der gleiche Hashwert ($h(x_1) = h(x_2)$) berechnet werden kann. Jene Eigenschaft ist schwerer einzuhalten, als bei der schwachen Kollisionsresistenz, da Angreifende zwei Freiheitsgrade haben: Beide Nachrichten können so verändert werden, dass sie ähnliche Hashwerte erzeugen. Das Geburtstagsparadoxon beschreibt den Effekt, bei dem die Anzahl an Personen, die benötigt wird, um mit hoher Wahrscheinlichkeit zwei Personen zu finden, die denselben Geburtstag haben, oft intuitiv unterschätzt wird. Es benötigt nur 23 Personen für eine Wahrscheinlichkeit von 50%. Das Geburtstagsparadoxon lässt sich zum

Finden von Kollisionen anwenden. Zur Verhinderung dieser Angriffe muss ein Hashwert mindestens 128 Bit verwendet werden [166].

Digitale Signaturen
Die Eigenschaft, nachzuweisen, dass eine bestimmte Person eine Nachricht erstellt hat, ist auch außerhalb des digitalen Bereichs von Bedeutung. Im realen, analogen Kontext wird dies durch handschriftliche Unterschriften erreicht, etwa beim Unterzeichnen eines Vertrags, wobei Empfangende vor Gericht beweisen können, dass die Unterschrift tatsächlich von der betreffenden Person stammt. Analog zu handschriftlichen Unterschriften muss auch bei digitalen Nachrichten nur die Person, die die Nachricht erstellt, in der Lage sein, eine gültige digitale Signatur zu erzeugen. Dies wird mit Public-Key-Kryptographie realisiert, bei dem Unterzeichnende ihren privaten Schlüssel verwenden, während Empfangende den entsprechenden öffentlichen Schlüssel zur Verifizierung nutzen. [166]

Der Prozess beginnt mit dem Signieren der Nachricht durch Bob, wobei der private Schlüssel k_{pr} von Bob die Signatur erzeugt. Danach wird diese der Nachricht hinzugefügt und an Alice gesendet. Eine digitale Signatur ist ohne die zugehörige Nachricht jedoch nutzlos. Um die Signatur zu überprüfen, benötigt Alice die Nachricht, die Signatur und Bobs öffentlichen Schlüssel k_{pub}. Dieser kann mit dem öffentlichen Schlüssel die Echtheit der Signatur prüfen. [166]

2.3.12 Bedrohungsmodellierung

Softwaresysteme sind gegenwärtig diversen Bedrohungen und Angriffen ausgesetzt und müssen sich aufgrund technologischer Fortschritte ständig weiterentwickeln. Diese Bedrohungen können sowohl aus dem Inneren einer Organisation als auch von externen Agierenden ausgehen und je nach Zielsetzung potenziell schwerwiegende Auswirkungen haben [167]. An dieser Stelle setzt die sogenannte Bedrohungsmodellierung (engl. *Threat Modeling*) an. Bedrohungsmodellierung wird eingesetzt, um Schwachstellen eines Systems und deren Konsequenzen strukturiert zu analysieren [167]. Shostack definiert Bedrohungsmodellierung als *"[...] the key to a focused defense. Without threat models, you can never stop playing*

whack-a-mole." [168]. Darüber hinaus hebt Shostack hervor, dass Bedrohungsmodellierung nicht als einzelne Aktivität zu verstehen ist, sondern als eine Abfolge von Schritten zur Erreichung von Teilzielen [168]. Typischerweise beinhaltet die Bedrohungsmodellierung die folgenden Schritte: (1) Erstellung einer Systemabstraktion, (2) die Identifikation von Bedrohungen und deren Auswirkungen und (3) die Entwicklung und Validierung von Gegenmaßnahmen zur Bewältigung der Bedrohungen [168, 169]. Es wurden bereits zahlreiche Methoden zur Durchführung der Bedrohungsmodellierung sowohl in der Wissenschaft als auch in der Industrie entwickelt, wie beispielsweise die STRIDE-Methode [168], die LINDDUN-Methode [170] sowie die PASTA-Methode [171]. Die Wahl einer geeigneten Methode zur Bedrohungsmodellierung sollte auf den spezifischen Anforderungen des Systems sowie den Bedürfnissen und Ressourcen der beteiligten Stakeholder basieren (z.B. anhand der verfügbaren Zeit und der vorhandenen Erfahrung in der Bedrohungsmodellierung) [169]. Aufgrund der Tatsache, dass die STRIDE-Methode im Rahmen der vorliegenden Arbeit verwendet wurde, und wegen ihrer Ausgereiftheit sowie der häufigen praktischen Anwendung, insbesondere im Hinblick auf die Sicherheitsmerkmale Integrität, Verfügbarkeit und Vertraulichkeit, wird im Folgenden diese Methode detailliert vorgestellt [169, 172].

Die STRIDE-Methode basiert auf einem Merkspruch zur Erfassung und Validierung von Sicherheitsbedrohungen und steht für die sechs Bedrohungs-Kategorien: **S**poofing, **T**ampering, **R**epudiation, **I**nformation Disclosure, **D**enial of Service und **E**levation of Privilege [168, 169]. Dabei sind die Bedrohungs-Kategorien laut Shostack wie folgt zu verstehen:

- **Spoofing** (*dt. Identitätsverschleierung*) bezeichnet das Vorgeben einer Identität oder Rolle, die nicht der eigenen entspricht [168].
- **Tampering** (*dt. Manipulation*) beschreibt das Ändern von Daten oder Systembestandteilen. Dies kann Netzwerkpakete, Daten auf Speichermedien oder Informationen im Arbeitsspeicher umfassen [168].
- **Repudiation** (*dt. Verleugnung*) bezeichnet das Abstreiten von Handlungen, unabhängig davon, ob diese durch eine Person erfolgt sind oder nicht [168].

- **Denial of Service** (DoS - *dt. Verweigerung des Dienstes*) bezeichnet Angriffe, die Systeme an der Bereitstellung ihrer Dienste hindern. Dies kann beispielsweise durch das Herbeiführen von Abstürzen, die erhebliche Verlangsamung der Systemleistung oder die vollständige Auslastung des verfügbaren Speichers erreicht werden [168].
- **Information Disclosure** (*dt. Offenlegung von Informationen*) bezeichnet die Offenlegung von Informationen an Personen, die keine Berechtigung zur Einsichtnahme haben [168].
- **Elevation of Privilege** (*dt. Rechteausweitung*) tritt auf, wenn ein Programm oder User technisch in der Lage ist, Dinge zu tun, welche diesem nicht erlaubt sind [168].

Das Ziel der STRIDE-Methode besteht unter der Zuhilfenahme dieser Bedrohungs-Kategorien darin, mögliche Angriffe und Bedrohungen auf ein System systematisch zu identifizieren. Hierzu stellt der erste Schritt der STRIDE-Methode die Erstellung einer Systemabstraktion (z.B. ein Datenflussdiagramm) dar. Darauf aufbauend erfolgt die Identifikation der Bedrohungen entsprechend der STRIDE-Kategorien. Hierfür existieren bereits Checklisten und Tabellen zur Unterstützung bei der Beschreibung von Bedrohungen sowie Bestimmung von typischen Opfern und Handlungsmustern von Angreifenden. Anschließend erfolgt die Ableitung, Dokumentation und Priorisierung von Sicherheitsmechanismen zur Verhinderung der identifizierten Bedrohungen [168, 169]. Ein Vorteil der STRIDE-Methode liegt laut Shevchenko et al. in ihrer einfachen Anwendbarkeit. Ein entsprechender Nachteil besteht jedoch im hohen Zeitaufwand, da die Anzahl der Bedrohungen bei komplexeren Systemen schnell ansteigt [169].

2.3.13 Interoperabilität

Um den Austausch von Gesundheitsinformationen zwischen verschiedenen Informationssystemen zu ermöglichen, stellt Interoperabilität ein zentrales Element dar. Die Healthcare Information and Management Systems Society (HIMSS) definiert

Interoperabilität im Gesundheitswesen als die Fähigkeit unterschiedlicher Informationssysteme, Geräte und Anwendungen, koordiniert auf Daten zuzugreifen, diese auszutauschen, zu integrieren und kooperativ zu nutzen, um eine nahtlose Portabilität von Informationen über organisatorische, regionale und nationale Grenzen hinweg zu gewährleisten und die Gesundheit von Einzelpersonen sowie Bevölkerungen weltweit zu optimieren [119]. Die HIMSS unterscheidet bei der Interoperabilität vier Ebenen: grundlegende, strukturelle, semantische und organisatorische Interoperabilität. Auf der Ebene der grundlegenden Interoperabilität geht es zunächst darum, dass Daten von einem System in ein Zielsystem übertragen, empfangen und gespeichert werden, ohne dass das Zielsystem in der Lage ist, diese Daten zu verarbeiten [119]. An dieser Stelle setzt die strukturelle Interoperabilität an, bei der die Datenformate, Datensyntax und Datenstruktur einheitlich definiert werden [119]. Unter semantischer Interoperabilität versteht man die Anwendung von Datenmodellen und Kodierungssystemen, die auf standardisierten Vokabularien basieren, um ein gemeinsames Verständnis der Daten zu gewährleisten und somit deren Bedeutung sowie Verarbeitungsmöglichkeit zwischen unterschiedlichen Systemen und Anwendungen sicherzustellen [119]. Schließlich beinhaltet die organisatorische Interoperabilität die Festlegung und Anwendung sozialer, rechtlicher und organisatorischer Richtlinien sowie Überlegungen, die eine sichere Kommunikation und Nutzung von Daten innerhalb und zwischen Organisationen sowie Individuen ermöglichen [119]. Zur Sicherstellung und Schaffung von Interoperabilität gibt es eine Vielzahl an Standards im Gesundheitswesen, welche in Abschnitt 2.1.6 näher beschrieben werden.

3 Stand der Technik und Wissenschaft

Im folgenden Kapitel wird der aktuelle Stand der Technik und Wissenschaft im Bereich des Blockchain-basierten Daten-, Zugriffs- und Identitätsmanagements im Gesundheitswesen untersucht, wobei eine umfassende Taxonomie entwickelt wird, um die unterschiedlichen Ansätze und Systeme in diesen Bereichen systematisch zu klassifizieren und zu analysieren.

Zur Erfassung der verschiedenen Ansätze gemäß dem Stand der Technik und Wissenschaft wurde eine strukturierte Literaturrecherche in drei Iterationen durchgeführt. Hierzu wurden die vier wissenschaftlichen Datenbanken ACM Digital Library, IEEE Xplore, EBSCOhost und ScienceDirect mit Hilfe von verschiedenen Suchstrings durchsucht. Die initiale Durchführung der ersten Iteration erfolgte im Jahr 2022 unter Verwendung des folgenden Suchstrings: *(Blockchain OR distributed ledger) AND (sensitiv* OR personal OR priva* OR confidential*) AND (data sharing OR data storage OR data exchange OR off-chain OR on-chain)*. Diese Suche führte zu einer Gesamtzahl von 257 Publikationen. In einem ersten Schritt wurden Duplikate, Nachrichtenartikel sowie Veröffentlichungen, die nicht in englischer oder deutscher Sprache verfügbar waren, ausgeschlossen. Nach diesem Ausschluss verblieben 211 Publikationen, die im Hinblick auf relevante Schlüsselwörter, Abstracts und Titel analysiert wurden. Diese Analyse reduzierte die Anzahl der relevanten Arbeiten auf 53. Im Rahmen dieser Analyse konnten insgesamt 18 Veröffentlichungen identifiziert werden, die als relevant für die Fragestellung der Arbeit erachtet wurden. Diese Publikationen enthielten insgesamt 19 unterschiedliche Ansätze, welche die Grundlage für die erste Iteration bildeten (siehe Tabelle A.2 und A.3). Aufgrund des längeren Bearbeitungszeitraums der

vorliegenden Arbeit wurde in einer zweiten Iteration der identische Suchstring erneut verwendet, jedoch auf den Zeitraum von 2022 bis zum 31. Dezember 2023 beschränkt. Dies sollte sicherstellen, dass auch neuere wissenschaftliche Arbeiten in die Analyse einbezogen wurden, die nach der ersten Iteration veröffentlicht wurden. Diese Suche führte zur Identifikation von 13 zusätzlichen Ansätzen, die in die entwickelte Taxonomie integriert wurden. Diese sind in den Abbildungen A.4 und A.5 dargestellt. Aufgrund der zunehmenden wissenschaftlichen Diskussion des Blockchain-basierten Identitäts- und Zugriffsmanagements zwischen der ersten und zweiten Iteration, wurde zeitgleich zur zweiten Iteration eine dritte Literatursuche durchgeführt. Hierbei wurde der folgende Suchstring verwendet, der diese spezifischen Themenfelder abbildet: *(decentral* OR Blockchain OR distributed ledger) AND identity AND (health* OR medic*) AND (data storage OR data donation OR access management)*. Diese dritte Iteration führte zur Identifikation von sechs weiteren relevanten Publikationen (siehe Abbildung A.6 und Abbildung A.7). Insgesamt lieferten die verschiedenen Iterationen 38 relevante Ansätze und Architekturen.

Im Folgenden werden die Unterschiede der verschiedenen Ansätze im Bereich des Daten-, Zugriffs- und Identitätsmanagements detailliert dargestellt, bevor die entwickelte Taxonomie präsentiert wird. Anschließend erfolgt eine Analyse der Architekturansätze hinsichtlich der Durchführung von Sicherheitsbetrachtungen, sowie ein Überblick über den Stand der Technik in Bezug auf Designprozesse für Blockchain-basierte Anwendungen und Entscheidungsmodelle. Auf dieser Basis wird die bestehende Forschungslücke identifiziert, die als Grundlage für die vorliegende Arbeit dient.

3.1 Datenmanagement

In der identifizierten Literatur kommen verschiedene Systeme und Ansätze zum Einsatz, um sensible Gesundheitsdaten zu speichern, auszutauschen und zu verwalten. Die Mehrheit dieser Systeme konzentriert sich auf die Speicherung von ePAs [173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187,

188, 189, 190]. Diese in der Literatur untersuchten Ansätze zielen überwiegend darauf ab, medizinische Patient*innendaten zwischen Ärzt*innen verschiedener Institutionen, Patient*innen und deren Zugehörigen auszutauschen, wobei Patient*innen unterschiedliche Grade an Autonomie bei der Entscheidungsfindung eingeräumt werden. Einige dieser Ansätze legen zudem besonderen Wert darauf, den Patient*innen möglichst viel Kontrolle über ihre eigenen Daten zu ermöglichen [178, 177, 191]. Darüber hinaus beinhalten die Ansätze von Thwin und Vasupongayya [192], Li et al. [183], Zhang et al. [193], Lee et al. [194], Zou, Lv und Zhao [195], Zaghloul et al. [196], Lee et al. [197], Zhao, Yu und Yan [198], Li, Yue und Wu [199] auch die Verwaltung von eGAs. Einige Systeme speichern auch medizinische Messungen von Sensoren und IoT-Geräten, z.B. Blutzuckermessungen [200, 193, 190, 173]. Zusätzlich wurden Systeme identifiziert, die darauf ausgelegt sind, Forschenden den Zugang zu umfangreichen medizinischen Datenmengen zu ermöglichen und somit Fortschritte im medizinischen Bereich zu fördern [176, 182, 184, 201, 202, 179, 198, 180, 181, 203, 195, 193, 204]. Dies gilt insbesondere im Hinblick auf das Potenzial, das Künstliche Intelligenz (KI) in diesem Bereich bietet, z.B. um die Behandlung und Diagnose von Krankheiten zu verbessern [184, 205, 179, 202, 206] oder um Trainingsdaten für maschinelle Lernalgorithmen bereitzustellen [180, 181]. Qin, Jin und Liu konzentrieren sich speziell auf die Behandlung von Schlaganfällen. Hawig et al. [200] implementieren zwei verschiedene Systeme, um einen On-Chain- und Off-Chain-Ansatz zu vergleichen.

Im folgenden Abschnitt werden die technischen Charakteristika der verschiedenen, in der wissenschaftlichen Literatur bewährten Ansätze zum Blockchain-basierten Datenmanagement erläutert. Die Unterabschnitte befassen sich mit den spezifischen Aspekten des Datenmanagements: dem Speicherort der Daten, den verschiedenen Blockchain-Typen, der Off-Chain-Speicherung und den Sicherheitsmechanismen zur Wahrung der Datensicherheit und -privatheit bei der Speicherung.

3.1.1 Speicherort

Allgemein lassen sich die identifizierten Ansätze aus der Literatur nach dem Speicherort der sensiblen Daten in die folgenden drei Kategorien einteilen: *On-Chain*, *Off-Chain* und *Hybrid*.

Bei **On-Chain** Speicheransätzen werden Daten entweder direkt oder, im Fall von sensiblen Informationen, verschlüsselt auf einer Blockchain gespeichert. Die Menge der auf einer Blockchain speicherbaren Daten ist oft durch maximale Transaktions- und Blockgrößen begrenzt, z.B. hat Bitcoin eine Begrenzung von 40 Bytes pro Transaktion [155] und 1 Megabyte pro Block [207]. Die Blockgröße könnte erhöht werden, was jedoch zu längeren Replikationszeiten führen würde [207, 155]. On-Chain-Daten profitieren von der Unveränderlichkeit und Dezentralisierung der Blockchain, was bedeutet, dass Daten vor Manipulation und Verlust geschützt sind, jedoch nicht gelöscht werden können [183]. Ein Problem beim Speichern sensibler Daten, welche Vertraulichkeit erfordern und nicht für jeden Teilnehmende sichtbar sein sollen, ist die Transparenz der Blockchain [155]. Die identifizierten Ansätze aus der Literatur verwenden Verschlüsselung, um dessen entgegen zu wirken und damit sicherzustellen, dass nur Teilnehmende mit einem geheimen Schlüssel auf gespeicherte Daten zugreifen können [184, 200]. Dies erfordert jedoch einen zusätzlichen Schlüsselaustausch und eine Schlüsselverwaltung außerhalb der Blockchain [155]. Als Faustregel gilt, dass Daten, die kleiner als ihr Hash-Wert sind, On-Chain gespeichert werden sollten, während größere Daten Off-Chain abgelegt werden sollten [155]. Darüber hinaus sollten Daten, die geändert oder gelöscht werden müssen, nicht On-Chain gespeichert werden. Ein On-Chain-Ansatz wird von Zhang und Lin [184] sowie Hawig et al. [200] verwendet.

Die Mehrheit der Ansätze speichert sensible Daten **Off-Chain** und nur Metadaten On-Chain. Diese Metadaten umfassen einen Verweis auf den Speicherort sowie einen Hash der gespeicherten Daten, um Manipulationen der Off-Chain-Daten zu erkennen [155]. Durch die On-Chain-Speicherung von Metadaten wird die Skalierbarkeit der Blockchain nicht beeinträchtigt, wenn große Datenmengen gespeichert werden. Allerdings profitieren nur die Metadaten von der Unveränderlichkeit

und Dezentralisierung der Blockchain. Während Manipulationen von Off-Chain-Daten anhand ihres Hash-Werts erkannt werden können, ist es nicht möglich, sie zu verhindern. Daher sind zusätzliche Sicherheitsmaßnahmen erforderlich, um Off-Chain-Daten vor unbefugtem Zugriff zu schützen [155]. Zusammenfassend eignet sich die Off-Chain-Speicherung besonders gut für große Datenmengen oder für Daten, die in Zukunft geändert oder gelöscht werden müssen. In der erfassten Literatur zur Speicherung von Gesundheitsdaten ist dies der häufigste Ansatz, da Gesundheitsdaten über den gesamten Lebenszyklus gesammelt werden und aufgrund der DSGVO auf Anfrage gelöscht werden können.

Die **hybride** Speicherung ist eine Kombination aus dem On-Chain und Off-Chain Speicheransatz. Statt sensible Daten ausschließlich On-Chain oder Off-Chain zu speichern, werden sensible Daten je nach Anforderungen entweder On-Chain oder Off-Chain gespeichert, was zu erhöhter Flexibilität und besserer Skalierbarkeit im Vergleich zu einem reinen On-Chain-Speicheransatz führt [183]. Um zu entscheiden, ob Daten On-Chain oder Off-Chain gespeichert werden sollen, ist ein Entscheidungskriterium erforderlich. Im Fall von Li et al. [183] werden Multimedia-Dateien, wie Bilder beispielsweise, Off-Chain und Textdateien On-Chain gespeichert. Daten könnten auch nach ihrer Größe unterschieden werden. Wenn weder ein On-Chain- noch ein Off-Chain-Ansatz ausreicht, z.B., wenn Textdateien unveränderlich und für immer erhalten bleiben sollen, aber auch große Bilddateien gespeichert werden müssen, sollte ein Hybridansatz in Betracht gezogen werden [183].

3.1.2 Blockchain-Typ

In der erfassten Literatur werden verschiedene Blockchain-Typen verwendet, die nach Art des Netzwerkmanagements und der Berechtigungen der Teilnehmenden den folgenden drei Kategorien zugeordnet werden können: *öffentliche Blockchain*, *private Blockchain* und *Konsortium-Blockchain*. Diese Kategorien lassen sich weiter in zwei Haupttypen unterteilen: *permissioned Blockchains* und *permissionless Blockchains*. Anstelle anonymer öffentlicher Teilnahme kann eine

Blockchain als **permissioned Blockchain** organisiert sein, bei der eine oder mehrere Autoritäten die Teilnahme kontrollieren. Dies bedeutet, dass Berechtigungen erforderlich sind, um dem Netzwerk beizutreten, Informationen aus der Blockchain zu lesen, Transaktionen zu initiieren oder Mining zu betreiben. Im Gegensatz dazu ermöglicht eine **permissionless Blockchain** jedem ohne Genehmigung die Teilnahme und das Ausführen dieser Aktivitäten. [155]

Öffentliche Blockchains, wie z.B. Bitcoin, sind dezentral, öffentlich und häufig permissionless d.h. nicht von einer zentralen Instanz kontrolliert. Sie verfügen über offene Netzwerke, in die jeder eintreten und Transaktionen ohne Genehmigung verifizieren kann [208]. Da sich die Netzwerkteilnehmenden nicht kennen und somit kein gegenseitiges Vertrauen besteht, werden Konsensmechanismen eingesetzt, welche den korrekten Betrieb der Blockchain sicherstellen und den Teilnehmenden Anreize bieten, sich korrekt zu verhalten und so eine Manipulation der Blockchain zu verhindern [209]. Oft wird dazu PoW verwendet, bei dem Teilnehmende für das Überprüfen von Transaktionen und das anschließende Erstellen korrekter Blöcke mit Kryptowährung belohnt werden [155]. Um einen Block anzuhängen, muss ein kryptographisches Puzzle gelöst werden, was die Rate der erzeugten Blöcke stark einschränkt. Diese begrenzte Blockrate führt auch zu einer begrenzten Transaktionsverarbeitungsrate, wodurch Transaktionen im Minutenbereich validiert werden müssen. Das Lösen des Puzzles erfordert zudem einen erheblichen Rechenaufwand, was zu einem hohen Energieverbrauch des Netzwerks führt. [155]

Eine öffentliche Blockchain wird z.B. verwendet, um Nutzende die Kontrolle über ihre eigenen Daten zu ermöglichen und Abhängigkeiten von Gesundheitseinrichtungen zu minimieren [178]. Wenn Dezentralisierung im Vordergrund steht, sind öffentliche Blockchains besonders geeignet.

Private Blockchains sind permissioned und werden von einer einzigen Organisation verwaltet und betrieben [208], was eine hohe Konfigurationsflexibilität ermöglicht. Die verantwortliche Organisation entscheidet, wer neue Blöcke erstellen kann, authentifiziert Teilnehmende und kontrolliert den Blockchain-Zugriff,

indem sie Berechtigungen an Netzwerkteilnehmende vergibt [155]. Da die Teilnehmenden verifiziert und autorisiert sind, neue Blöcke zu erzeugen, können effizientere Konsensalgorithmen wie PBFT (vgl. Abschnitt 2.3.8) verwendet werden [155]. In der untersuchten Literatur wird z. B. eine private Blockchain von Hanley und Tewari [180] verwendet, um eine Plattform für anonymisierte Maschinendaten zu erstellen. In diesem Fall fungiert die Regierung als einziger Anbieter und zentrale Kontrollinstanz.

Der dritte Typ von Blockchains sind **Konsortium-Blockchains** [208], welche ähnlich wie private Blockchains permissioned sind, jedoch von mehreren Organisationen gemeinsam betrieben werden. Die Erstellung und Validierung von Blöcken erfolgt durch vorautorisierte Netzwerkknoten. [155] Konsortium-Blockchains werden z. B. im System von Wang et al. [182] verwendet, das den Austausch von Gesundheitsdaten zwischen mehreren Gesundheitsdienstleistenden ermöglicht. Dafür stellt jeder Gesundheitsdienstleistende einen bevollmächtigten Leader-Knoten bereit, um Transaktionen und Blöcke zu verifizieren. Daraghmi et al. [207] schlagen einen Anreizmechanismus vor, der in den Proof of Authority (PoA) Konsensalgorithmus integriert ist, um zu entscheiden, welche Netzwerkknoten und damit welche Leistungserbringende für die Validierung verantwortlich sind und neue Knoten oder die Erstellung von Blöcken basierend auf der Qualität der ePAs der Leistungserbringenden hinzufügen. Zusammenfassend wird eine Konsortium-Blockchain verwenden, wenn die Blockchain und das Netzwerk durch mehrere Organisationen gemeinsam verwaltet werden sollen.

Diese drei Arten von Blockchains können entweder durch die **Anbindung an eine öffentliche und beliebte Blockchain** auf Transaktionsebene oder durch Verwendung mehrerer privater Blockchains kombiniert werden [155]. Bei der ersten Methode wird der Hashwert der verwendeten Blockchain regelmäßig auf einer öffentlichen und populären Blockchain (z.B. Ethereum) gespeichert, um von deren Vertrauen und Sicherheit zu profitieren [155]. Med-PPPHIS [202] und DeepLinQ [181] verwenden diese Methode, um Manipulationen ihrer privaten oder Konsortium-Blockchains zu erkennen. Zur Verbesserung der Skalierbarkeit können Transaktionen einer einzelnen Blockchain auf **mehrere private Blockchains** aufgeteilt und verteilt werden, die durch eine gemeinsame Blockchain verbunden

sind [155]. Dies wird z. B. von Zhang und Lin [184] angewendet, um mehrere Krankenhäuser zu verbinden, wobei jedes Krankenhaus eine eigene private Blockchain mit eigenen Daten und eine gemeinsame Konsortium-Blockchain als Index der Daten der privaten Blockchain pflegt. Jedes Krankenhaus behält dabei die Kontrolle über die eigenen Daten. Jener Ansatz eignet sich entsprechend, um bei einer On-Chain Speicherung die Skalierbarkeit zu verbessern, indem eine Aufteilung der Daten auf mehrere Blockchains erfolgt. Hierzu müssen die gespeicherten Daten jedoch trennbar sein.

3.1.3 Off-Chain-Speicherung

Die Verwendung einer Off-Chain-Speicherung erfordert die Auswahl eines geeigneten Speicherorts. In den untersuchten Ansätzen werden Daten entweder *dezentral, zentral* oder *verteilt* gespeichert. Ein **dezentraler Ansatz** wird hauptsächlich zur Speicherung von ePAs verwendet, da sie von individuellen Gesundheitsdienstleistenden innerhalb ihrer eigenen Infrastruktur erstellt und verwaltet werden [175]. Die Blockchain wird in diesem Fall verwendet, um diese bestehenden Datenspeicher der einzelnen Gesundheitsdienstleistenden zu verbinden. Gespeicherte Daten können dann in den Datenbanken der einzelnen Dienstleistenden durch Abfrage des Speicherorts der angeforderten Daten aus der Blockchain lokalisiert werden [175, 177, 179, 150]. Durch die Verwendung vorhandener Datenbanken entsteht kein neuer zentraler Angriffspunkt (*engl. Single Point of Failure*) und es entstehen keine zusätzlichen Speicherkosten [177]. Globale Abfragen aller Datenbanken der Leistungserbringenden können jedoch aufgrund von Unterschieden in Hardware und Software schwierig sein [209, 210].

Als Alternative zur dezentralen Speicherung kann ein **zentraler Datenspeicher** verwendet werden, um Gesundheitsdaten mehrerer Leistungserbringender zu speichern. Ein zentraler Datenspeicher ermöglicht eine einfache Einrichtung, Datenverwaltung und Zugriffsmanagement, stellt jedoch einen zentralen Angriffspunkt dar [180, 202]. Zentrale Speicher werden von einer einzigen Partei bereitgestellt, z.B. der Regierung oder Cloud-Anbietern [180]. Da diese auf die gespeicherten

Daten zugreifen können, müssen diese vertrauenswürdig sein [182]. ePAs, die Daten verschiedener Gesundheitsdienstleister enthalten und vom Datensubjekt selbst verwaltet werden, werden häufig in einem zentralen Cloud-Speicher gespeichert [192]. Allerdings können ePAs auch zentral gespeichert werden, um Unabhängigkeit von den Datengenerierenden zu erreichen, z.B. um eine staatlich kontrollierte Forschungsplattform aufzubauen, die anonymisierte Kopien von ePAs bereitstellt [180].

Für **verteilte Off-Chain-Speicherung** werden Peer-to-Peer-Netzwerke (P2P-Netzwerke) verwendet. In P2P-Netzwerken werden Daten im lokalen Speicher der einzelnen Netzwerkteilnehmenden gespeichert. Daten können dann zwischen diesen ohne zentralen Server verteilt werden [155]. Vorteile sind die Vermeidung eines zentralen Angriffspunktes, hohe Speicherdurchsatzraten und kurze Lesezeiten, was für den Austausch großer Datenmengen vorteilhaft ist [173, 200]. Verteilte Speicher haben jedoch tendenziell eine höhere Einrichtungs- und Verwaltungskomplexität [200]. Beispiele für P2P-Protokolle sind das InterPlanetary File System (IPFS) und Dat. IPFS ermöglicht das Abrufen von Daten über deren Hash und wird von Hawig et al. [200] verwendet, um medizinische Sensordaten zu speichern, mit dem Ziel, ein unabhängiges System zu schaffen, sowie von Nguyen et al. [173] zur Speicherung von ePAs. Um den Netzwerkzugang zu kontrollieren und die Löschung gespeicherter Daten durchzusetzen, muss ein privates IPFS-Netzwerk verwendet werden [200]. Anstatt ein separates P2P-Netzwerk zu betreiben, ist es auch möglich, Daten direkt im lokalen Speicher der Blockchain-Knoten zu speichern, z.B. indem Daten in Fragmente aufgeteilt und zwischen den Knoten verteilt werden [202].

3.1.4 Sicherheitsmechanismen zur Datenspeicherung

Eine Maßnahme zum Schutz sensibler Daten und Wahrung der Vertraulichkeit stellt die **Verschlüsselung** dar, da ein Schlüssel zur Verwendung der Daten notwendig ist und die Daten ohne diesen Schlüssel nicht zugänglich sind [211]. Die Daten können entweder **asymmetrisch** mit einem öffentlichen Schlüssel (PK)

des Empfangenden oder **symmetrisch** mit einem gemeinsamen Schlüssel verschlüsselt werden, den alle Personen, die auf die Daten zugreifen dürfen, teilen [155, 161]. Um asymmetrisch verschlüsselte Daten mit anderen Personen zu teilen, müssen sie entschlüsselt und dann mit dem öffentlichen Schlüssel des Empfangenden neu verschlüsselt werden. Im Fall von Nguyen et al. [173] wird diese Neuverschlüsselung auf Benutzendenanfrage von einem zentralen Server durchgeführt. **Hybride** Methoden, z.B. Elliptic Curve Integrated Encryption Scheme (ECIES), verwenden asymmetrische Methoden, um symmetrische Schlüssel auszutauschen [178, 183, 200, 202]. Symmetrische Verschlüsselungsverfahren sind aufgrund ihrer bis zu tausendfachen Geschwindigkeitsvorteile gegenüber asymmetrischen Verfahren und ihrer Resistenz gegen chosen-plaintext-Angriffe besonders geeignet für die Datenverschlüsselung [211]. Der verwendete Schlüssel kann effizient zwischen mehreren Parteien geteilt werden [201], was sowohl das Entschlüsseln gespeicherter Daten als auch die erneute Verschlüsselung mit dem Schlüssel des Empfangenden überflüssig macht. Vor der Nutzung muss jedoch ein geheimer Schlüssel vereinbart werden, um die Sicherheit der Daten zu gewährleisten [211].

Drei der untersuchten Ansätze [176, 180, 181] stellen medizinische Daten als Trainingsdaten für maschinelles Lernen bereit und verzichten daher bewusst auf Verschlüsselung, da maschinelle Lernalgorithmen nicht mit verschlüsselten Daten arbeiten können [180]. Zusammenfassend lässt sich sagen, dass Verschlüsselung nützlich ist, um die Vertraulichkeit sensibler Daten zu gewährleisten, wenn keine Berechnungen auf den Daten erforderlich sind. Jayasinghe et al. [186] verwenden **Passwortschutz** als einfachen Weg zum Schutz der gespeicherten Dateien.

3.2 Zugriffsmanagement

Zugriffsmanagement ist wichtig, um sensible Daten vor unbefugtem Zugriff zu schützen [184, 150]. In diesem Abschnitt werden die verschiedenen wissenschaftlich anerkannten Ansätze zum Zugriffsmanagement detailliert vorgestellt. Die

Unterabschnitte behandeln spezifische Aspekte des Zugriffsmanagements, einschließlich der Autorität zur Zugriffsverwaltung, der Zugriffskontrollstrategien und der Sicherheitsmechanismen für die Zugriffskontrolle.

3.2.1 Autorität zur Zugriffsverwaltung

Die in der Literatur identifizierten Ansätze zur Verwaltung der Zugriffsrechte können danach differenziert werden, welche Entität die Autorität hat, Entscheidungen über Zugriffe zu treffen. Dies kann beispielsweise der oder die Patient*in als *Datensubjekt*, *das System* oder eine *geteilte Autorität* sein. Zudem wird in der Literatur unterschieden, in Bezug auf welche Daten eine Entität die Autorität zur Zugriffsverwaltung hat, wobei differenziert wird, wer die Autorität beim allgemeinen Gesundheitsdatenmanagement (Autorität Allgemein) hat und wer bei der Verwaltung von Daten für die Forschung (Authorität Forschung und Entwicklung) zuständig ist. Die meisten der identifizierten Systeme zielen darauf ab, die Zugriffskontrollverwaltung vollständig dem **Datensubjekt** zu überlassen [207, 183, 184, 200, 180, 182, 202, 181, 175, 177, 179, 192, 173, 174, 196, 188, 190, 204, 186, 195, 185, 212, 193, 187, 191, 197]. In Dagher et al. [208] können Datensubjekte selbst entscheiden, wer Zugriff auf deren Psychotherapie-Notizen erhält. Zusätzlich können Dritte durch das Datensubjekt autorisiert werden, ebenfalls die Kontrolle über die Zugriffsverwaltung zu erhalten [208]. In diesem Fall handelt es sich um eine **geteilte Autorität**. Im Ansatz von Lee et al. [194] wird die Autorität der Datensubjekte geteilt, indem Ärzt*innen, welche die Daten erstellt haben, ein Re-Verschlüsselungsschlüssel ausgestellt wird. Dieser Schlüssel ermöglicht den Ärzt*innen den Zugriff auf die Daten, falls das Datensubjekt nicht ansprechbar ist [194]. Huang et al. [203] stellen einen Ansatz vor, bei dem das Maß an Zugriff auf Gesundheitsdaten eines Datensubjekts durch eine der vordefinierten Rollen bestimmt wird, welche einem Teilnehmenden im System zugewiesen sind [203]. In diesem Fall wird die Verwaltung vollständig vom **System** bestimmt. Gupta et al. [189] verfolgen einen ähnlichen Ansatz. Lin et al. [213] präsentieren einen Ansatz, bei dem den Krankenhausadministrierenden das Recht erteilt wird,

den Zugriff auf die in der Cloud gespeicherten Daten zu gewähren, da alle Daten mit dem privaten Schlüssel des Krankenhauses entschlüsselt werden können. Physiologische Informationen, die durch das Smartphone, die Smartwatch oder das private medizinische Gerät eines Datensubjekts generiert wurden, werden von den Datensubjekten selbst mit dem öffentlichen Schlüssel des Krankenhauses verschlüsselt, bevor sie in die Cloud des Krankenhauses hochgeladen werden [213].

3.2.2 Zugriffskontrollstrategie

Die Zugriffskontrollstrategie beschreibt, auf welcher Grundlage Zugriff gewährt wird. In der Literatur werden Zugriffskontrollstrategien in zwei Hauptgruppen unterteilt: *Discretionary Access Control (DAC)*, bei der die Zugriffskontrolle auf der Ebene einzelner Entitäten im System erfolgt, und *Non-Discretionary Access Control (NDAC)* oder einer *hybride Kombination* aus beiden [163].

DAC ist eine identitätsbasierte Zugriffskontrollstrategie, die es einem Verwaltenden ermöglicht, individuell Zugriffsrechte für andere Entitäten zuzuweisen. Diese Art von Zugriffskontrollstrategie wird üblicherweise mit einer Access Control List (ACL) implementiert und gewährt dem Datenverwaltenden die größte Flexibilität, Zugriffsrechte für Einzelpersonen zu gewähren und zu widerrufen. Diese Freiheit geht jedoch mit einem höheren Verwaltungsaufwand einher und ermöglicht keine Überwachung zur Sicherstellung einer sicheren Zugriffsverwaltung. [163] Ein Beispiel für eine solche Lösung ist der Ansatz von Azaria et al. [177] einer ePA. Jede ePA hat hierbei eine eigene ACL. Beim Zugriff auf Off-Chain-Daten wird dann eine signierte Anfrage an einen zentralen Server gesendet, der als Orakel fungiert und die notwendigen Smart-Contract-Funktionen aufruft, um zu überprüfen, ob der Anfragende autorisiert ist oder nicht. Dies ist der Fall, wenn die entsprechenden Berechtigungen für den Anfragenden auf der Blockchain existieren. Benutzende werden mit ihrer Blockchain-Adresse identifiziert, und die Signatur der Zugriffsanfrage garantiert, dass der Anfragende der Besitzende der Adresse ist. Die Blockchain-Adresse kann mit Hilfe eines zusätzlichen Smart

Contracts, der z. B. eine zugehörige von der Regierung ausgestellte Identifikationsnummer, wie eine Sozialversicherungsnummer, speichert [177], auf eine reale Identität abgebildet werden. Auch der Name der Person [181] oder der Wohnort [173] können zu diesem Zweck gespeichert werden. Um Berechtigungen zu widerrufen, kann die entsprechende Variable des Smart Contracts aktualisiert werden [177].

Beim **NDAC** werden Zugriffe erteilt, indem Regeln definiert werden oder Logik angewendet wird, um die notwendige Abstraktion von einzelnen Entitäten zu bilden [163]. Dies geschieht beispielsweise durch die Definition von Rollen und deren Zugriffsberechtigungen. Die Definition von solchen Regeln ermöglicht mehr Flexibilität, da die Berechtigungen nicht für jede Person einzeln vergeben werden müssen, sondern z. B. für alle Ärzt*innen in einem bestimmten Krankenhaus in einer einzigen Zugriffsrichtlinie. Die NDAC Strategien lassen sich ebenfalls in drei weitere Kategorien unterteilen: *Obligatorische Zugriffskontrolle* (MAC - *engl. Mandatory Access Control*), *Rollenbasierte Zugriffskontrolle* (RBAC - *engl. Role-Based Access Control*) and *Regelbasierte Zugriffskontrolle* (RuBAC - *engl. Rule-Based Access Control*)[163]. Beim Konzept des **MAC**, welches Lee et al. [194] und Zaghloul et al. [178] verwenden, wird jeder Datei und jeder Entität ein „Level“ zugewiesen. Der Zugriff wird nur gewährt, wenn das Level einer Entität gleich oder höher ist als das der Datei, auf die zugegriffen werden soll [163, 214]. **RBAC** definiert den Zugriff einer Entität basierend auf dessen zugewiesenen Rolle im System [163]. Die dritte Zugriffskontrollstrategie ist die **RuBAC**, bei der der Zugriff durch Regeln oder Attribute zugewiesen wird [163]. RuBAC umfasst demgemäß die attributbasierte Zugriffskontrolle [214, 163]. Dazu müssen Teilnehmende zuerst einen zertifizierten Registrar aufsuchen, der ihre Attribute überprüft und in einem Smart Contract speichert. Benutzenden können dann Zugriffsrichtlinien für ihre Daten in einem anderen Smart Contract festlegen. Diese Zugriffsrichtlinien geben an, welche Attribute für den Datenzugriff erforderlich sind. Im Falle eines Zugriffs überprüft der Smart Contract, ob der Zugriffnehmende die gemäß der Richtlinie erforderlichen Attribute besitzt. Wenn dies der Fall ist, wird der Zugriff genehmigt und ein Schlüsselherausgebender

erstellt einen Schlüssel basierend auf den Attributen [178]. ACLs und attributbasierte Zugriffskontrolle können verwendet werden, um den Zugriff auf bestimmte Operationen zu beschränken, z. B. um nur Lese- oder Schreiboperationen zu erlauben [174, 175]. Ein Problem bei beiden ist, dass ein vertrauenswürdiges Orakel erforderlich ist, um Smart-Contract-Funktionen mit den richtigen Eingabedaten aufzurufen oder die korrekten Transaktionen auf der Blockchain zu erstellen.

Lee et al. [194] und Zaghloul et al. [196] nutzen Ansätze, die MAC als ihre Zugriffskontrolllogik verwenden [194, 196]. Xiao et al. [175], Gupta et al. [189], Hu et al. [204], Jayasinghe et al. [186], Huang et al. [203], Lee et al. [197], Lomotey, Kumi und Deters [191] verwenden eine rollenbasierte Zugriffskontrollrichtlinie. Die verbleibende identifizierte Literatur verwendet eine regelbasierte Zugriffskontrollrichtlinie bei NDAC, da diese im Gegensatz zu anderen Ansätzen eine geringere Komplexität und weniger Aufwand bei der Verwaltung von Berechtigungen aufweist, insbesondere, wenn Zugriffe einer großen Anzahl an Entitäten erteilt werden sollen [215]. Bei einer **hybriden Zugriffskontrollstrategie** wird DAC mit NDAC kombiniert, um die Vorteile beider Ansätze zu nutzen. Dies ermöglicht eine effiziente Zugriffskontrollstrategie durch NDAC sowie eine feingranulare Zugriffskontrolle auf Identitätsebene durch DAC [163]. Einen hybriden Ansatz verwenden Daraghmi et al. [207], Zhang und Lin [184], Chang et al. [181], Lee et al. [197], Li, Yue und Wu [199] und Lomotey, Kumi und Deters [191].

3.2.3 Sicherheitsmechanismen für die Zugriffskontrolle

Die betrachteten Ansätze können durch den Mechanismus, der verwendet wird, um den Zugriff auf eine Ressource zu ermöglichen, differenziert werden. Zugriffsempfangende können in diesem Zusammenhang Zugang zu einem Geheimnis in Form eines Schlüssels, eines Passworts oder eines Zugriffstokens erhalten, ebenso wie eine Dateireferenz oder die Datei selbst. In seiner einfachsten Form wird der Zugriff dadurch gewährt, dass die Ressource direkt an den Empfangenden übertragen wird. Dieser Ansatz wird von Nguyen et al. [173] angewandt. Sabu et al. [190] tauschen lediglich ein einmaliges Passwort mit dem Zugriffsberechtigten

aus. Alle Arten von Zugriffsressourcen (Schlüssel, Token, Passwort, Dateiverweis und Datei) können zusätzlich verschlüsselt übertragen werden, was zusätzliche Sicherheit gewährleistet. Thwin und Vasupongayya [192] sowie Zaghloul et al. [178] verwenden eine Public-Key-Verschlüsselung und damit eine **asymmetrische Verschlüsselung**, bei der eine verschlüsselte Datei übertragen wird, die vom Empfangenden mit dem eigenen privaten Schlüssel entschlüsselt werden kann. Im Gegensatz dazu verwendet Zhao, Yu und Yan [198] eine symmetrische Verschlüsselung, bei der eine verschlüsselte Datei übertragen wird, die vom Empfangenden mithilfe eines zuvor ausgetauschten Schlüssel entschlüsselt werden kann. Darüber hinaus beschreiben andere Ansätze den Austausch eines verschlüsselten Schlüssels, der verwendet werden kann, um die gewünschte Datei zu entschlüsseln [200, 183, 201, 202]. Fortgeschrittene Verschlüsselungsmethoden in der Literatur umfassen Ciphertext-Policy Attribute-Based Encryption (CP-ABE), die die Integration von attributbasierten Zugriffspolitiken in Verschlüsselungen ermöglicht [178]; **Proxy Re-Encryption** (PRE), die die Delegation von Entschlüsselungsrechten an Dritte ohne Offenlegung der Klartextdaten erlaubt [192, 182, 202, 194]; und Searchable Encryption (SE), die es ermöglicht, dass Verschlüsselungen Schlüsselwörter enthalten, die ohne Entschlüsselung durchsucht werden können [182, 184, 177].

Bei der attributbasierten Zugriffskontrolle durch **CP-ABE** werden zunächst ein öffentlicher Schlüssel (PK) und ein Master-Schlüssel erstellt. Der Master-Schlüssel wird verwendet, um für jeden Nutzenden einen spezifischen Schlüssel zu generieren, der die individuellen Attribute des Nutzenden berücksichtigt. Bei der Verschlüsselung von Daten wird eine Zugriffsrichtlinie definiert, die die notwendigen Attribute angibt, die zur Entschlüsselung erforderlich sind. Das Entschlüsseln der Daten mit dem entsprechenden SK ist nur möglich, wenn der Nutzende über die erforderlichen Attribute verfügt [178]. Zur Verwaltung der Nutzendenattribute kann beispielsweise ein Smart Contract verwendet werden, der die Zuordnung der Nutzenden zu ihren Attributen speichert. [178]

Bei der Verwendung von PRE werden Daten mit dem öffentlichen Schlüssel (PK) des Empfangenden (z. B. eines Servers, der die Daten speichert) und dem geheimen Schlüssel (SK) des Sendenden (z. B. des Datenbesitzenden) verschlüsselt.

Der Empfangende kann dann mit seinem SK, den PK des Sendenden und dem PK einer dritten Partei einen Re-Encryption-Schlüssel erstellen. Dieser ermöglicht es der dritten Partei, das Chiffrat zu entschlüsseln, ohne dass der Server Zugang zum Inhalt der Daten erhält [182]. Das Schema erlaubt es also, eine für Partei A verschlüsselte Nachricht in eine für Partei B verschlüsselte Nachricht umzuwandeln, ohne den Inhalt der Nachricht offenzulegen. Es ist keine Ent- und Verschlüsselung zur Übertragung der Daten erforderlich. [182] Es ist auch möglich, ACL mit PRE zu kombinieren, indem zusätzlich Re-Verschlüsselungsschlüssel für autorisierte Personen gespeichert werden, die eine Re-Verschlüsselung der verschlüsselten Off-Chain-Daten ermöglichen [182, 192, 202, 207, 208].

Die Searchable Encryption (SE), wie von Wang et al. [182] und Zhang und Lin [184] beschrieben, ermöglicht es, verschlüsselte Schlüsselwörter in der Blockchain zu speichern. Diese verschlüsselten Schlüsselwörter können durchsucht werden, ohne dass sie zuvor entschlüsselt werden müssen. Ein Schlüsselwort w wird dabei mit einem PK verschlüsselt. Mit dem entsprechenden SK kann eine Hintertür für ein Schlüsselwort w' erzeugt werden, um zu überprüfen, ob $w = w'$, ohne den geheimen Schlüssel preiszugeben. [184]

Bei dem Ansatz von Liu et al.[174] und Zhang et al. [179] wird die Referenz auf den Speicherort der Off-Chain-Daten asymmetrisch verschlüsselt, wodurch nur der Besitzer des entsprechenden geheimen Schlüssels (SK) herausfinden kann, wo die Daten gespeichert sind. Ein Nachteil dieses Ansatzes besteht darin, dass für jeden zugriffsberechtigten Nutzenden eine eigene verschlüsselte Referenz gespeichert werden muss. Hierbei kann beispielsweise ein Smart Contract verwendet werden [179]. Allgemein kann die Blockchain zur Umsetzung einer Zugriffskontrolle verwendet werden, indem die Zugriffsberechtigungen in Smart Contracts abgelegt werden. [177, 181, 174, 173, 176, 184]. Falls die Blockchain keine direkte Unterstützung für Smart Contracts bietet, können diese Informationen alternativ in Transaktionen gespeichert werden [175, 182]. Thwin und Vasupongayya [192] verwenden keine Blockchain, sondern einen Server zur Verwaltung von Berechtigungen. Die Blockchain dient nicht nur zur Verwaltung von Zugriffsberechtigungen, sondern kann auch zur Protokollierung von Zugriffen genutzt werden

[177, 181, 184, 173, 174, 175, 176, 178, 182, 192, 202, 207]. Die dokumentierten Zugriffe sind durch die Blockchain unveränderbar und vor Manipulationen geschützt [216]. Diese Aufzeichnungen können verwendet werden, um Zugriffe auf Daten zu verfolgen und die Verantwortenden bei Missbrauch der Daten zu identifizieren [192]. Dies erweist sich insbesondere bei Datenleaks als nützlich [174].

FHIRChain [179] verwendet eine **tokenbasierte Zugriffskontrolle**, die kein vertrauenswürdiges Orakel benötigt. Um einer Person Zugriff auf die Daten zu gewähren, muss der Benutzende zunächst ein Zugriffstoken erstellen, indem er die Referenz zu den gewünschten Off-Chain-Daten signiert und sie mit dem öffentlichen Schlüssel (PK) der Person verschlüsselt. Die digitale Signatur stellt die Authentizität des Tokens sicher. Das Token wird dann in einem Smart Contract gespeichert. Autorisierte Personen können den Smart Contract aufrufen, um ihr gespeichertes Zugriffstoken zu erhalten und es mit ihrem privaten Schlüssel (SK) zu entschlüsseln, um den Speicherort zu erfahren. Der Smart Contract bietet auch ein unveränderliches Protokoll der Token-Erstellungen und -Verwendungen [184]. Das Widerrufen des Zugriffs ist jedoch schwieriger, da Tokens nicht von der Blockchain gelöscht werden können. Im Allgemeinen kann der Zugriff auf bestimmte Teile der Aufzeichnungen zum Zweck der Datenminimierung beschränkt werden [174, 176, 178, 177, 184, 181, 207], beispielsweise durch vordefinierte SQL-Abfragen [177] oder temporär [174, 175, 176, 184, 207], z. B. durch die Verwendung einer temporären URL [175]. Eine vollständige Aufhebung des Zugriffs nach der Autorisierung ist nicht möglich, da ein Abbild der Daten gemacht werden kann [150].

3.3 Identitätsmanagement (IdM)

In der untersuchten Literatur werden verschiedene Identitätsmanagementansätze betrachtet, die sich nach dem Ort und der Autorität unterscheiden, unter der das Benutzendenauthentifizierungsgeheimnis aufbewahrt wird [161]. Diese Ansätze

lassen sich mit Betrachtung des Stand der Technik in *zentrales Identitätsmanagement*, *föderiertes Identitätsmanagement*, *benutzerzentriertes Identitätsmanagement* und dezentrales Identitätsmanagement unterteilen, wobei das dezentrale Identitätsmanagement weiter in *Self-Sovereign Identity (SSI)* und *Decentralized Trusted Identity (DTI)* gegliedert wird [161].

Zentralisierte Identitätsmanagementsysteme bedienen sich an einem Client/Server-Modell, bei dem Benutzendeninformationen an einem zentralen Speicherort abgelegt und dort ebenfalls die Benutzendenauthentifizierung durchgeführt wird [161]. Benutzendenanmeldeinformationen, wie Passwort und Benutzendenname, werden in diesem zentralen Identitätsdienst gespeichert, wodurch die Kontrolle über die Identitäten beim Identitätsdienst liegt. Dies ermöglicht dem Identitätsdienst, den Benutzendenzugriff zu beschränken oder betrügerisch auf Benutzendendaten zuzugreifen. Aufgrund dieser Tatsache bieten zentrale Ansätze einen schwachen Datenschutz, da sie einen zentralen Angriffspunkt für böswillige Dritte darstellen. Benutzende haben daher wenig Kontrolle über ihre eigenen Informationen bezüglich dem Identitätsdienst. Zentralisierte IdM-Systeme eignen sich besonders zur Verwaltung einer großen Anzahl von Benutzenden und können auf verschiedene Weise implementiert werden, wie z.B. mittels Single Sign-On, Identifier Domain und Meta-identifier Domain. [161] Ein derartiges IdM wird von einer Vielzahl der in der Literatur beschriebenen Ansätze verwendet [182, 173, 192, 176, 188, 189, 190, 199, 206, 191]. Es ist zu beachten, dass selbst wenn die zugrundeliegenden Identitätsinformationen mittels Blockchain gespeichert werden, ein dezentrales Identitätsmanagementsystem als zentrales Identitätsmanagementsystem betrachtet werden muss, wenn der private Schlüssel eines Benutzenden durch ein System in einer zentralisierten Datenbank verwaltet wird [189, 173]. In diesem Fall liegt die Kontrolle über das Benutzendenkonto weiterhin beim System selbst und nicht bei den Benutzenden. Dies ist bei den Lösungen von Nguyen et al. [173] und Gupta et al. [189] der Fall.

Föderierte Identitätsmanagementsysteme teilen eine Identität, die von mehreren Dienstanbietenden einer Föderation verwendet wird. Benutzende können

sich bei einem der Partner der Föderation authentifizieren und werden automatisch bei allen anderen authentifiziert. Dies verringert die Menge an Authentifizierungsinformationen, die sich Benutzende merken müssen, und vereinfacht ihre Interaktionen mit diesen Dienstanbietenden. Dienstanbietende der Föderation verknüpfen die föderierte Identität der Benutzenden mit deren lokaler Identität, die in einem eigenen zentralen Speicherort gespeichert ist. Obwohl Benutzende von einer verbesserten Benutzendenerfahrung in der Föderation profitieren, haben diese dennoch nicht mehr Kontrolle über ihre Identität als in einem zentralisierten Identitätsmanagementsystem. Außerdem ist gegenseitiges Vertrauen hinsichtlich der Richtigkeit der Benutzendeninformationen notwendig. Mehrere Protokolle und IdM-Systeme implementieren das föderierte Modell, darunter die Security Assertion Markup Language (SAML) [217] und das Open-Source-Projekt Shibboleth [218]. [161]

Die Idee des **Benutzerzentrierten Identitätsmanagement** besteht darin, einen Mechanismus bereitzustellen, der es den Benutzenden ermöglicht, verschiedene Kennungen und Anmeldeinformationen an einem Ort zu speichern und dennoch die vollständige Kontrolle über ihre digitale Identität zu erhalten, sodass eine Identität von einem Dienst auf einen anderen übertragen werden kann. Hierzu kann ein persönliches Authentifizierungsgerät (*engl. personal authentication device - PAD*) verwendet werden. Dieses Gerät ermöglicht es Benutzenden, sich lokal zu authentifizieren, da es alle für verschiedene Dienstanbietende benötigten Authentifizierungsinformationen speichert. Die Authentifizierungsinformationen werden den einzelnen Dienstanbietenden nach erfolgreicher Authentifizierung übermittelt. Dieses Modell legt nicht fest, wie die Identitätsinformationen bei jedem Dienstanbietenden gespeichert werden, und kann daher unabhängig von der Art der Authentifizierung und des Zugriffs mit jedem Identitätsmodell kombiniert werden. Es verbessert die Benutzendenerfahrung und bietet deutlich mehr Kontrolle über die eigene Identität. Große Unternehmen wie Google und Facebook nutzen es unter dem Namen Identity 2.0. Jedoch weist das Modell weiterhin Schwächen im Hinblick auf Autonomie und Autorität auf, da die Benutzenden nach wie vor nicht die volle Kontrolle über ihre Identität haben und bei einer Löschung ihres Kontos durch den Dienstanbieter keinen Zugriff mehr auf ihre

Identität besitzen. [161]
In der Literaturrecherche wurden keine föderierten und benutzerzentrierten Ansätze in Kombination mit Blockchain identifiziert. Jene Ansätze stellen jedoch den Stand der Technik dar [161].

Wie vorab beschrieben sind die aktuellen Herausforderungen im Bereich des Identitätsmanagements eng mit dem Konzept des zentralisierten Designs verbunden, da mit der zunehmenden Datensammlung und -speicherung Sicherheits- und Datenschutzbedenken aufkommen [161]. Das Ziel der dezentralen Identität besteht darin, die Verantwortung den Benutzenden zu übertragen und ihnen die Befugnis und Fähigkeit zu geben, ihre eigene Identität zu kontrollieren und zu verwalten. Hierzu wird der Schlüssel zur Identitätsbeweisführung dezentral gespeichert. Dezentrale Identitätsmanagementsysteme wie **Decentralized Trusted Identity (DTI)** oder **Self-Sovereign Identity (SSI)** nutzen dementsprechend die Distributed-Ledger-Technologie zur Speicherung und Authentifizierung von Benutzendenidentitäten, wodurch eine dezentrale und manipulationssichere Datenspeicherung ermöglicht wird, bei der keine Organisation das Netzwerk kontrolliert oder es verändern kann. [161]

Im Kontext von SSI speichern Datensubjekte ihre digitale Identität und Nachweise selbst, etwa lokal in einer digitalen Brieftasche (*engl. wallet*). Diese lassen sich dann von Vertrauensdiensten überprüfen und gegenüber Diensten präsentieren. Dazu kommen Verifiable Credentials (VC), Decentralized Identifier (*engl. Decentralized Identifier - DID*) und DID-Dokumente zum Einsatz. Ein Verifiable Credential enthält neben einer Behauptung (*engl. claim*) über das Identitätssubjekt zusätzliche Informationen darüber, wann das Credential von wem für wen unter welchen Umständen ausgestellt wurde. Zur Verifizierung der Identität des Identitätssubjekts und des Credential-Ausstellenden (*engl. issuer*) wird ein dezentraler Identifikator (*engl. Decentralized Identifier - DID)* referenziert. Hinter diesem stehen in einem dezentralen Register Informationen, wie mit dem jeweiligen Akteur sicher kommuniziert werden kann und wie digitale Signaturen überprüft werden können. Die Verifikation erfolgt durch kryptographische Verfahren wie digitale Signaturen, wobei ein öffentlicher Schlüssel in dem DID-Dokument in einer Distributed-Ledger-Technologie (DLT) gespeichert wird. Die Credentials

selbst werden entsprechend dem Prinzip der Datensparsamkeit nicht in der DLT gespeichert. Ein solcher SSI-basierter Ansatz wird von Zhao, Yu und Yan [198] angewandt. [161]

Im Gegensatz dazu bleibt bei DTI der Identitätsausstellende zentralisiert und die Registrierung neuer Benutzende hängt von einer bestehenden vertrauenswürdigen sowie verifizierten Identität einer Drittpartei ab, wie z.B. einer Regierungs-ID aus dem Reisepass. Dadurch bietet SSI ein höheres Maß an Anonymität im Vergleich zu DTI. Darüber hinaus wird bei DTI die DLT verwendet, um die Identitätsbestätigung zur späteren Validierung durch Dritte zu speichern. Der Identitätsausstellende stellt der empfangenden Entität ein Zeugnis über die Gültigkeit der Daten zur Verfügung, während alle Anmeldeinformationen verschlüsselt und lokal durch das Identitätssubjekt selbst gespeichert werden. Im Gegensatz zu SSI ist hierbei kein dezentrales Register notwendig. Allerdings ist für die Identitätsprüfung stets eine dritte Instanz involviert. [161]

Die meisten bei der Literaturrecherche identifizierten Ansätze im Bereich des dezentralen Identitätsmanagements verwenden DTI [200, 179, 181, 175, 208, 207, 180, 178, 202, 204, 186, 195, 194, 185, 196]. Azaria et al. [177] und Chang et al. [181] verwenden eine offizielle Regierungs-ID als Referenz für die vertrauenswürdige Identität einer Drittpartei. Hu et al. [204] verallgemeinern diese Drittparteireferenz als realweltliche Identität. Jayasinghe et al. [186] schlagen vor, die Amazon Recognition API zur Validierung ihrer "Know-Your-CustomerVerifizierung zu nutzen. Lee et al. [194] erwähnen, dass ihre Zertifizierungsstelle eine Zuordnung zwischen der realen Identität und anonymen Benutzendenidentitätsschlüsseln innerhalb der Blockchain speichert, was das Identitätssystem als DTI klassifiziert. Nach der Registrierung bei dem vorgeschlagenen SPChain-System von Zou et al. [195] erhalten Patient*innen ihre eigene Blockchain-Adresse und ihr persönliches Blockchain-Schlüsselpaar. Die Registrierung erfordert die medizinische Aktennummer der Patient*innen, dessen Alter und weitere ergänzende Informationen, die beim ersten Besuch einer medizinischen Einrichtung gesammelt werden [195]. Die medizinische Aktennummer der Einrichtung dient als vertrauenswürdige Identifikation einer Drittpartei, wodurch dieses Identitätsmanagementsystem als DTI eingestuft wird. Huang et al. [203] stellen einen Ansatz vor,

bei dem die Registrierung von Patient*innen, Krankenhäusern und Forschungseinrichtungen mithilfe der Zertifizierungsstelle von Hyperledger Fabric und ihrer Blockchain-Adresse erfolgt. Es wird jedoch nicht spezifiziert, ob die Blockchain-Adresse auf eine realweltliche Identität abgebildet wird, weshalb dieser Ansatz ohne diese Verbindung weder als DTI- noch als SSI-basiert klassifiziert werden kann [203]. Dieser Sachverhalt ist in mehreren identifizierten Ansätzen zu finden, bei denen ausschließlich die Blockchain-Adresse zur Identifikation verwendet wird und der Verweis auf die realweltliche Identität nicht näher spezifiziert wird [173, 183, 184, 200, 174, 203, 212, 197, 193, 187, 213]. Li et al.. [183] verknüpft die Blockchain-Adresse und Benutzendeninformationen explizit nicht, um die Offenlegung der Identität zu verhindern und Privatsphäre zu gewährleisten. Der Ansatz von Nguyen et al. [173] hingegen lässt in seiner Beschreibung unklar, wie Smart Contracts Identitätssubjekte hinzufügen und löschen können und wer die Gesamtautorität darüber besitzt, diese Verträge auszulösen. Dagher et al. [208] erwähnen die Verwendung eines Konsens-Verifizierungsprozesses zur Authentifizierung durch vertrauenswürdige Behörden und Arbeitgebende, welche die Systemidentiäten (z.B. Patient*innen, Ärzt*innen) vorab im System überprüfen sollen. Hierdurch soll die Authentifizierung mittels realweltlicher Identität ermöglicht werden, trotz der Tatsache, dass im System ausschließlich die Blockchain-Adresse verwendet wird. Dementsprechend wird dieser Ansatz als DTI klassifiziert. Boumezbeur und Zarour [188] stellen ein Schema vor, bei dem ein Webportal als “erste Sicherheitsstufe” verwendet wird. Hier können sich Patient*innen mit Benutzendenname und Passwort anmelden, um die eigenen grundlegenden medizinischen Informationen einzusehen. Es ist unklar, wie die zur Interaktion und Identitätsbeweisführung innerhalb der Blockchain verwendeten Schlüssel gespeichert werden und wie diese beiden Identitäten miteinander verknüpft sind. Es wird angenommen, dass die Verwaltung dieser Schlüssel durch die Systemanbietenden erfolgt. Gemäß der Definition von Bouras et al. [161] widerspricht die Verwendung eines dezentralen Ansatzes, bei dem der private Schlüssel des Datensubjekts in einem zentralisierten Speicherort durch einen Dritten gespeichert wird, der Autorität des Datensubjekts, die der dezentrale Ansatz gewähren würde. Wenn ein privater Schlüssel nicht lokal und unabhängig vom Datensubjekt gespeichert wird, gibt das Datensubjekt die Kontrolle über die Schlüsselverwendung auf und überträgt

daher die Kontrolle über das Identitätsmanagement an den Dritten des zentralisierten Speicherorts. Daher garantiert allein die Verwendung von Blockchain und DLT nicht die Autorität der Datensubjekte. Die zentrale Speicherung der wesentlichen Authentifizierungsinformationen wandelt dementsprechend einen dezentralen Ansatz effektiv in einen zentralisierten um [188].

3.4 Taxonomie

Zur Klassifizierung und Clusterung der im Rahmen der Recherche erfassten Ansätze wurde eine Taxonomie entwickelt, basierend auf der Methode von Nickerson et al. [219]. Die Methodik von Nickerson et al. [219] umfasst die folgenden Schritte: *(1) Bestimmung von Meta-Charakteristiken, welche als Grundlage für die Ableitung der Charakteristiken der Taxonomie verwendet werden; (2) Definition der Abbruchkriterien; (3) Auswahl eines der beiden Ansätze: a.) Empirisch-zu-Konzeptionell:* Untersuchung von Objekten auf gemeinsame Merkmale, welche anschließend gruppiert werden; *b.) Konzeptionell-zu-Empirisch:* Merkmale und Dimensionen von Objekten werden zunächst konzeptuell erarbeitet; *(4) Rückkehr zu Schritt 3, sofern die Abbruchbedingungen nicht erfüllt sind.* [219]

Besonderes Augenmerk lag hierbei auf den Charakteristiken des Datenmanagements, Identitätsmanagements, Zugriffsmanagements und der Sicherheitsmechanismen, die daher als Meta-Charakteristiken gewählt wurden. Der Entwicklungsprozess galt als abgeschlossen, wenn alle Ansätze untersucht worden waren und mindestens ein Ansatz unter jede Charakteristik klassifiziert wurde. Durch die Anwendung eines empirisch-konzeptuellen Ansatzes sowie eines konzeptuell-empirischen Ansatzes wurden die folgenden Dimensionen und entsprechenden Charakteristiken identifiziert:

- *Blockchain-Typ*: Öffentlich, Privat, Konsortium
- *Speicherort*: On-Chain, Off-Chain, Hybrid
- *Off-Chain-Speicher*: Dezentral, Zentral und Verteilt

- *Identitätsmanagement*: Zentrales IdM, Föderiertes IdM, DTI, SSI, Benutzerzentriertes IdM
- *Autorität zur Zugriffsverwaltung im Allgemeinen (Allg.)*: System, Datensubjekt, Geteilt
- *Autorität zur Zugriffsverwaltung hinsichtlich Forschung und Entwicklung (F&E)*: System, Datensubjekt, Geteilt
- *Zugriffskontrollstrategie*: DAC, NDAC, Hybrid
- *NDAC Logik*: MAC, Rollenbasiert, Regelbasiert
- *Sicherheitsmechanismen zur Zugriffskontrolle*: Symmetrische und Asymmetrische Verschlüsselung, Proxy-Re-Encryption, CP-ABE, Tokenisierung
- *Sicherheitsmaßnahmen zur Datenspeicherung*: Symmetrische, Asymmetrische, Hybride Verschlüsselung, Passwortschutz.

Die Zuordnung der Ansätze zu identifizierten Charakteristiken und Dimensionen ist in den Tabellen A.2, A.3, A.4, A.5, A.6, A.7 in Form von Taxonomien dargestellt. In diesem Fall zeigen die Tabellen A.2 und A.3 die Zuordnung der ersten Iteration, die Tabellen A.4 und A.5 die Zuordnung in der zweiten Iteration und die Tabellen A.6 und A.7 die Zuordnung der dritten Iteration. Die Klassifizierungen basieren auf den Informationen aus den jeweiligen Veröffentlichungen und wurden nach bestem Wissen und Gewissen durchgeführt. Da jede geprüfte Veröffentlichung unterschiedliche Schwerpunkte setzte und nicht immer alle Aspekte klar und ausführlich erklärte, wurde bei Unklarheiten oder fehlenden Charakteristiken auf eine Klassifizierung verzichtet.

3.5 Sicherheitsbetrachtungen bei der Konzeption Blockchain-basierter Gesundheitsdatenmanagementanwendungen

Der Stand der Technik im Bereich Blockchain-basierter Gesundheitsdatenmanagementanwendungen zeigt deutliche Lücken hinsichtlich der systematischen Durchführung von Sicherheitsanalysen auf. Tabelle 3.1 fasst zusammen, welche Sicherheitsanalysen in den identifizierten Ansätzen der Literatur durchgeführt wurden. Dabei liegt der Fokus vieler Arbeiten darauf, spezifische Sicherheitsbedenken aufzuzeigen, die durch die vorgeschlagenen Blockchain-basierten Anwendungen adressiert werden sollen [177, 200, 175, 189, 188, 204, 193, 212, 190, 197, 199, 191, 198]. Jedoch fehlt es oftmals an systematischen Untersuchungen der zugrundeliegenden Sicherheitsmechanismen.

Zheng et al. [201] beschreiben ausschließlich, inwiefern ihr Systemkonzept die Anforderungen an Datensicherheit erfüllt. Einige Veröffentlichungen, wie beispielsweise Daraghmi et al. [207], Dagher et al. [208] und Chang et al. [181], stellen Blockchain-basiertes Systemkonzepte vor, die den Anforderungen des Health Insurance Portability and Accountability Act (HIPAA) entsprechen sollen. Allerdings mangelt es an detaillierten Ausführungen, welche spezifischen HIPAA-Anforderungen in welcher Weise adressiert werden. Dagher et al. [208] integrieren vier technische Schutzmaßnahmen, um die HIPAA-Anforderungen zu erfüllen, doch auch hier fehlt eine umfassende, systematische Analyse der zugrundeliegenden Sicherheitsmechanismen.

Einige Arbeiten gehen gezielt auf Sicherheitsanforderungen ein. Lin et al. [213] beschreiben ihr Sicherheitsmodell, indem sie Sicherheitsaspekte wie Ciphertext-Privatsphäre und Keyword-Privatsphäre im Kontext eines Proxy-Re-Encryption-Ansatzes analysieren. Liu et al. [174] bewerten die Sicherheit ihres vorgeschlagenen Systems BPDS hinsichtlich der Aspekte Manipulationssicherheit, Privatsphäre sowie sichere Speicherung und den sicheren Austausch von Daten. Huang et al. [203] analysieren die Sicherheitsanforderungen ihres Systems und belegen durch theoretische Beweise die Einhaltung der definierten Sicherheitskriterien.

Die Studie von Zhang et al. [179] verwendet die Interoperabilitäts-Roadmap des "Office of the National Coordinator for Health Information Technology"(ONC) [220], einer Abteilung des US-amerikanischen Gesundheitsministeriums, um deren Auswirkungen auf die Systemgestaltung zu diskutieren. Dennoch erfolgt keine umfassende Analyse über diese definierte Anforderungsliste hinaus. Lee et al. [194] greift ebenfalls auf diese Anforderungen zurück und führt genauso wie Li et al. [183], Nguyen et al. [173, 187], Jayasinghe et al.[186] eine entsprechende Angreifendenanalyse durch, wobei ein systematischer oder formalisierter Ansatz zur Angreifendenmodellierung fehlt.

Die meisten Arbeiten beschränken sich auf die Untersuchung spezifischer Angriffsvektoren. Beispielsweise analysieren Lee et al. [194] vier interne und externe Angreifendenszenarien (Abhören, Denial-of-Service (DoS), abnormaler Zugriff, Datenfälschung) und untersuchen einzelne jeweils damit verbundene Bedrohungen. Li et al. [183] diskutieren ebenfalls vier Hauptbedrohungen, leiten daraus spezifische Designziele für ihr Systemkonzept ab und ergänzen diese durch Sicherheitsanalysen der zugrundeliegenden Blockchain-Protokolle. Nguyen et al. [173, 187] analysieren die Sicherheit ihres Systems anhand von zwei definierten Angreifendenszenarien. Jayasinghe et al.[186] untersuchen, inwiefern das vorgestellte System Schutzmaßnahmen gegen Datenmanipulation, DDoS, SQL-Injection und Spoofing implementiert. Theodouli et al. [176] führen eine grundlegende Sicherheitsanalyse durch, indem sie die beteiligten Entitäten, deren Anreize zur Nutzung des Systems, das Vertrauensniveau zwischen diesen sowie das zugrunde liegende Blockchain-Modell untersuchen. Darüber hinaus werden potenzielle bösartige Bedrohungen identifiziert und im Kontext des vorgeschlagenen Systemkonzepts diskutiert. Thwin und Vasupongayya [192] analysieren bösartige Zugriffe sowie Replay-Angriffe auf ihr Systemkonzept und leiten daraus entsprechende Erkennungs- und Abwehrmaßnahmen ab. Wang et al. [182] sowie Zhang und Lin [184] definieren ein Angreifendenmodell, aus dem sie spezifische Sicherheitsziele ableiten, die durch ihr Systemkonzept adressiert werden. Zaghloul, Li und Ren [178] analysieren den Schutz gegen Replay-Angriffe und die Kollusionsresistenz ihres Smart Contracts. Darüber hinaus erweitern diese ihre Analyse in

Zaghloul et al. [196], indem sie weitere Angriffsszenarien wie Datenmanipulation und DDoS-Angriffe untersuchen. Ähnlich fokussieren sich Zhou, Li und Zhao [202], Wang und He [185] sowie Qin, Jin und Liu [206] auf die Analyse ihrer Konzepte und Schemas hinsichtlich potenzieller Angriffsszenarien und deren Abwehr. Zou et al. [195] bewerten die Widerstandsfähigkeit ihres Systems speziell gegen Blockchain-Netzwerkangriffe.

Zusammenfassend lässt sich feststellen, dass die existierende Literatur zwar häufig Sicherheitsaspekte adressiert, jedoch meist auf systematische Ansätze zur Analyse und Modellierung verzichtet. Dieses Defizit unterstreicht eine wesentliche Schwäche in der bisherigen Forschung: Es bleibt weitgehend unklar, inwieweit Forschende systematische Methoden eingesetzt haben, um Sicherheitsbedenken umfassend zu analysieren und angemessen zu adressieren [221]. Diese Erkenntnis weist auf eine signifikante Forschungslücke hin, die geschlossen werden muss, um die Sicherheit Blockchain-basierter Gesundheitsdatenmanagementanwendungen fundiert bewerten und gewährleisten zu können.

3.6 Designprozess für Anwendungen auf Basis von Blockchain

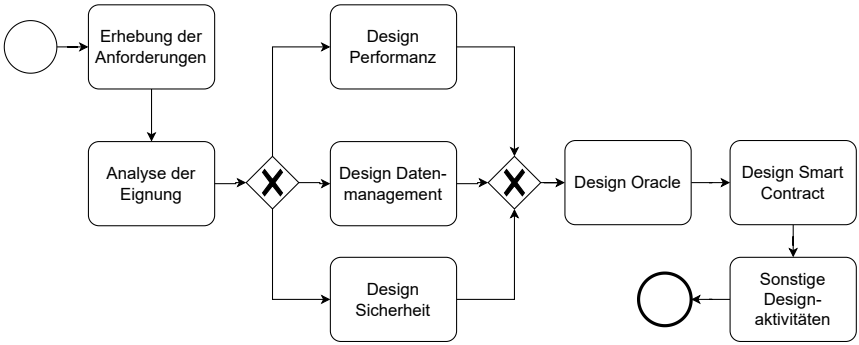

Abbildung 3.1: Entwurfsprozess zur Entwicklung von Blockchain-basierten Anwendungen laut Xu et al. (2021) [150].

Tabelle 3.1: Analyse der identifizierten Literatur in Bezug auf durchgeführte Sicherheitsbetrachtungen bei der Konzeption Blockchain-basierter Systeme.

Sicherheitsbetrachtungen	**Referenz**	**Anzahl**
Adressierung von Sicherheitsbedenken	Azaria et al. [177], Dagher et al. [208], Zhang et al. [179], Lee et al. [194], Li et al. [183], Daraghmi et al. [207], Chang et al. [181], Hanley und Tewari [180], Hawig et al. [200], Liu et al. [174], Nguyen et al. [173, 187], Theodouli et al. [176], Thwin und Vasupongayya [192], Wang et al. [182], Xiao et al. [175], Zaghloul et al. [178, 196], Zhang und Lin [184], Zheng et al. [201], Zhou, Li und Zhao [202], Gupta et al. [189], Boumezbeur und Zarour [188], Hu et al. [204], Jayasinghe et al.[186], Zhang et al. [193], Lee et al. [212], Lin et al. [213], Sabu et al. [190], Wang und He [185], Zou et al. [195], Huang et al. [203], Lee et al. [197], Li, Yue und Wu [199], Lomotey, Kumi und Deters [191], Qin, Jin und Liu [206], Zhao, Yu und Yan [198]	36 von 37
Analyse der Sicherheitsanforderungen	Dagher et al. [208], Zhang et al. [179], Lee et al. [194], Li et al. [183], Liu et al. [174], Nguyen et al. [173], Theodouli et al. [176], Thwin und Vasupongayya [192], Wang et al. [182], Zaghloul et al. [178, 196], Zhang und Lin [184], Zheng et al. [201], Zhou, Li und Zhao [202], Jayasinghe et al.[186], Lin et al. [213], Wang und He [185], Zou et al. [195], Huang et al. [203], Qin, Jin und Liu [206]	20 von 37
Angreifendenanalyse durchgeführt	Lee et al. [194], Li et al. [183], Nguyen et al. [173, 187], Theodouli et al. [176], Thwin und Vasupongayya [192], Wang et al. [182], Zaghloul et al. [178, 196], Zhang und Lin [184], Zhou, Li und Zhao [202], Jayasinghe et al.[186], Wang und He [185], Zou et al. [195], Qin, Jin und Liu [206]	15 von 37
Systematische und formalisierte Angreifendenmodellierung	keine	0 von 37

Aufgrund ihrer grundlegenden Eigenschaften und Einschränkungen eignen sich Blockchains nicht für alle Anwendungsfälle. Daher sollte vor der Konzeption eines Systems die Eignung der Blockchain in Bezug auf die spezifischen Bedürfnisse und Anforderungen vorab abgewogen werden [155]. Dementsprechend sieht Xu et al. [155] in deren 2019 vorgestellten Designprozess für Blockchain-basierte Systeme die Evaluation der Eignung der Blockchain als ersten Schritt vor. In der weiterführenden Veröffentlichung von Xu et al. aus dem Jahr 2021 [150] zur Einordnung der vorgestellten Entscheidungsmodelle (siehe für Details zu dem Entscheidungsmodell Abschnitt 3.7) wird explizit die Anforderungs- und Bedarfserhebung vor dieser Abwägung genannt. Jener Designprozess von Xu et al. [150] ist in Abbildung 3.1 dargestellt. Gefolgt wird diese Eignungsanalyse von der Entscheidung, welche Daten auf der Blockchain gespeichert werden (On-Chain) und welche Off-Chain verbleiben sollen. Aufgrund der Tatsache, dass mit einer On-Chain-Speicherung eine begrenzte Rechen- und Datenspeicherkapazität sowie Sicherheitsaspekte einhergehen, ist nicht nur das Design des Datenmanagements relevant, sondern ebenfalls Performanz- und Sicherheitsbetrachtungen. Dementsprechend gehen jene Entwurfsentscheidungen zum Datenmanagement, Performanz und Sicherheit Hand in Hand und werden im Entwurfsprozess als parallele Schritte definiert. Das Design von Oracles, also jener Komponente, welche die Blockchain mit der Außenwelt verbindet, ist unmittelbar von der Wahl des Datenmanagements abhängig. Ebenso basiert das Design von Smart Contracts auf den zuvor getroffenen Entwurfsentscheidungen, beispielsweise das Kommunikationsverhalten eines Oracles zwischen On- und Off-Chain-Datenspeichern. Demzufolge stellen diese Schritte die nächsten sequentiellen Phasen in dem Designprozess von Xu et al. dar [150]. Laut Xu et al. [150] können weitere Entwurfsaktivitäten folgen, welche ebenfalls auf den Smart Contracts aufbauen und den vorgestellten Entwurfsprozess abschließen.

Der Design-Prozess von Xu et al. aus 2017 [222] und 2019 [155] sieht zwischen der Eignung der Blockchain und der Wahl des Datenspeicherorts das Treffen von Entwurfsentscheidungen in Bezug auf die Dezentralisierung von Vertrauen vor (siehe Abbildung 3.2). Diese Entscheidungen umfassen die Abwägung, in welchem Maße Funktionen und Komponenten des Systems dezentralisiert oder zentralisiert

von einer oder mehreren Autoritäten bzw. einem Konsortium betrieben werden sollen. Diese Entscheidungen sind gemäß Xu et al. [155] von grundlegender Bedeutung, da die Wahl der Vertrauensverteilung nach deren Aussage die Sicherheit, Performanz, Interoperabilität und regulatorische Konformität beeinflusst. Im weiteren Verlauf des Designprozesses folgt nach der Entscheidung über den Datenspeicherort die Auswahl eines geeigneten Blockchain-Frameworks sowie die Festlegung der konkreten Blockchain-Konfiguration. Unter der Blockchain-Konfiguration verstehen Xu et al. [222] die Klärung des Bedarfs an einer oder mehrerer Blockchains, die Wahl des Blockchain-Typs, der Datenstruktur, des Konsensusprotokolls sowie die Festlegung der Blockgröße und Blockfrequenz. Im Anschluss sehen Xu et al. [222] weitere Designentscheidungen vor, welche vor dem Deployment und Betrieb der Blockchain getroffen werden müssen. Dazu gehören die Klärung der (1) Incentivierung der Netzwerteilnehmenden sowie die (2) Notwendigkeit von Anonymitätsmechanismen, wie beispielsweise Zero-Knowledge-Proofs. Abschließend erfolgt das Deployment und der Betrieb der Blockchain, wobei die Wahl des Deployment-Ortes, wie beispielsweise die Nutzung eines privaten virtuellen Netzwerks, eines Blockchain-as-a-Service Modells oder die Speicherung bei einem Cloud-Provider, von Bedeutung für die Systemqualität ist. Blockchain-basierte Systeme können beim Betrieb zudem schwieriger zu modifizieren sein als konventionelle Systeme, da die Software auf mehreren unabhängig operierenden Knoten läuft, was die Koordination von Updates sowohl administrativ als auch physisch herausfordernd macht. Zudem ist das Blockchain-Ledger aufgrund seiner Unveränderlichkeit nicht rückwirkend anpassbar, was eine nachträgliche Systemmodifikation erschwert [155].

In der wissenschaftlichen Literatur wurde ein weiterer Designprozess identifiziert (siehe Abbildung 3.3). Betzwieser et al. [149] kombinieren einen entwickelten Designprozess mit einem Entscheidungsmodell. Dementsprechend wird dieser im folgenden Abschnitt 3.7 zu Entscheidungsmodellen umfassend beschrieben.

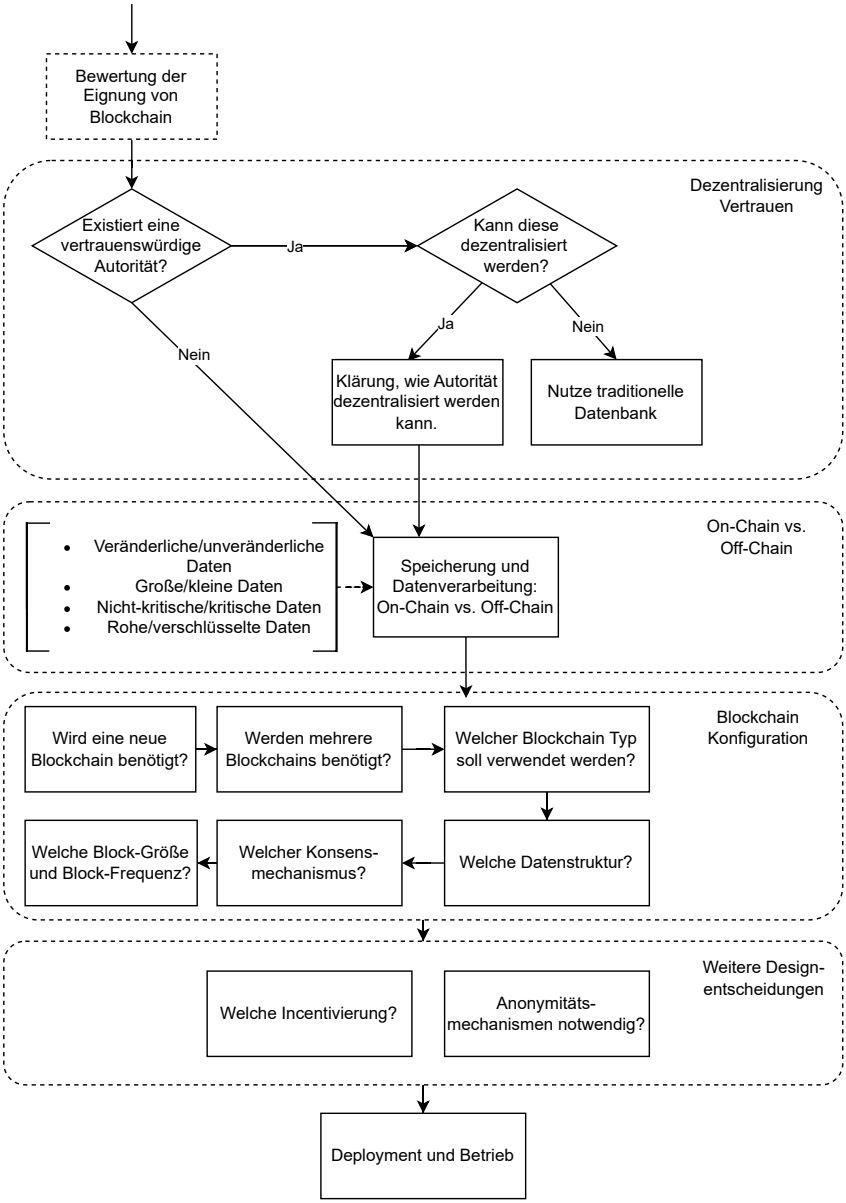

Abbildung 3.2: Designprozess für Blockchain-basierte Systeme laut Xu et al. [222, 155].

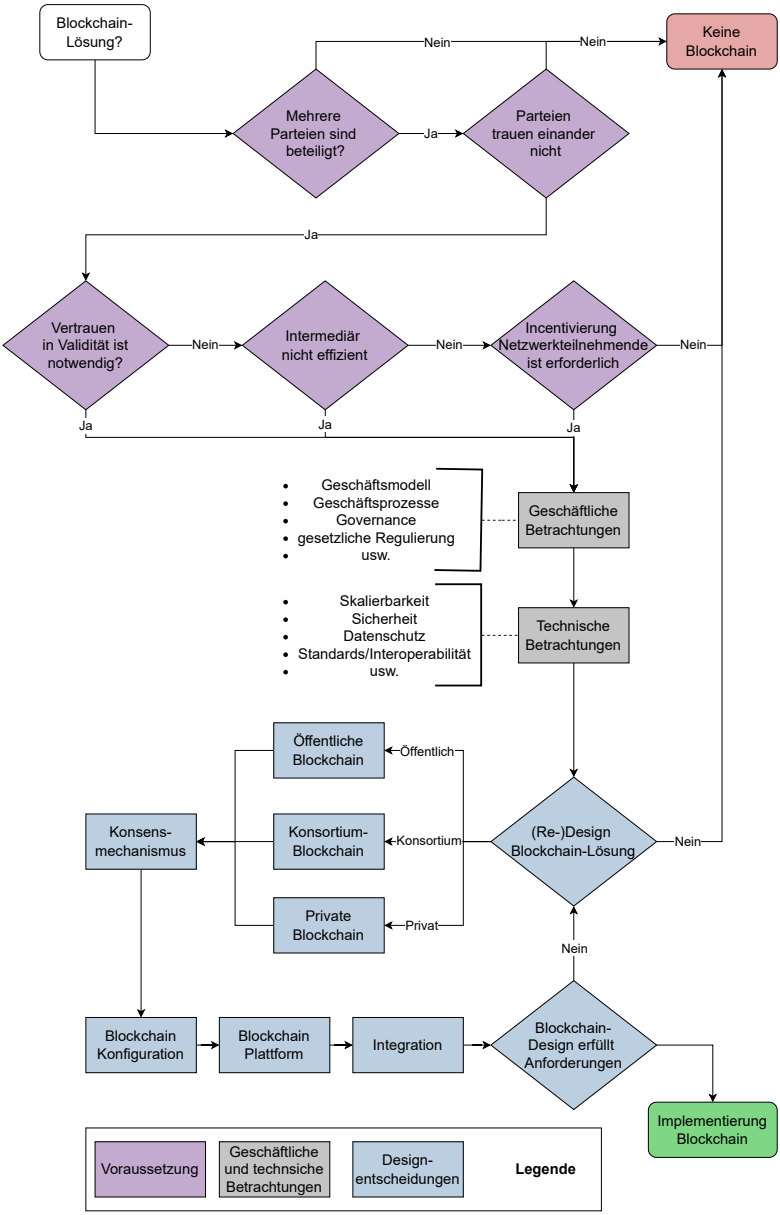

Abbildung 3.3: Entscheidungsmodell nach Betzwieser et al. [149].

3.7 Entscheidungsmodelle in der Entwicklung von Blockchain-basierten Systemen

Im Kontext von Blockchain-basierten Systemen existiert eine Vielzahl an Entscheidungsmodellen, welche darauf abzielen, bei der Beantwortung der Frage zu unterstützen, ob die Blockchain als technische Lösung für den gewählten Kontext oder Anwendungsbereich im Vergleich zu etablierten Technologien, wie zentrale oder dezentrale Datenbanken, geeignet ist [210, 223, 224, 225, 149, 226]. Mit Hilfe dieser Entscheidungsmodelle wird beispielsweise geprüft, ob mehrere Parteien mit Schreibrechten involviert sind, ob Vertrauen zwischen den beteiligten Akteuren besteht oder ob vertrauenswürdige dritte Parteien (*engl. trusted third party*) an dem Prozess beteiligt sind. Nach Prüfung dieser Bedingungen empfehlen beispielsweise Wüst et al. [210] sowie Pahl, Ioini und Helmer [223], je nach Ergebnis entweder auf den Einsatz einer Blockchain zu verzichten oder eine *permissionless*, öffentliche *permissioned* oder private *permissioned* Blockchain einzusetzen. Koens und Poll [224] sowie Lo et al. [226] inkludieren in deren Technologieempfehlungen nicht nur Blockchains oder Distributed Ledgers sondern auch zentrale, geteilte oder verteilte Datenbanken.

Betzwieser et al. [149] erweitern diese Modelle, indem sie über die reine Evaluierung der Einsatzmöglichkeit von Blockchain hinausgehen und sowohl technische als auch geschäftliche Betrachtungen einbeziehen. Das von ihnen vorgeschlagene Entscheidungsmodell stellt dabei nicht nur einen Bewertungsansatz bereit, sondern entwickelt zudem einen strukturierten Prozess, der die Konzeption und Umsetzung eines Blockchain-basierten Systems unterstützt (siehe Abbildung 3.3). Hierzu lässt sich das Entscheidungsmodell von Betzwieser et al. [149] in drei Ebenen unterteilen: (1.) Die Klärung der Voraussetzungen, die erfüllt sein müssen, damit eine Blockchain-basierte Lösung Vorteile bietet und der Einsatz der Technologie gerechtfertigt ist, (2.) Geschäftliche und technische Betrachtungen, die die Analyse der relevanten Rahmenbedingungen und Anforderungen an das System (z.B. bestehende Geschäftsmodelle und -prozesse) sowie die potenziellen Herausforderungen umfassen, und (3.) Die Festlegung von Designentscheidungen sowie Integrationsaspekten, die den Designprozess eines Blockchain-basierten Systems

umfassen. Die dritte Ebene beinhaltet zunächst die Wahl eines Blockchain-Typs (Öffentliche, Private oder Konsortium-Blockchain) je nach Zugriffsrechten, anschließend die Wahl des Konsensmechanismus, der Blockchain-Konfiguration, der Blockchain-Plattform sowie der Integrationsmöglichkeit in bestehende Systeme. Nach Abschluss dieser Schritte sieht die dritte Ebene einen Re-Design-Zyklus vor, der zum erneuten Durchlaufen der Schritte führt, falls die Anforderungen nach Beendigung der Designphase noch nicht vollständig erfüllt sind. Nach diesen Iterationen kann die Implementierung der Blockchain-Lösung erfolgen und das Entscheidungsmodell endet. Grundlegend unterstützt das Entscheidungsmodell neben der Frage nach dem Einsatz der Blockchain auch das prozessuale Vorgehen. Technische Aspekte werden dabei diskutiert und liefern einen groben Rahmen für den Designprozess. Konkrete technische Konsequenzen, spezifische Ausgestaltungsvorschläge oder Muster für das Blockchain-basierte System werden jedoch, abgesehen von der Wahl des Blockchain-Typs, nicht detailliert vorgegeben. Sicherheits- und datenschutzrelevante Herausforderungen, die mit Blockchain-Lösungen verbunden sind, werden insbesondere in der technischen Betrachtung angesprochen, einschließlich kritischer Problemstellungen wie der 51%-Attacke. Jedoch fehlen auch hier detaillierte Handlungsanweisungen oder konkrete Vorschläge für Schutzmaßnahmen.

In diesem Zusammenhang setzen die Entscheidungsmodelle von Xu et al. [150] an. Die Entscheidungsmodelle verfolgen das Ziel, Entwickelnde und Architekturschaffende von Blockchain-basierten Anwendungen bei der Auswahl geeigneter Muster zu unterstützen. Die Auswahl basiert auf Grundlage der Charakteristiken der Anwendungsfälle sowie der Trade-Offs, welche mit den vorgeschlagenen Mustern einhergehen. Durch die Verbindung von Mustern und Qualitätsanforderungen in graphisch dargestellten Entscheidungsmodellen unterstützen die BPMN-Modelle von Xu et al. [150] die systematische Entscheidungsfindung bei der Architekturgestaltung von Blockchain-basierten Anwendungen. Hierbei decken dedizierte Entscheidungsmodelle die folgenden Themenschwerpunkte ab: (1) On-Chain Datenmanagement und Performanz, (2) Off-Chain-Zugriffskontrolle, (3) Authentifizierung, (4) Autorisierung, (5) Interaktion mit der Außenwelt sowie

(6) die Entwicklung mehrerer Smart Contracts. Die von Xu et al. [150] verwendeten Muster werden in Tabelle A.1 vorgestellt und beschrieben. Grundlegend stellt die Kombination der Entscheidungsmodelle ein universelles Framework für die Entwicklung von Blockchain-basierten Anwendungen dar, berücksichtigt jedoch nicht die Wechselwirkungen der Muster mit stark regulierten Branchen wie das Gesundheitswesen sowie die damit einhergehenden Herausforderungen. Diese Domäne stellt spezifische Anforderungen an die Architektur, wie die Einhaltung gesetzlicher Vorschriften (z.B. Einhaltung der DSGVO), die Interoperabilität zwischen bestehenden informationstechnischen Systemen und verschiedenen Akteuren im Gesundheitssystem, die jeweils unterschiedliche Anforderungen und Zugriffskontrollen erfordern, sowie die effiziente Verwaltung hochsensibler und zugleich großer Datenmengen.

3.8 Fazit

Der in den vorangegangenen Abschnitten dargestellte Stand der Technik und Wissenschaft sowie die begleitende Literaturrecherche zu Blockchain-basierten Anwendung in der Forschung verdeutlicht technische Entwicklungen, Tendenzen und Herausforderungen. Im Bereich des Datenmanagements im Gesundheitswesen dominiert primär der Einsatz von Off-Chain-Ansätzen, während im Bereich des Identitätsmanagements ein klarer Fokus auf dezentrale und selbstsouveräne Konzepte zu erkennen ist. Dies wird in der Literatur häufig mit deren Vorteilen in Bezug auf Sicherheits- und Datenschutzaspekte begründet. Allerdings wird dabei übersehen, dass zentrale und föderierte Identitätsmanagementsysteme den aktuellen Stand der Technik abseits von Blockchain-basierten Anwendungen darstellen [161]. Beispielsweise in der TI2.0 ist ein föderiertes Identitätsmanagement vorgesehen [56]. Föderierte Ansätze, welche die Blockchain-Technologie in diesem Kontext anwenden, konnten im Rahmen der durchgeführten Literaturrecherche nicht identifiziert werden.

Zusammenfassend wird zudem deutlich, dass bei der Konzeption und Architektur Blockchain-basierter Systeme eine systematische Durchführung und Berücksichtigung von Sicherheitsanalysen häufig unzureichend behandelt oder gänzlich fehlen. Obwohl universelle Entscheidungsmodelle existieren, die Sicherheitsaspekte adressieren und insbesondere bei der Wahl von Mustern beim Design unterstützen, bleibt deren Unterstützung häufig unkonkret. Zudem konnte kein Entscheidungsmodell identifiziert werden, das sich explizit mit den Spezifika des Gesundheitswesens auseinandersetzt und diese in den Designprozess eines Blockchain-basierten Systems einbezieht. Diese Forschungslücke bildet den Ausgangspunkt für die vorliegende Arbeit, die sich der in Abschnitt 1.2 formulierten zentralen Forschungsfrage widmet:

> **Wie kann die Konzeption und Architektur dezentraler Infrastrukturen auf Basis von Blockchain-Technologien im Gesundheitswesen zur souveränen Verwaltung und effektiven Nutzung von Gesundheitsdaten unterstützt werden?**

Das übergeordnete Ziel dieser Arbeit besteht demgemäß darin, ein Entscheidungsmodell zu entwickeln, das auf die besonderen Anforderungen des Gesundheitswesens zugeschnitten ist und eine systematische Unterstützung bei der Gestaltung und Implementierung von Blockchain-basierten Systemen bietet. Auf diese Weise soll die identifizierte Forschungslücke geschlossen werden.

4 Entwicklung eines Entscheidungsmodells

Unter Berücksichtigung der in Kapitel 3 erfassten Literatur und Charakteristiken der verschiedenen Blockchain-basierten Systeme zum Gesundheitsdatenmanagement wurden die folgenden zehn Fragestellungen abgeleitet:

- Welcher Speicherort soll verwendet werden?
- Welcher Blockchain-Typ sollte eingesetzt werden?
- Welcher Speicher eignet sich zur Off-Chain Speicherung?
- Welches Identitätsmanagementsystem sollte verwendet werden?
- Mit welchen Parteien sollen Gesundheitsdaten geteilt werden?
- Sollen die Daten für Zwecke des maschinellen Lernens benutzt werden?
- Wer hat die Autorität zur Verwaltung der Zugriffskontrollstrategie?
- Welche Zugriffskontrollstrategie sollte verwendet werden?
- Welche Sicherheitsmaßnahmen zur Zugriffskontrolle sollen ergriffen werden?
- Welche zusätzlichen Sicherheitsmaßnahmen sind für die Datenspeicherung notwendig?

Diese Fragestellungen sind für die Gestaltung von Blockchain-basierten Systemen zum Gesundheitsdatenmanagement von Bedeutung und spiegeln die relevanten

Designentscheidungen wider. Auf Basis der in den folgenden Abschnitten näher erläuterten Fragestellungen wird anschließend ein Entscheidungsmodell in Form eines Entscheidungsbaums entwickelt. Dieses Modell soll als Unterstützung für Software- und Systemarchitekten bei der Konzeption von Blockchain-basierten Gesundheitsdatenmanagementsystemen dienen.

Der Prozess zur Entwicklung dieses Entscheidungsmodells erfolgte iterativ in zwei Phasen, entsprechend den Literaturrecherche-Iterationen (siehe Abschnitt 3). Ein erstes Entscheidungsmodell wurde auf Basis der ersten Literaturrecherche-Iteration erstellt und veröffentlicht [11]. Der Fokus lag hierbei auf der Datenspeicherung, weshalb dieser durch die zweite und dritte Literaturrecherche-Iteration um die Aspekte des Zugriffs- und Identitätsmanagement ergänzt wurde. Zusätzlich wurden Entscheidungsfaktoren modifiziert, um das Vertrauen zwischen den Akteur*innen im System zu berücksichtigen. Jenes Entscheidungsmodell wurde anschließend in Form eines Journalartikels veröffentlicht [2]. Das daraus resultierende finale Entscheidungsmodell wird in Abbildung 4.1 dargestellt, welches zur besseren Lesbarkeit in mehrere Teilbäume unterteilt ist (siehe die Abbildungen 4.2, 4.3, 4.4, 4.5, 4.6, 4.7).Bevor das Entscheidungsmodell angewendet wird, sollte geprüft werden, ob eine Blockchain-basierte Lösung für den gewählten Anwendungsfall geeignet ist. Bestehende Arbeiten, wie beispielsweise von Wüst und Gervais [210], können zu diesem Zweck herangezogen werden.

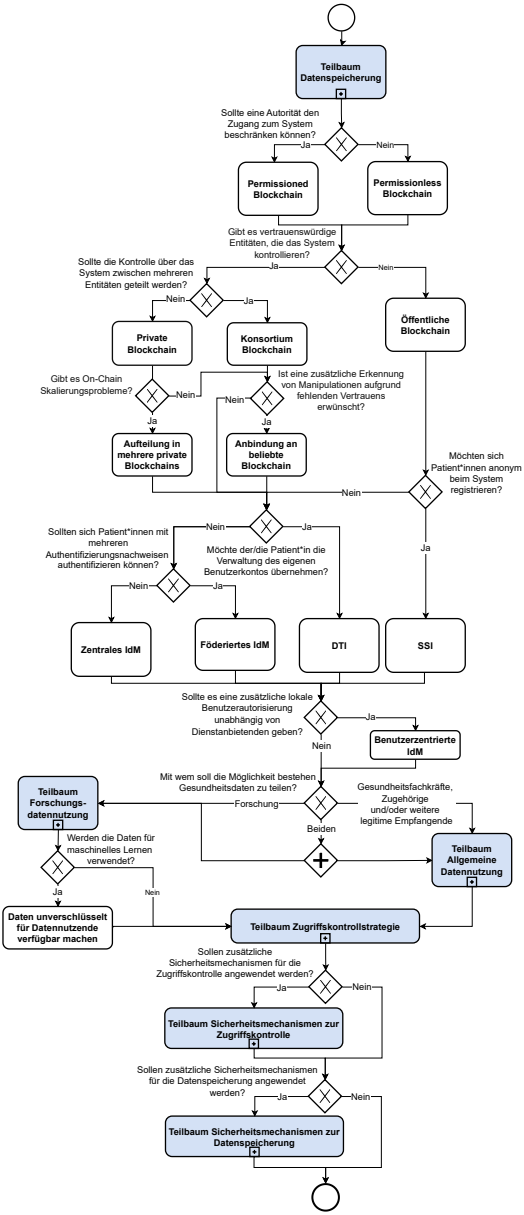

Abbildung 4.1: Das entwickelte Entscheidungsmodell.

4.1 Welcher Speicherort soll verwendet werden?

Es gibt drei Ansätze zur Speicherung sensibler Daten in Blockchain-basierten Systemen: die Speicherung direkt auf der Blockchain (On-Chain), die Speicherung außerhalb der Blockchain (Off-Chain) oder eine hybride Speicherung (siehe Abschnitt 3.1.1 und das Teilmodell in Abbildung 4.2). On-Chain-Speicherung eignet sich insbesondere für kleine Datenmengen sowie für Daten, die in der Zukunft unverändert und ungelöscht bleiben müssen. Für große Datenmengen sowie für Daten, die gelöscht oder verändert werden müssen, wird die Verwendung eines Off-Chain-Speichers empfohlen. Als Faustregel gilt, dass Daten, die kleiner als ihr Hashwert sind, On-Chain und größere Daten Off-Chain gespeichert werden sollten [155]. Generell zeigen zahlreiche Gesundheitsdienstleistende eine Zurückhaltung gegenüber der Datenfreigabe an Dritte [13, 18]. In Fällen, in denen kein ausreichendes Vertrauen in das Speichersystem besteht, sollte daher eine Off-Chain-Speicherlösung in Betracht gezogen werden. Sollte weder eine On-Chain- noch eine Off-Chain-Speicherung ausreichen und eine höhere Flexibilität erforderlich sein, kann auch eine hybride Speicherstrategie in Erwägung gezogen werden. Ein hybrider Ansatz erlaubt es, Metadaten wie den Hash der Daten auf der Blockchain zu speichern, um von deren Unveränderlichkeit zu profitieren [155].

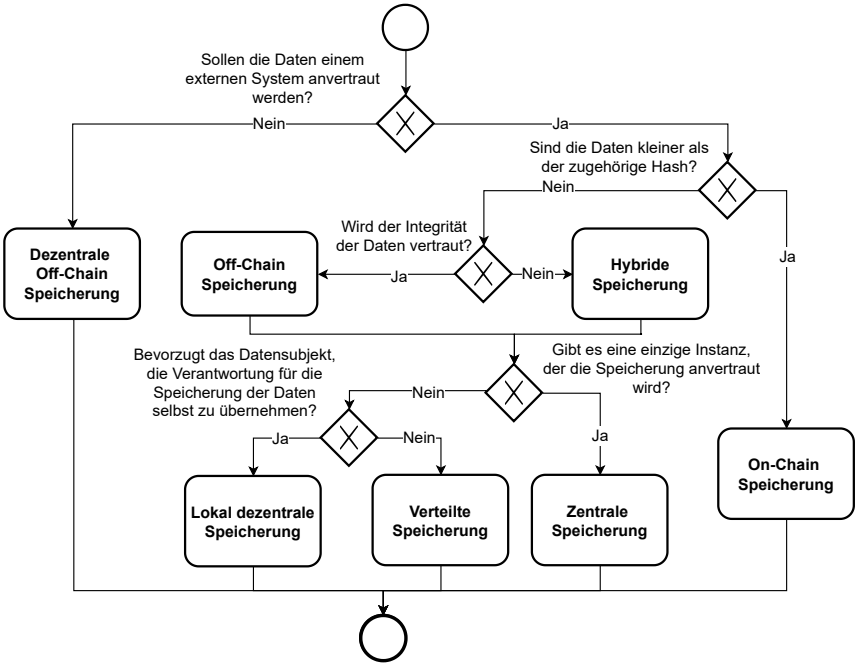

Abbildung 4.2: Das Teilmodell für die Datenspeicherung.

4.2 Welcher Blockchain-Typ sollte eingesetzt werden?

Die Wahl des Blockchain-Typs hängt hauptsächlich davon ab, ob das System dezentral, zentral von einem einzelnen Gesundheitsdienstleistenden oder gemeinsam von mehreren Gesundheitsdienstleistenden betrieben und kontrolliert werden soll (siehe Entscheidungsmodell 4.1). Öffentliche Blockchains eignen sich für dezentral betriebene Systeme, die sich auf keinen Anbieter verlassen wollen. Für Systeme, die von einem einzelnen Gesundheitsdienstleistenden oder einer staatlichen Institution verwaltet werden, ist eine private Blockchain geeignet, und für Systeme, die gemeinsam von mehreren Gesundheitsdienstleistenden verwaltet werden, ist

eine Konsortium-Blockchain sinnvoll. Unter bestimmten Umständen kann auch der Einsatz mehrerer Blockchains von Vorteil sein. Um Manipulationen durch die verantwortlichen Organisationen in privaten oder Konsortium-Blockchains zu erkennen, können die Transaktionen dieser Blockchains in einer weit verbreiteten öffentlichen Blockchain verankert werden. Darüber hinaus können Skalierbarkeitsprobleme bei der On-Chain-Speicherung in einer privaten Blockchain durch die Aufteilung der Daten auf mehrere private Blockchains verringert werden.

4.3 Welcher Speicher eignet sich zur Off-Chain Speicherung?

Im Falle eines Off-Chain-Speichers muss ein geeigneter Speichertyp für die Off-Chain Daten gewählt werden. Entsprechend der Literatur können Daten entweder dezentral, zentral oder verteilt gespeichert werden (siehe Abbildung 4.2). Dezentrale Speicherlösungen sind für die Speicherung von elektronischen Patientenakten (ePAs) zu bevorzugen, da die Daten bereits in der bestehenden Infrastruktur der jeweiligen Gesundheitsdienstleistenden vorliegen [175]. Außerdem können aufgrund ihrer Fürsorgepflicht gegenüber den Patient*innen und Anforderung der Integrität die Gesundheitsdienstleistenden zögern, die sensiblen Daten in externe Systeme zu übertragen. Bei mangelndem Vertrauen der teilnehmenden Gesundheitsdienstleistenden wäre dementsprechend die naheliegende Option, die gespeicherten Daten dezentral innerhalb der ursprünglichen Speicherorte der Gesundheitsdienstleistenden zu belassen.

Wenn das Ziel darin besteht, die Unabhängigkeit von den Systemen der datenerzeugenden Einrichtungen zu erreichen, kann ein zentraler oder verteilter Off-Chain Speicher verwendet werden. Dies ist insbesondere bei Daten aus elektronischen Gesundheitsakten (eGAs) und sensorischen Messungen relevant. Hierfür muss die Bereitschaft und das Vertrauen der Gesundheitsdienstleistenden in Dritte vorhanden sein. Wenn einer einzelnen Instanz für die Speicherung vertraut wird, kann eine zentrale Speicherlösung wie ein Cloud-Server gewählt werden. Wenn keiner einzelnen Instanz vertraut wird, können die Daten entweder lokal

bei jedem Datensubjekt gespeichert werden, was einen administrativen Aufwand für diesen darstellen würde, oder in verteilter Weise durch P2P-Systeme, wie im InterPlanetary File System (IPFS).

4.4 Welches Identitätsmanagementsystem sollte verwendet werden?

Identitätsmanagementsysteme lassen sich in zentralisiert, föderiert, benutzerzentriert und dezentralisiert unterteilen, wobei Letztere in SSI und DTI unterteilt sind (siehe Abbildung 4.1). Bei dezentralen IdM-Systemen werden die Identitäten auf einer Blockchain statt auf einem zentralen Server gespeichert, wodurch sie nicht von einer einzigen Entität kontrolliert werden. SSI und DTI unterscheiden sich dabei in ihrer Registrierungsweise. Während SSI es ermöglicht, einem System vollständig anonym beizutreten, erfordert die Registrierung bei DTI eine Identitätsreferenz durch eine vertrauenswürdige Drittpartei für jede neue Entität. Der Vorteil der völligen Anonymität wird jedoch eingeschränkt, wenn SSI auf einer privaten oder Konsortium-Blockchain umgesetzt wird, da hier der Zugang nach eigenen Kriterien einschränkt und gefiltert werden kann. Zentralisierte und föderierte IdM speichern Identitäten zentral auf Servern, die von den Dienstanbietenden kontrolliert werden. Zentralisierte Systeme sind besonders für Organisationen mit einer großen Nutzendenbasis geeignet, bieten den Entitäten jedoch keine Kontrolle über ihre eigenen Identitätsdaten. Diese Systeme sind weit verbreitet und erfordern von den Entitäten, für jedes System separate Konten zu erstellen und eine Vielzahl an unterschiedlichen Anmeldedaten zu verwalten. Im Gegensatz dazu bieten föderierte IdMs eine verbesserte Benutzendenerfahrung, indem sie eine einfachere Authentifizierung durch die gemeinsame Nutzung von Authentifizierungsdaten ermöglichen. Durch die Integration einer zusätzlichen Authentifizierungsebene, etwa mittels eines persönlichen Authentifizierungsgeräts, könnte jedes der genannten IdM-Systeme die Funktionalität eines benutzerzentrierten IdM-Systems erreichen.

4.5 Mit welchen Parteien sollen Gesundheitsdaten geteilt werden?

Je nach Bereitschaft des Datensubjekts, die eigenen Daten zu teilen, kann ein System den Austausch von Gesundheitsdaten zwischen Gesundheitsdienstleistenden, Angehörigen/Zugehörigen und/oder Forschungseinrichtungen ermöglichen (siehe Abbildung 4.1). Für jede dieser Gruppen müssen spezifische Zugriffskontrollstrategien (siehe Abbildung 4.5) und Autoritäten zur Zugriffsverwaltung (siehe Abbildungen 4.3 und 4.4) basierend auf den Fragen aus Abschnitt 4.8 und 4.8 festgelegt werden.

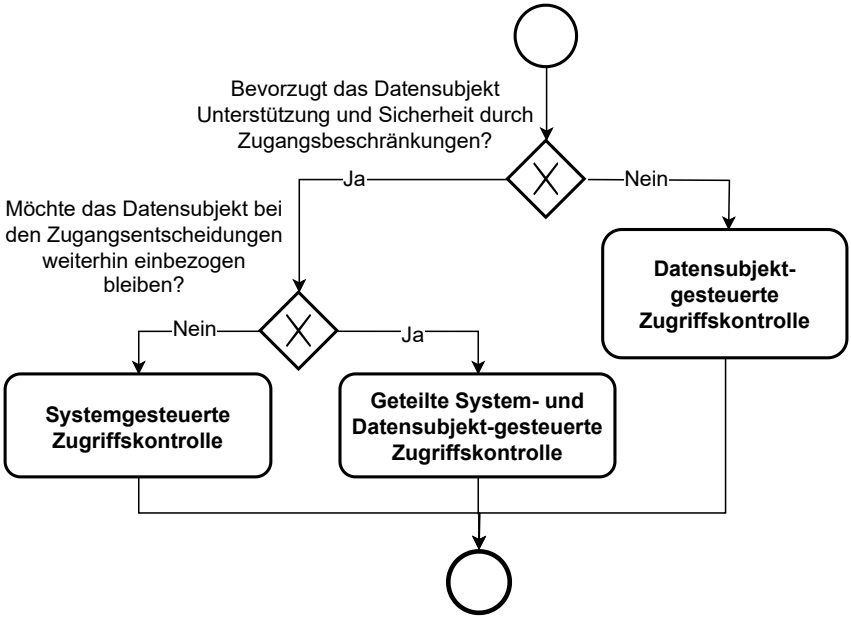

Abbildung 4.3: Das Teilmodell für die allgemeine Datennutzung.

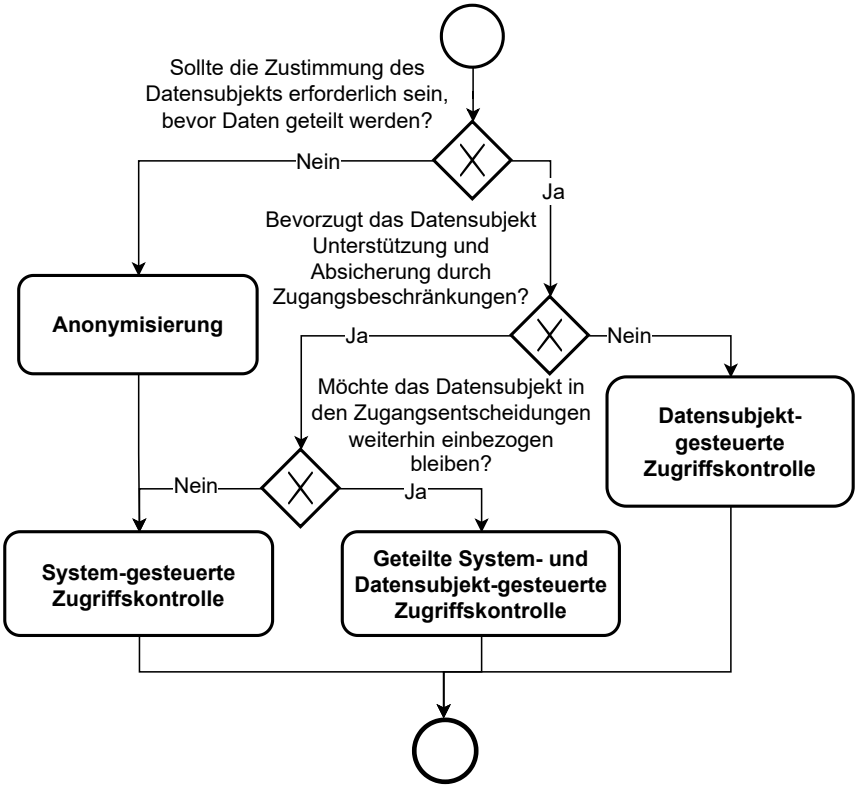

Abbildung 4.4: Das Teilmodell für die Forschungsdatennutzung.

4.6 Sollen die Daten für Zwecke des maschinellen Lernens benutzt werden?

Bei der Weitergabe von Gesundheitsdaten an die Forschungsgemeinschaft ist es erforderlich, dass die potenziellen Datennutzenden Zugang zu den unverschlüsselten Daten erhalten, um die effektive Anwendung maschineller Lernverfahren zu gewährleisten. Dies liegt daran, dass Berechnungen auf verschlüsselten Daten, insbesondere unter Verwendung homomorpher Verschlüsselung, in der Regel mit

einer erheblichen Rechenlast verbunden sind und sich nur für einfache Operationen eignen [180]. Daraus lässt sich ableiten, dass Verschlüsselung zur Sicherstellung der Vertraulichkeit sensibler Daten sowohl On-Chain als auch Off-Chain eine relevante Sicherheitsfunktion darstellt, solange keine Berechnungen auf diesen verschlüsselten Daten erforderlich sind. Daher ist es wichtig zu überprüfen, ob die Daten für maschinelles Lernen verwendet werden sollen oder nicht (siehe Abbildung 4.1).

4.7 Wer hat die Autorität zur Verwaltung der Zugriffskontrollstrategie?

Die Autorität zur Verwaltung der Zugriffskontrollstrategie kann entweder beim Datensubjekt, beim System oder durch eine gemeinsame Verwaltung beider Parteien liegen (siehe Abbildungen 4.3 und 4.4). Das Übertragen der Kontrolle über die Zugriffskontrollstrategie an ein System erfordert einen sorgfältig durchdachten und umfassend getesteten Ansatz, der die Datensubjekte universell vor eigenen Fehlern oder uninformierten Entscheidungen schützt. Allerdings entzieht dieser Ansatz dem Datensubjekt die Entscheidungsbefugnis und beraubt ihn der Möglichkeit, seine persönlichen Daten eigenständig zu verwalten. Im Gegensatz dazu ermöglicht die ausschließliche Kontrolle der Zugriffskontrollstrategie durch das Datensubjekt, dass jene die volle Verantwortung für die persönlichen und sensiblen Daten übernimmt. Dieser Ansatz birgt jedoch das Risiko, unbeabsichtigt böswilligen Zugriff durch externe Parteien zuzulassen. Ein ausgewogener Ansatz, bei dem das System gewisse Einschränkungen und Aufsicht auferlegt, während dem Datensubjekt gleichzeitig ein gewisses Maß an Selbstverwaltung gewährt wird, bietet der Kompromiss einer geteilten Autorität.

Im Hinblick auf die Weitergabe von Daten an die Forschungsgemeinschaft kann die Notwendigkeit der Zustimmung der Datensubjekte zur pseudonymisierten Nutzung ihrer Daten vermieden werden, indem die betreffenden Daten anonymisiert werden. In diesem Fall kann das System die gesamte Zugriffsverwaltung übernehmen.

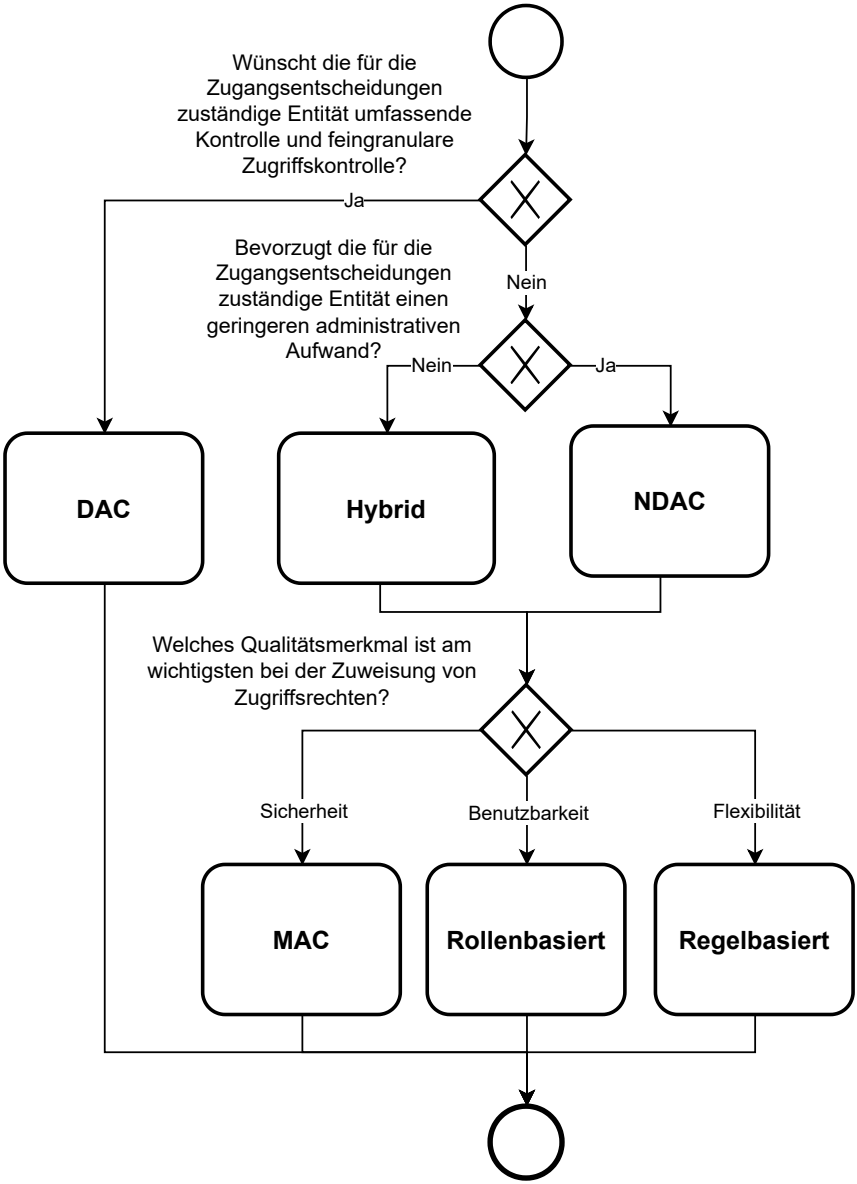

Abbildung 4.5: Das Teilmodell für die Zugriffskontrollstrategie.

4.8 Welche Zugriffskontrollstrategie sollte verwendet werden?

Das Teilmodell in Abbildung 4.5 veranschaulicht die Auswahl einer geeigneten Zugriffskontrollstrategie in Verbindung mit der entsprechenden Zugriffskontrolllogik. DAC ermöglicht ein feingranulares Zugriffsmanagement und eine erleichterte Überwachung, geht jedoch mit einer administrativen Belastung der für die Zugriffskontrolle verantwortlichen Autorität einher, sei es das System, das Datensubjekt oder beide. Im Gegensatz dazu vereinfacht NDAC das Zugriffsmanagement und reduziert die administrative Last, was allerdings zu einer geringeren Granularität bei der Zugriffskontrolle führen kann. In bestimmten Szenarien könnte ein hybrider Ansatz mehr Flexibilität bieten, obwohl zu beachten ist, dass das Ausnutzen von Lücken in der Zugriffspolitik Sicherheitslücken im System einführen könnte.

Die Wahl der NDAC-Zugriffskontrolllogik sollte auf der Grundlage des bevorzugten Qualitätsmerkmals erfolgen. Im Rahmen der MAC wird das Qualitätsmerkmal Sicherheit priorisiert, da die Zugriffskontrolle strikte Regeln und Sicherheitsrichtlinien durchsetzt, die Manipulationen oder Änderungen durch Benutzer*innen weitgehend ausschließen. Bei der regelbasierten Zugriffskontrolllogik wird die Flexibilität favorisiert, da sie eine feingranulare und spezifischere Steuerung von Zugriffsrechten ermöglicht und somit auf unterschiedliche Kontexte und Anforderungen zugeschnitten werden kann. Hingegen steht bei der rollenbasierten Zugriffskontrolllogik die Benutzbarkeit im Vordergrund, da sie eine effizientere und einfachere Verwaltung von Zugriffsrechten ermöglicht, indem Benutzer*innen in vordefinierte Rollen eingeteilt werden, was die administrativen Aufwände reduziert und dynamische Zugriffsverteilungen ermöglicht.

4.9 Welche Sicherheitsmaßnahmen zur Zugriffskontrolle sollen ergriffen werden?

Die zusätzlichen Sicherheitsmaßnahmen zur Zugriffskontrolle gewährleisten, dass die Entität, der die Zugriffsberechtigung auf eine Ressource erteilt wird, auch die empfangende Person dieses Zugriffs ist (siehe das Teilmodell in Abbildung 4.6). Um die Berechtigung in ihrer einfachsten Form zu gewähren, könnte der Zugriff durch die Ausgabe eines symmetrischen Schlüssels erfolgen, welcher zur Verschlüsselung der Ressource verwendet wurde. Diese Methode stellt jedoch nicht sicher, dass die Zugriffsberechtigung an spezifische Empfangende gebunden ist, da derselbe Schlüssel an mehrere Entitäten ausgegeben werden kann. Jede Entität, die im Besitz dieses Schlüssels ist, wäre in der Lage, auf die verschlüsselte Ressource zugreifen.

Sollte der Zugriff jedoch an eine spezifische Entität gebunden werden, könnte dies durch ein Attributset erfolgen, das jene Entität besitzt, oder auf individueller Einzelfallbasis definiert werden. Die Verschlüsselung mit CP-ABE würde es einer Entität, welche die erforderlichen Attribute besitzt, ermöglichen, auf eine Ressource zuzugreifen. Diese Methode erlaubt damit einem breiten Kreis von Entitäten den Zugang zu einer Ressource. Allerdings muss die Zugriffsrichtlinie vor der Gewährung des Zugriffs festgelegt werden, was die Flexibilität jener Methode einschränkt.

Die Zugriffssicherheit für einen auf individueller Basis definierten Empfangenden kann entweder durch eine spezifische Identität, die in den Sicherheitsmechanismus integriert ist, oder ohne eine solche erfolgen. Die Tokenisierung ermöglicht es dem individuellen Empfangenden eines Tokens, auf eine Ressource zuzugreifen. Tokens werden ausschließlich auf individueller Basis verteilt. Dies bedeutet, dass pro Entität ein Token vergeben wird. Jeder Inhabende dieses Tokens, kann bei einer Verifizierung eine digitale Signatur verwenden, um zusätzlich die Authentizität der Entität nachzuweisen. Hierdurch wird eine zusätzliche Sicherheitsebene hinzuzufügt und die Identität des Empfangenden vor der Gewährung des Zugriffs

überprüfbar. Bei der Tokenisierung allein wird jedoch die spezifische Identität des Identitätssubjekts nicht spezifiziert.

Sollte die spezifische Identität mit der Zugriffssicherheit verknüpft werden, kommt das Paar aus öffentlichem und privatem Schlüssel einer Entität zum Einsatz. Durch die Anwendung asymmetrischer Verschlüsselung als Methode der Zugriffssicherheit kann ausschließlich der vorgesehene Empfangende, welcher den privaten Schlüssel besitzt, Zugang zu der Ressource erhalten.

Falls die sensiblen Daten bei einem nicht vertrauenswürdigen Speicherort gespeichert werden, kann Proxy Re-Encryption ermöglichen, einer bestimmten Identität den Zugang zu gewähren. Hierzu wird ein Re-Verschlüsselungsschlüssel ausgegeben, der es dem potenziell böswilligen Speicherort ermöglicht, die bereits verschlüsselten Daten erneut zu verschlüsseln, ohne die Daten offenzulegen. Der Empfangende kann dann die Daten mit dem privaten Schlüssel entschlüsseln. Werden die Daten hingegen an einem vertrauenswürdigen Speicherort abgelegt, würde eine einfache Verschlüsselung mit einem öffentlichen Schlüssel ausreichen, um den Zugriff für eine spezifische Identität zu sichern.

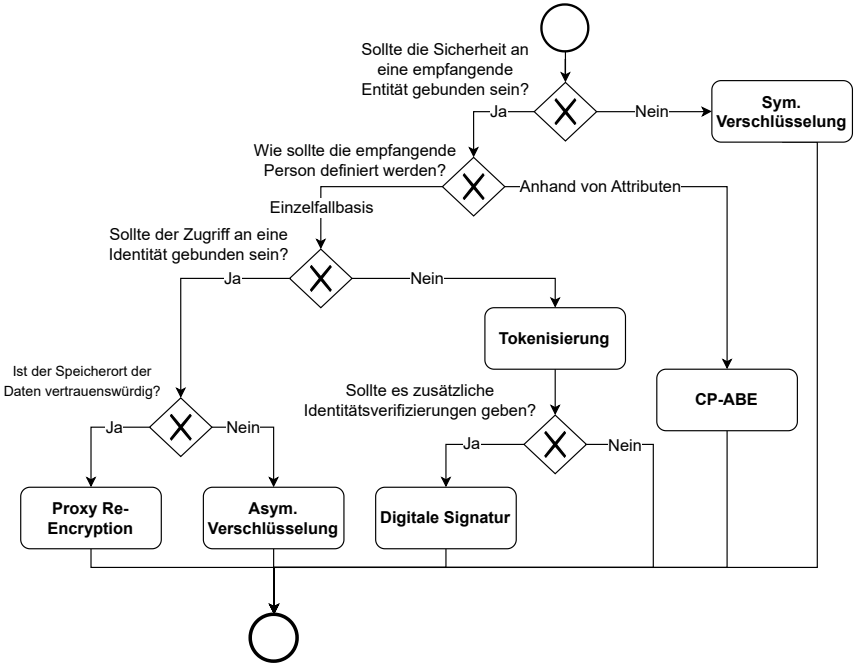

Abbildung 4.6: Das Teilmodell für die Sicherheitsmechanismen zur Zugriffskontrolle.

4.10 Welche zusätzlichen Sicherheitsmaßnahmen sind für die Datenspeicherung notwendig?

Daten können durch Verschlüsselung oder Passwortschutz gesichert werden (siehe Abbildung 4.7). Im Vergleich zur kryptographischen Verschlüsselung bietet der Passwortschutz jedoch lediglich eine oberflächliche Sicherheit. Insbesondere erweist sich Passwortschutz als unzureichend, wenn das Risiko eines externen

Hackerangriffs besteht. Für die Verschlüsselung existieren zwei grundlegende Arten: die symmetrische und asymmetrische Verschlüsselung. Asymmetrische Verschlüsselung bietet aufgrund der Verwendung von zwei verschiedenen Schlüsseln für Verschlüsselung und Entschlüsselung ein höheres Maß an Sicherheit. Diese Methode erfordert jedoch einen höheren Ressourcenaufwand und ist daher hauptsächlich für kleinere Daten, wie z.B. Textdateien, geeignet. Für größere Daten wird hingegen die symmetrische Verschlüsselung empfohlen, da sie einen effizienteren Verschlüsselungsprozess bietet. Jedoch ist diese Methode im Vergleich zur asymmetrischen Verschlüsselung weniger robust. Eine hybride Verschlüsselungsmethode, die die symmetrische Verschlüsselung zur Sicherung der Dateien selbst mit der asymmetrischen Verschlüsselung zur Sicherung des symmetrischen Schlüssels kombiniert, stellt eine effektive Lösung dar. Diese hybride Methode ermöglicht eine sichere Verschlüsselung bei gleichzeitiger Effizienz und reduziert das Risiko der Offenlegung von Schlüsseln.

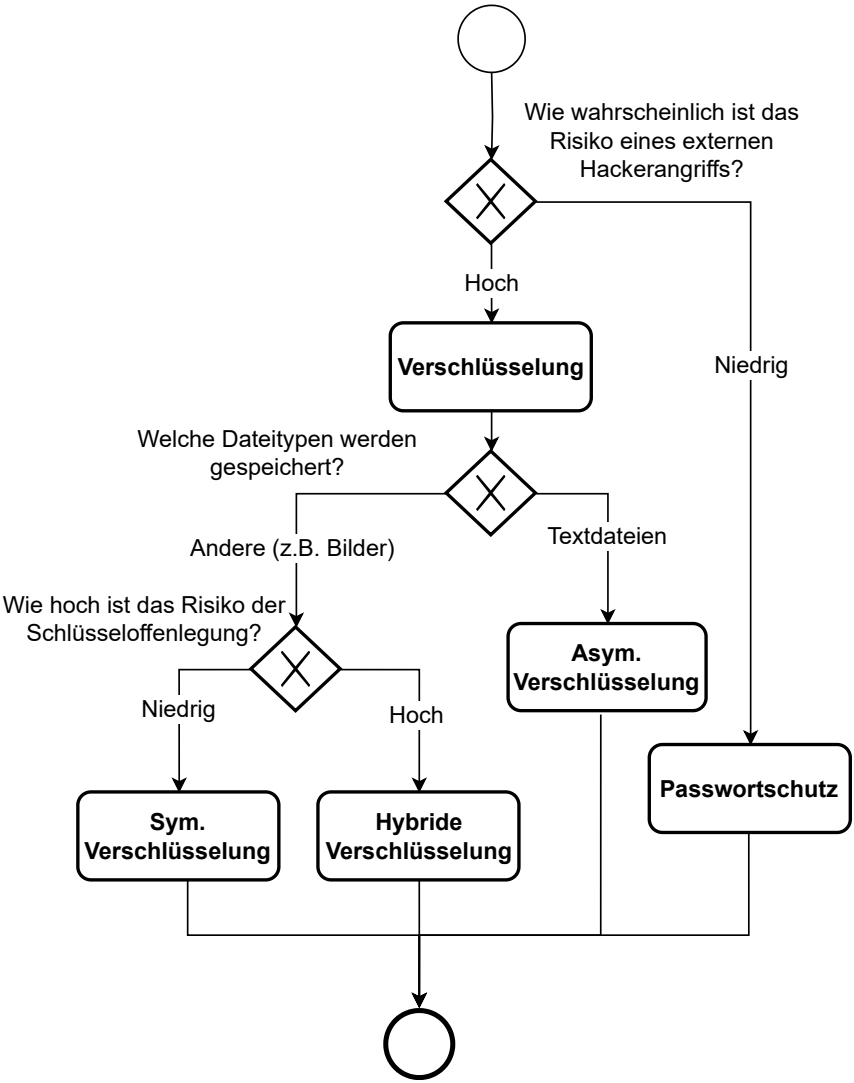

Abbildung 4.7: Das Teilmodell für die Sicherheitsmechanismen zur Datenspeicherung.

5 Anwendungsfälle

Wie bereits zuvor in den Grundlagen (siehe Kapitel 2) beschrieben, umfasst das Gesundheitswesen eine Vielzahl von Akteur*innen, Informationssystemen und regulatorischen Vorschriften. In der wissenschaftlichen Literatur wird der Mehrwert von Blockchain-basierten Systemen dementsprechend für unterschiedliche Zwecke und in verschiedenen Anwendungsfällen im Gesundheitssektor gesehen. Laut einer Literaturrecherche von Shanshan et al. [227] lassen sich sechs relevante Anwendungsfälle von DLT und Blockchain-basierten Systemen im Gesundheitswesen zusammenfassen: Patient*innen-zentriertes Gesundheitsdatenmanagement, Verwaltung elektronischer Gesundheitsakten, Fernüberwachung von Patient*innen, biomedizinische Forschung, Lieferkettenmanagement für pharmazeutische Produkte oder medizinische Geräte sowie die Kontaktverfolgung und Infektionskontrolle bei Pandemien [227]. Zur realweltlichen Evaluation des Entscheidungsmodells wurden aus diesen sechs Anwendungsfällen zwei wesentliche und einschlägige Anwendungsfälle ausgewählt: (1) das Patient*innen-zentrierte Gesundheitsdatenmanagement in der medizinischen Versorgung und (2) die Sekundärdatennutzung für die medizinische Forschung und Entwicklung. Die beiden Anwendungsfälle wurden gezielt aufgrund ihrer wissenschaftlichen Relevanz und praktischen Bedeutung für die Optimierung des Gesundheitsdatenmanagements sowie für die Förderung medizinischer Forschung und Innovation ausgewählt. Der EHDS adressiert spezifisch jene Anwendungsfälle der Primär- und Sekundärnutzung von Gesundheitsdaten, wodurch deren Relevanz unterstützt wird [105]. Im folgenden Kapitel werden diese Anwendungsfälle detailliert analysiert und es wird eine Systemarchitektur ohne Anwendung des Entscheidungsmodells für diese Anwendungsfälle entwickelt. Ergänzend hierzu werden die perspektivische

Einbettung sowie mögliche Skalierungsmöglichkeiten der entwickelten Systemkonzepte diskutiert. Dieses Kapitel stützt sich auf den Veröffentlichungen für Anwendungsfall 1 [[13, 14, 17]] sowie für Anwendungsfall 2 [[9, 2, 18]], bei denen Christina Erler als Autorin beteiligt war.

5.1 Anwendungsfall 1: Patient*innen-zentriertes Gesundheitsdatenmanagement in der medizinischen Versorgung

Das Ziel des Anwendungsfalls besteht darin, Patient*innen in den Mittelpunkt des Gesundheitsdatenmanagements zu stellen, indem ihnen die volle Kontrolle über ihre medizinischen Daten, welche entlang des gesamten Behandlungs- und Therapiepfades erhoben werden, ermöglicht wird. Patient*innen sollen jederzeit in der Lage sein, ihre in verschiedenen medizinischen Einrichtungen erzeugten und verwalteten Gesundheitsdaten einzusehen und aktiv zu verwalten. Durch die Bereitstellung einer digitalen Anwendung erhalten Patient*innen die Möglichkeit, detaillierte Zugriffsrichtlinien für ihre Daten zu definieren und selbst zu entscheiden, wer unter welchen Bedingungen Zugriff auf welche Dokumente erhält. Dies gewährleistet nicht nur Transparenz und Autonomie für die Patient*innen, sondern auch die Einhaltung der Datenschutzgrundverordnung (DSGVO), indem Patient*innen das Recht haben, die Löschung ihrer Daten in den verschiedenen Einrichtungen zu beantragen.

Der Anwendungsfall zielt darauf ab, eine nahtlose und sichere Integration der Gesundheitsdaten aus unterschiedlichen IT-Systemen der medizinischen Einrichtungen zu erreichen. Gleichzeitig wird sichergestellt, dass medizinische Leistungserbringende und vertrauenswürdige Zugehörige, die von den Patient*innen autorisiert wurden, Zugriff auf relevante Informationen zu geben, um die bestmögliche Versorgung zu gewährleisten. Durch die klare Trennung und Kontrolle

der Zugriffsrechte wird das Vertrauen der Patient*innen gestärkt, und eine personalisierte, patient*innenzentrierte Gesundheitsversorgung wird unterstützt.

5.1.1 Systemkontext

Dieser Unterabschnitt beschreibt den Systemkontext der geplanten Gesundheitsdatenmanagementanwendung. Es umfasst die beteiligten Akteure, einschließlich der Patient*innen, medizinischen Leistungserbringenden, Systemadministrierenden und Zugehörigen, sowie die wesentlichen bestehenden technologischen Komponenten, mit denen jenes System interagieren soll. Des Weiteren werden die Systemgrenzen beschrieben, die für die Integration und den Betrieb des Systems von Bedeutung sind.

Beteiligte Akteure und Rollen:
Patient*innen repräsentieren die Datensubjekte, deren Gesundheitsdaten in verschiedenen medizinischen Einrichtungen erzeugt und verwaltet werden. Jene Datensubjekte sollten in der Lage sein, Einblick in ihre über verschiedene Einrichtungen gesammelten Gesundheitsdaten zu erhalten und zu entscheiden, wer Zugriff auf diese Daten haben darf. Zu diesem Zweck soll ihnen eine digitale Anwendung zur Verfügung gestellt werden, die es ihnen ermöglicht, ihre Gesundheitsdaten (z.B. Daten aus ePA und eGA) sowie deren Metadaten (Datenerstellende, Erstellungszeitpunkt usw.) einzusehen. In dieser Anwendung sollen Patient*innen in der Lage sein, die Zugriffsrichtlinien für einzelne Dokumente granular zu definieren, d.h., festzulegen, wer unter welchen Bedingungen auf welches Dokument zugreifen darf. Empfangende solcher Zugriffsrechte können in der Zukunft sowohl medizinische Einrichtungen, die an der Behandlung und Therapie beteiligt sind, als auch Zugehörige sein. Daten aus der Selbstbewertung der Patient*innen können durch diese Anwendung erfasst und den medizinischen Einrichtungen zur Verfügung gestellt werden, die diese gemäß den Wünschen der Patient*innen an die anderen Akteure weitergeben. In Übereinstimmung mit der DSGVO können Patient*innen über die Anwendung das Recht auf Löschung ausüben, indem sie

die Löschung ihrer Gesundheitsdaten in den verschiedenen medizinischen Einrichtungen beantragen.

Leistungserbringende sind medizinische Fachkräfte, die in einer medizinischen Einrichtung arbeiten und medizinische Daten über Patient*innen nach oder während einer medizinischen Dienstleistung erstellen. Als Verfassende der medizinischen Dokumente können Leistungserbringende Daten lesen, bearbeiten oder löschen. Die verfassten Daten werden in internen IT-Systemen der medizinischen Einrichtungen gespeichert. Neben ihrer Rolle als Datenanbietende können Leistungserbringende auch als Konsumierende von Gesundheitsdaten anderer medizinischer Einrichtungen auftreten, wenn ihnen von Patient*innen Zugriffsrechte gewährt wurden.

Die **Systemadministrierende** der medizinischen Einrichtungen sind verantwortlich für den Betrieb der Software- und Hardwarekomponenten aller Systembestandteile in ihrer medizinischen Einrichtung. Alle von den Patient*innen in den medizinischen Einrichtungen erzeugten Daten werden über standardisierte Schnittstellen aus den bestehenden IT-Systemen an das Zielsystem bereitgestellt.

Die letzte Gruppe von Akteuren sind **Zugehörige**. Zugehörige können verwandte Personen oder aus Sicht der Patient*innen vertrauenswürdige Personen sein. Zugehörige nehmen nur am System teil, wenn Patient*innen ihnen Zugang gewährt hat. Sobald sie Zugang haben, können Zugehörige die mobile App nutzen, um die ihnen zur Verfügung gestellten Daten der Patient*innen einzusehen. Darüber hinaus können Patient*innen Zugehörige auch als Administratoren ihrer Gesundheitsdaten benennen. Mit diesen Rechten haben Zugehörige dieselben Privilegien bezüglich der Daten wie die Rolle der Patient*innen.

Bestehende technische Komponenten:
Interne IT-Systeme bezeichnen die technischen Systeme, die innerhalb der medizinischen Einrichtungen betrieben werden, wie beispielsweise Krankenhausinformationssysteme (KIS). Diese Systeme sind für die Speicherung, Verwaltung und Verarbeitung der Gesundheitsdaten von Patient*innen zuständig. Sie sollen die Daten bereitstellen, die in der geplanten Gesundheitsdatenmanagementanwendung verwaltet und weitergegeben werden sollen.

Systemkontext und -grenzen:
Das Systemkontextdiagramm in Abbildung 5.1 entsprechend des C4-Modells (siehe Abschnitt 2.3.2) veranschaulicht die Systemgrenzen und die Interaktionen mit den Akteuren und externen Komponenten. Das zu entwickelnde Gesundheitsdatenmanagementsystem ist in hellblau dargestellt und umfasst sämtliche Funktionen, die direkt der Verwaltung und Kontrolle der Gesundheitsdaten durch die Patient*innen dienen. Die Akteure, einschließlich Patient*innen, Leistungserbringende, Systemadministrierende und Zugehörige, sind in dunkelblau dargestellt. Die externen Systeme, insbesondere die internen IT-Systeme der medizinischen Einrichtungen, welche von außen mit dem System interagieren, sind in grau dargestellt.

Juristische Rahmenbedingungen:
Die juristischen Rahmenbedingungen für die Schaffung einer einrichtungsübergreifenden Patient*innenakte, wie im ersten Anwendungsfall vorgesehen, umfassen die Einhaltung der Datenschutz- und Datensicherheitsvorgaben gemäß DSGVO und BDSG sowie die spezifischen Regelungen des PDSG. Darüber hinaus müssen die gesetzlichen Grundlagen berücksichtigt werden, in die Infrastrukturen wie die TI eingebettet sind, darunter das GMG, das E-Health-Gesetz, das TSVG und das DVG sowie die Verordnung zum EHDS. Die detaillierten juristischen Rahmenbedingungen, welche sich aus diesen Gesetzgebungen ergeben, werden im Abschnitt 2.2 ausführlich beschrieben.

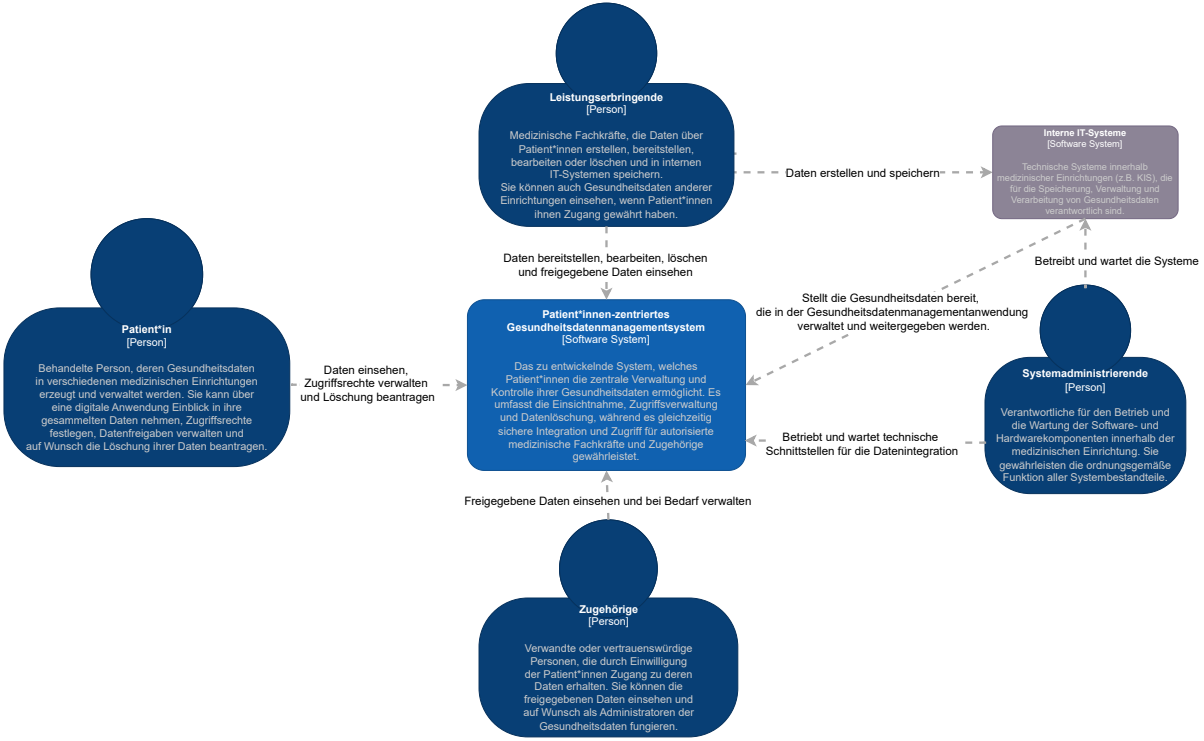

Abbildung 5.1: Systemkontextdiagramm des zu entwickelnden Patient*innen-zentrierten Gesundheitsdatenmanagementsystems.

5.1.2 Anforderungserhebung

Die Anforderungserhebung bildet einen essenziellen Schritt in der Entwicklung benutzendenzentrierter Systeme und trägt entscheidend zur Sicherstellung der Nutzendenzufriedenheit bei. Für den vorliegenden Anwendungsfall wurden detaillierte Anforderungen in Form von User Stories entwickelt, welche die funktionalen Bedürfnisse der relevanten Akteur*innen widerspiegeln. Zu diesem Zweck wurden in digitalen Workshops gemeinsam mit Vertreter*innen unterschiedlicher Fachdisziplinen – darunter Ärzt*innen, Forscher*innen, Vertreter*innen der Pflegewissenschaften und -einrichtungen, Expert*innen im Entlassmanagement, Softwarehersteller von Informationssystemen im Gesundheitswesen, technische Blockchain-Expert*innen sowie Informatiker*innen – die User Stories erarbeitet und konkretisiert. Die im Rahmen dieser Workshops entwickelten User Stories wurden anschließend mithilfe der MoSCoW-Methode [228] kategorisiert. Diese Methode erlaubt es, die Anforderungen nach ihrer Priorität zu ordnen, indem sie zwischen den Kategorien „Must have“, „Should have“, „Could have“ und „Won't have“ unterscheidet. Diese systematische Priorisierung stellt sicher, dass die wichtigsten Anforderungen frühzeitig in den Entwicklungsprozess einfließen und die weniger dringlichen oder nicht umzusetzenden Anforderungen entsprechend zurückgestellt werden.

Zusätzlich wurden Vertreter*innen von Patient*innenorganisationen, insbesondere der „Haus der Krebsselbsthilfe“, aktiv in den Anforderungsprozess einbezogen, um die Sichtweise der Betroffenen angemessen zu berücksichtigen. Diese direkte Einbindung von Patient*innenvertretungen ermöglichte eine Validierung der erarbeiteten User Stories aus der Perspektive der Endnutzer*innen. Die Rückmeldungen dieser Patient*innenvertretenden lieferte wertvolle Impulse, um den Anforderungskatalog kritisch zu überprüfen und gegebenenfalls anzupassen. Dadurch wurde sichergestellt, dass die finalen Anforderungen die tatsächlichen Bedürfnisse und Erwartungen der betroffenen Nutzendengruppen widerspiegeln. Die so priorisierten und validierten Anforderungen bildeten die Grundlage für die technische Entwicklung und wurden als Product Backlog kontinuierlich in

den Entwicklungsprozess integriert. In den folgenden Abschnitten werden sowohl die funktionalen als auch die nicht-funktionalen Anforderungen in Form von Qualitätsanforderungen kurz vorgestellt. Hierbei werden ausschließlich die „Must have"-Anforderungen dargelegt, da diese die zentralen Anforderungen des Anwendungsfalls beschreiben.

5.1.2.1 Funktionale Anforderungen

In diesem Abschnitt werden die funktionalen Anforderungen dargestellt, die auf Grundlage der definierten Systemakteure abgeleitet wurden.
Patient*innen:

- PFA-1: Als Patient*in möchte ich das die Verbindung zwischen meiner Identität, meinen Behandler*innen sowie sonstigen Gesundheitsdienstleistenden und Gesundheitsdaten für Dritte, die keinen Zugriff haben sollen, nicht ersichtlich wird, um meine Privatsphäre zu gewährleisten.
- PFA-2: Als Patient*in möchte ich meine über mich bei verschiedenen Institutionen gespeicherten Gesundheitsdaten übersichtlich abrufen können, um selbst einen umfassenden Überblick über die gespeicherten medizinischen Informationen zu erhalten.
- PFA-3: Als Patient*in möchte ich, dass mir die Daten verständlich, in angemessener Form und ethisch-konform zur Verfügung gestellt werden, um eine Überforderung zu vermeiden, Ängste zu reduzieren und eine zielführende Verwaltung meiner Gesundheitsdaten zu ermöglichen.
- PFA-4: Als Patient*in möchte ich die Möglichkeit haben, medizinische Einrichtungen oder zugehörige Fachpersonen als zugriffsberechtigte Akteure hinzuzufügen und die Zugriffsrechte für meine Gesundheitsdaten entsprechend meiner Bedürfnisse (feingranular und grobgranular) definieren zu können, damit ich festlegen kann, wer unter welchen Bedingungen auf welche Dokumente zugreifen darf.

- PFA-5: Als Patient*in möchte ich die Möglichkeit haben, anderen Ärzt*innen und Leistungserbringenden meine Behandlungsinformationen zur Verfügung zu stellen, damit durch das lückenlose und organisationsübergreifende Teilen der Daten eine optimale Behandlung und Therapie gewährleistet werden kann.
- PFA-6: Als Patient*in möchte ich die Möglichkeit haben, einzusehen, an welche Institutionen oder Personen ich Zugriff auf Gesundheitsdaten erteilt habe, um diese Freigabe zu validieren und einen Überblick über diese zu erhalten.
- PFA-7: Als Patient*in möchte ich, dass die Wahrscheinlichkeit einer unbefugten Nutzung meiner Gesundheitsdatenmanagementanwendung durch Dritte minimiert wird, um den Schutz meiner Gesundheitsdaten zu gewährleisten.
- PFA-8: Als Patient*in möchte ich die Möglichkeit haben, die Metadaten meiner Gesundheitsdokumente, einschließlich des Erstellungszeitpunkts und der datenerstellenden Person, einzusehen, um die Herkunft und den Kontext meiner Daten (z.B. Erstellungszeitpunkt) besser verstehen und nachvollziehen zu können.
- PFA-9: Als Patient*in möchte ich Dokumente, die in der Gesundheitsdatenmanagementanwendung hinterlegt wurden, suchen und filtern können, um entsprechende Dokumente schnell finden zu können.
- PFA-10: Als Patient*in möchte ich die Möglichkeit haben, selbst erhobenen Gesundheitsdaten (z.B. meine Selbsteinschätzungen zum täglichen Befinden über ein elektronisches Tagebuch) als Patient Reported Outcomes zu dokumentieren und Leistungserbringenden Freigaben zu Daten von Wearables, Health Apps und anderen Sensoren zu erteilen, damit ich meine Gesundheitsentwicklung verfolgen, relevante Informationen für meine Behandlung bereitstellen und die Behandlung unterstützen kann.
- PFA-11: Als Patient*in möchte ich die Möglichkeit haben, zu sehen, wann und welche Gesundheitsdaten ich selbst erhoben habe, um einen Überblick über die Zeitpunkte der Erhebung und den Datenumfang zu erhalten.

- PFA-12: Als Patient*in möchte ich das Recht haben, meine in der Gesundheitsdatenmanagementanwendung verfügbaren personenbezogenen Gesundheitsdaten, Selbsteinschätzungen und die erteilten Zugriffe jederzeit löschen zu können, um die Kontrolle über meine Daten gemäß der DSGVO zu gewährleisten und sicherzustellen, dass meine Privatsphäre respektiert wird (z.B. auch bei einem Arztwechsel oder Vertrauensverlust zu Behandelnden).
- PFA-13: Als Patient*in möchte ich die Möglichkeit haben, eine Systemvollmacht anzulegen und diese entsprechend meiner Wünsche zu befüllen, festzulegen, welcher Zugehörende bestimmte Rechte durch die Systemvollmacht erhält, sowie die Dauer der Vollmacht zu definieren, um die Autorisierung meiner Zugehörenden zu ermöglichen und die Übersicht über alle von mir ausgestellten Systemvollmachten zu behalten, damit ich diese verwalten und einsehen kann.

Leistungserbringende:

- LFA-1: Als Leistungserbringende möchte ich Gesundheitsdaten während oder nach einer medizinischen Dienstleistung erstellen können, um die Behandlung der Patient*innen effizient zu dokumentieren.
- LFA-2: Als Leistungserbringende möchte ich von mir erhobene Gesundheitsdaten einfach und schnell ablegen sowie Patient*innen bereitstellen können, um mich auf die medizinische Betreuung/Behandlung zu konzentrieren und nicht Zeit mit Dokumentation, Verknüpfung und Bereitstellung von Informationen zu verschwenden.
- LFA-3: Als Leistungserbringende möchte ich, dass die Gesundheitsdatenmanagementanwendung eine einfache Integration und Bereitstellung von Daten aus bestehenden und etablierten Systemen in meiner Einrichtung ermöglicht, um Informationen effizient an relevante Dritte weiterzugeben.
- LFA-4: Als Leistungserbringende möchte ich die Identität anderer Teilnehmenden einsehen und validieren können, um Missbrauch durch Dritte zu vermeiden und sicherzustellen, dass relevante medizinische Einrichtungen sowie von mir

behandelte Patient*innen Zugang zu den von mir erhobenen Gesundheitsdaten erhalten.

- LFA-5: Als Leistungserbringende möchte ich Zugriff auf die Gesundheitsdaten meiner Patient*innen erhalten, die von anderen Institutionen bereitgestellt wurden, sofern ich die entsprechende Erlaubnis von diesen Institutionen sowie die Zustimmung der Patient*innen erhalten habe, um eine umfassende und koordinierte medizinische Versorgung gewährleisten zu können sowie fundierte medizinische Entscheidungen treffen zu können.
- LFA-6: Als Leistungserbringende möchte ich das externe Daten vertrauenswürdig und korrekt sind, um valide Daten in weitere Behandlungs- und Therapieschritte einbeziehen zu können.
- LFA-7: Als Leistungserbringende möchte ich die Möglichkeit haben, relevante Dokumente, die ich hinterlegt habe, zu aktualisieren, um Anpassungen vorzunehmen und die Korrektheit der Informationen zu gewährleisten, die den aktuellen medizinischen Anforderungen entsprechen.
- LFA-8: Als Leistungserbringende möchte ich die Möglichkeit haben, relevante Dokumente, die ich hinterlegt habe, zu löschen, um Fehler bei Freigaben korrigieren zu können sowie meiner Löschpflichten gegenüber meiner Patient*innen nachkommen zu können.
- LFA-9: Als Leistungserbringende möchte ich über die Gesundheitsdatenmanagementanwendung schnell und einfach auf die relevanten medizinischen Informationen meiner Patient*innen zugreifen können, um eine informierte und qualitativ hochwertige medizinische Versorgung sicherzustellen.

Systemadministrierende:

- SFA-1: Als Systemadministrierende möchte ich den Betrieb der Software- und Hardwarekomponenten der IT-Systeme sicherstellen, um einen reibungslosen Betrieb der medizinischen Einrichtungen zu gewährleisten.

- SFA-2: Als Systemadministrierende möchte ich sicherstellen, dass alle von den Patient*innen in den Einrichtungen erzeugten Daten über standardisierte Schnittstellen an das Zielsystem bereitgestellt werden, um die Interoperabilität, Integration und effiziente Weitergabe der Daten zu ermöglichen.
- SFA-3: Als Systemadministrierende möchte ich die Möglichkeit haben, meine Dienste mit der Gesundheitsdatenmanagementanwendung zu koppeln, um den Nutzer*innen die Synergien beider Plattformen als Mehrwert zu bieten, insbesondere durch die Integration und den Austausch relevanter Daten.
- SFA-4: Als Systemadministrierende möchte ich Zugriff auf umfassende Protokolle und Berichte über Systemnutzung und -leistung haben, um mögliche Probleme frühzeitig zu erkennen und die Effizienz der IT-Systeme zu optimieren.
- SFA-5: Als Systemadministrierende möchte ich die Benutzendenverwaltung für alle IT-Systeme der medizinischen Einrichtung durchführen können, um den Zugriff auf sensible Daten zu steuern und die Sicherheit der Systeme zu gewährleisten.
- SFA-6: Als Systemadministrierende möchte ich regelmäßige Software-Updates und -Patches einspielen können, um die Systeme vor Sicherheitsrisiken zu schützen und die Funktionsfähigkeit der Software zu gewährleisten.

Zugehörige:

- ZFA-1: Als Zugehörige, der von Patient*innen als Administrierende im Rahmen einer Systemvollmacht benannt wurde, möchte ich dieselben Privilegien wie die Patient*innen haben, um ihre Gesundheitsdaten zu verwalten und Zugriffsrechte zu steuern, damit ich die mir nahestehende Person in der Verwaltung unterstützen kann, wenn sie dies z.B. nicht mehr selbstständig tun kann oder meine Meinung dazu wünscht.
- ZFA-2: Als Zugehörige möchte ich die von Patient*innen freigegebenen Gesundheitsdaten einsehen können, um den Gesundheitszustand der Patient*innen nachvollziehen zu können.

- ZFA-3: Als Zugehörige möchte ich sehen können, wer eine Systemvollmacht auf mich ausgestellt hat, um diese sinnvoll verwalten zu können.
- ZFA-4: Als Zugehörige möchte ich die Dauer einer Systemvollmacht einsehen können, um deren Laufzeit abschätzen zu können.
- ZFA-5: Als Zugehörige möchte ich meine durch die Systemvollmacht abgedeckten Pflichten einsehen können, um diese zu verinnerlichen und zu berücksichtigen.

5.1.2.2 Qualitätsanforderungen

In diesem Abschnitt werden die nicht-funktionalen Anforderungen beschrieben, die auf den Qualitätsmerkmalen des ISO-Standards 25010 basierend abgeleitet wurden.

- NFA-1 - Bedienbarkeit: Die Gesundheitsdatenmanagementanwendung muss benutzendenfreundlich und intuitiv bedienbar sein, sodass die Nutzer*innen ohne umfangreiche Schulung ihre Gesundheitsdaten effizient einsehen und verwalten können.
- NFA-2 - Zeitverhalten: Die Gesundheitsdatenmanagementanwendung muss innerhalb von 10 Sekunden auf Nutzendeneingaben reagieren, um Frustration zu vermeiden.
- NFA-3 - Sicherheit: Die Gesundheitsdatenmanagementanwendung muss höchste Anforderungen an Datenschutz und Datensicherheit erfüllen, um den Schutz der Nutzendendaten gemäß den Vorgaben der DSGVO sicherzustellen.
- NFA-4 - Datenintegrität: Die Gesundheitsdatenmanagementanwendung muss sicherstellen, dass alle erfassten, übermittelten und gespeicherten Daten vollständig, unverändert und vertrauenswürdig sind.

- NFA-5 - Vertraulichkeit: Die Gesundheitsdatenmanagementanwendung muss sicherstellen, dass alle persönlichen und medizinischen Daten streng vertraulich behandelt werden.
- NFA-6 - Skalierbarkeit: Die Gesundheitsdatenmanagementanwendung muss eine hohe Skalierbarkeit gewährleisten, um eine kontinuierliche und effiziente Verarbeitung von Nutzer*inneninteraktionen zu ermöglichen. Laut der gematik wird im Rahmen des Mengengerüsts der TI von etwa 73 Millionen potenziellen Nutzenden und 173.149 Konnektoren im Produktivbetrieb ausgegangen [229]. Die Nutzenden setzen sich aus 70 Millionen gesetzlich Versicherten und 266.600 medizinischen Leistungserbringenden zusammen [229].
- NFA-7 - Interoperabilität: Die Gesundheitsdatenmanagementanwendung muss eine umfassende Interoperabilität gewährleisten, indem sie alle relevanten Systeme und Standards im Gesundheitswesen unterstützt und integriert. Dies schließt die nahtlose Kommunikation mit etablierten IT-Systemen in medizinischen Einrichtungen sowie die Einhaltung international anerkannter Standards wie HL7, FHIR und DICOM ein.
- NFA-8 - Verfügbarkeit: Die Gesundheitsdatenmanagementanwendung muss eine hohe Verfügbarkeit sicherstellen, um einen kontinuierlichen und zuverlässigen Zugriff auf die Gesundheitsdaten der Nutzer*innen zu ermöglichen. Dies erfordert robuste Ausfallsicherheitsmechanismen, redundante Systemarchitekturen und eine effiziente Fehlerbehebung, um Systemausfälle zu minimieren und Betriebsunterbrechungen auf ein absolutes Minimum zu reduzieren.
- NFA-9 - Authentizität: Die Gesundheitsdatenmanagementanwendung muss sicherstellen, dass alle gespeicherten Gesundheitsdaten authentisch und nachvollziehbar sind. Patient*innen sollen jederzeit transparent einsehen können, welche Daten erfasst wurden, wer Zugriff darauf hat, und wann sowie durch wen Änderungen vorgenommen wurden.
- NFA-10 - Modularität: Die Gesundheitsdatenmanagementanwendung muss sicherstellen, dass die erfassten Daten am ursprünglichen Standort verbleiben

und nahtlose Schnittstellen zwischen diesen Datenquellen durch eine modulare Architektur geschaffen werden.

5.1.3 Systemabstraktion und Sicherheitsbetrachtungen

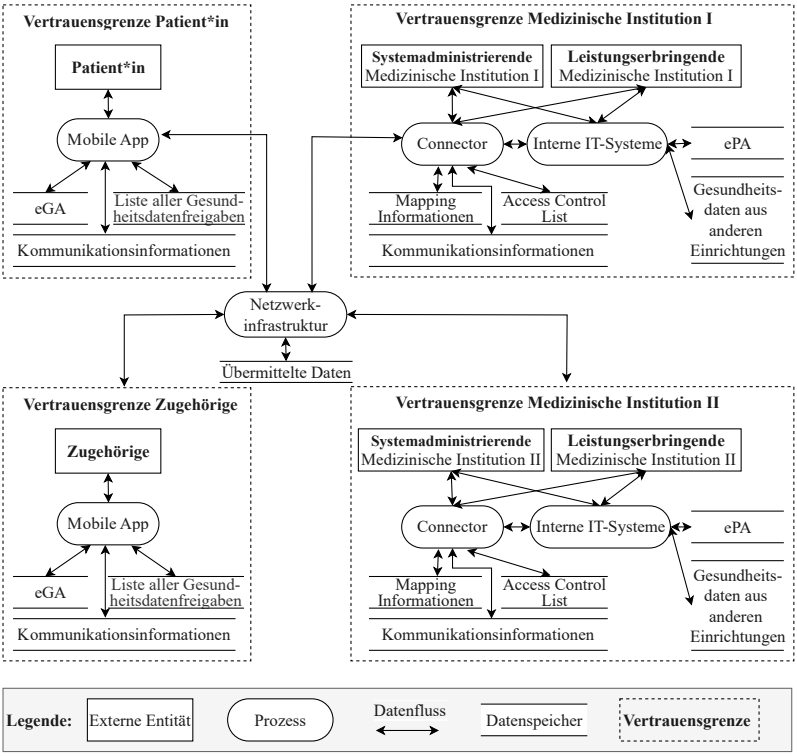

Abbildung 5.2: Datenflussdiagramm des zu entwickelnden Gesundheitsdatenmanagementsystems.

Die wesentlichen technischen Komponenten und Datenflüsse des zu entwickelnden Gesundheitsdatenmanagementsystems werden in einem Datenflussdiagramm

(DFD) als Systemabstraktion dargestellt (siehe Abbildung 5.2). Das System umfasst drei Hauptkomponenten: die mobile App, die Netzwerkinfrastruktur und den Connector. Diese Komponenten sind in den im Abschnitt 5.1.1 beschriebenen Systemkontext integriert. Im Folgenden werden die einzelnen Komponenten sowie deren Datenflüsse detailliert beschrieben, und es erfolgt eine Betrachtung der erforderlichen Sicherheitsaspekte, einschließlich der Vertrauensannahmen und der schützenswerten Güter (*engl. assets*).

Mobile App: Die zentrale Plattform für Patient*innen zur Einsichtnahme und Verwaltung ihrer Gesundheitsdaten. Die App ermöglicht es, Daten (wie ePA und eGA) und deren Metadaten einzusehen sowie Zugriffsrichtlinien zu verwalten.

Connector: Ein Systemelement, das die Kommunikation zwischen den internen IT-Systemen der medizinischen Einrichtungen und der Netzwerkinfrastruktur ermöglicht. Der Connector sorgt für den sicheren Austausch und die Verknüpfung der Daten (z.B. Identifier) im Netzwerk und den angebundenen internen IT-Systemen.

Interne IT-Systeme: Die Systeme innerhalb der medizinischen Einrichtungen, wie Krankenhausinformationssysteme (KIS), in denen die Gesundheitsdaten gespeichert und verwaltet werden.

Netzwerkinfrastruktur: Die zugrundeliegende P2P-Infrastruktur, welche die Kommunikation und Datenübertragung zwischen der mobilen App, dem Connector und den internen IT-Systemen ermöglicht.

Vertrauensannahmen: Bedrohungen für ein System können von internen, autorisierten Agenten oder von externen, unbefugten Dritten mit bösartigen Absichten ausgehen. Da interne Agenten Zugang zu bestimmten Systemressourcen haben, ist der Schutz gegen sie aufwendiger als gegen externe Agenten und hängt von deren Zugang, Wissen, Rechten, Fähigkeiten, Risiken, Taktiken und Motivation ab [230]. Die Annahmen zu den Vertrauensebenen für alle vier internen Rollen und die entsprechenden Komponenten sind wie folgt beschrieben:

- **VA1:** Patient*innen können nur über die mobile App auf die bereitgestellten Daten zugreifen, die im Kontext ihrer Versorgung über sie generiert und gesammelt wurden (einschließlich Metadaten), auf die patient*innenbezogenen Daten, die lokal in ihrer eigenen mobilen App generiert wurden, sowie auf die Systeminformationen, die für die sichere Übertragung der Daten über die Netzwerkinfrastruktur erstellt wurden (z.B. Identifier, Schlüsselmaterial, Berechtigungsnachweise). Sie dürfen keine (Meta-)Daten anderer Systemnutzenden einsehen, es sei denn, sie haben Zugang von der betroffenen Person erhalten. Patient*innen könnten bösartig (*engl. malicious*) handeln, um auf Funktionen, Berechtigungen und Daten zuzugreifen, zu denen sie keinen Zugang haben. Aufgrund eines häufigen Mangels an IT-Expertise und persönlicher Interessen an der Plattform wird davon ausgegangen, dass Patient*innen größtenteils nicht bösartig handeln, sondern ehrlich aber neugierig (*engl. honest-but-curious*).
- **VA2:** Leistungserbringende haben Zugang zu allen Gesundheitsdaten ihrer eigenen Patient*innen, die von ihrer eigenen Institution im Rahmen vertraglicher Vereinbarungen für Behandlungs- und Dokumentationsanforderungen generiert wurden, oder zu Daten, die von anderen Institutionen mit Zustimmung der Patient*innen diesen bereitgestellt wurden. Leistungserbringende haben keinen Zugriff auf nicht geteilte Daten von anderen Leistungserbringenden sowie auf nicht geteilte Selbsteinschätzungsdaten. Ähnlich wie bei Patient*innen können Leistungserbringende böswillig handeln. Ihre IT-Expertise kann sowohl gering als auch ausgeprägt sein. Es wird angenommen, dass gesetzliche und soziale Rahmenbedingungen den Missbrauch von Daten durch Leistungserbringende unwahrscheinlich machen.
- **VA3:** Zugehörige sehen nur die Daten der Patient*innen, auf die sie durch dessen Freigabe Zugang erhalten haben, und können nur die Daten der Patient*innen verwalten, für die sie eine Autorisierung erhalten haben. In Bezug auf bösartiges Verhalten, Motivation und IT-Fähigkeiten können ähnliche Vertrauensannahmen wie für die Patient*innen getroffen werden.
- **VA4:** Systemadministrierende haben Zugriff auf alle IT-Systeme in ihrer eigenen medizinischen Einrichtung, um deren Funktionalität zu gewährleisten. Sie

sollten jedoch keinen direkten Zugriff auf die patient*innenbezogenen Daten haben. Als Administrierende des Connectors haben sie Einblick in die Informationen zur Zuordnung von Patient*innenidentitäten mit den Identifiern in der Netzwerkinfrastruktur. Es wird davon ausgegangen, dass die von den Systemadministrierenden verwalteten Systeme den Protokollen folgen und dementsprechend vertrauenswürdig sind. Neben den Gesundheitsdaten, die in ihrer eigenen Einrichtung generiert wurden, speichern die Systeme auch zuverlässig Daten, die von anderen medizinischen Einrichtungen bereitgestellt wurden. In Bezug auf den Connector wird erwartet, dass die im internen IT-Systemen gespeicherten Daten für Datenanfragen im Netzwerk bereitgestellt werden. Der Connector setzt die von den Patient*innen festgelegten Zugriffsrichtlinien durch. Die Systemadministrierenden verfügen über fundierte IT-Kompetenzen. Es gibt jedoch vertragliche und gesetzliche Rahmenbedingungen, die einen aktiven Missbrauch der Daten unwahrscheinlich erscheinen lassen. Schlecht geschulte Systemadministrierende können zu unbeabsichtigten Bedrohungen führen.

- **VA5:** Die Systemkomponenten, auf denen die verschiedenen Prozesse außerhalb der medizinischen Einrichtungen ablaufen (z.B. mobile Geräte, Netzwerkinfrastruktur), folgen den Protokollen. Bedrohliches Verhalten von externen Identitäten kann nicht ausgeschlossen werden.

Schützenswerte Güter: Schützenswerte Güter werden im Rahmen dieser Sicherheitsbetrachtung als Güter definiert, die innerhalb des Systems gegen Missbrauch geschützt werden müssen. Die für die Gesundheitsdatenmanagementanwendung identifizierten schützenswerten Güter sind in Tabelle 5.1 beschrieben.

Tabelle 5.1: Die identifizierten schützenswerten Güter der Gesundheitsdatenmanagementanwendung und deren Beschreibung.

Komponente	ID	Name des schützenswerten Guts	Beschreibung des Guts
Mobile App	A1	Patient*innen-bezogene Daten aus Sicht der Patient*innen	Alle lokal von der mobilen App der Patient*innen im Kontext der Selbsteinschätzung gesammelten Daten sowie alle von verschiedenen Gesundheitsinstitutionen bereitgestellten und empfangenen Daten über die dazugehörigen Patient*innen (ePA/eGA-Daten).
	A2	Mit Zugehörigen geteilte patient*innen-bezogene Daten	Alle von Patient*innen mit einem Zugehörigen geteilten Daten.
	A3	Kommunikationsinformationen der Nutzenden	Netzwerkinterne Identifier der mobilen Geräte-Nutzenden und Netzwerkadressen, die für den Austausch von Nachrichten und Daten über die Netzwerkinfrastruktur erforderlich sind.
	A4	Liste aller Gesundheitsdatenfreigaben	Informationen über bestehende Kommunikationsbeziehungen mit den bereitstellenden und konsumierenden Parteien sowie deren Austauschinhalt und Zugriffsrechte.
Netzwerkinfrastruktur	A5	Patient*innen-bezogene Daten	Die patient*innenbezogenen Daten und Nachrichten, die über die Netzwerkinfrastruktur ausgetauscht werden sollen.

	A6	Identifizierende Kommunikationsinformationen	Netzwerkinterne Identifier von Nutzenden und Netzwerkadressen, die für den Austausch von Nachrichten und Daten erforderlich sind.
Connector	A7	Zugriffsprotokolle	Liste aller Zugriffe, Status über den Erfolg des Zugriffs und die zugehörigen Daten und Metadaten des Anfordernden (z.B. Netzwerkadresse, Datum und Uhrzeit des Zugriffs).
	A8	Access Control List (ACL)	Die ACL definiert das Ausmaß, in dem einzelne Nutzenden und Systeme Zugang zu den internen Objekten in den medizinischen Einrichtungen haben (z.B. Dienste, Dateien usw.).
	A9	Zuordnungsinformationen	Interne Zuordnung der Identitäten verschiedener Nutzenden im Netzwerk zu denen in den internen IT-Systemen.
Interne IT-Systeme	A10	Patient*innenbezogene Daten	Die in internen IT-Systemen im Kontext der internen Behandlung gesammelten Daten sowie die von anderen Leistungserbringenden bereitgestellten Daten.

Technische Annahmen: Neben den beteiligten Rollen, Komponenten und Rechten wurden folgende technische Annahmen getroffen:

- **TA1:** Es sollten Prinzipien der Datenminimierung angewendet werden, um keine redundanten Datensilos zu schaffen.
- **TA2:** Grundlegende Änderungen am internen IT-Systemen, abgesehen von den Schnittstellen zum Connector, sollten nicht vorgenommen werden.

- **TA3:** Es kann nicht davon ausgegangen werden, dass ein mobiles Gerät eines Nutzenden stets verfügbar ist. Auch auf Seiten der medizinischen Einrichtungen kann eine ständige Verfügbarkeit der Server nicht immer erwartet werden.
- **TA4:** Die medizinische Einrichtung, welche die Daten generiert hat, stellt diese über die gesamte Dauer der Datenspeicherung in ihrem Zuständigkeitsbereich bereit.

5.1.4 Sicherheits- und Bedrohungsmodellierung

In der vorangegangenen Betrachtung wurde die grundlegende Systemidee sowie die vorgesehenen Datenflüsse für das zu entwickelnde Systemkonzept umfassend dargestellt. Der nächste Schritt gemäß der STRIDE-Methode (siehe Abschnitt 2.3.12) besteht darin sicherheitsrelevante Schwachstellen und potentielle Bedrohungen, die im Rahmen des definierten Systems auftreten können, zu identifizieren und zu analysieren. Abbildung 5.3 bietet einen Überblick über die potenziellen Bedrohungen, die im Rahmen des vorliegenden Anwendungsfalls systematisch mithilfe der STRIDE-Methode identifiziert wurden. In diesem Zusammenhang wurden spezifische Angriffsvektoren berücksichtigt, die charakteristisch für Systeme sind, die auf Blockchain-Technologie basieren. Dazu zählen Sybil-Angriffe, egoistisches Mining, 51% -Angriffe und Wallet-Diebstahl [231].

Auf Grundlage der identifizierten Bedrohungen wurden potenzielle Schutzmechanismen abgeleitet, wobei die von Shostack [168] beschriebenen Gegenmaßnahmen und Minderungstaktiken berücksichtigt wurden. Diese Schutzmechanismen sind in der rechten Spalte von Abbildung 5.3 aufgeführt. Beim Aufbau eines dezentralen Systems sollte besonderes Augenmerk auf die Überprüfung der Identitäten der Kommunikationspartner sowie auf die Integrität des Systems und der Kommunikationskanäle gelegt werden, um die Offenlegung sensibler Gesundheitsdaten zu verhindern. Im medizinischen Kontext geht oft ein persönlicher Kontakt mit dem Gesundheitsdienstleistenden einer Speicherung von Gesundheitsdaten

voraus. Dieser Kontaktpunkt kann als Basis für eine vertrauenswürdige Identifizierung und den Zugang zum System dienen, wobei entsprechende Identifikatoren ohne persönliche Referenzen und kryptografische Schlüssel im direkten Kontakt ausgetauscht werden können. Fehlt dieser direkte Kontakt, wird es potenziellen Angreifenden erschwert, einen Kommunikationskanal aufzubauen. Die Authentizität der über diesen sicheren Kommunikationskanal übertragenen Daten kann durch digitale Signaturen gewährleistet werden.

Sobald der Kommunikationskanal eingerichtet ist, sollte es Angreifenden unmöglich gemacht werden, Informationen aus dem Netzwerkverkehr und -inhalt zu extrahieren. Daher sollten Ende-zu-Ende-Verschlüsselung und sichere Netzwerkprotokolle implementiert werden. Zudem ist der Einsatz von Mix-Netzwerken erforderlich, um die Identitäten der beiden Kommunikationspartner sowie die Frequenz ihrer Datenübertragungen zu verschleiern. Neben dem Datentransfer sollte innerhalb der Zielsysteme sichergestellt werden, dass vertrauliche Daten verschlüsselt und Passwörter gehasht werden. Um den Zugriff Dritter im Falle eines Verlusts oder Diebstahls dieser Daten zu verhindern, sollten die Anwendungen durch Passwortschutz gesichert sein.

Die Prinzipien von Privacy-by-Design und Privacy-by-Default sollten implementiert werden, um potenzielle Bedrohungen, die aus fehlerhaften Benutzendeninteraktionen resultieren könnten, zu minimieren. Darüber hinaus sollten geeignete Zugangskontrollmechanismen eingesetzt werden, um den Zugriff ausschließlich den Akteuren zu gestatten, die von Patient*innen autorisiert wurden. Eine umfassende Protokollierung von Authentifizierungsaktionen, Zugriffsberechtigungen, Datenanforderungen und Datenabfragen ist notwendig, um die Abstreitbarkeit der durchgeführten Aktionen zu gewährleisten.

Abbildung 5.3: Die identifizierten Sicherheitsbedrohungen im ersten Anwendungsfall und deren Beschreibung sowie die Sicherheitsmaßnahmen zur Abwehr dieser Bedrohungen.

5.1.5 Entwickelte Systemarchitektur

Der folgende Abschnitt beschreibt die entwickelte Systemarchitektur, die auf dem zuvor erläuterten Bedrohungsmodellierungsprozess basiert. Im Zentrum dieser Architektur steht die Netzwerkinfrastruktur, welche die dezentralen Peers (mobile Apps und die internen Krankenhausinformationssysteme (KIS) medizinischer Einrichtungen) miteinander verbindet und die Kommunikation sowie den Datenaustausch zwischen Patient*innen, Zugehörigen und/oder medizinischen Leistungserbringenden ermöglicht. Die Netzwerkinfrastruktur bildet ein dezentrales Peer-to-Peer-Kommunikationsnetzwerk, das durch Blockchain-Technologie ergänzt wird. Dieses Netzwerk entspricht dem modernen Open-Source-Standard DIDComm[1] (*SM2, SM4, SM7, SM8, SM24*) und basiert auf den Open-Source-Implementierungen von Hyperledger Aries und Indy[2], die von der Hyperledger Foundation bereitgestellt werden.

Die Verbindung, Kommunikation und der Datentransfer innerhalb der Netzwerkinfrastruktur erfolgen über das dezentrale DIDComm-Netzwerk, das aus Aries-Agenten[3] besteht. Ein solcher Agent ist eine Softwarelösung zur Speicherung privater kryptografischer Schlüssel, Verbindungen und Berechtigungsnachweise in einem digitalen Wallet, die für die Ende-zu-Ende-Verbindung zu anderen Agenten und Blockchain-Knoten sowie für die verschlüsselte Nachrichtenübermittlung zwischen Edge-Agenten unter Verwendung des DIDComm-Protokolls erforderlich ist (*SM16-18*). Edge-Agenten werden in diesem Kontext als die in mobilen Apps integrierten Aries Mobile Agents sowie die von medizinischen Einrichtungen verwendeten Aries Cloud Agents definiert. Die sogenannten Aries Mediator Agents[4] fungieren dabei als eine Art Postfach, das eine asynchrone und verschlüsselte Kommunikation zwischen Edge-Agenten ermöglicht, wodurch die kontinuierliche Verfügbarkeit mobiler Agenten nicht erforderlich ist. Der Mediator-Agent

1 https://identity.foundation/didcomm-messaging/spec/
2 https://www.hyperledger.org/use
3 https://github.com/hyperledger/aries
4 https://github.com/hyperledger/aries-rfcs/blob/main/concepts/0046-mediators-and-relays/README.md

wird vom Edge-Knoten konfiguriert, besitzt eigene kryptografische Schlüssel und stellt Nachrichten erst nach Entschlüsselung einer äußeren, kryptographischen Umschlagsschicht (*engl. envelope*) zu, um Weiterleitungsanfragen (*engl. forward-requests*) zu erkennen.

Zur Vermeidung von Korrelationen, die auf Beziehungen zwischen Agenten basieren, kann jeder Edge-Agent eine Konfiguration mehrerer Mediator-Agenten als Mix-Netzwerke nutzen, die dem Absender nicht vollständig bekannt sind. Dadurch wird sichergestellt, dass eingehende Nachrichten anonymisiert über das Netzwerk empfangen werden können (*SM19*). Darüber hinaus kann der Edge-Agent zur Verschleierung eigener Netzwerkaktivitäten verschiedene Kommunikationswege über die vermittelnden Mediator-Agenten nutzen, um Nachrichten zu senden und zu empfangen (*SM20*). Zusätzlich wird das DIDComm-Netzwerk durch eine Blockchain für Selbstbestimmte Identitäten (SSI-Blockchain) ergänzt, die es den Datensubjekten ermöglicht, ihre eigenen Identitätsdaten zu besitzen und zu kontrollieren, und somit als zusätzlicher Vertrauensanker für die Netzwerkteilnehmenden dient.

Um die Kontrolle der Datensubjekte über ihre Identitätsdaten sicherzustellen, stellt die öffentliche SSI-Blockchain eine dezentrale öffentliche Schlüsselinfrastruktur bereit. Diese Infrastruktur veröffentlicht kryptografische digitale Schlüssel und Netzwerkadressen als eindeutige, dezentrale Identifikatoren, sogenannte public Decentralised Identifiers (public DIDs) [232]. In der Systemarchitektur werden die public DIDs von medizinischen Einrichtungen und Zugehörigen in der permissioned Indy SSI-Blockchain registriert, wodurch Patient*innen und andere medizinische Einrichtungen neue Agenten-zu-Agenten-Verbindungen aufbauen und vertrauenswürdig auf Basis verifizierter digitaler Identitäten (*SM27*) kommunizieren können. Darüber hinaus speichert die SSI-Blockchain Datenmodelle (*engl. schemas*) und Berechtigungsnachweisdefinitionen (*engl. credential definitions*) zur Ausstellung und Überprüfung zertifizierter und digital signierter Identitätsnachweise (*engl. credentials*) über das DIDComm-Netzwerk. Diese Credentials werden beispielsweise von staatlichen Stellen, medizinischen Verbänden oder medizinischen Institutionen ausgestellt (z.B. ein Versicherungsnachweis der Krankenkassen oder ein Heilberufsausweis).

Zusätzlich stellt die Blockchain ein Widerrufsregister (*engl. revocation registry*) zur Verfügung, mit dem die Gültigkeit eines Credentials aufgehoben werden kann. Das DIDComm-Messaging-Protokoll bietet die notwendigen Standards für die Ausstellung und den Austausch solcher digitalen Credentials. Mithilfe dieser verifizierbaren Berechtigungsnachweise kann die Authentizität einer Identität einer Person sowohl technisch als auch manuell bestätigt werden, wenn zwei Edge-Agenten vor Ort eine neue Verbindung aufbauen. Zur Minimierung der Korrelation von Informationen und Identitäten wird für jede private Verbindung zwischen zwei Agenten im DIDComm-Netzwerk ein paarweiser Peer-DID[5] generiert, der ausschließlich den beiden beteiligten Parteien bekannt ist. Zusätzlich besteht der Connector der medizinischen Einrichtung aus einem API-Gateway, einer Metadaten-Datenbank und einem Web-Frontend. Das API-Gateway dient als zentrale Schnittstelle für Verbindung, Kommunikation, Zugriffs- und Datenmanagementfunktionen der medizinischen Einrichtungen und ist über RESTful Webservices zugänglich. Systemadministrierende medizinischer Einrichtungen können über dieses API-Gateway medizinische Daten aus den internen IT-Systemen ins Netzwerk einpflegen sowie Daten von anderen Einrichtungen empfangen. Die Metadaten-Datenbank speichert Metadaten, die für die Systemfunktionen relevant sind, wie beispielsweise die Zuordnung von Peer-DIDs zu Benutzenden-IDs aus den internen IT-Systemen, eine Liste der behandelnden Fachkräfte, zeitlich begrenzte Zugriffsrichtlinien oder Zugriffsprotokolle (*SM10, SM12, SM13, SM25, SM26*).

Das Web-Frontend ermöglicht medizinischen Leistungserbringenden die Interaktion mit der Geschäftslogik des Systems, etwa zur Herstellung neuer Verbindungen zu Patient*innen oder zur Bereitstellung von Daten aus den internen IT-Systemen der medizinischen Einrichtung. Die mobile App ist eine passwortgeschützte SSI-fähige Anwendung, die auf einem Aries Mobile Agent aufbaut. Sie stellt die vorgesehenen Verbindungs-, Kommunikations-, Zugriffs- und Datenmanagementfunktionen bereit und ermöglicht Benutzendeninteraktionen über eine Benutzendenoberfläche. Darüber hinaus umfasst die App die verschlüsselte

[5] https://identity.foundation/peer-did-method-spec/

Speicherung vertraulicher Daten sowie die Protokollierung von Authentifizierungsaktionen (*SM3, SM6, SM9, SM11, SM14-16, SM21-23*).

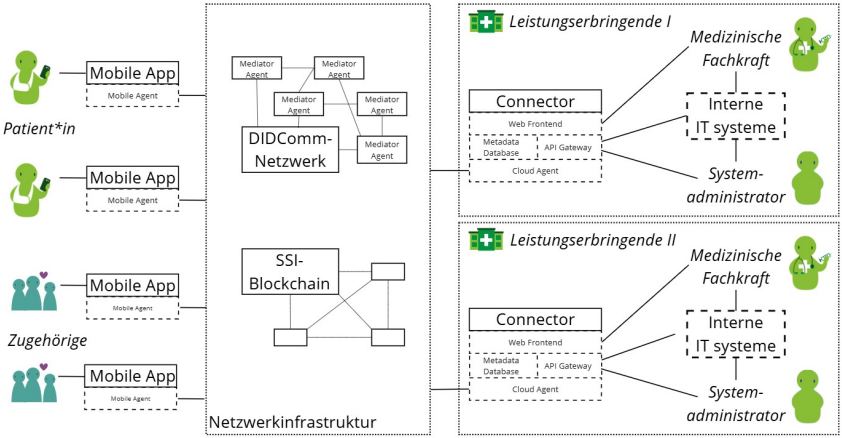

Abbildung 5.4: Die entwickelte Systemarchitektur der Gesundheitsdatenmanagementanwendung.

Chronologische Beschreibung des Kommunikations- und Datenverarbeitungsprozess innerhalb der Systemarchitektur: Im Folgenden wird der Prozess beschrieben, wie ein Datensubjekt (z.B. ein Patient) mit einer medizinischen Einrichtung verknüpft werden kann und wie die durch diese Verknüpfung zugänglich gemachte Daten sowohl dem Datensubjekt selbst als auch weiteren Datenkonsumenten (z.B. anderen medizinischen Einrichtungen oder Zugehörigen) bereitgestellt werden können. In Abbildung 5.5 ist dieser Kommunikationsablauf grafisch dargestellt.

1.) Registrierung und Aufbau einer neuen Verbindung: Um als Patient*in Gesundheitsdaten von einer Klinik abzurufen oder Freigaben für Datenkonsumenten zu erstellen, ist eine einmalige persönliche Registrierung vor Ort erforderlich (*SM1*). Zur Initiierung der Registrierung scannt das Datensubjekt vor Ort einen Einladung-QR-Code mit der mobilen App, welche von der Klinik bereitgestellt

wird. Dieser QR-Code dekodiert die öffentlich auflösbare DID der datenbereitstellenden Klinik. Die Endpunktinformationen und der öffentliche Schlüssel in dieser öffentlichen DID werden daraufhin verwendet, um eine verschlüsselte Verbindungsanfrage zurück an die Klinik zu senden, um eine Verbindung herzustellen und entsprechende Peer DIDs zu generieren und auszutauschen, die nur zwischen den beiden Akteuren bekannt sind (z.B. DID_{OP} des Datensubjekts für eine Verbindung mit der Klinik mit der DID_{PO} oder DID_{OC} des Datensubjekts für eine Verbindung mit dem Datenkonsumenten mit der DID_{CO}). Nach erfolgreicher Verbindung befinden sich beide Akteure in der Liste der registrierten Kontakte des jeweils anderen und können in der mobilen App oder der Webanwendung der Leistungserbringenden angezeigt werden (*SM5*). Zusätzlich wurden die für die weitere Kommunikation notwendigen kryptografischen Schlüssel in ihrer passwortgeschützten digitalen Wallet abgelegt. Als zusätzlichen Identitätsnachweis können die beiden Akteure digitale Berechtigungsnachweise austauschen, um diese mit physischen Ausweisdokumenten im persönlichen Kontakt abzugleichen und zu überprüfen (*SM1*).

2.) Datenbereitstellung an das Datensubjekt durch eine Klinik: Der Leistungserbringende der Klinik speichert die zu teilenden Gesundheitsdaten in seinem internen IT-System und teilt sie über das Web-Frontend mit dem Datensubjekt in dessen Kontaktliste. Das API-Gateway wird verwendet, um Meta-Informationen über die angeforderten Daten im internen IT-System abzurufen. Das Datensubjekt wird dann automatisch in der mobilen App mit einer Nachricht über das DIDComm-Netzwerk informiert, dass neue Gesundheitsdaten mit den entsprechenden Meta-Informationen und der Ressourcen-ID (z.B. ID_1) zum Abruf bereitstehen. Das Datensubjekt kann dann auf die Gesundheitsdaten über die mobile App zugreifen, indem er die Daten von der Klinik ebenfalls über das DIDComm-Netzwerk und den Connector abruft und eine lokale Kopie in seiner mobilen App speichert.

3.) Erteilung des Datenzugriffs für Datenkonsumenten durch das Datensubjekt: Um Daten freizugeben, muss bereits eine Registrierung zwischen dem Datensubjekt und dem Datenkonsumenten sowie zwischen dem Datensubjekt und dem Leistungserbringenden in der Klinik erfolgt sein. Über die mobile App wählt

das Datensubjekt die ihm verfügbaren Daten aus bereits verlinkten medizinischen Einrichtungen zur Freigabe aus. Da keine direkte Verbindung zwischen der Klinik und dem Datenkonsumenten besteht, leitet die mobile App des Datensubjekts zunächst die öffentliche DID (DID_P) des Leistungserbringenden und die Ressourcen-ID (ID_1) über den etablierten Kanal der Netzwerkinfrastruktur an den Datenverbraucher weiter. Zusätzlich generiert die mobile App des Datensubjekts ein kollisionsfreies Authentifizierungstoken, das an den Datenkonsumenten und die Klinik weitergegeben wird und den Datenkonsumenten bei Vorlage zur Abholung der Daten beim Leistungserbringenden autorisiert (*SM12*). Mithilfe der öffentlichen DID kann der Datenkonsument (DID_{PC}) einen sicheren Kommunikationskanal zur leistungserbringenden Klinik (DID_{CP}) herstellen. Durch Vorlage des Authentifizierungstokens weist der Datenkonsument nach, dass er berechtigt ist, die Daten von der Klinik abzurufen. Nach erfolgreicher Überprüfung der Zugriffsrechte in der Metadaten-Datenbank überträgt die Klinik die Gesundheitsdaten über dies zuvor hergestellte sichere Verbindung an den Datenkonsumenten.

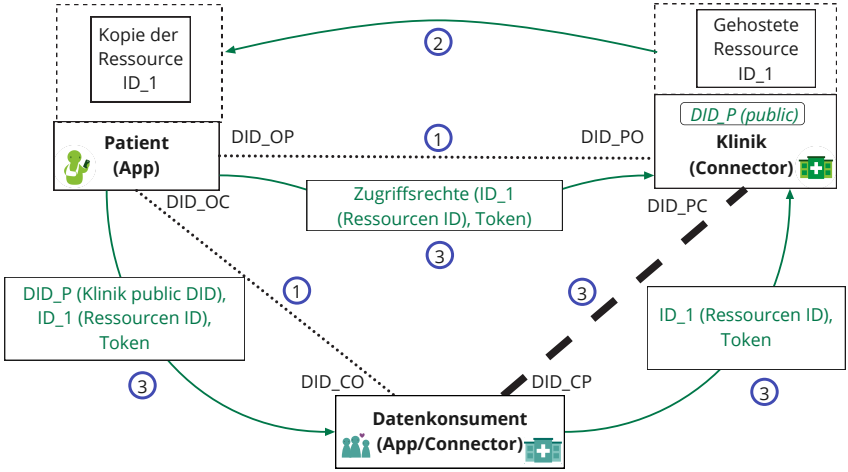

Abbildung 5.5: Kommunikations- und Datenverarbeitungsprozess innerhalb der Systemarchitektur der Gesundheitsdatenmanagementanwendung.

Insgesamt erfüllt die vorgestellte Systemarchitektur sämtliche in Abbildung 5.3 abgeleiteten Sicherheitsmechanismen. Durch den Einsatz eines dezentralen Peer-to-Peer-Netzwerks entfällt die Notwendigkeit eines zentralen Vermittlers, was potenzielle Risiken für Datenmanipulationen oder unbefugte Offenlegung minimiert. Der Datenaustausch sowie die Übermittlung von Nachrichten erfolgen vollständig verschlüsselt über das DIDComm-Protokoll, das die Vertraulichkeit und Integrität der übertragenen Informationen sicherstellt. Die initiale Registrierung im persönlichen Kontakt bildet die Grundlage für eine vertrauliche Netzwerkkommunikation. Darüber hinaus wird für jede neue Verbindung zufällige IDs (peerDIDs) erzeugt, um die Preisgabe von Informationen auf ein Minimum zu beschränken. Diese sichere Netzwerkinfrastruktur wird durch benutzendenfreundliche und sicherheitsoptimierte Anwendungen ergänzt, die eine einfache und zugleich sichere Interaktion ermöglichen.

5.1.6 Perspektive der Architektur im deutschen Gesundheitswesen

Im Rahmen des Projekts BloG^3 (vgl. Abschnitt 1.3) wurde durch das Konsortium ein Forschungsprototyp der vorgestellten Architektur entwickelt und evaluiert [16, 233]. Designs des Forschungsprototyps werden in Abbildung A.1 veranschaulicht. Die Evaluation des Prototyps umfasst unter anderem eine Untersuchung von Danelski et al. [14] zu potenziellen Geschäftsmodellen und perspektivischen Chancen sowie eine Analyse der technischen Performanz- und Sicherheitsvalidierung durch Erler et al. [16], um die praktische Anwendbarkeit und perspektivischen Chancen der entwickelten Blockchain-basierten Lösung zu bewerten. Ergänzend führte die Autorin in Erler et al. [21] eine Evaluation zur Akzeptanz und Gebrauchstauglichkeit des SSI-basierten Ansatzes durch. Die Ergebnisse jener Untersuchungen, an denen die Autorin der vorliegenden Arbeit beteiligt war, bilden die Grundlage für die nachfolgende Diskussion der Perspektive der Architektur.

Danelski et al. [14] führte von September 2022 bis Januar 2023 eine Expert*innenbefragung mittels der Delphi-Methode mit 14 Expert*innen durch. Bei der Delphi-Methode handelt es sich um ein systematisches, mehrstufiges Befragungsverfahren mit Rückkopplungsschleifen, dessen Ziel es ist, zukünftige Ereignisse, Trends oder technische Entwicklungen mithilfe von Expert*innen zu identifizieren und zu bewerten [234]. Die Teilnehmenden an den dreistufigen Expert*innenbefragungen von Danelski et al. setzen sich aus neun Themenexpert*innen, darunter Vertretende der gematik, der Techniker Krankenkasse, der kv.digital und des Betriebskrankenkassen Dachverband sowie fünf Patient*innenvertretenden aus Selbsthilfegruppen zusammen. In der ersten Runde wurden schriftlich die Perspektiven der Expert*innen abgefragt. In der zweiten Runde wurde durch Danelski et al. ein Workshop durchgeführt, bei dem die Stimmungsbilder aus der ersten Runde vorgestellt, reflektiert, weiterentwickelt und diskutiert wurden. Die abschließende dritte Runde umfasste eine finale schriftliche Einschätzung der Expert*innen basierend auf den erarbeiteten Erkenntnissen aus den vorherigen Runden. [14]

Die Autorin der vorliegenden Arbeit war insbesondere an der Aufbereitung, Vorstellung und Diskussion der technischen Architektur und der Blockchain-Technologie in den verschiedenen Delphi-Runden beteiligt. Eine zentrale Erkenntnis der Delphi-Befragung war, dass die Expert*innen die Perspektive der Architektur in verschiedenen Settings und über unterschiedliche Zeithorizonte hinweg (kurz-, mittel- und langfristig) empfehlen. Kurz- und mittelfristig sehen sie einen Mehrwert der sicherheitsorientierten Architekturlösung in einem abgegrenzten Setting, wie beispielsweise eine ländliche Modellregion mit wenigen ambulanten Einrichtung (z.B. Praxen von Hausärzt*innen und Fachärzt*innen), wo die Vernetzung und Digitalisierung aktuell noch nicht flächendeckend durchgedrungen ist und eine selbstsouveräne Lösung Mehrwert bieten kann. Der Vorteil, den die Expert*innen hierin sehen, besteht darin, dass die Lösung ihren Mehrwert entfaltet, ohne dass eine Anbindung an bestehende Digitalisierungsvorhaben, wie etwa die TI, erforderlich ist. Auch die Fokussierung auf spezifische Teilfunktionen, wie etwa die Bereitstellung von Daten für die Forschung, wurde als zielführend eingeschätzt. Die Expert*innen identifizierten in diesem Szenario insbesondere

die mobile App als potenzielle onkologische Therapiebegleitung. Dabei wurden Geschäftsmodelle im Rahmen einer Business-to-Consumer- und Business-to-Patient-Lösung diskutiert. Konkret wurde die Implementierung der App als digitale Pflegeanwendung (DiPA) oder digitale Gesundheitsanwendung (DiGA) über den Abschluss eines Selektivvertrags mit einer Krankenkasse als geeigneter direkter Weg in den ersten Gesundheitsmarkt bewertet. Mehrwerte lägen bei einer DiPA beispielsweise in erster Linie in der Verbesserung der Kommunikation mit den beteiligten Leistungserbringenden im Rahmen der ambulanten pflegerischen Versorgung durch die Möglichkeit der selbstständigen sicheren Verwaltung und Freigabe medizinischer Dokumente durch Patient*innen. Darüber hinaus könnte auch durch die Möglichkeit des Self-Monitorings mithilfe der Tagebuch-Funktion in der mobilen App eine erweiterte Therapiebegleitung im Zusammenspiel mit den Leistungserbringenden im pflegerischen Alltag ermöglichen. Ein weiterer Ansatz, der von den Expert*innen diskutiert wurde, ist der indirekte Weg über ein Geschäftsmodell als Business-to-Business-IT-Dienstleister für medizinische Einrichtungen oder Krankenkassen. Dabei könnten White-Label-Lösungen oder Infrastruktur- beziehungsweise Software-as-a-Service-Modelle implementiert werden. Der Mehrwert solcher Ansätze würde vor allem für Nutzendengruppen mit besonderen Anforderungen an Souveränität und Sicherheit gesehen, da diese spezifischen Bedürfnisse durch die entwickelte Architektur adressiert wird. Insgesamt wurde die Einbindung von Zugehörigen von den Expert*innen als eine wünschenswerte Funktion bewertet. Diese Einbindung von Zugehörigen auf Wunsch des Datensubjekts wird auch im EHDS perspektivisch vorgesehen [105]. Ferner sieht die EHDS-Verordnung allgemeine Regelungen vor, darunter die Bereitstellung eines Zugangsdienstes zu u.a. den Daten der ePA, das Recht auf Datenübertragung an weitere Institutionen sowie die Möglichkeit zur Beschränkung des Zugangs durch das Datensubjekt, welche durch die Systemarchitektur abgebildet werden. Langfristig betonten die Expert*innen die Notwendigkeit, politische und andere Digitalisierungsvorhaben in die strategische Verwertung der Architektur einzubeziehen. Dabei wurde insbesondere die mögliche Integration von Systemelementen in bestehende und geplante Anwendungen sowie Ausbaustufen der ePA und TI hervorgehoben. Ein Beispiel hierfür wäre die Nutzung

der Blockchain-basierten Netzwerkinfrastruktur zur Implementierung selbstsouveräner digitaler Identitäten und Verifiable Credentials, welche eine wertvolle Ergänzungsfunktion der ePA/TI darstellen könnten. Speziell jene Systemkomponenten können dazu beitragen, die in der TI 2.0 und im EHDS angestrebten Ziele hinsichtlich der Einführung digitaler Identitäten und elektronischer Identifizierungsmerkmale zu unterstützen [105, 56]. Über die Geschäftsmodelloptionen in den kurz- und mittelfristigen Szenarien hinaus, sahen die Expert*innen insbesondere die Option der Aufnahme der Systemelemente der Architektur als Zentraler Dienst der gematik als ein mögliches Geschäftsmodell. [14]

Die Autorin führte in Erler et al. [16] eine Sicherheits- und Performanzevaluation des Forschungsdemonstrators durch. Die Sicherheitsevaluation bestand in der Überprüfung der Umsetzung der in den Anwendungsfällen definierten Sicherheitsmechanismen. Dabei stellte sich heraus, dass die mobile App 52% und die Webanwendung der Connectoren 38,5% der aus der Angreifendenmodellierung abgeleiteten Sicherheitsmechanismen bereits implementieren. Zur Performanzevaluation wurde ein System-Kurzzeittest mit vier Connectoren auf verteilten Servern durchgeführt. Diese Connectoren wurden in verschiedenen Einrichtungen betrieben, darunter Betreibende eines Pflegeinformationssystems, Betreibende einer Datenauswerteplattform, Betreibende einer Plattform zur Erhebung von Daten aus Fitnesstrackern sowie eine Forschungseinrichtung. In diesem System-Kurzzeittest wurde auf die Connectoren die Last der Kategorie 4 aus der Spezifikation der TI der gematik angewendet, wobei die Antwortzeiten der Webanwendungen der Connectoren während der Belastung erfasst wurden [229]. Der Forschungsdemonstrator erfüllte bei den Operationen Dokument erstellen (10 KB), Dokument herunterladen (10 KB, Dokument herunterladen (100 KB) und Login vollständig die Reaktionszeiten der gematik. Während des System-Kurzzeittests zeigte sich jedoch, dass 51 % der Versuche bei der Operation „Dokument erstellen (100 KB)“ fehlschlugen. Ursache war ein fehlerhaft konfigurierter Webserver eines Connectors, der Anfragen oberhalb einer bestimmten Nutzlast ablehnte. Zusätzlich wiesen 21% der Ausführungen der Operation „Patient*in: Dateien abrufen“ Fehler auf. Diese resultierten aus einer Bad-Gateway-Fehlermeldung, die typischerweise bei einer gestörten Kommunikation zwischen Frontend-Server

und weiteren Systemkomponenten auftritt. Mögliche Ursachen hierfür waren ein unerwartetes Beenden des Backend- oder Aries-Containers oder Implementierungsfehler bei den Pfaden. Ein Aries-Container beendet sich automatisch, wenn die Verbindung zur Blockchain unterbrochen ist, während das Backend des Containers abstürzt, wenn die Kommunikation mit dem Aries-Agenten fehlschlägt. Zudem wurde festgestellt, dass die Antwortzeiten beim Abrufen von Dateien linear zur Datenmenge ansteigen. Dies führte zur Überschreitung der von der gematik vorgegebenen Zeitlimits. Das Problem resultiert daraus, dass beim Abrufen oder Auflisten gespeicherter Daten alle Dateiinhalte übertragen werden müssen. Da während des Tests kontinuierlich weitere Daten gespeichert wurden, stieg die Menge der zu übertragenden Informationen, was die Antwortzeiten weiter erhöhte. Durch die Behebung der identifizierten Fehler wird davon ausgegangen, dass eine Verbesserung der Performanz erzielt werden kann. Dennoch zeigte der Vergleich mit einem zentralen Prototyp, dass der bestehende dezentrale Forschungsdemonstrator aufgrund der erhöhten Kommunikationsaufwände innerhalb des dezentralen Systems eine geringere Performanz erzielte. Aufgrund von zeitlichen Gegebenheiten war eine erneute Testung in dem genannten Setting nicht umsetzbar. Eine ursprünglich geplante Anbindung einer Klinik konnte ebenfalls aufgrund von Firewall-Problematiken, die beim Demonstratoraufbau und -betrieb auftraten, nicht realisiert werden. Solche Herausforderungen und die Ergebnisse der Evaluation verdeutlichen, dass der Aufbau und Betrieb dezentraler Systeme mit zusätzlichen Aufwänden einhergeht. Dazu zählen insbesondere der erhöhte Aufwand für den verteilten Aufbau, den Betrieb und die Wartung der technischen Infrastruktur sowie die Erkennung und Behebung von Fehlern über die verschiedenen Systeme. Die Auslastung der Indy-Blockchain wurde ebenfalls während des Tests überprüft. Dabei zeigten sich keine Engpässe bei den Transaktionen, da die vorherrschenden Lastspitzen von 10 Transaktionen pro Sekunde durch die Indy-Blockchain bewältigt werden konnten. [16]

Darüber hinaus führte die Autorin in Erler et al. [21] eine Evaluation der Gebrauchstauglichkeit und Akzeptanz der SSI-basierten Anwendung der Systemarchitektur basierend auf Online-Interviews und dem System Usability Scale (SUS) [235] durch. An der Evaluation nahmen 13 Proband*innen im Alter von 20 bis

75 Jahren teil. Der im Rahmen der Evaluation ermittelte durchschnittliche SUS-Wert für die Anwendung beträgt 78,27, was gemäß den Interpretationsrichtlinien des SUS einer "guten" Gebrauchstauglichkeit entspricht. Die Evaluation zeigte zudem die Bedeutung des Vertrauens in die Betreibenden von Identitätsdiensten und Wallet-Lösungen auf, wobei staatliche oder zertifizierte Lösungen bevorzugt werden, um Bedenken gegenüber privaten Anbietern zu minimieren. Um die Akzeptanz insbesondere bei älteren und weniger technikaffinen Personengruppen zu fördern, ist eine umfassende Aufklärungsarbeit zu den technischen Lösungen und deren Mehrwert erforderlich. Alle 13 Proband*innen der Befragung gaben an, dass sie ein erhöhtes Maß an Kontrolle über ihre Daten mit den Lösungen besitzen. Zudem stimmten sie der Aussage zu, dass der zusätzliche Aufwand für die gesteigerte Kontrolle gerechtfertigt sei, jedoch betonten sie, dass dieser Aufwand im angemessenen Verhältnis zum daraus resultierenden Nutzen stehen müsse. [21]

Zusammenfassend zeigen die vorgestellten Ergebnisse der Untersuchungen, dass dezentrale Ansätze mit höheren Aufbau-, Koordinations- und Fehlerbehebungsaufwänden sowie einer geringeren Performanz im Vergleich zu zentralen Lösungen verbunden sind. Dies verdeutlicht, dass die Blockchain-Technologie nur in spezifischen, klar definierten Anwendungsfällen als ergänzendes Werkzeug einen Mehrwert bieten kann. Eine Betrachtung der Blockchain als universelle Gesamtlösung ist weder praktikabel noch zielführend. Insbesondere muss bei der Planung neuer Anwendungen stets sorgfältig abgewogen werden, ob nicht performantere Alternativen zur Verfügung stehen. Zukünftig sollte der Fokus daher auf klar abgegrenzte Anwendungsfälle gelegt werden, die den spezifischen Mehrwert dieser Technologie transparent darstellen. Parallel dazu könnte die weitere Forschung im Bereich Performanzoptimierung dazu beitragen, bestehende Nachteile der Blockchain zu reduzieren und somit den Anwendungsbereich sinnvoll zu erweitern. Im Kontext der durchgeführten Delphi-Befragung wurde deutlich, dass die Architektur aus dem ersten Anwendungsfall nicht als Konkurrenz zur TI oder zur bestehenden Implementierung der ePA gemäß EHDS verstanden werden sollte. Vielmehr bieten Teilelemente als ergänzender Dienst eine wertvolle Option für Zielgruppen mit höheren Ansprüchen an Sicherheits- und Souveränitätsbedürfnis. Darüber

hinaus zeigt sich insbesondere in weniger stark digitalisierten Regionen oder in Settings, in denen der direkte Austausch durch Fachkräftemangel eingeschränkt ist, wie beispielsweise in Hausärzt*innenetzwerken in ländlichen Gebieten, ein potenzieller Mehrwert. Nichtsdestotrotz stellt ein benutzendenfreundliches Design der Anwendungen in Kombination mit umfassender Aufklärungsarbeit ein zentrales Element dar, um die Akzeptanz digitaler Lösungen im Gesundheitsdatenmanagement zu stärken und die Nutzenden in die Lage zu versetzen, diese digitalen Werkzeuge effektiv anzuwenden.

5.2 Anwendungsfall 2: Sekundärdatennutzung für die medizinische Forschung und Entwicklung

Gesundheitsdatenmanagementanwendungen, wie im ersten Anwendungsfall beschrieben, schaffen die Grundlage für eine umfassende Nutzung digitaler Gesundheitsdaten. Dementsprechend ist die Bereitstellung dieser Daten nicht nur für die medizinische Versorgung, sondern auch für die Forschung ein logischer nächster Schritt. Die freiwillige Bereitstellung von in solchen Systemen gesammelten Daten könnte die medizinische Forschung durch personalisierte Medizin erheblich beschleunigen, die Entwicklung neuer therapeutischer Ansätze fördern und die Nutzung für digitale Gesundheitsanwendungen erweitern [236]. Neben dem politisch motivierten Ziel der digitalen Souveränität für Einzelpersonen bestehen die Interessen von Organisationen, datenintensive Algorithmen in Produkte und Dienstleistungen zu integrieren, die nachweislich Vorteile für Patient*innen und medizinische Fachkräfte bieten. Als potenzielle Lösung zur Überwindung isolierter Gesundheitsdatenmanagementanwendungen sowie Vermeidung von privatwirtschaftlichen Datenmonopolen werden aktuell Datentreuhänder diskutiert. Diese können als rechtlich konforme und neutrale Instanzen zwischen Datengebenden und -nutzenden fungieren [237]. Unabhängig von staatlich geförderten

Initiativen (siehe Abschnitt 2.1.4 und 2.1.5) verfolgen derzeit auch privatwirtschaftliche Unternehmen, motiviert durch die oben genannten Interessen, den Zugang zu repräsentativen Gesundheitsdaten aus der Versorgung. Die Bereitschaft der Bevölkerung, Gesundheitsdaten mit privatwirtschaftlichen Unternehmen zu teilen, ist jedoch im Vergleich zur Datenweitergabe an öffentliche Forschungseinrichtungen deutlich geringer ausgeprägt [236], was in der Konzeption zukünftiger Systemarchitekturen berücksichtigt werden muss. Das Ziel dieses Anwendungsfalls ist es daher, eine Systemarchitektur für ein zukunftsfähiges, transparentes und vertrauenswürdiges Datentreuhandsystem zu entwickeln, welches das freiwillige Teilen von Gesundheitsdaten durch Einzelpersonen sicher gestaltet sowie eine ausgewogene und sichere Sekundärnutzung der Daten sowohl im öffentlichen als auch im privaten Sektor ermöglicht und dabei die Akzeptanz sowie das Vertrauen der Bevölkerung in datenbasierte Anwendungen für Forschung und Versorgung langfristig stärken.

5.2.1 Systemkontext

Im folgenden Unterabschnitt wird der Systemkontext des geplanten Datentreuhandsystems erläutert. Das Datentreuhandsystem umfasst fünf wesentliche Akteure, die jeweils spezifische Rollen übernehmen und mit dem System interagieren: Datengebende, Datenerzeugende, Datennutzende, Reviewende und die Rechtsvetretende Person der Datentreuhand.

Beteiligte Akteure und Rollen:
Datengebende repräsentieren die in der Routineversorgung behandelten Personen oder Proband*innen von Forschungsstudien, welche ihre Gesundheitsdaten durch eine informierte Einwilligung dem Datentreuhandsystem zur Verfügung stellen. Sie haben Einsicht in eine Zusammenfassung ihrer verfügbaren Daten in den medizinischen Einrichtungen durch eine technische Verknüpfung und können ihre Einwilligungen zur Sekundärnutzung gemäß festgelegten Bedingungen erteilen, bearbeiten oder entziehen. Zudem erhalten sie Transparenz über die

Nutzung ihrer Daten, indem sie z.B. über Forschungsprojekte des Datentreuhandsystems informiert werden. Zudem können sie durch Datenerzeugende über das System benachrichtigt werden, wenn Zufallsbefunde durch Datennutzende festgestellt worden sind.

Datenerzeugende sind medizinische Einrichtungen oder deren Leistungserbringende, wie Ärzt*innen und Pflegepersonal, welche medizinische Daten über Datengebende erzeugen und diese mit der ausdrücklichen Erlaubnis der Datengebenden zur Nutzung im Datentreuhandsystem zur Verfügung stellen. Vor der Registrierung im System informieren sie die Datengebenden in einem persönlichen Gespräch über das Datentreuhandsystem, dessen Prozesse sowie den Nutzen der Datennutzung für die Forschung. Diese Information dient als Grundlage für die informierte Einwilligung der Datengebenden. Datenerzeugende fungieren als Vertrauensstelle, entfernen den direkten Personenbezug und besitzen als alleinige Instanz die direkte Personenzuordnung zur Ermöglichung der Rückkopplung von Zufallsbefunden. Hierbei nehmen Datenerzeugende Zufallsbefunde von Forschenden entgegen, validieren diese und leiten diese gegebenenfalls an die betroffenen Datengebenden weiter. Sie erhalten ebenfalls Einblick in die Einwilligungen der Datengebenden sowie die weitergegebenen Daten. Sie können zusätzlich ihre eigenen Bedürfnisse als Erzeugende bezüglich der Daten ausdrücken.

Die **Datennutzenden** sind öffentliche oder privatwirtschaftliche Forschende, welche die im Datentreuhandsystem zur Verfügung stehenden de-identifizierten Gesundheitsdaten im Rahmen von Experimenten zur Beantwortung ihrer Forschungsfragen oder für Untersuchungen und Entwicklungen innovativer Gesundheitsanwendungen verwenden möchten. Datennutzende können Forschungsprojekte anlegen, Nutzungsabsichten definieren und entsprechende Nutzungsanfragen ausführen. Basierend auf diesen Angaben können Sie einen Nutzungsvertrag mit dem Datentreuhandsystem schließen, welche diese zur Datennutzung autorisiert. Bei einem erfolgreichen Vertragsabschluss können Datennutzende im Datentreuhandsystem Experimente (z.B. mittels deskriptiver Statistik oder Maschinellen Lernens) anlegen, konfigurieren, durchführen und evaluieren. Im Rahmen der Experimente erfasste Zufallsbefunde können über das Treuhandsystem gemeldet und Forschungsergebnisse geteilt werden. Zur Sicherstellung der Reproduzierbarkeit

und Validität von Forschungsergebnissen haben Datennutzende die Möglichkeit, auf die Forschungsergebnisse anderer Datennutzender zuzugreifen und diese einzusehen.

Die **Rechtsvetretende Person der Datentreuhand** fungiert als zentrale juristische Instanz im Datentreuhandsystem und wird von den angeschlossenen Einrichtungen benannt. Ihre Hauptverantwortung liegt in der Klärung und Verwaltung der rechtlichen Rahmenbedingungen für die Datenverarbeitung und -nutzung. Dazu zählt insbesondere die Überprüfung, Genehmigung oder Ablehnung von Registrierungen potenzieller Datennutzender im System, um sicherzustellen, dass die gesetzlichen und ethischen Standards eingehalten werden. Zusätzlich übernimmt die rechtsvertretende Person die juristische Verwaltung und Abwicklung der Nutzungsverträge. Dies umfasst die Formulierung und Freigabe der Vertragsbedingungen, die Sicherstellung der Einhaltung dieser Vereinbarungen und die rechtliche Dokumentation der Datennutzungsrechte und -pflichten aller Beteiligten. Ein weiterer wichtiger Aufgabenbereich ist die Durchführung von Audits. Diese regelmäßigen Überprüfungen dienen dazu, die Konformität der Datenverarbeitungsprozesse mit den geltenden rechtlichen Anforderungen sicherzustellen und potenzielle Missbrauchs- oder Compliance-Risiken frühzeitig zu identifizieren und zu beheben. Die rechtsvertretende Instanz prüft dabei insbesondere die Rechtmäßigkeit der Datenzugriffe und -nutzungen sowie die Einhaltung der datenschutzrechtlichen Vorgaben, etwa der DSGVO. Zu diesem Zweck überwacht sie auch die Protokollierung von Zugriffen und die Nutzung von Daten durch Datennutzende, um die Nachvollziehbarkeit und Integrität aller Systemaktivitäten zu gewährleisten.

Reviewende sind unabhängige externen Personen oder Organisationen, welche die rechtmäßige und regelkonforme Nutzung der Daten im Datentreuhandsystem überprüfen. Ein wesentlicher Bestandteil ihrer Aufgaben ist die Validierung der Reproduzierbarkeit von Forschungsexperimenten, die mit den Daten aus dem Treuhandsystem durchgeführt wurden. Dabei überprüfen die Reviewenden, ob die in den Nutzungsverträgen der Datennutzenden angegebenen Analysen oder Experimente anhand der bereitgestellten Algorithmen und Daten reproduziert werden

können und ob die eingesetzten Methoden mit den Angaben der Forschenden übereinstimmen. Diese Überprüfung trägt maßgeblich dazu bei, die Glaubwürdigkeit und Qualität der Forschungsergebnisse sicherzustellen.

Bestehende technische Komponenten:
Die **Informationssysteme der medizinischen Einrichtungen** umfassen die in medizinischen Einrichtungen eingesetzten technischen Systeme zur Speicherung, Verwaltung und Verarbeitung von Gesundheitsdaten der Patient*innen, wie etwa Krankenhausinformationssysteme (KIS). Diese Systeme sind für das Datenmanagement innerhalb der jeweiligen Einrichtung verantwortlich und ermöglichen den sicheren und strukturierten Zugriff auf Patient*innendaten für diagnostische, therapeutische und administrative Zwecke. Im Kontext des Datentreuhandsystems fungieren die Informationssysteme als zentrale Quelle für die Bereitstellung von Gesundheitsdaten, sofern dies im Einklang mit den Einwilligungen und Wünschen der Datengebenden geschieht.

Systemkontext und -grenzen:
Das Systemkontextdiagramm in Abbildung 5.6 entsprechend des C4-Modells (siehe Abschnitt 2.3.2) illustriert die Systemgrenzen sowie die Interaktionen mit den beteiligten Akteuren und externen Komponenten. Das zu entwickelnde Datentreuhandsystem ist in hellblau hervorgehoben und umfasst alle Funktionen zur Verwaltung, Vermittlung und Verarbeitung der treuhänderisch zu betreuenden Gesundheitsdaten. Die Akteure, darunter Datengebende, Datennutzende, Datenerzeugende, Reviewende und die Rechtsvertretende Person der Datentreuhand, sind in dunkelblau dargestellt. Die externen Systeme, insbesondere die bereits vorhandenen Informationssysteme der medizinischen Einrichtungen, die zur Datenübermittlung mit dem Treuhandsystem in Wechselwirkung treten, sind in grau markiert.

Juristische Rahmenbedingungen:
Für Datentreuhandsysteme ergeben sich die juristischen Rahmenbedingungen insbesondere aus den Vorgaben zur Einhaltung von Datenschutz und Datensicherheit gemäß DSGVO, BDSG und PDSG. Zusätzlich sind die gesetzlichen Grundlagen zu berücksichtigen, in die Infrastrukturen wie das FDZ-Gesundheit eingebettet

sind, einschließlich der DaTraV, des DigiG, des GDNG und des DGA sowie der Verordnung zum EHDS. Eine detaillierte Darstellung der juristischen Anforderungen, die sich aus diesen Regelungen ableiten, erfolgt im Abschnitt 2.2.

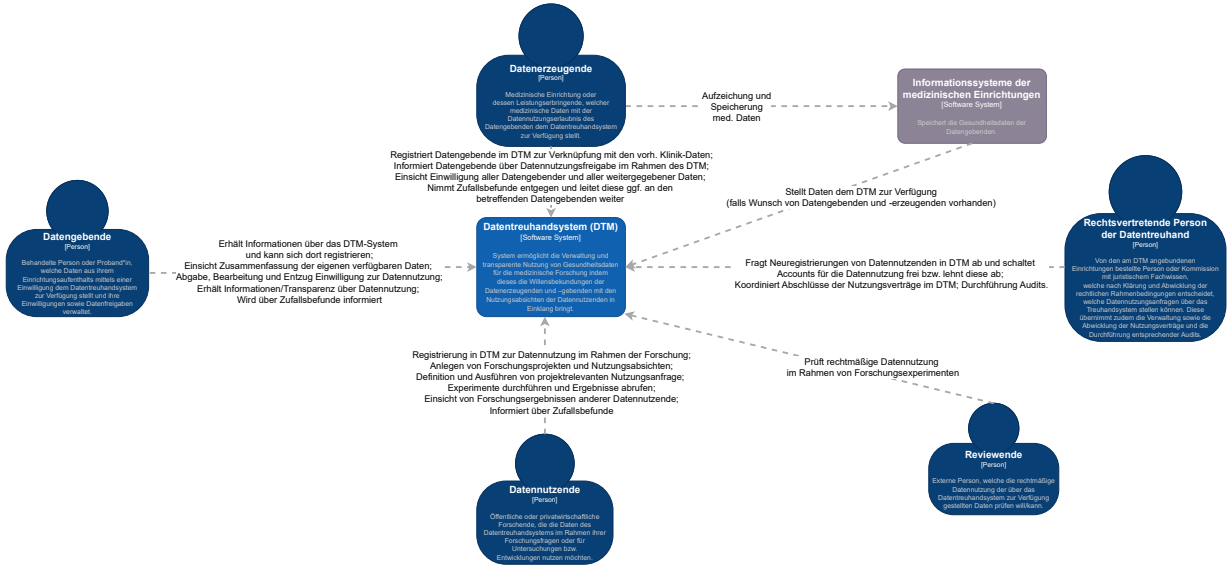

Abbildung 5.6: Systemkontextdiagramm des zu entwickelnden Datentreuhandsystems.

5.2.2 Anforderungserhebung

Im Rahmen der Entwicklung eines Datentreuhandsystems wurden sowohl funktionale als auch Qualitätsanforderungen ermittelt, um die unterschiedlichen Bedürfnisse und Erwartungen der relevanten Akteur*innen zu adressieren. Zunächst erfolgte eine umfassende Literaturrecherche, um bestehende Datentreuhand- und Datenspendesysteme sowie bereits in der wissenschaftlichen Diskussion behandelte Anforderungen zu identifizieren. Die gewonnenen Erkenntnisse stützen sich vor allem auf einer Publikation der Autorin [9], in der diese Anforderungen sowie die Literaturrecherche zusammengefasst wurden. Aufbauend auf den Ergebnissen der Literaturrecherche wurden detaillierte Anforderungen formuliert, die in Form von User Stories die funktionalen Bedürfnisse der relevanten Akteur*innen widerspiegeln. Um eine praxisorientierte und fundierte Entwicklung dieser Anforderungen zu gewährleisten, wurden digitale Expert*innenworkshops mit Vertreter*innen aus den Disziplinen Medizin, Forschung, Ethik, Recht und Informatik durchgeführt [19]. In den Workshops wurden die User Stories gemeinsam erarbeitet, konkretisiert und weiterentwickelt. Anschließend erfolgte die Priorisierung der Anforderungen mithilfe der MoSCoW-Methode, wie sie im ersten Anwendungsfall beschrieben ist. In den nachfolgenden Abschnitten werden die „Must-have“-Anforderungen aus dem zweiten Anwendungsfall, unterteilt in funktionale und Qualitätsanforderungen, dargelegt.

5.2.2.1 Funktionale Anforderungen

In diesem Abschnitt werden die funktionalen Anforderungen an das zu entwickelnde Datentreuhandsystem systematisch dargelegt, die auf Basis des festgelegten Systemkontexts und der beteiligten Systemakteure abgeleitet wurden.

Datengebende:

- DGFA-1: Als Datengebende möchte ich in einer für mich verständlichen Art und Weise eine Übersicht über meine erhobenen Gesundheitsdaten erhalten, um evaluieren zu können, welche Daten in den verschiedenen Einrichtungen über

mich gespeichert sind und welche Daten für eine Datenfreigabe zur Verfügung stehen.

- DGFA-2: Als Datengebende möchte ich eine zentrale Verwaltungsanwendung des Datentreuhandsystems, um meine Daten einfach verwalten zu können und mich nicht bei mehreren Systemen anmelden zu müssen.
- DGFA-3: Als Datengebende möchte ich mich über Forschungsprojekte, Forschungsergebnisse und Datennutzungen informieren können, um die potenziellen Vorteile einer Datenfreigabe besser zu verstehen, eine fundierte Entscheidung treffen zu können oder allgemein Transparenz zu erhalten.
- DGFA-4: Als Datengebende möchte ich mich zuvor über Chancen, Risiken und Konsequenzen einer Datenfreigabe informieren, um so eine informierte Einwilligung geben zu können und den Nutzen und die Auswirkungen einer Freigabe zu verstehen.
- DGFA-5: Als Datengebende möchte ich spezifische Einwilligungen zur Datennutzung erteilen, entziehen oder ändern können, damit ich die Kontrolle über meine Daten behalte und sicherstellen kann, dass meine Daten nur im Einklang mit meinen aktuellen Wünschen und Bedingungen genutzt werden.
- DGFA-6: Als Datengebende möchte ich die Bedingungen einer Datennutzung festlegen können (z.B. ob öffentliche oder private Forschung unterstützt werden soll), um diesen gezielt Zugriff auf meine Daten zu gewähren.
- DGFA-7: Als Datengebende möchte ich benachrichtigt werden, wenn relevante medizinische Zufallsbefunde aus Forschungsprojekten vorliegen, um gegebenenfalls eine notwendige medizinische Behandlung einleiten oder geeignete Maßnahmen ergreifen zu können.
- DGFA-8: Als Datengebende möchte ich, dass meine freigegebenen Daten nicht durch Datennutzende auf mich zurückgeführt werden können, damit mir als Person auch bei einem ungewollten Datenleck keine persönlichen Nachteile entstehen.

- DGFA-9: Als Datengebende möchte ich die Möglichkeit haben, meine Daten auch über meinen Tod hinaus zur Verfügung zu stellen, um die fortlaufende Unterstützung der Forschung sicherzustellen.
- DGFA-10: Als Datengebende möchte ich einen wahrnehmbaren Nutzen aus der Bereitstellung meiner Daten erhalten, beispielsweise in Form von personalisierten Behandlungserkenntnissen basierend auf kuratierten Forschungsalgorithmen oder durch den Zugang zu Forschungsergebnissen, sodass ich das Gefühl habe, die Forschung aktiv unterstützt zu haben.

Datenerzeugende:

- DEFA-1: Als Datenerzeugende möchte ich sicherstellen, dass durch mich erzeugte Daten gemäß rechtlicher und wirtschaftlicher Anforderungen (z. B. Urheberrecht, Verwertungsrechte, Fürsorgepflicht, Datenschutz) an ein Datentreuhandsystem weitergegeben und genutzt werden, um meine Verpflichtungen und wirtschaftlichen Interessen zu wahren.
- DEFA-2: Als Datenerzeugende möchte ich Datengebende in einem persönlichen Gespräch über das Datentreuhandsystem informieren, um eine informierte Einwilligung sicherzustellen.
- DEFA-3: Als Datenerzeugende möchte ich eine Vereinbarung zur Datenverarbeitung, Datenweitergabe und Datennutzung mit dem Datentreuhänder vereinbaren können, um eine rechtliche Absicherung erhalten zu können.
- DEFA-4: Als Datenerzeugende möchte ich festlegen können, welche Daten und Nutzungsregeln (z.B. öffentliche vs. private Forschung) für die von mir generierten Daten gelten, um die Kontrolle über die Nutzung zu behalten.
- DEFA-5: Als Datenerzeugende möchte ich sicherstellen, dass alle Daten, insbesondere schützenswerte medizinische Daten, DSGVO-konform verwendet werden, damit Patient*innen und Proband*innen kein Nachteil entsteht.
- DEFA-6: Als Datenerzeugende möchte ich Daten nur bei Bedarf und nicht auf Vorrat weitergeben, um Datenschutz und Datenminimierung sicherzustellen.

- DEFA-7: Als Datenerzeugende möchte ich, dass Daten nach Ablauf einer bestimmten Frist vom Datentreuhandsystem gelöscht werden, um Datenschutzanforderungen zu erfüllen.
- DEFA-8: Als Datenerzeugende möchte ich einsehen können, welche Datennutzenden auf meine Daten zugreifen, um die Rechtmäßigkeit und Zweckmäßigkeit der Freigabe nachvollziehen zu können.
- DEFA-9: Als Datenerzeugende möchte ich die Einwilligungen von Datengebenden verwalten und deren Wünsche bei der Datennutzung berücksichtigen, um eine den Vorgaben entsprechende Datennutzung sicherzustellen.
- DEFA-10: Als Datenerzeugende möchte ich, dass das Datentreuhandsystem direkt auf bestehende Informationssysteme zugreift, wobei die Klinik hinsichtlich dabei notwendiger Schnittstellen die Kontrolle behält und die Datenprozesse sowie Speicherorte für die Datenaufbereitung selbst bestimmen kann (z.B. durch On-Premise-Lösungen). Dadurch möchte ich sicherstellen, dass der manuelle Aufwand reduziert wird, die Standardprozesse meiner Klinik effizient bleiben und um zusätzlichen Aufwand zu minimieren.
- DEFA-11: Als Datenerzeugende möchte ich sicherstellen, dass die Identität der Patient*innen nur in de-identifizierter Form an das Treuhandsystem übermittelt wird, um deren Privatsphäre zu schützen.
- DEFA-12: Als Datenerzeugende möchte ich Zufallsbefunde von Datennutzenden erhalten und deren Relevanz bewerten können, um eine ethische und forschungsgerechte Zusammenarbeit sicherzustellen und den Bedürfnissen der Patient*innen gerecht zu werden.
- DEFA-13: Als Datenerzeugende möchte ich den Datengebenden (falls von diesem gewünscht) über von mir geprüfte Zufallsbefunde informieren können, um den ethischen Anforderungen gerecht zu werden.
- DEFA-14: Als Datenerzeugende möchte ich Daten gemäß den Einwilligungen und Wünschen der Datengebenden zur Verfügung stellen, um wissenschaftliche

Forschung zu ermöglichen, die Reputation meiner Institution zu steigern und gleichzeitig Forschung im eigenen Haus mit Daten Dritter zu fördern.

Datennutzende:

- DNFA-1: Als Datennutzende möchte ich die Möglichkeit haben, eine Anfrage zur Nutzung von Gesundheitsdaten für Forschungsprojekte zu stellen (z.B. anhand von mir festgelegter Filterkriterien), um auf für mich relevante Daten zugreifen zu können.
- DNFA-2: Als Datennutzende möchte ich Forschungsprojekte und -experimente im System anlegen, die für meine Forschung notwendig sind, und diese nach Freigabe ausführen zu können und meine Algorithmen zu optimieren.
- DNFA-3: Als Datennutzende möchte ich am Datentreuhandsystem ohne zusätzlichen technischen Aufwand teilnehmen können, um keine teure Infrastruktur bei mir selbst betreiben zu müssen bzw. diese aufbauen zu müssen.
- DNFA-4: Als Datennutzende möchte ich eine Vereinbarung zur Datennutzung mit dem Datentreuhandsystem schließen können, um Zugriff auf die Daten zu erhalten und eine rechtliche Absicherung erhalten zu können.
- DNFA-5: Als Datennutzende möchte ich die Metadaten eines Datensatzes vor einem Vertragsabschluss einsehen können, um deren Relevanz für mein Forschungsvorhaben einschätzen zu können.
- DNFA-6: Als Datennutzende möchte ich standardisierte Daten (z.B. in Form von FHIR-Ressourcen) erhalten, um diese einfach in meiner Forschung verwenden zu können und eine effiziente Datennutzung zu gewährleisten.
- DNFA-7: Als Datennutzende möchte ich Zugang zu longitudinalen Daten erhalten, die kontinuierlich über einen längeren Zeitraum hinweg gesammelt wurden, um eine fundierte Analyse und Auswertung der Entwicklungen des Gesundheitszustands von Patient*innen durchführen zu können.

- DNFA-8: Als Datennutzende möchte ich die Sicherheit haben, dass eine Vereinbarung zur Datennutzung nicht während der Dauer meiner Forschung widerrufen wird, um diese abschließen zu können.
- DNFA-9: Als Datennutzende möchte ich die Möglichkeit haben Zufallsbefunde an den Datenerzeugenden gemäß den Wünschen des Datengebenden zurückzuspiegeln, um den damit einhergehenden ethischen Anforderungen gerecht zu werden und einer weiteren Prüfung durch med. Personal zu ermöglichen.
- DNFA-10: Als Datennutzende möchte ich einen Datensatz, welcher anhand einer meiner Datenanfragen erstellt wurde und den ich in meinem Experiment verwendet habe, eindeutig identifizieren können, um diesen auch in wissenschaftlichen Publikation referenzieren zu können.
- DNFA-11: Als Datennutzende möchte ich Forschungs- und Studienergebnisse (z.B. wiss. Publikationen über DOI) zu vorhandenen Projekten/Studien hinzufügen können, um Datengebende über neue Ergebnisse zu informieren.
- DNFA-12: Als Datennutzende möchte ich die Möglichkeit haben, Ergebnisse anderer Forschenden einzusehen, um die Reproduzierbarkeit und Vergleichbarkeit von publizierten Forschungsergebnissen zu unterstützen.
- DNFA-13: Als Datennutzende möchte ich Methodik, Daten, Algorithmen teilen können, um eine Reproduzierbarkeit meiner Ergebnisse für andere Forschende ermöglichen zu können.

Reviewende:

- RFA-1: Als Reviewende möchte ich verifizieren können, dass einsehbare Experimente von Organisationen im Rahmen der DT erfolgt sind, um im Zweifel nachweisen zu können, dass die Daten (nicht) zweckmäßig verwendet wurden.
- RFA-2: Als Reviewende möchte ich die Experimente von Datennutzenden einsehen können, um mich über deren Forschungsarbeiten und -algorithmen informieren zu können.

- RFA-3: Als Reviewende möchte ich über ein Webinterface verifizieren können, dass einsehbare Experimente von Organisationen im Rahmen der DT erfolgt sind, um im Zweifel nachweisen zu können, dass die Daten (nicht) zweckmäßig verwendet wurden.
- RFA-4: Als Reviewende*r möchte ich sicherstellen, dass die Forschungsexperimente durch die freigegebenen Algorithmen und Daten reproduzierbar sind, um die Integrität der Forschungsergebnisse zu gewährleisten.
- RFA-5: Als Reviewende*r möchte ich über das Datentreuhandsystem widersprüchliche Angaben zwischen den erklärten Verwendungszwecken der Datennutzenden und deren tatsächlicher Datennutzung melden können, um sicherzustellen, dass alle Datennutzungen den vorgegebenen Richtlinien entsprechen.

Rechtsvertretende Person:

- RVFA-1: Als Rechtsvertretende*r der Datentreuhand möchte ich Registrierungsanfragen von Datennutzenden prüfen und freigeben oder ablehnen, um sicherzustellen, dass nur berechtigte Forschende Zugang zum Datentreuhandsystem erhalten.
- RVFA-2: Als Rechtsvertretende*r der Datentreuhand möchte ich die rechtlichen Verträge und Nutzungsvereinbarungen verwalten, um die Nutzung der Daten innerhalb der vereinbarten Grenzen zu gewährleisten.
- RVFA-3: Als Rechtsvertretende*r der Datentreuhand möchte ich regelmäßige Audits durchführen, um sicherzustellen, dass alle Prozesse im Datentreuhandsystem rechtlich und ethisch einwandfrei sind.
- RVFA-4: Als Rechtsvertretende*r der Datentreuhand möchte ich Meldungen rund um die Datennutzung bearbeiten und bei Bedarf rechtliche Schritte einleiten, um die Interessen der Datengebenden und Datenerzeugenden zu schützen.

5.2.2.2 Qualitätsanforderungen

In diesem Abschnitt werden die nicht-funktionalen Anforderungen beschrieben, die auf den Qualitätsmerkmalen des ISO-Standards 25010 basierend abgeleitet wurden.

- NFA-1 - Bedienbarkeit: Das Datentreuhandsystem muss eine hohe Gebrauchstauglichkeit aufweisen, indem es benutzendenfreundliche, patient*innenzentrierte Tools zur Visualisierung und Verwaltung von Gesundheitsdaten sowie den zugehörigen Einwilligungen bereitstellt. Die Benutzendenoberfläche sollte intuitiv gestaltet sein, um die Einstiegshürde für nicht-technische Anwendende zu minimieren und die Akzeptanz der Nutzung zu fördern.
- NFA-2 - Vertraulichkeit: Die Vertraulichkeit eines Datentreuhandsystems muss durch Privacy-by-Design-Architekturen gewährleistet werden, um die Privatsphäre der Datengebenden gemäß nationalen Vorschriften zu schützen. Dies umfasst die Möglichkeit der Datenlöschung durch den Datengebenden sowie die De-Identifikation von medizinischen Daten durch Anonymisierung oder Pseudonymisierung. Zudem sollte das System datenschutzfreundliche Analysealgorithmen unterstützen, die Daten lokal verarbeiten, ohne sie zu übertragen. Die Zugriffs- und Nutzungsverwaltung muss fein abgestufte Kontrolle über den Datenzugriff ermöglichen und vom Datengebenden sowie Datenerzeugenden autorisiert werden. Schließlich muss eine Benutzendenauthentifizierung sichergestellt sein, um nur autorisierten Akteuren Zugang zu gewähren.
- NFA-3 - Sicherheit: Das Datentreuhandsystem muss die Sicherheit der verarbeiteten Daten durch die konsequente Anwendung des Security-by-Design-Ansatzes in allen Phasen seines Lebenszyklus sicherstellen. Hierbei sind insbesondere Maßnahmen zur Verschleierung sensibler Daten durch Verschlüsselung erforderlich, um die Vertraulichkeit der Daten zu gewährleisten. Zudem muss die Genauigkeit der bereitgestellten Daten garantiert sein, um deren Richtigkeit zu gewährleisten. Es ist sicherzustellen, dass die Daten unveränderlich sind und durch geeignete Mechanismen vor unbefugten Änderungen geschützt werden,

um das Vertrauen in ihre Integrität zu stärken. Darüber hinaus muss die Verfügbarkeit der Daten für alle autorisierten Akteure zu jeder Zeit gewährleistet sein. Schließlich muss das System eine nachvollziehbare und überprüfbare Verknüpfung der Daten mit der Zustimmung zur Nutzung bieten, um die rechtmäßige Verwendung der Daten sicherzustellen.

- NFA-4 - Datenintegrität: Das Datentreuhandsystem muss sicherstellen, dass die bereitgestellten Daten während ihres gesamten Lebenszyklus unverändert und korrekt bleiben. Dies umfasst Mechanismen zur kontinuierlichen Überprüfung und Validierung der Datenintegrität, um sicherzustellen, dass keine unbefugten Änderungen oder Datenmanipulationen vorgenommen werden. Die Integrität der Daten muss durch kryptographische Verfahren, wie z. B. digitale Signaturen oder Hash-Werte, gewährleistet werden, um die Authentizität und Richtigkeit der Daten zu überprüfen. Außerdem muss das System in der Lage sein, jede Änderung an den Daten lückenlos nachzuvollziehen und eine vollständige Historie aller Datenoperationen bereitzustellen, um das Vertrauen in die Daten zu stärken und die Verlässlichkeit ihrer Herkunft zu bestätigen.
- NFA-5 - Vertraulichkeit: Die Vertraulichkeit stellt per Definition eine zentrale Anforderung in einem Datentreuhandsystem dar. Datengebende, die ihre Daten über ein Datentreuhandsystem bereitstellen, müssen sowohl der technologischen Sicherheit der Plattform zur Datenfreigabe und -speicherung als auch der Integrität und Zuverlässigkeit der anderen Akteure vertrauen. Es muss sichergestellt sein, dass die Daten ausschließlich in Übereinstimmung mit den Vorgaben und Einwilligungen der Datengebenden verwendet werden. Gleichzeitig müssen Datennutzende darauf vertrauen können, dass die bereitgestellten Daten korrekt, vollständig und vertrauenswürdig sind. Dieses Vertrauen muss nicht nur gewährleistet, sondern auch für alle Beteiligten erkennbar und nachvollziehbar sein.
- NFA-6 - Interoperabilität: Das Datentreuhandsystem muss eine umfassende Interoperabilität gewährleisten, um die Zusammenarbeit zwischen verschiedenen Organisationen, Systemen und Datenquellen zu ermöglichen. Dies umfasst:

  - **Organisatorische Interoperabilität**: Unterstützung organisationsübergreifender Prozesse, Identitäten und Rechte. Dazu zählt beispielsweise die Integration von Mechanismen, die es ermöglichen, dass Datennutzende und -gebende nicht für jede einzelne datenerzeugende Einrichtung oder verschiedene Treuhandsysteme ein separates Konto anlegen müssen.
  - **Semantische Interoperabilität**: Einsatz international anerkannter Standards und Terminologien (z. B. HL7 FHIR, DICOM, SNOMED CT) für konsistente Datenbeschreibungen.
  - **Strukturelle Interoperabilität**: Ermöglichung von Datenflüssen zwischen Systemen und heterogenen Datensilos durch standardisierte Schnittstellen und Datenformate.

- NFA-7 - Nachweisbarkeit: Die Nachweisbarkeit muss gewährleistet sein, sodass die Datennutzung durch den Datengebenden nachvollzogen werden kann. Dazu muss es für Datengebende möglich sein, Forschungsprojekte einzusehen und deren jeweilige Zwecke. Darüber hinaus schließt dies die Rückmeldung der Forschungsergebnisse an die Datengebenden sowie an betroffene Personen mit ähnlichen Krankheiten und Symptomen ein. Alle Aktionen, die im Zusammenhang mit den bereitgestellten Daten stehen, sind transparent und vollständig zu protokollieren. Dies gewährleistet eine lückenlose Rückverfolgbarkeit des gesamten Lebenszyklus der Daten und trägt dazu bei, Missbrauch oder unbefugte Verwendung zu verhindern.
- NFA-8 Nachhaltigkeit: Die Nachhaltigkeit eines Datentreuhandsystems erfordert Finanzierungsmechanismen, die den Aufwand für Aufbereitung, Pflege und Austausch von Forschungsdaten abdecken und die langfristige Verfügbarkeit sowie Integrität der Daten sichern. Das System sollte technologieneutrale, offene Formate verwenden, um eine effiziente Migration zwischen Technologie-Stacks zu ermöglichen. Um Vendor-Lock-ins zu vermeiden soll auf offene Standards und Open-Source-Software gesetzt werden.
- NFA-9 Funktionale Eignung: Die funktionale Eignung des Datentreuhandsystems bezieht sich auf die Fähigkeit, die Anforderungen an die Datenqualität

zu erfüllen. Dies umfasst die Erfassung, Messung und Dokumentation der Datenqualität während des gesamten Lebenszyklus der Daten, einschließlich ihrer Aggregation und Qualitätsverbesserung.

- NFA-10 Leistungseffizienz: Das Datentreuhandsystem muss so ausgelegt sein, dass es große Mengen komplexer Daten effizient verarbeitet. Dabei soll es eine optimale Nutzung der verfügbaren Ressourcen gewährleisten und gleichzeitig hohe Reaktions- und Verarbeitungszeiten auch bei steigender Datenmenge und Nutzendenzahl aufrechterhalten. Die Systemarchitektur muss in der Lage sein, die Leistung dynamisch anzupassen, um eine konstante Effizienz bei der Datenverarbeitung und -bereitstellung sicherzustellen, ohne die Systemressourcen unnötig zu belasten.
- NFA-11 Modifizierbarkeit: Die IT-Architektur des Datentreuhandsystems muss so konzipiert sein, dass Änderungen, Erweiterungen oder Anpassungen mit minimalem Aufwand und ohne signifikante Auswirkungen auf die bestehenden Funktionen durchgeführt werden können. Das System sollte eine klare Trennung von Modulen und Komponenten aufweisen, um eine einfache Wartung und Anpassung an sich verändernde Anforderungen zu ermöglichen. Darüber hinaus sollte es ein transparentes und dokumentiertes System zur Fehlerbehebung und Aktualisierung bieten, um eine langfristige Flexibilität und die schnelle Implementierung neuer Anforderungen zu gewährleisten.

5.2.3 Systemabstraktion und Sicherheitsbetrachtung

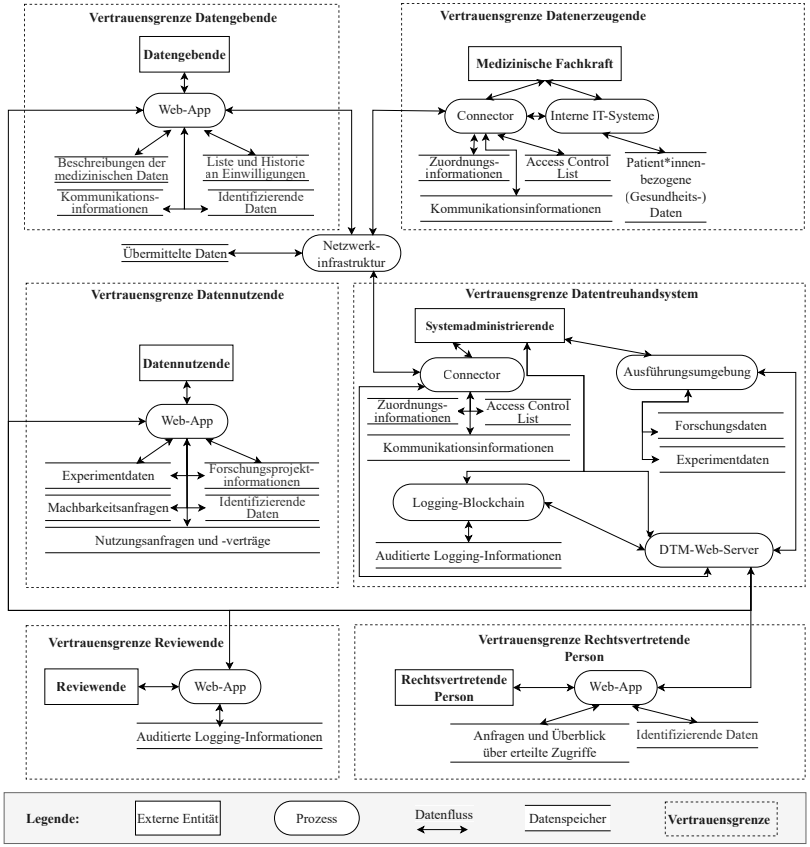

Abbildung 5.7: Datenflussdiagramm des zu entwickelnden Datentreuhandsystems.

Die zentralen technischen Bestandteile sowie die Datenflüsse des geplanten Datentreuhandsystems werden durch das Datenflussdiagramm in Abbildung 5.7 visualisiert. Das Datentreuhandsystem besteht aus sechs Hauptkomponenten: den Webanwendungen mit dem dazugehörigen Web-Server, der Netzwerkinfrastruktur, den verschiedenen Connectoren, der Ausführungsumgebung für Experimente sowie der Logging-Blockchain. Für die Entwicklung der Systemarchitektur wurden diese Komponenten in den in Abschnitt 5.6 dargelegten Systemkontext eingebettet. Im folgenden Abschnitt werden die einzelnen Komponenten und deren Datenflüsse detailliert erläutert. Zudem werden die erforderlichen Sicherheitsaspekte untersucht, einschließlich der zugrundeliegenden Vertrauensannahmen sowie der Identifikation der zu schützenden Güter.

Web-App inkl. Web-Server: Die Web-Apps fungieren als zentrale Benutzendenschnittstelle für die verschiedenen Akteure des Datentreuhandsystems. Datengebende haben hier die Möglichkeit, ihre verfügbaren Gesundheitsdaten einzusehen, ihre Einwilligungen zur Sekundärnutzung dieser Daten zu erteilen, zu verwalten oder zu widerrufen. Zudem wird Transparenz hinsichtlich der Datennutzung geschaffen, und die Nutzenden können sich über laufende Forschungsprojekte informieren. Datennutzende können hingegen ihre Forschungsprojekte registrieren, Machbarkeitsanfragen und Nutzungsanfragen stellen sowie Nutzungsverträge abschließen. Darüber hinaus steht ihnen die Experimentierplattform über deren dedizierte Web-Anwendung zur Verfügung, mit der sie Algorithmen entwickeln, anhand der Daten des Datentreuhandsystems optimieren und validieren können. Zusätzlich können sie ihre Forschungsergebnisse teilen und Zufallsbefunde melden. Reviewende haben die Möglichkeit, die auf der Blockchain protokollierten Informationen zur Datennutzung benutzendenfreundlich zu validieren. Die rechtsvertretende Person ist befugt, Zugriffsberechtigungen für Datennutzende zu erteilen sowie die Nutzungsverträge zu prüfen, zu verwalten und deren Abwicklung und Umsetzung zu überwachen.

Connector: Ein Systemelement, das die Kommunikation zwischen den internen IT-Systemen der medizinischen Einrichtungen und der Netzwerkinfrastruktur ermöglicht. Der Connector sorgt für den sicheren Austausch und die Verknüpfung der Daten (z.B. Identifier) im Netzwerk und den angebundenen internen IT-Systemen.

Netzwerkinfrastruktur: Die zugrundeliegende P2P-Infrastruktur bildet die Grundlage für die Kommunikation und den Datentransfer zwischen den verschiedenen Connectoren. Sie gewährleistet die dezentrale Struktur zur Anfrage und Bereitstellung von Daten aus verschiedenen datenerzeugenden Einrichtungen an das Datentreuhandsystem ohne die Notwendigkeit einer zentralen Datenspeicherung. Zudem werden die Bedingungen zur Datennutzung der Datengebenden aus deren Einwilligungen übermittelt sowie relevante Systeminformationen zur Gewährleistung der Datentreuhandfunktionen in jenem verteilten Setting.

Interne IT-Systeme: Die Systeme innerhalb der medizinischen, datenerzeugenden Einrichtungen, wie beispielsweise Krankenhausinformationssysteme (KIS), in denen Gesundheitsdaten gespeichert und verwaltet werden, welche für die Sekundärnutzung im Rahmen des Datentreuhandsystems vorgesehen werden sollen.

Ausführungsumgebung für Experimente: Die Ausführungsumgebung für Experimente ist eine technische Systemkomponente des Datentreuhandsystems, die Datennutzenden eine vertrauensvolle Umgebung (*engl. trusted execution enviroment*) für die Durchführung derer Forschungsexperimente und Analysen bereitstellt.

Logging-Blockchain: Die Logging-Blockchain dient der lückenlosen Auditierung sämtlicher Zugriffe und Verarbeitungsschritte, die im Rahmen von Experimenten innerhalb der Ausführungsumgebung des Datentreuhandsystems durchgeführt werden. Dies legt die Grundlage für eine transparente und manipulationssichere Überprüfung der Rechtmäßigkeit und Konformität der Datennutzung mit den geltenden Vorgaben und Einwilligungen.

Vertrauensannahmen: Die zugrundeliegenden Annahmen hinsichtlich der Vertrauensebenen für die Systemakteure des Datentreuhandsystems sowie der zugehörigen Komponenten werden im Folgenden dargelegt:

- **VA1:** Datengebende haben über den Web-App-Zugang des Datentreuhandsystems ausschließlich Zugriff auf ihre eigenen Gesundheitsdaten, Metadaten sowie auf systemrelevante Informationen, die für die sichere Datenübertragung innerhalb der Netzwerkinfrastruktur erforderlich sind (z. B. Identifier, Schlüsselmaterial, Berechtigungsnachweise). Sie können nur Einwilligungen verwalten, die ihre eigenen Daten betreffen, und haben keinen Zugriff auf (Meta-)Daten von anderen Systemteilnehmenden. Um die Nutzungsfreundlichkeit zu erhöhen und das notwendige Vertrauen der Datengebenden in das Datentreuhandsystem zu fördern, werden digitale Identitätsnachweise sowie die Datenverwaltungsinstrumente nicht lokal auf den Endgeräten der Datengebenden gespeichert, sondern auf den Systemen des Datentreuhandsystems zentral verwaltet. Im Interesse der ethischen Sensibilität soll darauf verzichtet werden, eine direkte Rückkopplung zu spezifischen Forschungsprojekten vorzusehen, die personenbezogene Daten der Datengebenden verwenden. Dies wurde unter anderem entschieden, um mögliche Ängste oder Missverständnisse zu vermeiden, etwa bei Projekten wie der Demenzforschung, bei der die Datengebenden ohne Diagnose als Teil einer Kontrollgruppe einbezogen werden könnten. Stattdessen erhalten die Datengebenden in der Web-App eine allgemeine Übersicht über alle Forschungsprojekte, die im Rahmen des Datentreuhandsystems durchgeführt werden. Ähnlich wurde bei der Handhabung von Zufallsbefunden entschieden, diese zunächst an medizinische Fachkräfte der datenerzeugenden Einrichtungen weiterzuleiten. Die Fachkräfte prüfen die Validität der Befunde und kontaktieren die Datengebenden gegebenenfalls, um sicherzustellen, dass diese die Befunde korrekt einordnen und darauf reagieren können. Es wird angenommen, dass Datengebende potentiell bösartig (*engl. malicious*) handeln könnten, um unautorisierten Zugriff auf Funktionen, Berechtigungen oder Daten zu erhalten. Dennoch wird unterstellt, dass die technische Expertise der Datengebenden im Allgemeinen begrenzt ist. Aufgrund persönlicher Interessen an der Plattform und eines grundlegenden Vertrauens in das Datentreuhandsystem wird angenommen, dass

Datengebende überwiegend ehrlich, aber neugierig (*engl. honest-but-curious*) agieren.

- **VA2:** Datenerzeugende haben Zugriff auf Gesundheitsdaten, die sie selbst im Rahmen vertraglicher Vereinbarungen zur medizinischen Versorgung generiert haben. Zudem können sie ausschließlich auf die Einwilligungen der Datengebenden zugreifen, die sich auf die von ihnen erzeugten oder verwalteten Daten beziehen. Der Zugriff auf die Verknüpfungen zwischen den Identifiern des Datentreuhandsystems und den internen Identifiern ihrer IT-Systeme ist nur nach einer expliziten Autorisierung möglich. Diese Verknüpfungen sind notwendig, um die Rückführung von Daten zu ermöglichen, etwa für die Rückkopplung von Zufallsbefunden an die Datengebenden. Es wird angenommen, dass Datenerzeugende größtenteils im Einklang mit den gesetzlichen und sozialen Rahmenbedingungen handeln, die böswilliges Verhalten effektiv minimieren. Während die Möglichkeit eines bösartigen Verhaltens nicht vollständig ausgeschlossen werden kann, wird es als unwahrscheinlich betrachtet. Ihre IT-Kenntnisse werden als mäßig bis gering eingestuft, was das Risiko unabsichtlicher Fehlbedienungen durch mangelnde technische Expertise erhöhen könnte. Dennoch wird unterstellt, dass Datenerzeugende überwiegend in gutem Glauben handeln und die ihnen gewährten Zugriffsrechte und Funktionen verantwortungsvoll nutzen.
- **VA3:** Datennutzende greifen nur auf die im Datentreuhandsystem bereitgestellten, de-identifizierten Daten zu, die durch Nutzungsverträge zur Beantwortung spezifischer Forschungsfragen oder zur Entwicklung innovativer Gesundheitsanwendungen autorisiert worden sind. Sie können keine direkten Rückschlüsse auf Einzelpersonen oder medizinische Einrichtungen ziehen. Darüber hinaus ist es Datennutzenden nicht gestattet, Informationen zu erhalten, die Geschäftsgeheimnisse der Datenerzeugenden offenlegen könnten. Alle Analysen auf den bereitgestellten Daten dürfen ausschließlich innerhalb der vertrauensvollen Ausführungsumgebung des Datentreuhandsystems durchgeführt werden. Diese technische Einschränkung dient nicht nur dem Schutz der sensiblen Daten, sondern auch der Verhinderung unbeabsichtigter Sicherheitsverletzungen, etwa durch die versehentliche Veröffentlichung von Forschungsergebnissen

nach einem Angriff auf persönliche Geräte, wie einem kompromittierten Arbeitslaptop eines Forschenden. Datennutzende verfügen in der Regel über sehr hohe technische Expertise und könnten theoretisch versuchen, bestehende Sicherheitsvorkehrungen zu umgehen oder unautorisierte Rückschlüsse aus den bereitgestellten Daten zu ziehen. Dennoch wird davon ausgegangen, dass sie überwiegend ehrlich, aber neugierig (*engl. honest-but-curious*) handeln und die ihnen gewährten Zugriffsrechte im Rahmen der vorgesehenen Nutzungsbedingungen verwenden.

- **VA4:** Reviewende sind externe, unabhängige Akteure, welche die Einhaltung der rechtlichen Vorgaben und die Reproduzierbarkeit der Forschung sicherstellen. Es wird davon ausgegangen, dass Reviewende neutral und unabhängig handeln. Reviewende verfügen in der Regel über eine hohe fachliche und technische Expertise, was ein potenzielles Risiko für die Umgehung bestehender Sicherheitsvorkehrungen darstellt. Dennoch wird angenommen, dass sie überwiegend im Interesse der Wahrung ihrer beruflichen Integrität und des Systems handeln und die ihnen gewährten Zugriffsrechte ausschließlich für den vorgesehenen Zweck nutzen. Die Möglichkeit bösartiger Absichten (*engl. malicious*) seitens der Reviewenden kann jedoch nicht vollständig ausgeschlossen werden. Dieses Risiko wird durch gezielte Beschränkungen minimiert. Reviewende erhalten nur die für ihre Aufgabe notwendigen Informationen, um die Transparenz und Datenintegrität zu gewährleisten. Direkter Zugriff auf personenbezogene Daten, Geschäftsgeheimnisse oder vertrauliche Informationen der Systembeteiligten ist nicht vorgesehen.

- **VA5:** Systemadministrierende der zentralen und dezentralen Komponenten des Datentreuhandsystems sind entscheidende Akteure, die für die technische Verwaltung, Wartung und Sicherstellung des störungsfreien Betriebs des Systems verantwortlich sind. Es wird erwartet, dass sie gemäß ihren beruflichen Verpflichtungen handeln und hohe Standards der Vertraulichkeit, Integrität und Verfügbarkeit der verwalteten Daten wahren. Sie besitzen weitreichenden Zugriff auf die technischen Komponenten des Systems und verfügen über sehr hohe IT-Kompetenzen zur Konfiguration und Wartung von technischen Systemen. Es wird davon ausgegangen, dass die von den Systemadministrierenden verwalteten

Systeme etablierte Sicherheitsprotokolle einhalten und daher als vertrauenswürdig gelten. Zudem minimieren gesetzliche und vertragliche Rahmenbedingungen das Risiko eines aktiven Missbrauchs durch die Systemadministrierenden. Trotz dieser Annahmen können schlecht geschulte Systemadministrierende potenzielle Bedrohungen darstellen, insbesondere durch unbeabsichtigte Fehler. Um dies zu adressieren, werden strenge Sicherheitsprotokolle implementiert und regelmäßige Audits durchgeführt, um die Integrität des Systems sicherzustellen. Insgesamt wird angenommen, dass Systemadministrierende überwiegend ehrlich, aber neugierig (*engl. honest-but-curious*) handeln und das System so konzipiert ist, dass Missbrauchsmöglichkeiten durch technische und organisatorische Maßnahmen minimiert werden.

- **VA5:** Rechtsvertretende Personen übernehmen eine zentrale juristische Rolle. Ihre Arbeit basiert auf dem Vertrauen, dass sie unabhängig, neutral und ihre juristischen Kompetenzen ausschließlich im Einklang mit den geltenden rechtlichen und ethischen Vorgaben einsetzen. Sie handeln mit der primären Absicht, die Interessen der beteiligten Parteien – einschließlich der Datengebenden, Datenerzeugenden und Datennutzenden – zu schützen und die Funktionsfähigkeit des Datentreuhandsystems sicherzustellen. Ihre Entscheidungen, beispielsweise bei der Überprüfung von Registrierungsanfragen oder der Verwaltung von Nutzungsverträgen, basieren auf objektiven juristischen Kriterien und verfolgen keine eigenen oder unzulässigen Interessen. Die Annahme ist, dass die Rechtsvertretenden überwiegend ehrlich, aber neugierig (engl. honest-but-curious) und in gutem Glauben handeln. Während das Risiko von Fehlverhalten oder Missbrauch durch Rechtsvertretende aufgrund ihrer Schlüsselrolle im System nicht vollständig ausgeschlossen werden kann, wird davon ausgegangen, dass dies durch vertragliche, organisatorische und gesetzliche Maßnahmen effektiv minimiert wird. Zusätzlich wird davon ausgegangen, dass die Rechtsvertretenden nur über ein begrenztes technisches Wissen verfügen. Dies könnte ihre Fähigkeit einschränken, technische Details oder potenzielle Risiken, die sich aus komplexen IT-Systemen ergeben, vollständig zu bewerten. Deshalb wird erwartet, dass sie sich bei technischen Fragestellungen auf die Expertise anderer Fachleute innerhalb des Datentreuhandsystems stützen.

- **VA6:** Es wird angenommen, dass die technischen Komponenten insbesondere die dezentralen Connectoren, die Netzwerkinfrastruktur sowie die technischen Systeme und Server des Datentreuhandsystems (z.B. Web-Apps, Logging-Blockchain, Ausführungsumgebung) den festgelegten Protokollen folgen. Bedrohliches Verhalten durch externe Angreifende kann nicht ausgeschlossen werden und wird durch geeignete Sicherheitsmaßnahmen adressiert.

Schützenswerte Güter: Die Identifikation schützenswerter Güter ist ein zentraler Bestandteil beim Aufbau eines Datentreuhandsystems, das als neutraler Vermittler zwischen Datengebenden, Datenerzeugenden und Datennutzenden agiert. In dieser Rolle ist das Vertrauen aller Beteiligten von entscheidender Bedeutung, da der Datentreuhänder die Verantwortung trägt, sensible Daten sicher und im Einklang mit rechtlichen sowie ethischen Vorgaben zu verwalten. Die schützenswerten Güter umfassen dabei nicht nur die sensiblen Daten selbst, sondern auch die Mechanismen, die Vertraulichkeit, Integrität und Verfügbarkeit sicherstellen. Die schützenswerten Güter, die für das Datentreuhandsystem identifiziert wurden, sind in Tabelle 5.2 aufgeführt und beschrieben.

Tabelle 5.2: Die identifizierten schützenswerten Güter des Datentreuhandsystems und deren Beschreibung.

Komponente	ID	Name des schützenswerten Guts	Beschreibung des Guts
Web-App des Datengebenden	A1	Beschreibungen zu in den medizinischen Einrichtungen vorhandenen medizinischen Daten	Die von verschiedenen datenerzeugenden Gesundheitseinrichtungen bereitgestellten Beschreibungen der Art von vorhandenen Daten als Grundlage zur Erteilung/ Nicht-Erteilung einer Einwilligung durch den Datengebenden.

	A2	Liste und Historie an Einwilligungen	Informationen über bestehende Kommunikationsbeziehungen mit verbundenen Connectoren sowie deren Austauschinhalte und Zugriffsrechte basierten auf erteilten Einwilligungen und deren Historie.
	A3	Informationen zur Kommunikation des Datengebenden	Netzinterne Kennung des Webservers der Datengebenden und Netzadressen, die für den Austausch von Nachrichten und Daten über die Netzinfrastruktur zwischen Connectoren erforderlich sind.
	A4	Identifizierende Daten	Die DTM-ID als Pseudonym/ Unique Identifier des Datengebenden sowie die optional angebbare E-Mailadresse.
Web-App des Datennutzenden	A5	Identifizierende Daten	Identifier der Datennutzenden sowie dessen Accountdaten (Name, Institution, Adresse Institution, geschäftliche E-Mail-Adresse).
	A6	Forschungsprojektinformationen	Informationen zu Forschungsprojekten (Projektkürzel, Projekttitel, Forschungsziel, Wissenschaftlicher Hintergrund, Forschungsmethode, Abschätzung Machbarkeit, Nutzungsabsicht, Datenschutzmaßnahmen, Start- und Enddatum Vorhaben, Kontaktinformation des Projektleitenden).

	A7	Machbarkeitsanfragen	Die Inhalte von Nutzungsanfragen sowie die dazugehörig festgelegten Ein- und Ausschlusskriterien.
	A8	Nutzungsanfragen und -verträge	Inhalte von Nutzungsanfragen (Machbarkeitsanfrage, Informationen zu bestehendem Ethikvotum, Verwertungsziele, Schutzrechte) und dazugehörige ausgearbeitete Nutzungsverträge.
	A9	Experimentdaten	Schützenswerte Daten und Algorithmen von Datennutzenden für Experimente (Dateien, Titel Algorithmus, Beschreibung Algorithmus, Experimentergebnisse, Forschungsergebnisse, Publikationsinformationen).
Web-App des Reviewenden	A10	Auditierte Logging-Informationen	Die zu validierenden Logging-Information bei Datennutzungen im Rahmen des Treuhandsystems (Organisationsname, Zeitstempel, ProjektURL, ExperimentURL, Hash des Experiments/ Algorithmus).
Web-App der Rechtsvertretenden Person	A11	Identifizierende Daten	Identifier der Rechtsvertretenden Personen sowie deren Accountdaten.
	A12	Anfragen und Überblick über erteilte Zugriffe	Anfragen und Historie über Entscheidungen zur Annahme von Datennutzenden in der Datentreuhandplattform.

Logging-Blockchain	A13	Auditierte Logging-Informationen	Logging-Information bei der Datennutzung (Organisationsname, Zeitstempel, ProjektURL, ExperimentURL, Hash des Experiments/ Algorithmus), welche zur Nachvollziehbarkeit in der Blockchain gespeichert werden.
Ausführungsumgebung für Experimente	A14	Pseudonymisierte Forschungsdaten	Während der Ausführung der Experimente im Datentreuhänder verwendete Forschungsdaten, welche nach Abschluss dieser wieder gelöscht werden.
	A15	Experimentdaten	Algorithmen, Machine-Learning-Modelle und Konfigurationsdaten von Datennutzenden für die Ausführung von Experimenten.
Netzwerkinfrastruktur	A16	Pseudonymisierte Forschungsdaten	Die pseudonymisierten (Gesundheits-) Daten, die über die Netzinfrastruktur ausgetauscht werden.
	A17	Kommunikationsinhalte und Nachrichten	Nachrichten, Berechtigungen und Anfragen, welche über die Netzwerkinfrastruktur ausgetauscht werden.
	A18	Identifizierende Kommunikationsdaten	Netzinterne Kennungen von Benutzer*innen (DIDs) und Netzadressen, die für den Austausch von Nachrichten und Daten erforderlich sind.

Connector	A19	Zugriffsprotokolle	Liste aller Zugriffsanfragen, Status über den Erfolg der Anfrage und die zugehörigen Daten und Metadaten des Antragstellers (z. B. Netzwerkadresse, Datum und Uhrzeit des Zugriffs).
	A20	Access Control List (ACL)	Die ACL legt auf Grundlage der Einwilligungen der Datengebenden fest, in welchem Umfang einzelne Benutzer*innen und Systeme Zugriff auf die internen Objekte der Datenerzeugenden (wie z. B. Dienste, Dateien) haben.
	A21	Zuordnungsinformationen	Interne Zuordnung der Identitäten der Datengebenden im Netz/ Datentreuhandsystem zu denen in den internen IT-Systemen der Datenerzeugenden.
Interne IT Systeme der medizinischen, datenerzeugenden Einrichtungen	A22	Personenbezogene medizinische Daten	Die in den internen Informationssystemen der medizinischen Einrichtungen im Rahmen der Versorgung und Studien aufgezeichneten Gesundheitsdaten, welche dem Datentreuhandsystem zur Verfügung gestellt werden können.

Technische Annahmen: Es wurden die folgenden technischen Annahmen für das Datentreuhandsystem formuliert:

- **TA1:** Das Prinzip der Datenminimierung wird angewendet, wodurch die langfristige Bildung zentralisierter und potenziell unsicherer Datensilos

vermieden wird. Zudem werden ausschließlich die für die jeweilige Forschungsnutzung erforderlichen Daten an die Datennutzenden weitergegeben.

- **TA2:** Es sollten keine grundlegenden Änderungen an den internen IT-Systemen der Datenerzeugenden vorgenommen werden, mit Ausnahme der Anpassungen der Schnittstellen zum Connector.
- **TA3:** Die medizinische Einrichtung, welche die Daten generiert hat, stellt diese dem Datentreuhandsystem entsprechend der Einwilligung des Datengebenden zur Verfügung.
- **TA4:** Es verlassen keine Daten die internen IT-Systeme, für die keine Einwilligung des Datengebenden vorliegt.
- **TA5:** Daten werden nur für den vorgesehenen Zeitraum der Experimentausführung im Datentreuhandsystem gespeichert und danach entsprechend gelöscht.

5.2.4 Sicherheits- und Bedrohungsmodellierung

Im Kontext des zweiten Anwendungsfalls wurde in den vorangegangenen Abschnitten das zugrundeliegende Systemkonzept sowie die entsprechenden Datenflüsse des Datentreuhandsystems detailliert beschrieben. Darauf aufbauend erfolgt gemäß der STRIDE-Methode (siehe Abschnitt 2.3.12) eine systematische Identifikation und Analyse sicherheitsrelevanter Schwachstellen und potenzieller Bedrohungen. Gleichzeitig werden geeignete Sicherheitsmechanismen abgeleitet, die dazu dienen, diese Schwachstellen zu adressieren und die identifizierten Bedrohungen zu mitigieren. Eine Übersicht der im Rahmen dieses Anwendungsfalls identifizierten Bedrohungen und zugehörigen Sicherheitsmechanismen ist in Abbildung 5.3 dargestellt.

Da die Netzwerkinfrastruktur des Datentreuhandsystems technologisch auf derjenigen des Gesundheitsdatenmanagements aus Anwendungsfall 1 basiert, ergeben

sich vergleichbare Angriffsvektoren und Sicherheitsmechanismen, insbesondere solche, die für Blockchain-basierte Systeme charakteristisch sind (siehe Abschnitt 5.1.4). Im Folgenden wird eine weiterführende Analyse relevanter Sicherheitsaspekte beschrieben, die über diese bereits genannten Angriffsszenarien und Schutzmechanismen hinausgehen.

Ein zentraler Bestandteil beim Aufbau eines Datentreuhandsystems, das als neutraler Vermittler agiert, ist die Verifizierung der Identitäten aller Kommunikations- und Systemakteure sowie die Sicherstellung der Integrität des Systems, der Kommunikationskanäle und der bereitzustellenden Daten, um die unbeabsichtigte Offenlegung personenbezogener Daten wirksam zu verhindern und gleichzeitig repräsentative Forschung auf validen Gesundheitsdaten zu ermöglichen. Im Hinblick auf die Wahrung der Privatsphäre und den Schutz personenbezogener Daten spielt die De-Identifikation eine zentrale Rolle. Durch die Ersetzung personenbezogener Identifikatoren durch pseudonyme Identifikatoren wird die direkte Zuordnung von Gesundheitsdaten zu einer bestimmten Person verhindert, wodurch das Risiko der Offenlegung personenbezogener Daten verringert wird. Zudem wird eine strikte Trennung der Identifier vorgenommen: Es werden getrennte Datentreuhand-Identifier verwendet, um sicherzustellen, dass keine Korrelationen zwischen den Datengebenden und ihren Versorgungsentitäten möglich sind. Dabei bleibt dem Datentreuhandsystem das Mapping zwischen den Identifikatoren unbekannt, da diese ausschließlich in den Connectoren der datenerzeugenden Einrichtungen abgespeichert werden. Insofern wird ein zusätzliches Schutzniveau gegenüber der Re-Identifikation geschaffen. Grundlegend könnte ebenfalls eine Anonymisierung als Methode der De-Identifikation in Betracht gezogen werden. Zur Rückkopplung von Zufallsbefunden wurde sich jedoch im Falle des medizinischen Datentreuhänders für die Pseudonymisierung entschieden.

Neben der De-Identifikation sollen ebenfalls Dienste zur Normalisierung und Qualitätsmessung/-verbesserung im Datentreuhandsystem vorgesehen werden. Dieser Prozess beinhaltet die Modifikation der Daten. Dementsprechend sollte eine Versionierung der Daten integriert werden sowie Testungen und Simulationen

der eingesetzten Verfahren, um die Sicherheit und Robustheit der Verarbeitungsfunktionen vor dem produktiven Einsatz zu überprüfen und eine kontinuierliche Testung während des Betriebs sicherzustellen.

Der Einstiegspunkt für Datengebende im Datentreuhandsystem erfolgt während ihres Aufenthalts in einer medizinischen Einrichtung, in deren Rahmen im persönlichen Kontakt auf das Datentreuhandsystem und dessen Web-App hingewiesen und Informationen bereitgestellt werden können, die als Grundlage für eine informierte Einwilligung dienen. Dieser Kontaktpunkt bildet die technische Basis für eine vertrauenswürdige Identifizierung und den Zugang zum System, wobei entsprechende Identifikatoren ohne persönliche Referenzen und kryptografische Schlüssel im direkten Kontakt ausgetauscht werden können. Auf dieser Grundlage können anschließend die verfügbaren Daten an die Datengebenden über einen sicheren Kommunikationskanal zurückgespiegelt werden. Eine detaillierte Sichtung der konkreten Gesundheitsdaten sowie die Erteilung der Einwilligung kann im Anschluss, unter Berücksichtigung eines zeitlichen Abstands zur Behandlung, erfolgen, sodass kein Druck besteht und die Entscheidung freiwillig und nach umfassender Überlegung getroffen wird. Für die Vorhaltung und Weitergabe der dezentralen Daten durch Datenerzeugende an das Datentreuhandsystem ist ebenfalls ein sicherer Kommunikationskanal erforderlich. Die Gewährleistung der Authentizität der über den sicheren Kommunikationskanal übertragenen Daten kann durch den Einsatz digitaler Signaturen erfolgen.

Nach der Einrichtung des Kommunikationskanals muss es für potenzielle Angreifer unmöglich sein, Informationen aus dem Netzwerkverkehr und -inhalt zu extrahieren. Zu diesem Zweck sind die Implementierung von Ende-zu-Ende-Verschlüsselung sowie die Verwendung sicherer Netzwerkprotokolle erforderlich. Darüber hinaus ist der Einsatz von Mix-Netzwerken erforderlich, um sowohl die Identitäten der Kommunikationspartner als auch die Frequenz ihrer Datenübertragungen im Netzwerk zu verschleiern.

Neben dem sicheren Datentransfer muss innerhalb der Zielsysteme, Connectoren und Web-Apps sichergestellt werden, dass vertrauliche Daten verschlüsselt und Passwörter gehasht werden. Um den Zugriff Dritter im Falle eines Verlusts oder

Diebstahls dieser Daten zu verhindern, sollten die webbasierten Anwendungen durch Passwortschutz gesichert sein. Die Prinzipien von Privacy-by-Design und Privacy-by-Default sollten implementiert werden, um potenzielle Bedrohungen, die aus fehlerhaften Benutzendeninteraktionen resultieren könnten, zu minimieren. Ein wesentlicher Aspekt dabei ist die benutzendenfreundliche Verwaltung von Einwilligungen. Diese sollte so gestaltet sein, dass Nutzende diese intuitiv verwenden können, einschließlich der Bereitstellung von Plausibilitätsprüfungen und Warnhinweisen, um Fehler bei der Eingabe oder Verarbeitung von Einwilligungen zu vermeiden. Darüber hinaus ist sicherzustellen, dass angemessene Fehlermeldungen in den Web-Apps verwendet werden, die weder sensible Daten noch Metadaten enthalten, um potenzielle Sicherheitsrisiken durch die Preisgabe solcher Informationen auszuschließen.

Ein besonderer Fokus sollte ferner auf die Bereitstellung einer umfassenden Widerrufsmöglichkeit gelegt werden, die es Nutzenden erlaubt, erteilte Einwilligungen, gespeicherte Daten sowie Zugriffsrechte auf die Plattform jederzeit zurückzuziehen. Im Falle eines Widerrufs der Einwilligung können Daten, die bereits im Rahmen eines geschlossenen Nutzungsvertrags durch Datennutzende verwendet werden, weiterhin zur Durchführung bestehender Forschungsprojekte genutzt werden, um deren Fortführung zu gewährleisten. Neue Machbarkeitsanfragen oder Datenabrufe durch andere Datennutzende sind jedoch nach dem Widerruf nicht mehr möglich. In diesem Zusammenhang wurde in unserem Fall eine maximale Gültigkeit von Einwilligungen auf fünf Jahren festgelegt. Innerhalb dieses Zeitraums können Datennutzende Anfragen auf Datenzugriff stellen. Zugriffsberechtigungen auf die Daten sind für das Datentreuhandsystem auf eine maximale Dauer von zehn Jahren begrenzt. Dementsprechend sollten Einladungen zur Datennutzung, erteilte Einwilligungen sowie Zugriffsrechte auf pseudonymisierte Daten zeitlich durch eine entsprechende Nutzungsvereinbarung begrenzt werden, um die Einhaltung datenschutzrechtlicher und ethischer Vorgaben sicherzustellen.

Darüber hinaus sollten geeignete Zugangskontrollmechanismen implementiert werden, um den Zugriff ausschließlich den Akteuren zu gestatten, die von den Datengebenden über eine Einwilligung autorisiert wurden. Eine umfassende Protokollierung von Authentifizierungsaktionen, Zugriffsberechtigungen, Datenan- und

-abfragen sowie Datennutzungen ist notwendig, um die Abstreitbarkeit der durchgeführten Aktionen zu gewährleisten. Insbesondere zur Protokollierung der Datennutzungen kann eine öffentliche, manipulationssichere, permissioned Blockchain als Auditierungsmechanismus eingesetzt werden. Diese ermöglicht eine transparente und unveränderbare Dokumentation aller Datenverarbeitungen im Rahmen des Datentreuhandsystems. Hierbei sollte jedoch darauf geachtet werden, dass keine vertraulichen Daten direkt in den Logdaten auf der Blockchain gespeichert werden. Darüber hinaus können Zero-Knowledge-Proofs (ZKPs) genutzt werden, um die entsprechenden Nachweise zu erbringen, ohne dass sensible Informationen offengelegt werden. Die Transparenz der Datennutzung stärkt das Vertrauen in den Datentreuhänder und ermöglicht es reviewenden Dritten, die Rechtmäßigkeit der Datennutzung zu überprüfen. Darüber hinaus bietet die Blockchain der wissenschaftlichen Community die Möglichkeit, zu überprüfen, ob ein mit den Daten des Treuhandsystems entwickelter Algorithmus tatsächlich mit diesen gespeist und trainiert wurde, um die Reproduzierbarkeit der Ergebnisse zu ermöglichen.

Abbildung 5.8: Die identifizierten Sicherheitsbedrohungen im zweiten Anwendungsfall und deren Beschreibung sowie die Sicherheitsmaßnahmen zur Abwehr dieser Bedrohungen.

5.2.5 Entwickelte Systemarchitektur

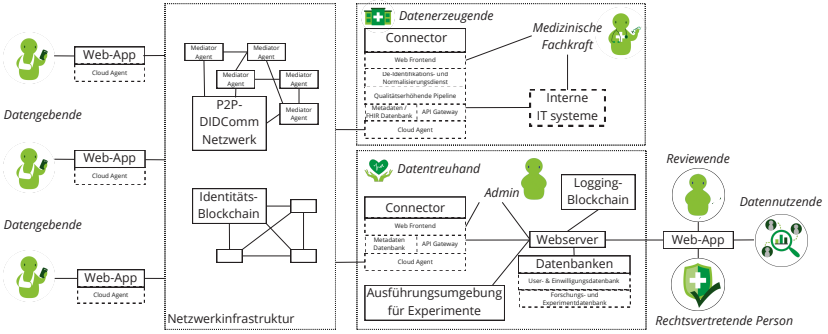

Abbildung 5.9: Die entwickelte Systemarchitektur des Datentreuhandsystems.

Der folgende Abschnitt beschreibt die vorgeschlagene Systemarchitektur des Datentreuhandsystems (siehe Abbildung 5.9), die auf den zuvor entwickelten Konzepten, Anforderungen, Annahmen und Modellierungen basiert. Im Zentrum der Systemarchitektur steht die Netzwerkinfrastruktur, die hinsichtlich ihrer grundlegenden Eigenschaften und Konzepten vergleichbar zur Netzwerkinfrastruktur aus Anwendungsfall 1 ist (siehe Abschnitt 5.1.5) und auf dem Open-Source-Protokoll DIDComm[6] (*SM18-20*) sowie den Open-Source-Implementierungen von Hyperledger Indy und Aries[7] (*SM2, SM4, SM8, SM9, SM29*) basiert. Im vorliegenden Anwendungsfall dient die Netzwerkinfrastruktur zur Kommunikation und zum Datenaustausch zwischen den Datengebenden, Datenerzeugenden und dem Datentreuhandsystem, welches als dezentrales, durch Blockchain-Technologie gesichertes Peer-to-Peer-Kommunikationsnetzwerk aufgespannt wird.

Zur Umsetzung dieses Peer-to-Peer-Netzwerkes betreibt das Datentreuhandsystem einen Connector mit integrierten Aries Cloud-Agent für die Web-Apps der

6 https://identity.foundation/didcomm-messaging/spec/
7 https://www.hyperledger.org/use

Datengebenden sowie einen weiteren Connector der Datentreuhand zur Kommunikation mit den Connectoren der beteiligten Datenerzeugenden sowie dem Datengebenden-Connector. Mediator Agents ermöglichen dabei eine asynchrone und verschlüsselte Kommunikation zwischen den Connectoren mit den dazugehörigen Web-Apps, wobei diese als Mix-Netzwerke konfiguriert werden können (*SM21, SM23*). Ergänzt wird das DIDComm-Netzwerk ebenfalls durch eine Identitäts-Blockchain, die den Netzwerkteilnehmenden durch die Veröffentlichung eines manipulationssicheren Registers für digitale Signaturen und Netzwerkadressen als Vertrauensanker dient und so eine vertrauenswürdige Kommunikation auf Basis verifizierter digitaler Identitäten zulässt.

Die in diesem Kontext verwendeten Identitäten, bestehend aus Peer-DIDs und public DIDs, sind ausschließlich für die Kommunikation im Netzwerk vorgesehen und sind getrennt von den Identitäten innerhalb der datenerzeugenden Einrichtungen (z.B. Patient*innen-, Mitarbeitenden- oder Fall-IDs) sowie den Pseudonymen oder Identifikatoren im Datentreuhandsystem (*SM25*). Die public DIDs der Datenerzeugenden und der Datentreuhand werden von diesen in der permissioned Indy Blockchain registriert, wodurch neue Agenten-zu-Agenten-Verbindungen durch eine persönliche Registrierung aufgebaut und verifizierbare Berechtigungsnachweise zur technischen Authentifizierung ausgetauscht sowie in einer passwortgeschützten Wallet gespeichert werden können (*SM1, SM5, SM6, SM33*).

Um die Korrelation von Informationen und Identitäten zu minimieren, wird für jede private Verbindung zwischen zwei Agenten im DIDComm-Netzwerk ein einzigartiges Peer-DID-Paar [8] erstellt, welches ausschließlich den beiden beteiligten Parteien bekannt ist. Die Web-App dient als zentrale Benutzendenoberfläche, über die Datengebende Verbindungen zu Connectoren der datenerzeugenden medizinischen Einrichtungen herstellen, Metadaten sowie Beschreibungen ihrer dort gespeicherten Gesundheitsdaten einsehen und eine informierte Einwilligung zur

[8] https://identity.foundation/peer-did-method-spec/

Nutzung dieser Daten für Forschungszwecke durch das Datentreuhandsystem erteilen können (*SM28*).

Nach der persönlichen Registrierung mit einem Datenerzeugenden wird dessen Connector das Datentreuhand-Pseudonym des Datengebenden übermittelt, welches in der Metadatenbank des Connectors gespeichert, mit den Identifikatoren in den bestehenden IT-Systemen verlinkt und mit zukünftigen Einwilligungen zum Zugriffsmanagement assoziiert. Die Metadaten-Datenbank dient demgemäß der Speicherung von Metadaten, die für die Funktionsweise des Systems von Bedeutung sind. Dazu gehören unter anderem die Zuordnung von Peer-DIDs zu Patient*innen-IDs aus den internen IT-Systemen, eine Übersicht der behandelnden Fachkräfte, temporäre Zugriffsrichtlinien sowie Protokolle über Zugriffsereignisse (*SM13, SM14, SM22, SM31, SM32*). Darüber hinaus umfasst die Metadatenbank die verschlüsselte Speicherung vertraulicher Daten in einer SSI-Wallet sowie die Protokollierung von Authentifizierungsaktionen (*SM3, SM7, SM12, SM15, SM16-18, SM24, SM28*).

Der Connector der Datenerzeugenden ist eine dezentrale Komponente und Schnittstelle des Datentreuhandsystems, welche lokal (On-Premise) von den Datenerzeugenden betrieben wird. Hierdurch erfolgt die Speicherung der Gesundheitsdaten, der Einwilligungen für das Zugriffsmanagement und die Verknüpfungen zu den realen Identitäten weiterhin dezentral bei den Datenerzeugenden. Dabei erhält das Datentreuhandsystem diese sensiblen personenbezogenen Daten entweder gar nicht oder ausschließlich in einem strikt begrenzten Umfang. Darüber hinaus umfasst der Connector ein API-Gateway, einen Dienst zur De-Identifikation und Normalisierung, eine FHIR-Datenbank, eine qualitätserhöhende Pipeline sowie ein Web-Frontend für medizinische Fachkräfte. Das API-Gateway fungiert als zentrale Schnittstelle für die Kommunikation mit internen IT-Systemen der Datenerzeugenden sowie mit der Netzwerkinfrastruktur. Es ermöglicht die sichere Übertragung von Daten mittels der definierten Netzwerkprotokolle und die Verwaltung von Anfragen gemäß den definierten Zugriffsrichtlinien basierend auf den Einwilligungen. Zudem ist das API-Gateway über RESTful Webservices zugänglich, wodurch es eine standardisierte und interoperable Integration in bestehende IT-Landschaften gewährleistet.

Ein Dienst zur De-Identifikation und Normalisierung sorgt dafür, dass die personenbezogenen Gesundheitsdaten vor der Weitergabe an das Datentreuhandsystem pseudonymisiert oder anonymisiert werden (*SM10, SM11, SM26*). Zudem standardisiert der Dienst die Daten in Übereinstimmung mit dem FHIR-Standard [9], um eine Interoperabilität mit Daten aus verschiedenen datenerzeugenden Einrichtungen sicherzustellen. Die Standardisierung wird jedoch obsolet, wenn die Daten in den internen IT-Systemen der Datenerzeugenden bereits konform zum FHIR-Standard gespeichert sind. In solchen Fällen kann der Dienst zur Normalisierung deaktiviert werden, was den Verarbeitungsaufwand reduziert und die Effizienz des Systems steigert. Die FHIR-Datenbank dient zur Speicherung der Gesundheitsdaten im FHIR-Format und trägt insbesondere dann zur Verbesserung der Performanz bei, wenn die Daten in den internen IT-Systemen der Datenerzeugenden noch nicht gemäß dem FHIR-Standard vorliegen. Durch die Vorhaltung standardisierter Daten in der FHIR-Datenbank entfällt in solchen Fällen die Notwendigkeit einer erneuten Standardisierung bei jeder Anfrage oder Datenübermittlung an das Datentreuhandsystem. Die qualitätserhöhende Pipeline bewertet die Datenqualität anhand bestimmter Metriken, wie etwa Vollständigkeit, Aktualität und Widerspruchsfreiheit, und verbessert diese vor der Übermittlung und Nutzung. Zu ihren wesentlichen Funktionen zählen die Validierung von Datenformaten, die Korrektur von Fehlern und, soweit möglich, die Ergänzung fehlender Informationen, um eine hohe Datenintegrität zu gewährleisten. Zusätzlich besteht die Möglichkeit, qualitätssteigernde und kuratierte Algorithmen, die von der wissenschaftlichen Gemeinschaft entwickelt wurden, in das System zu integrieren. Diese Algorithmen können die Erfassung zusätzlicher Parameter ermöglichen und somit eine erweiterte Datenbasis für die Nutzung bereitstellen.

Das integrierte Web-Frontend für medizinische Fachkräfte stellt eine benutzendenfreundliche Oberfläche bereit, die den Zugriff auf die Funktionen des Connectors ermöglicht, ohne dass diese auf tiefgehende technische Vorkenntnisse angewiesen sind. Zu den Hauptfunktionen gehören die persönliche Registrierung der

[9] https://build.fhir.org/

medizinischen Fachkraft mit den Datengebenden, die Bereitstellung von Informationen zu den erhobenen Gesundheitsdaten sowie weiterer relevanter Informationen, die als Grundlage für die Einholung einer informierten Einwilligung dienen (*SM16, SM17*). Darüber hinaus bietet das Frontend die Möglichkeit, die erteilten Einwilligungen und an das Datentreuhandsystem bereitgestellte Daten einzusehen. Nach dem Erhalt der Einwilligung des Datengebenden über die Netzwerkinfrastruktur durch den Connector der Datengebenden, ermöglicht dieser die Anfrage und Bereitstellung der Daten an das Datentreuhandsystem, sofern die Nutzungsabsichten der Datennutzenden mit der erteilten Einwilligung übereinstimmen und diese nicht widerrufen wurde (*SM7, SM15*). Voraussetzung für diesen Prozess ist die initiale Verbindung zwischen den Connectoren der Klinik und dem Connector des Datentreuhandsystems, die erfolgt, sobald die Einrichtungen auf organisatorischer, rechtlicher und technischer Ebene beschlossen haben, am Datentreuhandsystem teilzunehmen. Dieser Schritt umfasst die Schaffung vertraglicher Vereinbarungen, die technische Anbindung der Connectoren an die bestehenden IT-Systeme der Einrichtung sowie die Implementierung und Akzeptanzförderung organisatorischer Prozesse, welche durch entsprechende Schulungsmaßnahmen in den klinischen Alltag integriert werden.

Nachdem die Verknüpfung und Etablierung einer P2P-Verbindung zwischen Datenerzeugenden und dem Datentreuhandsystem sowie eine Einwilligung zur Datennutzung durch den Datengebenden erfolgt ist, können Datennutzende nach erfolgreicher Registrierung über deren Web-App Anfragen über die Verfügbarkeit von Daten stellen (*SM28*). Hierzu können Datennutzende forschungsrelevante Kriterien (Ein- und Ausschlusskriterien) hinsichtlich der benötigten Daten definieren, die als sogenannte Machbarkeitsanfragen bezeichnet werden (*SM30*). Zusätzlich müssen diese vorab Informationen zu ihrem Forschungsprojekt und -ziel angeben, um deren Nutzungsabsichten auszudrücken und generell von einer rechtsvertretenden Person über deren Webanwendung für die Nutzung des Systems berechtigt zu werden. Daraufhin wird die Anfrage an die Connectoren der verknüpften Datenerzeugenden weitergeleitet und ein automatisierter Abgleich der Nutzungsabsichten mit den vorhandenen Daten und Einwilligungen erfolgt. Bei positiver Übereinstimmung von Nutzungsabsichten und Nutzungsbedingungen wird zunächst eine

aggregierte Rückmeldung (Information zur Anzahl der vorhandenen Daten bei dem die Kriterien matchen) der vorhandenen Datensätze von allen Connectoren an den Datennutzenden über den Datentreuhand-Connector weitergeleitet. Darauf basierend kann der Datennutzende einen Nutzungsvertrag mit dem Datentreuhandsystem über die rechtsvertretende Person abschließen, wenn dies Anzahl der Daten ausreichend für die Beantwortung seiner Forschungsfragen sind. Nach positivem Bescheid des Vertragsabschlusses durch die rechtsvertretende Person, können die Daten dem Datennutzenden im Rahmen der sicheren Ausführungsumgebung für Experimente der Datentreuhand in pseudonymisierter Form durch die verschiedenen Datenerzeugenden bereitgestellt werden. Über dessen Web-App kann der Datennutzende seine Algorithmen und Experimente hochladen, verwalten und in der Ausführungsumgebung ausführen (*SM28*).

Durch verschiedene Ebenen der Dateneinsicht und Implementierung mehrere Sicherheitsmechanismen wird sichergestellt, dass die Datennutzung ausschließlich bei Bedarf und unter Einhaltung entsprechender Berechtigungen sowie vertraglicher Vereinbarungen erfolgt. Dies umfasst beispielsweise die Bereitstellung von nur für die jeweilige Anfrage relevanten Daten und die Möglichkeit, den Zugriff in Echtzeit zu überwachen. Zusätzlich wird durch die ausschließliche Nutzung der Daten in der Ausführungsumgebung auf sicheren Servern des Datentreuhandsystems gewährleistet, dass das Risiko einer unbefugten Veröffentlichung erheblich reduziert wird. Alle Datenverarbeitungsprozesse werden in einer isolierten, geschützten Umgebung durchgeführt, die einen unautorisierten Zugriff auf die Daten verhindert.

Eine transparente Auditierung aller Zugriffe und Verarbeitungsschritte im Rahmen solcher Experimente durch eine öffentliche Logging-Blockchain ermöglichen die Protokollierung, Kontrolle und Nachvollziehbarkeit über die Rechtmäßigkeit der Datennutzung (*SM27*). Reviewende können anschließend die in der Blockchain geloggten Informationen auf einfache Weise über eine Web-App validieren, wodurch die Transparenz und Rechtmäßigkeit der Datenverarbeitung gewährleistet wird.

Im Rahmen des Datentreuhandsystems werden dedizierte Web-Server betrieben, die für die Webanwendungen der verschiedenen Akteure, einschließlich der Datengebenden, zur Verfügung stehen. Diese Web-Server sind an relevante Datenbanksysteme angebunden, welche die Systemfunktionen ermöglichen. Beispielsweise umfasst die User-Datenbank die identifizierenden Daten der Akteure, wie E-Mail-Adressen der User, während die Forschungs- und Experiment-Datenbank Informationen zu den Datennutzenden, deren Forschungszwecken, den verwendeten Algorithmen sowie den Konfigurationen und Ausführungsergebnissen enthält. Alle Datenbanken sind durch Sicherheitsmechanismen wie Verschlüsselung, Zugangskontrollen und regelmäßige Audits gesichert, um sicherzustellen, dass nur autorisierte Parteien Zugang zu sensiblen Informationen erhalten. Dies fördert die Integrität des Systems und schützt die Vertraulichkeit der gespeicherten Daten. Die vorgestellte Systemarchitektur gewährleistet die Umsetzung aller in Abbildung 5.8 definierten Sicherheitsmechanismen.

5.2.6 Perspektive der Architektur im deutschen Gesundheitswesen

Ebenso wie im ersten Anwendungsfall wurde auch für die Architektur des Datentreuhandsystems zur Sekundärnutzung von schlafmedizinischen Daten ein Forschungsdemonstrator im Rahmen des Forschungsprojekts SouveMed (vgl. Abschnitt 1.3) durch das Konsortium entwickelt und evaluiert [19, 18]. Die Autorin war ebenfalls maßgeblich in die Arbeit des Konsortiums eingebunden. Bestandteil der Evaluation waren unter anderem die Befragungen von Burmeister et al. [18], zur Bewertung des Datentreuhandmodells aus Sicht von Datengebenden, Datennutzenden und Datenerzeugenden hinsichtlich Vertrauen, Gebrauchstauglichkeit und inhaltlichen Verständnis des finalen Prototyp (siehe Abbildungen A.2 und A.3) sowie eine Zwischenevaluation der Universitätsmedizin Freiburg zur allgemeinen Bereitschaft zur Datenbereitstellung für die Sekundärnutzung und Bewertung eines ersten interaktiven Prototyps [19]. Die relevanten Aspekte hinsichtlich der Verwertung und Perspektive des Konzepts des Datentreuhandsystems, die aus dem durch Burmeister et al. und dem SouveMed-Konsortium erhobenen Feedback

hervorgehen, werden im Folgenden kurz zusammenfasst und diskutiert. Grundlegend besteht das Interesse Daten für die Sekundärnutzung selbstsouverän zu verwalten. Die Kombination eines persönlichen Aufklärungsgesprächs vor Ort durch medizinisches Personal mit einer begleitenden digitalen App unterstützt die Verständlichkeit der vorgesehenen Datenflüsse zur Sekundärnutzung und die Erteilung einer informierten Einwilligung durch Datengebende. Hierdurch haben Patient*innen das Gefühl, dass ihre Interessen im Einklang mit der DSGVO berücksichtigt werden und sie in den Entscheidungs- und Freigabeprozess eingebunden sind. Dieses Vorgehen erfordert jedoch zusätzliche Ressourcen für den Prozess der Aufklärung und der Förderung des Verständnisses in den datenerzeugenden Einrichtungen. [18, 19]

Aus diesem Grund wird im Rahmen des EHDS ein Opt-Out-Verfahren propagiert [105]. Um jedoch auf die Bedürfnisse verschiedener Interessengruppen einzugehen und gleichzeitig eine möglichst hohe Beteiligung an der Sekundärnutzung von Gesundheitsdaten zu gewährleisten, wäre es denkbar, künftig eine kombinierte Herangehensweise zu etablieren. Personen mit einem hohen Sicherheitsbewusstsein könnten die Möglichkeit erhalten, der pauschalen Weitergabe ihrer Daten zu widersprechen und stattdessen mithilfe eines selbstsouveränen Datentreuhandsystems ausschließlich jene Daten freizugeben, die sie explizit für die Forschung bereitstellen möchten. Diese feingranulare Freigabemöglichkeit könnte dazu beitragen, auch Personen mit einem ausgeprägten Sicherheitsbewusstsein zu motivieren, zumindest einen Teil ihrer Daten für die Forschung zur Verfügung zu stellen. Die übrigen Interessengruppen könnten weiterhin vom Opt-Out-Verfahren Gebrauch machen.

Im Rahmen der Befragungen von Burmeister et al. sowie dem SouveMed-Konsortium ergaben sich Erkenntnisse zur perspektivischen Verwertung hinsichtlich der Wahl einer geeigneten Rechts- und Organisationsform. Bei den Befragungen der Datengebenden sprachen sich 8 von 10 Personen für eine Non-Profit-Organisation als Betreibende des Datentreuhandsystems aus. Als wesentlichen Grund nannten die befragten Personen, dass die Verarbeitung und Nutzung sensibler medizinischer Daten in einer gewinnorientierten Struktur das Vertrauen

und die Bereitschaft zur Freigabe der Daten verringern könnte. In diesem Zusammenhang gaben 14 von 20 Datengebenden an, das Datentreuhandsystem nur dann nutzen zu wollen, wenn die Nutzung kostenfrei ist. Bei den Datennutzenden und -erzeugenden lehnten ausschließlich 3 von 20 Personen eine kostenpflichtige Nutzung ab. In diesem Zusammenhang würde ein Geschäftsmodell, bei dem die Kosten vorwiegend von den Datennutzenden getragen werden, eine praktikable Option darstellen. Dabei könnten verschiedene Ansätze wie ein Abonnementmodell oder eine nutzungsabhängige Bezahlung pro abgerufenem Datensatz in Betracht gezogen werden. Eine nutzungsabhängige Bezahlung pro Datensatz könnte zwar fair sein, birgt jedoch das Risiko, die Unabhängigkeit des Treuhänders durch einen Anreiz zur Maximierung von Datenfreigaben zu beeinträchtigen. Ein Abonnementmodell bietet hingegen die Möglichkeit einer unabhängigen Finanzierung. [18, 19]

Aktuell ist festzustellen, dass im Bereich von Gesundheitsdatentreuhandsystemen in Deutschland bislang keine vollständig tragfähigen Modelle etabliert wurden [135]. Bestehende Initiativen, wie beispielsweise die MII, nehmen derzeit noch staatliche Förderungen in Anspruch [83]. Die Akzeptanz aller beteiligten Stakeholder ist ein zentraler Aspekt für den erfolgreichen Aufbau eines Datentreuhandsystems [18, 19]. In diesem Zusammenhang ist es essenziell, dass in datenerzeugenden Einrichtungen keine vollständig neuen Prozesse etabliert werden müssen. Stattdessen sollten digitale Systeme nach Möglichkeit an bestehende Arbeitsabläufe angepasst und integriert werden. Die Grundlage für die Nutzung von Datentreuhandsystemen stellt die Digitalisierung und Standardisierung in den datenerzeugenden Einrichtungen voraus, welche derzeit im deutschen Gesundheitswesen noch Defizite aufweist [17]. Folglich sollten insbesondere dezentrale Systeme mit einer Vielzahl an beteiligten Akteuren durch gezielte Anstrengungen und angemessene finanzielle Unterstützung gefördert werden, um eine nachhaltige Grundlage für deren Implementierung und Betrieb zu schaffen. Die vorgestellte Systemarchitektur fördert zudem die aktive Einbindung der Datenerzeugenden und gewährleistet den Schutz ihrer Rechte am geistigen Eigentum sowie ihrer Geschäftsgeheimnisse (EHDS-Verordnung, Artikel 52) durch die Implementierung

von Machbarkeitsanfragen. Auf Seite der Datennutzenden bestehen die wesentlichen Herausforderungen in bürokratischen und rechtlichen Prozessen, welche für Außenstehende von medizinischen Einrichtungen häufig intransparent sind und insbesondere für kleine Unternehmen und Startups schwer zu bewältigen sind [17]. Hierbei kann ein Datentreuhandkonzept, wie das hier vorgestellte, unterstützen, welches einen zentralen Zugangs- und Nutzungspunkt für Datennutzende trotz der föderierten Datenhaltung und Verwaltung bereitstellt und damit eine niedrigschwellige Datenbeantragung und -nutzung in einer vertrauensvollen Ausführungsumgebung ermöglicht. Das Konzept steht im Einklang mit den Regularien des EHDS, indem es eine sichere Verarbeitungsumgebung sowie eine manipulationssichere Protokollierung von Zugriffen und Verarbeitungen vorsieht [105]. Durch die Modellierung potenzieller Angriffsszenarien und die daraus abgeleiteten Sicherheitsmechanismen wird eine umfassende Sicherheitsbetrachtung durchgeführt, welche die Grundlage für technische und organisatorische Maßnahmen zur Umsetzung der Sicherheits- und Interoperabilitätsanforderungen des EHDS bildet [105]. Zusätzlich wird der Aspekt der Datenminimierung unterstützt, da die Daten ausschließlich für den angegebenen Zweck bereitgestellt werden, während sie ansonsten in den medizinischen Einrichtungen verbleiben können. Durch den modularen Aufbau der Architektur des Datentreuhandsystems wäre es ebenso wie im ersten Anwendungsfall denkbar, Teilelemente des Konzepts als ergänzende Dienste in den nationalen und internationalen Datenraumbestrebungen zu integrieren [19]. Beispielsweise könnten Teile in das FDZ-Gesundheit des BfArM oder in das im Rahmen des EHDS geplante HealthData@EU integriert werden. Dies würde der Schaffung von Doppelstrukturen entgegenwirken. Gleichzeitig könnte die Systemarchitektur ebenso in ihrer Gesamtheit Anwendung finden. Perspektivisch sollte ein Datentreuhandsystem das voraussichtlich ab 2027 eingeführte europäische Austauschformat unterstützen, um eine nahtlose Integration und Interoperabilität im Rahmen des europäischen Datenraums zu gewährleisten [105].

6 Evaluation des Entscheidungsmodells

Im folgenden Kapitel werden das Vorgehen und die Ergebnisse der Evaluation des Entscheidungsmodells präsentiert. Die Evaluation des entwickelten Entscheidungsmodells umfasst zwei zentrale Aspekte: Zum einen wird die Konsistenz der Architekturentscheidungen der herkömmlichen Systemarchitektur mit den Empfehlungen des Entscheidungsmodells überprüft. Zum anderen erfolgt eine Beurteilung des Modells durch Expert*innen aus den Bereichen Informatik, Informationssicherheit und dem Gesundheitswesen. Beide Evaluationen dienen der Sicherstellung der Praktikabilität und Validität des Entscheidungsmodells bei der Anwendung.

6.1 Evaluation der Konsistenz von Architekturentscheidungen mit dem Entscheidungsmodell

Das in Kapitel 4 entwickelte Entscheidungsmodell wird im folgenden Abschnitt herangezogen, um die Architekturentscheidungen für die in Kapitel 5 beschriebenen Anwendungsfälle zu erfassen. Dabei wird evaluiert, ob das Entscheidungsmodell unter Zuhilfenahme der definierten Anforderungen und Rahmenbedingungen aus den Anwendungsfällen zu denselben Architekturentscheidungen führt wie die herkömmliche Systemkonzeption, die ohne den Einsatz eines Entscheidungsmodells durchgeführt wurde.

6.1.1 Anwendungsfall 1: Patient*innen-zentriertes Gesundheitsdatenmanagement in der medizinischen Versorgung

In diesem Abschnitt wird der erste Anwendungsfall (Patient*innen-zentriertes Gesundheitsdatenmanagement in der medizinischen Versorgung) im Kontext des Entscheidungsmodells analysiert. Dabei werden im Folgenden die Fragestellungen, die sich aus dem Entscheidungsmodell ergeben, systematisch durchgegangen und beantwortet. Die Architekturentscheidungen werden dabei schrittweise entlang des Entscheidungsmodells hergeleitet und nachvollziehbar dargelegt. Grundlage dieser Analyse bilden die in Abschnitt 5.1 definierten Anforderungen und Rahmenbedingungen, die spezifisch für diesen ersten Anwendungsfall formuliert wurden.

Welcher Speicherort soll verwendet werden?

Gemäß der nicht-funktionale Anforderung NFA-10 sollen die Gesundheitsdaten an den ursprünglichen Speicherorten verbleiben und entsprechende Schnittstellen zu diesen Datenquellen geschaffen werden. Das Entscheidungsmodell leitet daraus die Nutzung einer Off-Chain-Speicherung ab, da die Daten innerhalb der IT-Systeme der medizinischen Leistungserbringenden verbleiben und somit kein externes System für die Datenspeicherung erforderlich ist, dem vertraut werden müsste.

Welcher Blockchain-Typ sollte eingesetzt werden?

Für das Identitätsmanagement sollte eine Autorität den Zugang zum System beschränken können, sodass ausschließlich vertrauenswürdige Identitätsanbietende Identitätsnachweise ausstellen und die dazugehörigen Daten auf der Blockchain speichern können. Im Einklang mit dem Schutzmechanismus SM33 aus der Bedrohungsmodellierung empfiehlt das Entscheidungsmodell die Verwendung einer permissioned Blockchain. Da jedoch keine übergeordnet vertrauenswürdige Instanz zur Kontrolle des Identitätsmanagementsystems und zur Verwaltung der Identitäten vorhanden ist, wäre eine öffentliche Blockchain geeignet.

Welcher Speicher eignet sich zur Off-Chain Speicherung? Ein geeigneter Speicher für die Off-Chain-Speicherung sind die bestehenden IT-Systeme der medizinischen Leistungserbringenden und demgemäß eine dezentrale Off-Chain Speicherung. Diese dezentrale Lösung vermeidet die Notwendigkeit eines externen Speichersystems, dem vertraut werden müsste.

Welches Identitätsmanagementsystem sollte verwendet werden? Grundsätzlich verbleibt die Hoheit über sämtliche Daten, einschließlich der Identitätsdaten, bei den Patient*innen. Dies entspricht dem Prinzip der Datensouveränität. Darüber hinaus wird angestrebt, dass die Verbindung zwischen den Identitäten der Patient*innen und den medizinischen Leistungserbringenden nicht nachvollziehbar ist, um die Privatsphäre der Patient*innen zu gewährleisten (PFA-1). Die persönliche Registrierung, wie im Sicherheitsmechanismus SM1 vorgesehen, bildet hierfür die Grundlage. Das Entscheidungsmodell schlägt demgemäß Self-Sovereign Identity (SSI) als Identitätsmanagementsystem vor. Verifiable Credentials ermöglichen es den Patient*innen, ihre Anonymität zu wahren und sich dennoch sicher zu registrieren, indem die Nachweise ausschließlich mit den relevanten Einrichtungen geteilt werden. Zusätzlich wird eine lokale Benutzenden-Authentifizierung angestrebt, um das lokale Credential-Wallet mittels eines Passworts zu sichern (SM5). Diese Authentifizierung soll unabhängig von Dienstanbietenden erfolgen, weshalb ein benutzerzentriertes Identitätsmanagement (Benutzerzentriertes IdM) als ergänzende Maßnahme vorgeschlagen wird.

Mit welchen Parteien sollen Gesundheitsdaten geteilt werden?

Das übergeordnete Ziel ist es, Gesundheitsdaten primär mit medizinischen Einrichtungen, den zugehörigen Gesundheitsfachkräften sowie den Zugehörigen der Patient*innen zu teilen.

Sollen die Daten für Zwecke des maschinellen Lernens benutzt werden?

Nein, die Gesundheitsdaten sollen nicht für Zwecke des maschinellen Lernens verwendet werden. Die primäre Zielsetzung liegt darin, die Daten ausschließlich für den unmittelbaren medizinischen Bedarf und die Versorgung der Patient*innen zu

nutzen. Eine Nutzung der Daten für maschinelles Lernen innerhalb der Einrichtungen selbst ist grundsätzlich denkbar, fällt jedoch außerhalb der Systemgrenzen des entwickelten Systems.

Wer hat die Autorität zur Verwaltung der Zugriffskontrollstrategie?

Die Autorität zur Verwaltung der Zugriffsrechte liegt bei den Patient*innen, also den Datensubjekten selbst. In Übereinstimmung mit dieser Zielsetzung empfiehlt das Entscheidungsmodell die Implementierung einer Datensubjekt-gesteuerten Zugriffskontrolle, bei der die Entscheidungsbefugnis vollständig bei den Datensubjekten verbleibt (in Tabelle 6.1 mit Autorität Allgemein bezeichnet). Diese vollständig souveräne Lösung berücksichtigt dabei, dass dies möglicherweise mit einer geringeren Unterstützung und potenziell reduzierter Sicherheit bei der Zugangsbeschränkung verbunden sein kann.

Welche Zugriffskontrollstrategie sollte verwendet werden?

Die für die Zugangsentscheidungen zuständige Entität, der/die Patient*in, strebt eine umfassende Kontrolle sowie eine feingranulare Zugriffskontrolle an. In Übereinstimmung mit diesen Anforderungen empfiehlt das Entscheidungsmodell die Verwendung der Zugriffskontrollstrategie der Discretionary Access Control (DAC). Diese Strategie ermöglicht es den Datensubjekten, den Zugriff auf ihre Daten flexibel und feingranular zu steuern, basierend auf individuellen Entscheidungen und Präferenzen.

Welche Sicherheitsmaßnahmen zur Zugriffskontrolle sollen ergriffen werden?

Für den Zugriff auf die Daten sollten zusätzliche Sicherheitsmaßnahmen ergriffen werden, um spezifischen medizinischen Einrichtungen und den dazugehörigen Leistungserbringenden einen einzelfallbasierten Zugriff auf die Gesundheitsdaten zu gewähren. Zur Gewährleistung einer sicheren Identifikation und Authentifizierung der Leistungserbringenden soll eine zusätzliche Identitätsverifizierung erfolgen. Daher wird die Kombination von Tokenisierung und digitalen Signaturen als geeignete Sicherheitsmechanismen durch das Entscheidungsmodell vorgeschlagen.

Welche zusätzlichen Sicherheitsmaßnahmen sind für die Datenspeicherung notwendig?

Zusätzliche Sicherheitsmaßnahmen für die Datenspeicherung innerhalb der internen Systeme der medizinischen Einrichtungen sollten durch die Anwendung von Datenverschlüsselung gemäß dem Sicherheitsmechanismus SM16 sichergestellt werden. Dies ist insbesondere notwendig, da das Risiko eines externen Hackerangriffs aufgrund der Sensibilität der Gesundheitsdaten und der Bedeutung der medizinischen Einrichtungen als kritische Infrastruktur als hoch eingestuft wird. Da sämtliche Arten von Gesundheitsdaten gespeichert werden und das Risiko einer Schlüsseloffenlegung aufgrund des sicheren Schlüsselmanagements innerhalb der medizinischen Einrichtungen sowie der ausschließlichen Nutzung der Schlüssel in diesen Einrichtungen als gering betrachtet wird, empfiehlt das Entscheidungsmodell die Verwendung symmetrischer Verschlüsselung zur Sicherung der Daten.

6.1.2 Anwendungsfall 2: Sekundärdatennutzung für die medizinische Forschung und Entwicklung

Im Anschluss an die Analyse des ersten Anwendungsfalls wird in diesem Abschnitt der zweite Anwendungsfall (Sekundärdatennutzung für die medizinische Forschung und Entwicklung) untersucht. Auch hier werden die Fragestellungen des Entscheidungsmodells systematisch analysiert, basierend auf den in Abschnitt 5.2 definierten Anforderungen und Rahmenbedingungen.

Welcher Speicherort soll verwendet werden?

Es wird angenommen, dass die Datenerzeugenden von Gesundheitsdaten bevorzugen, ihre Daten in ihren eigenen internen Speichersystemen zu behalten und sie bei Bedarf dem Datentreuhänder bereitzustellen. Diese Vorgehensweise entspricht der Anforderung DEFA-6, die vorsieht, Daten ausschließlich bedarfsorientiert und nicht auf Vorrat weiterzugeben, um den Datenschutz zu gewährleisten und das Prinzip der Datenminimierung einzuhalten. Entsprechend schlägt das

Entscheidungsmodell die Nutzung einer dezentralen Off-Chain Speicherung vor. Die pseudonymisierten Daten hingegen sollen dem externen Datentreuhandsystem bereitgestellt werden. Diese Daten weisen einen größeren Umfang als ihren zugehörigen Hash auf und deren Integrität wird durch das Vertrauen in die Datenerzeugenden gewährleistet. Dementsprechend ist der vorgeschlagene Speicherort ebenfalls Off-Chain. Im Hinblick auf die Logging-Daten, die zur Transparenz der Datennutzung beitragen sollen, wird eine Speicherung in einem externen System empfohlen. Da diese Daten kleiner als ihr zugehöriger Hash sind, wird eine On-Chain-Speicherung als geeignete Lösung durch das Entscheidungsmodell vorgeschlagen.

Welcher Blockchain-Typ sollte eingesetzt werden?

Im Hinblick auf das Identitätsmanagement ergeben sich für den Blockchain-Typ die gleichen Architekturentscheidungen wie im ersten Anwendungsfall (siehe Abschnitt 6.1.1), nämlich die Wahl einer öffentlichen, permissioned Blockchain. Da sensible Gesundheitsdaten ausgetauscht werden sollen, sollte der Zugang zum System eingeschränkt und entsprechend eine permissioned Blockchain genutzt werden. Um Vertrauen und Transparenz zu gewährleisten, insbesondere in Bezug auf die Datennutzung durch externe Reviewende über einen Logging-Mechanismus (RFA-1 bis RFA-4), sollte keine Entität zur Kontrolle des Blockchain-Netzwerks benötigt werden. Eine öffentliche Blockchain wird entsprechend vom Entscheidungsmodell vorgeschlagen.

Welcher Speicher eignet sich zur Off-Chain Speicherung?

Die Speicherung der Gesundheitsdaten bei den Gesundheitsdienstleistenden erfolgt dezentral, da diese Daten keinem externen System anvertraut werden. Für die pseudonymen Daten übernimmt die Datentreuhand die Rolle einer zentralen vermittelnden Instanz und stellt die einzige autorisierte Entität dar, der diese Daten zur kurzzeitigen Speicherung anvertraut werden. Entsprechend dem Entscheidungsmodell ist daher eine zentrale Off-Chain-Speicherung für die pseudonymen Daten vorgesehen. Da die Logging-Daten On-Chain gespeichert werden, ist eine Off-Chain-Speicherung für diese Art von Daten nicht erforderlich.

Welches Identitätsmanagementsystem sollte verwendet werden?

Um die Pseudonymisierung und damit die Nutzung von Längsschnittdaten gemäß DNFA-7 zu ermöglichen, ist eine sichere Verwaltung der Identität der Datensubjekte erforderlich, die auch die Möglichkeit einer anonymen Registrierung unter einem Pseudonym einschließt. Das Entscheidungsmodell schlägt daher die Verwendung von SSI als geeigneten Identitätsmanagementansatz vor. Im Gegensatz zum ersten Anwendungsfall ist in diesem Kontext keine zusätzliche lokale Benutzenden-Authentifizierung unabhängig vom Dienstanbietenden, hier der Datentreuhand, erforderlich. Stattdessen wird das Vertrauen in die Datentreuhand als zentrale Instanz vorausgesetzt. Diese verwaltet auch die digitale Wallet für Identitätsnachweise in ihren Systemen, was eine erhöhte Gebrauchstauglichkeit gewährleistet und den Fokus auf Benutzendenfreundlichkeit (NFA-1) gegenüber vollständiger Selbstsouveränität legt.

Mit welchen Parteien sollen Gesundheitsdaten geteilt werden?

Das Datentreuhandsystem verfolgt das Ziel, den Austausch von Gesundheitsdaten mit der Forschungsgemeinschaft zu ermöglichen, im Einklang mit den Anforderungen DGFA-2 bis DGFA-6.

Sollen die Daten für Zwecke des maschinellen Lernens benutzt werden?

Die pseudonymen Forschungsdaten, die innerhalb des Datentreuhandsystems verwaltet werden, sollen für maschinelles Lernen genutzt werden. Dementsprechend sollten laut Entscheidungsmodell die Daten in unverschlüsselter Form für Datennutzende im Treuhandsystem verfügbar gemacht werden.

Wer hat die Autorität zur Verwaltung der Zugriffskontrollstrategie?

Das Datensubjekt, als Datengebende, sollte vollständige Kontrolle haben und entsprechend durch eine Einwilligung die Zustimmung erteilen, dass die Daten für die Forschung verwendet werden können (DGFA-5 und DGFA-6). Eine zusätzliche Unterstützung oder Absicherung durch systemgesteuerte Zugangsbeschränkungen ist dabei nicht vorgesehen. Demgemäß wird eine Datensubjekt-gesteuerte Zugriffskontrolle vorgeschlagen (in Tabelle 6.1 mit Autorität F&E bezeichnet).

Welche Zugriffskontrollstrategie sollte verwendet werden?

Wie im Entscheidungsmodell vorgeschlagen, sollte für die Freigabe von Forschungsdaten eine patient*innengesteuerte Zugriffsverwaltung implementiert werden, um den Datengebenden umfassende Kontrolle darüber zu ermöglichen, mit welchen Akteuren ihre Daten geteilt werden. Zur Unterstützung dieser Governance-Struktur wird die Zugriffskontrolllogik des Discretionary Access Control (DAC) empfohlen, da sie eine feingranulare Vergabe von Zugriffsrechten und eine weitreichende Kontrolle durch die Datengebenden gewährleistet.

Welche Sicherheitsmaßnahmen zur Zugriffskontrolle sollen ergriffen werden?

Für den Zugriff auf die Daten sollten zusätzliche Sicherheitsmaßnahmen ergriffen werden, um spezifischen Datennutzenden einen Zugriff zu den Daten zu gewähren. Die Zugriffsvergabe soll jedoch nicht auf Einzelfallbasis erfolgen. Zur Gewährleistung einer sicheren Identifikation und Authentifizierung der Akteure des Datentreuhandsystems (Datenerzeugende, Datengebende und Datennutzende) soll eine zusätzliche Identitätsverifizierung erfolgen. Das Entscheidungsmodell schlägt daher die Kombination aus Tokenisierung und digitalen Signaturen als geeignete Sicherheitsmechanismen vor.

Welche zusätzlichen Sicherheitsmaßnahmen sind für die Datenspeicherung notwendig?

Da die personenbezogenen Gesundheitsdaten außerhalb des Datentreuhandsystems gespeichert werden, obliegt deren Verwaltung und Sicherheit vollständig den Datenerzeugenden. Im Gegensatz dazu liegt der Schutz des Speicherorts für die pseudonymisierten Daten in der Zuständigkeit des Datentreuhandsystems. Obwohl die pseudonymisierten Daten keinen direkten Personenbezug mehr aufweisen, besteht dennoch das Risiko eines externen Hackerangriffs, da diese Forschungsdaten aufgrund ihres potenziellen Werts Ziel solcher Angriffe sein könnten. Das Risiko einer Offenlegung der Verschlüsselungsschlüssel wird jedoch als gering eingestuft, da dem Datentreuhandsystem ein hohes Maß an Vertrauen entgegengebracht wird. Daher wird eine symmetrische Verschlüsselung der

gespeicherten Daten durch das Datentreuhandsystem als geeignete Sicherheitsmaßnahme vorgeschlagen. Für die Nutzung der Daten und deren Verarbeitung im Rahmen von maschinellem Lernen ist es erforderlich, dass die Daten durch das Datentreuhandsystem entschlüsselt werden.

Tabelle 6.1: Zusammenfassung der durch das Entscheidungsmodell vorgeschlagenen Architekturentscheidungen für die beiden Anwendungsfälle.

	Datenmanagement													**Identitätsmanagement**				
	Blockchain-Typ					**Speicherort**			**Off-Chain Speicher**			**Blockchain Feature**						
Anwendungsfall	Öffentliche	Private	Konsortium	Permissioned	Permissionless	On-Chain	Off-Chain	Hybrid	Dezentral	Zentral	Verteilt	Mehrere private BC	Anbindung an beliebte BC	Zentrale IdM	Föderiertes IdM	DTI	SSI	Benutzerzentrierte IdM
1	✓			✓			✓		✓								✓	✓
2	✓			✓		✓	✓		✓	✓							✓	

	Zugriffsmanagement													**Sicherheitsmechanismen**								
	Autorität Allgemein			**Autorität F&E**			**Strategie**			**NDAC Logik**				**Zugriffskontrolle**					**Speicherung**			
Anwendungsfall	System	Datensubjekt	Geteilt	System	Datensubjekt	Geteilt	DAC	Hybrid	NDAC	MAC	Rollenbasiert	Regelbasiert		Asym. Verschlüsselung	Sym. Verschlüsselung	Proxy Re-Encryption	Tokenisierung	CP-ABE	Sym. Verschlüsselung	Hybride Verschlüsselung	Asym. Verschlüsselung	Passwortschutz
1		✓					✓										✓		✓			
2					✓		✓										✓		✓			

6.1.3 Fazit

In Tabelle 6.1 werden die Architekturentscheidungen für die beiden zuvor beschriebenen Anwendungsfälle, basierend auf dem Entscheidungsmodell, zusammengefasst. Diese Entscheidungen entsprechen denen der herkömmlichen Systemkonzeption und können zu denselben Systemarchitekturen führen, wie sie in den Abschnitten 5.1.5 und 5.2.5 dargelegt sind.

6.2 Expert*innenevaluation

Zur Validierung des entwickelten Entscheidungsmodells wurde ein zweistündiger Online-Workshop mit einer Expert*innengruppe durchgeführt, die aus sieben Expert*innen aus dem Bereich der Informatik und der Informationssicherheit bestand. Vier der sieben Expert*innen verfügen über Erfahrungen mit der Realisierung von Blockchain-Projekten, sowohl aus konzeptioneller als auch technologischer Perspektive. Die Expertise der verbleibenden Expert*innen ist überwiegend im Bereich der Datentreuhandsysteme im Gesundheitswesen sowie in den damit verbundenen Aspekten der De-Identifikation und der Datenqualität für die Sekundärnutzung verortet. Diese ausgewogene Zusammensetzung der Expert*innen aus unterschiedlichen Themengebieten ermöglicht eine umfassende und interdisziplinäre Perspektive auf die Anwendung des Entscheidungsmodells im Gesundheitswesen sowie auf die technische Bewertung des Entscheidungsmodells im Hinblick auf seine zweckmäßige Verwendung für die Konzeption Blockchain-basierter Systeme.

Der Expert*innenworkshop war wie folgt strukturiert: Zunächst erfolgte die Präsentation des Entscheidungsmodells, gefolgt von einer Demonstration seiner Anwendung anhand des ersten Anwendungsfalls. Anschließend hatten die Expert*innen die Möglichkeit, Fragen zu stellen und Anmerkungen zum Entscheidungsmodell in einer offenen Diskussion zu äußern. Die im Anschluss durch die

Workshopleitenden abgefragten Kriterien zur Bewertung des Entscheidungsmodells waren Vollständigkeit, Konsistenz und Aktualität sowie die logische Nachvollziehbarkeit und Plausibilität des Entscheidungsmodells im Hinblick auf seine Anwendung im Kontext Blockchain-basierter Systeme im Gesundheitswesen. Diese Dimensionen wurden gewählt, um sicherzustellen, dass das Entscheidungsmodell alle relevanten Faktoren abdeckt, klare und logische Entscheidungsstrukturen bietet und die neuesten Entwicklungen in der Blockchain-Technologie sowie im Gesundheitssektor berücksichtigt. Konkrete Metriken oder Leitlinien zur Bewertung wurden den Expert*innen jedoch nicht bereitgestellt. Die Beurteilung des Entscheidungsmodells erfolgte ausschließlich auf der Grundlage einer qualitativen Bewertung durch die Expert*innenmeinungen. Darüber hinaus wurden die Designentscheidungen für das deutsche Gesundheitssystem diskutiert.

Das erhaltene Feedback umfasste mehrere spezifische Vorschläge zur Verbesserung und Erweiterung des Entscheidungsmodells:

1. **Positionierung des Datenspeicherorts:** Es wurde angemerkt, dass die Entscheidung über den Datenspeicherort vor der Auswahl des Blockchain-Typs getroffen werden sollte. Entsprechend wurde vorgeschlagen, dieses Teilmodell an den Anfang des Entscheidungsmodells zu setzen.

2. **Sprachliche Bereitstellung:** Das Entscheidungsmodell wurde in englischer Sprache präsentiert. Die Expert*innen regten an, eine deutschsprachige Version bereitzustellen, insbesondere für Anwendungen im deutschen Gesundheitssystem, obwohl davon auszugehen ist, dass Softwareentwickler*innen in der Regel über ausreichende Englischkenntnisse verfügen.

3. **Designentscheidung bei dezentraler Speicherung:** Es wurde darauf hingewiesen, dass bei der Entscheidung für eine dezentrale Speicherung klargestellt werden sollte, dass dies in der Regel eine Off-Chain-Speicherung impliziert.

4. **Vertrauen in die Datenintegrität:** Ein weiterer Aspekt war die Frage, ob der Integrität der Daten vertraut wird. Während staatliche Gesundheitseinrichtungen in Deutschland grundsätzlich als vertrauenswürdig angesehen

werden, könnte es bei Daten von kommerziellen Gesundheitsanwendungen zu einem geringeren Vertrauen kommen. Ein vergleichbares Szenario könnte in Ländern bestehen, in denen das Vertrauen in staatliche (Gesundheits-) Einrichtungen insgesamt weniger stark ausgeprägt ist. Dieser Punkt sollte bei der Entscheidungsstruktur berücksichtigt werden.

5. **Frage zum Datenspeicherort:** Die Formulierung der Entscheidungsfrage *Vertraut das Krankenhaus seine Daten einem externen System an?* (*engl. Does the hospital entrust their data to an external system?)* wurde von einem Experten als zu eng gefasst bewertet. Es wurde vorgeschlagen, die Frage allgemeiner zu formulieren, z.B. *Sollen die Daten einem externen System anvertraut werden?*, da nicht nur Kliniken, sondern auch Datensubjekte (z. B. Patient*innen) sowie weitere Stakeholder wie Universitätskliniken mit Urheberrechten darüber entscheiden könnten.

6. **Umgang mit verschiedenen Datenarten bei der Datenspeicherung:** Die Expert*innen wiesen darauf hin, dass verschiedene Arten von Daten unterschiedliche Designentscheidungen erfordern könnten. Daher sollte das Entscheidungsmodell eine Schleife enthalten, um solche unterschiedlichen Speicherentscheidungen abzubilden. Die Expert*innen empfahlen, entweder eine Schleife im Entscheidungsmodell aufzunehmen oder das Entscheidungsmodell für jeden Datentyp separat durchzugehen. Nach Abwägung der Vor- und Nachteile entschied man sich für die zweite Option, da eine Schleife potenzielle Komplexitäts- und Konsistenzprobleme aufwerfen könnte.

7. **Permissioned vs. Permissionless Blockchains:** Neben der Unterscheidung zwischen privater, öffentlicher und Konsortium-Blockchain sollte auch die Entscheidung berücksichtigt werden, ob es sich um eine *permissioned* oder *permissionless* Blockchain handelt. Dies betrifft insbesondere die Frage, wer Schreibrechte besitzt und ob eine zentrale Instanz zur Vergabe der Rechte vorhanden sein sollte.

8. **Welche Entscheidungen sind zu berücksichtigen, wenn Gesundheitsdaten mit anderen als Forschungseinrichtungen, Gesundheitsfachkräften oder Zugehörigen geteilt werden sollen?** Die Expert*innen empfahlen, das

Teilmodell zur allgemeinen Datennutzung so zu gestalten, dass er ebenfalls auf weitere potenzielle Empfänger von Gesundheitsdaten anwendbar ist.

9. **Verschlüsselung und Analyseverfahren:** Da entsprechend der Expert*innen-Meinung privatsphärewahrende Analyseverfahren auf verschlüsselten Daten derzeit in ihrer Anwendbarkeit und Skalierbarkeit begrenzt sind, wurde vorgeschlagen, die Daten für die Nutzung unverschlüsselt bereitzustellen, während sie bei der Speicherung verschlüsselt abgelegt werden können.

10. **Teilprozesse für Sicherheits- und Speichermechanismen:** Es wurde empfohlen, das Entscheidungsmodell um zwei spezifische Teilprozesse zu erweitern: einen für Sicherheitsmechanismen im Zugriffsmanagement und einen für die Sicherheitsmechanismen bei der Datenspeicherung.

Das erhaltene Feedback wurde anschließend verwendet, um das Entscheidungsmodell iterativ anzupassen und weiter zu verfeinern (siehe Kapitel 4 für das finale Entscheidungsmodell).

7 Methodische Einbettung des Entscheidungsmodells

Übergeordnet soll das Entscheidungsmodell in den Softwarearchitektur-Prozess von Toth [143] eingebettet werden, da dieser Prozess eine strukturierte und iterative Herangehensweise für die Entwicklung von Softwarearchitekturen bietet. Im Rahmen dieses Prozesses bilden der Architekturzyklus (Analyse, Architektur, Reflexion) und der Implementierungszyklus (Analyse, Umsetzung, Review, Auslieferung) den Kern der Architekturbrezel, die mit der Anforderungserhebung initiiert wird. Da das Entscheidungsmodell bei Architekturentscheidungen unterstützen soll, wird dessen Anwendung als Teil des Architektur-Schrittes im iterativen Architekturzyklus verortet. Der Anwendung des Entscheidungsmodells geht stets die Anwendungsfallbeschreibung und Anforderungserhebung voraus, da diese Rahmenbedingungen für die fundierte Ableitung von Designentscheidungen unverzichtbar sind.

Wie bereits eingehend erläutert, eigenen sich Blockchains nicht für alle Anwendungsfälle. Aus diesem Grund ist eine vorgelagerte Evaluation erforderlich, um zu prüfen, ob der Einsatz einer Blockchain in Bezug auf den spezifischen Anwendungsfall und dessen Anforderungen geeignet ist. Hierfür existieren, wie in Abschnitt 3.7 dargestellt, verschiedene Entscheidungsmodelle, die zu diesem Zweck entwickelt wurden. Zu dieser Art von Modellen zählen unter anderem jenes von Wüst und Gervais [210], Pahl, Ioini und Helmer [223], Koens und Poll [224], Betzwieser et al. [149] sowie Lo, Chiam und Lu [226]. Demgemäß beginnt der Architekturprozess mit der Überprüfung der Eignung der Blockchain.

Im Designprozess von Betzwieser et al. [149] sowie in der Entwicklung der Softwarearchitekturen im Rahmen dieser Arbeit betrachteten Anwendungsfälle (siehe Kapitel 5) gehen der Architekturarbeit technische Betrachtungen voraus, insbesondere in Bezug auf Sicherheit und Datenschutz. In den Anwendungsfällen wurden Sicherheitsanalysen und Angreifendenmodellierungen unter Anwendung der STRIDE-Methode [168] durchgeführt. Die STRIDE-Methode umfasst eine Reihe von Schritten, die sich auf die folgenden drei grundlegenden Teilaufgaben unterteilt:

(1) Erstellung einer Systemabstraktion durch die Analyse von schützenswerten Gütern, beteiligten Entitäten sowie Vertrauensgrenzen und -beziehungen;

(2) Identifikation von Bedrohungen und deren Auswirkungen, basierend auf den STRIDE-Kategorien (Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service und Elevation of Privilege);

(3) Entwicklung von Gegenmaßnahmen zur Bewältigung der identifizierten Bedrohungen [168, 169].

In Anbetracht der Tatsache, dass die in der ersten Teilaufgabe vorgesehene Systemabstraktion und Analyse der Vertrauensbeziehungen und Entitäten, die Architekturentscheidungen beeinflussen können, sollte dieser Schritt vor der Anwendung des Entscheidungsmodells erfolgen. Die zweite und dritte Teilaufgabe können anschließend optional nach der Anwendung des Entscheidungsmodells zur detaillierten Bedrohungsmodellierung durchgeführt werden. Ergänzend können weitere Design-Aktivitäten zur Verfeinerung des Konzepts durchgeführt werden, die nicht durch das Entscheidungsmodell abgebildet werden oder keine schwer änderbaren Architekturentscheidungen betreffen. Abschließend kann optional eine Reflexion der getroffenen Entwurfsentscheidungen gemäß Toth [143] erfolgen oder es kann ein weiterer Architekturzyklus bzw. Implementierungszyklus basierend auf einer weiterführenden Analyse initiiert werden.

Die zuvor beschriebene Einbettung des Entscheidungsmodells sowie die damit zusammenhängenden Prozessschritte werden in Abbildung 7.1 veranschaulicht.

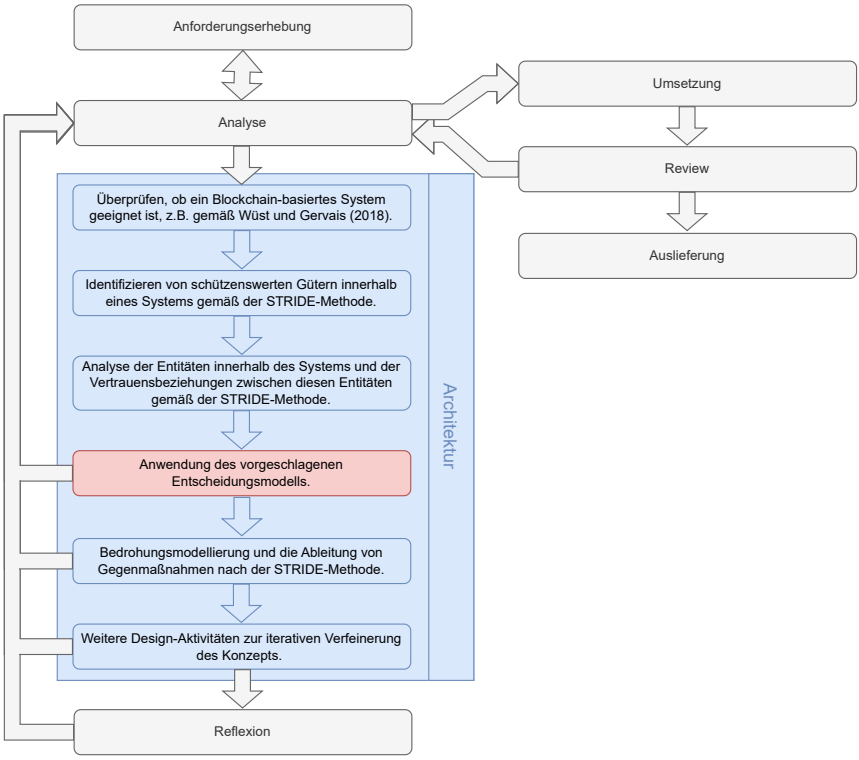

Abbildung 7.1: Methodische Einbettung des Entscheidungsmodells in den Architekturprozess von Toth [143] zur Unterstützung fundierter Architekturentscheidungen, z.B. gemäß Wüst und Gervais [210] sowie STRIDE-Methode [168].

8 Zusammenfassung und Ausblick

Das nachfolgende Kapitel bietet einen Abriss des Verlaufs und der zentralen Ergebnisse sowie eine Erläuterung der Adressierung der formulierten Forschungsfragen. Ferner wird ein Ausblick auf potentielle offene Forschungsfelder gegeben, welche aus den Erkenntnissen dieser Arbeit resultieren.

8.1 Zusammenfassung und Fazit

Zusammenfassend lässt sich feststellen, dass die Konzeption und Architektur von dezentralen Infrastrukturen auf Basis der Blockchain-Technologie für das Gesundheitswesen eine komplexe Herausforderung darstellt. Dies ist darauf zurückzuführen, dass eine Vielzahl an Designentscheidungen getroffen werden müssen, um den diversen Bedürfnissen und Anforderungen der Akteure im Gesundheitswesen gerecht zu werden und gleichzeitig eine allgemein akzeptierte Lösung zu entwickeln. Durch die fortlaufende Digitalisierung und verschiedene Initiativen, wie auch die in den letzten Jahren verabschiedeten regulatorischen Rahmenbedingungen, handelt es sich zudem um ein hochgradig dynamisches Feld. Die sich ständig weiterentwickelnden Anforderungen und Systeme machen es für Softwarearchitekten besonders herausfordernd adäquate und integrierte Lösungen zu entwickeln. Vor allem bei Blockchain-basierten Ansätzen ist es sorgfältig zu prüfen, ob der Einsatz der Technologie tatsächlich zielführend ist. Eine Blockchain sollte als ergänzendes Werkzeug betrachtet werden, welches insbesondere einen Mehrwert bietet, wenn ihre Eigenschaften der Dezentralisierung, Transparenz und Unveränderlichkeit von zentraler Bedeutung sind.

Aus diesem Grund ist eine Entscheidungsunterstützung erforderlich, die die Auswahl eines geeigneten Ansatzes für die Speicherung, Zugriffsverwaltung, Übertragung und Nutzung sensibler Gesundheitsdaten basierend auf verifizierten Identitäten unter Berücksichtigung der spezifischen Anforderungen im Gesundheitswesen erleichtert. Ein geeignetes Instrumentarium in der Softwarearchitektur stellen Architekturmuster in Kombination mit Entscheidungsmodellen dar. Diese haben sich in der Wissenschaft bei der Entwicklung Blockchain-basierter Systeme bereits bewährt, wurden jedoch bislang nicht speziell auf die Anforderungen, Vertrauensannahmen und Eigenschaften des Gesundheitswesens ausgerichtet. Im Rahmen dieser Arbeit wurde deshalb der Entwicklung eines solchen Entscheidungsmodells nachgegangen.

Durch eine strukturierte Literaturrecherche werden Ansätze und technische Implementierungen für Blockchain-basierte Gesundheitsdateninfrastrukturen identifiziert und in einen iterativen Taxonomie-Entwicklungsprozess eingebettet. Die so entwickelte Taxonomie fasst die technischen Charakteristika und Muster des Stands der Technik und Wissenschaft zusammen und zeigt die Punkte auf, an denen essentielle Designentscheidungen getroffen werden müssen. Diese Entscheidungen betreffen unter anderem den Speicherort und den Blockchain-Typ, die Auswahl geeigneter Off-Chain-Speicherlösungen, die Gestaltung des Identitätsmanagements, die Definition der Parteien, mit denen Gesundheitsdaten geteilt werden, sowie die Nutzung der Daten für maschinelles Lernen. Darüber hinaus müssen Strategien zur Zugriffskontrolle und Sicherheitsmaßnahmen sowohl für die Datenspeicherung als auch für den Zugriff definiert werden, einschließlich der Festlegung, welche Instanzen die Autorität über die Zugriffsverwaltung besitzen. Damit wird die Teilforschungsfrage **TFF1** umfassend adressiert, indem die notwendigen Designentscheidungen für die Architektur von dezentralen Dateninfrastrukturen im Gesundheitswesen unter Berücksichtigung der Blockchain-Technologie in der Taxonomie aufgezeigt werden. Insbesondere im Gesundheitswesen werden Off-Chain-Ansätze als geeignete Methoden für das Datenmanagement identifiziert. Allgemein sollte abgewogen werden, dass lediglich Daten kleiner als ihr Hash On-Chain gespeichert werden, während umfangreichere oder

sensible Gesundheitsdaten extern gespeichert werden sollten. Dadurch wird gewährleistet, dass nur Daten, deren unveränderliche und transparente Speicherung keine kritischen Implikationen hinsichtlich Datensicherheit und den Datenschutz mit sich bringt, direkt auf der Blockchain abgelegt werden. Jene Abwägungen aus der Taxonomie dienen anschließend als Hilfsmittel für die Entwicklung des Entscheidungsmodells.

Zur Erfassung und Analyse der spezifischen Rahmenbedingungen im Gesundheitswesen werden zwei repräsentative Anwendungsfälle gewählt: (1) das Patient*innen-zentrierte Gesundheitsdatenmanagement in der medizinischen Versorgung und (2) die Sekundärdatennutzung für die medizinische Forschung und Entwicklung. Im Zuge dieser Analyse werden die relevanten Rahmenbedingungen für diese Anwendungsfälle systematisch hergeleitet und beschrieben. Dieser Prozess folgt der klassischen Methodik der Softwarearchitektur und umfasst die systematische Definition des jeweiligen Systemkontexts, der spezifischen Anforderungen, regulatorischen Rahmenbedingungen sowie der beteiligten Akteure. Zusätzlich wird für jeden Anwendungsfall eine umfassende Sicherheitsanalyse durchgeführt, wobei die STRIDE-Methode zur Identifikation potenzieller Bedrohungen und zur Ableitung von Sicherheitsmechanismen Anwendung findet. Die daraus resultierenden Sicherheitsmaßnahmen fließen anschließend in die Entwicklung der Systemarchitekturen ein, die den spezifischen Anforderungen der Anwendungsfälle gerecht werden. Mit dieser Herangehensweise wird die Teilforschungsfrage **TFF2** beantwortet, da die (Sicherheits-)Anforderungen, regulatorischen Vorgaben und technischen Rahmenbedingungen dezentraler Dateninfrastrukturen im Gesundheitswesen für die beiden exemplarischen Anwendungsfälle analysiert und beschrieben werden.

Überdies werden die Systemarchitekturen hinsichtlich deren Skalierbarkeit analysiert und deren Anwendbarkeit im deutschen Gesundheitswesen mit Themenexpert*innen sowie potentiellen Nutzenden diskutiert. Für beide Anwendungsfälle bzw. die entsprechenden Systemarchitekturen wird der Mehrwert insbesondere für Personen mit einem ausgeprägten Sicherheitsbedürfnis gesehen, die ihre Gesundheitsdaten selbstsouverän verwalten möchten. Isolierte Einzellösungen werden als

weniger zielführend eingestuft. Dementsprechend weist das Szenario, Teilkomponenten der Systemarchitekturen, wie beispielsweise das digitale Identitätsmanagement, nahtlos in bestehende nationale oder internationale Dateninfrastrukturinitiativen (z. B. ePA/TI, EHDS, FDZ-Gesundheit, FDPG) zu integrieren, ein perspektivisches Potential zur Förderung der Digitalisierung und Vernetzung des deutschen Gesundheitswesens auf.

Nach der Evaluation des Entscheidungsmodells durch Expert*innen wurden die Rückmeldungen in einer abschließenden Iteration berücksichtigt, um das in dieser Arbeit vorgestellte endgültige Entscheidungsmodell zu entwickeln. Im Anschluss erfolgt eine Evaluierung der Anwendbarkeit des Modells, indem es anhand der Anforderungen der Anwendungsfälle überprüft wurde. Dabei wurde geprüft, ob die Ergebnisse der Designentscheidungen mit denen der Softwarearchitektur, die mittels der STRIDE-Methode und dem Softwarearchitekturprozess nach Toth [143] erarbeitet wurden, übereinstimmen. Die abschließende Evaluation des Entscheidungsmodells bestätigte dessen Anwendbarkeit, indem es ähnliche Designentscheidungen wie bei den entwickelten Architekturen in den Anwendungsfällen vorschlug.

Basierend auf den Feststellungen der Konzeption der Systemarchitekturen in den Anwendungsfällen wird das entwickelte Entscheidungsmodell zusammen mit der STRIDE-Methode zur Sicherheitsbetrachtung in den Softwarearchitekturprozess von Toth [143] eingebettet. Dies dient dazu, aufzuzeigen, inwiefern diese sich in der vorliegenden Arbeit als sinnvoll erwiesenen methodischen Ansätze zukünftig gezielt im Architekturentwicklungsprozess angewendet werden können. Auf diese Weise wird die Teilforschungsfrage **TFF3** beantwortet.

Insgesamt leistet die vorliegende Arbeit durch die Beantwortung der Teilfragen einen wesentlichen Beitrag zur Beantwortung der übergeordneten Forschungsfrage, indem sie aufzeigt, wie methodische Ansätze und Entscheidungsmodelle die Konzeption und Architektur von Blockchain-basierten Dateninfrastrukturen im Gesundheitswesen unterstützen können. Im Zentrum stehen die Stärkung der Datensouveränität und die Gestaltung praxisorientierter Architekturen für die nachhaltige und sichere Nutzung von Gesundheitsdaten.

8.2 Ausblick

Im Rahmen dieser Arbeit wurde ein Entscheidungsmodell erforscht, das die Architekturentwicklung dezentraler und Blockchain-basierter Dateninfrastrukturen im Gesundheitswesen unterstützt. Die zugrundeliegende Literaturrecherche sowie die iterative Entwicklung einer Taxonomie dienten dabei als wesentliche methodische Grundlage. Das daraus abgeleitete Entscheidungsmodell reflektiert dementsprechend den aktuellen Stand der Technik und Wissenschaft im Kontext der definierten Suchstrategie und der untersuchten Literatur. Da die Blockchain-Technologie und die Digitalisierung im Gesundheitswesen einem dynamischen und schnell wachsenden Umfeld unterliegen, besteht ein kontinuierlicher Bedarf an der Weiterentwicklung des Entscheidungsmodells. Zukünftige Forschungsarbeiten könnten darauf abzielen, das Entscheidungsmodell durch die systematische Einbeziehung zusätzlicher Expert*innenmeinungen sowie die Berücksichtigung neuer technologischer Entwicklungen und innovativer Ansätze auf diesem Gebiet zu aktualisieren und zu erweitern, um dessen Relevanz und Anwendbarkeit kontinuierlich sicherzustellen. Die Anwendbarkeit des Entscheidungsmodells wurde im Rahmen von zwei exemplarischen Anwendungsfällen demonstriert. Die Erweiterung und Erprobung des Entscheidungsmodells in weiteren Anwendungsfällen im Gesundheitswesen könnte dessen Generalisierbarkeit erhöhen. Ferner könnte die Adaption und Weiterentwicklung des Modells in anderen Domänen ebenfalls einen Mehrwert bieten.

Ein weiterer Fokus zukünftiger Arbeiten sollte auf der praktischen Implementierung, Testung und Skalierung der vorgeschlagenen Architekturen liegen, um deren Anwendung über Forschungsprototypen hinaus in realen Versorgungsszenarien zu beweisen und zu evaluieren. Besonders in aktuellen und geplanten Digitalisierungs- und Interoperabilitätsvorhaben in Deutschland und Europa bietet sich die Gelegenheit, die methodischen Ansätze der Softwarearchitekturentwicklung sowie Teilkomponenten der Architekturen weiter zu erproben. Derzeit gibt es wenige Blockchain-basierte Systeme im Gesundheitswesen, welche über

Forschungsprototypen hinausgehen [238]. Die praktische Demonstration und Validierung solcher Systeme könnten dazu beitragen, die Anwendbarkeit der Technologie im Gesundheitswesen weiter voranzutreiben.

Der Trend in den aktuellen politischen Bestrebungen liegt in der Förderung von Opt-Out-Lösungen, bei denen Daten standardmäßig geteilt werden, sofern keine explizite Ablehnung erfolgt. Dieser Ansatz kann wesentlich zur Erhöhung der Datenverfügbarkeit beitragen und damit die Grundlage für eine verbesserte Nutzung von Gesundheitsdaten schaffen. Gleichzeitig ist es entscheidend, dass auch Personen mit einem erhöhten Sicherheits- und Souveränitätsbedürfnis angemessen berücksichtigt werden. Eine Kombination aus selbstsouveränen Ansätzen und den vorgesehenen Opt-Out-Lösungen könnte dabei den Weg für nachhaltige, patient*innenzentrierte Dateninfrastrukturen ebnen und zu einer höheren Datenverfügbarkeit führen.

A Anhang

A.1 Muster für Blockchain-basierte Anwendungen

Tabelle A.1: Übersicht der in den Entscheidungsmodellen von Xu et al. [150] verwendeten Muster für Blockchain-basierte Anwendungen und deren Bescheibung (übersetzt aus Xu et al. [150]).

Kategorie	**Muster**	**Beschreibung**
On-Chain Daten-management	Verschlüsselung von On-Chain Daten	Sicherstellung der Vertraulichkeit der auf der Blockchain gespeicherten Daten durch Verschlüsselung.
	Genesis-Block festlegen	Setzen des States des Genesis-Blocks der Blockchain (z.B. Verteilung der nativen Token).
	Juristisches und Smart-Contract Vertragspaar	Kombiniere eine rechtliche Vereinbarung und den entsprechenden Smart Contract, der die rechtliche Vereinbarung codiert.
	On-Chain Daten-speicherung	Speicherung der Rohdaten direkt auf der Blockchain.
	Tokenisierung	Verwendung von Token auf der Blockchain zur Darstellung übertragbarer digitaler oder physischer Assets oder Services.

Kategorie	Muster	Beschreibung
	Token-Burning	Dauerhafte Entfernung von Smart Contracts oder States von der Blockchain.
Zugriffs-kontrolle für Off-Chain Daten	Zeitlich begrenzter Zugriff	Teilen eines Links, welcher in einem definierten Zeitfenster auf die Daten verweist.
	Selektive Zugriffssteuerung	Individualisieren des Zugriffs auf Daten gemäß der von Dateneigentümer*innen festgelegten Anforderungen.
	Einmaliger Zugriff	Teile einen Link, der nur einmalig auf den Inhalt weiterleitet.
Performanz	Hashing von On-Chain Daten	Speicherung des Hashs eines beliebig großen Datensatzes (der möglicherweise nicht in eine Blockchain-Transaktion passt) auf der Blockchain.
	Speichern des Wurzelhashes eines Merkle-Baums	Speichere einzelne Teile der Daten Off-Chain und erzeuge daraus einen Merkle-Baum. Speichere anschließend den Wurzelhash des Merkle-Baums On-Chain.
	State-Aggregation	Aggregiere eine Reihe von Zuständen in einen einzelnen (oder wenige) Zustand(e).
	State Channel	Mikrozahlungen erfolgen zwischen zwei Parteien über einen Off-Chain Zahlungskanal. Nur die Transaktionen zum Öffnen des Kanals und die endgültige Abrechnung (z. B. nach mehreren Zahlungen) werden in der Blockchain gespeichert.

Kategorie	Muster	Beschreibung
Authentifizierung	Identifier-Register	Verwalte die Zuordnungen zwischen einem Identifier und den entsprechenden Identitätsattributen mithilfe eines Registers.
	Mehrfache Registrierung	Verwende für jede Transaktionsbeziehung einen seperaten Identifier.
	Aktualisierung durch eine delegierte Person	Bestimmung einer vertrauenswürdigen Gruppe von Personen oder Entitäten, die im Falle eines Identitätsverlustes die Wiederherstellung dieser Identität unterstützen können.
Autorisierung	Eingebettete Berechtigung	Beschränke die Ausführung einzelner Funktionen im Smart Contract auf eine berechtigte Gruppe von Konten.
	Mehrfache Autorisierung	Erlaube einer Teilmenge vordefinierter Blockchain-Adressen, Transaktionen zu signieren.
	Off-Chain-Geheimnis-unterstützte dynamische Autorisierung	Verwende ein Off-Chain erstelltes Geheimnis, um die Autorität für eine Transaktion dynamisch zu binden.
	Einzelne Autorisierung	Verwende eine vordefinierte Blockchain-Adresse, um Transaktionen zu signieren.
Interaktion mit der Außenwelt	Zentralisierter Oracle	Einführung des Zustands externer Systeme in die geschlossene Blockchain-Ausführungsumgebung.
	Dezentralisierter Oracle	Einführung des Zustands externer Systeme in die Blockchain durch mehrere unabhängige Oracles.

Kategorie	Muster	Beschreibung
	Historischer Oracle	Bereitstellung historischer Zustände zusätzlich zum neuesten Zustand.
	Pull-basierter Eingangs-Oracle	Ermöglicht es der On-Chain-Komponente, den Off-Chain-Zustand von einer Off-Chain-Komponente anzufordern.
	Pull-basierter Ausgangs-Oracle	Ermöglicht es der Off-Chain-Komponente, den On-Chain-Zustand von einer On-Chain-Komponente abzurufen.
	Push-basierter Eingangs-Oracle	Ermöglicht es der Off-Chain-Komponente, den Off-Chain-Zustand an eine On-Chain-Komponente zu senden.
	Push-basierter Ausgangs-Oracle	Ermöglicht es der Off-Chain-Komponente, den On-Chain-Zustand von einer On-Chain-Komponente zu polen.
	Voting	Ermöglicht es einer Gruppe von Blockchain-Nutzenden oder Oracles, eine kollektive Entscheidung zu treffen.
Smart Contract	Contract Registry	Verwendung eines Registers, um die Zuordnung zwischen dem Vertragskennzeichen, der Version und seiner Adresse zu speichern.
	Datenvertrag	Speichern von Daten in einem separaten Smart Contract.
	Eingebettet in andere Funktionen	Einbetten zusätzlicher Funktionen in reguläre Funktionen, die aufgerufen werden müssen, um die Dienste zu nutzen.

Kategorie	Muster	Beschreibung
	Factory Contract	Verwenden eines On-Chain-Template-Vertrags als Fabrik, um Vertragsinstanzen aus der Vorlage zu generieren.
	Incentive Execution	Bieten Sie dem Aufrufenden einer Vertragsfunktion eine Belohnung für deren Ausführung.
	Proxy Contract	Verwendung eines Proxy-Smart-Contract, um Transaktionen an die neueste Version der Vertragslogik weiterzuleiten.

A.2 Visualisierung der Prototypen für die beiden Anwendungsfälle

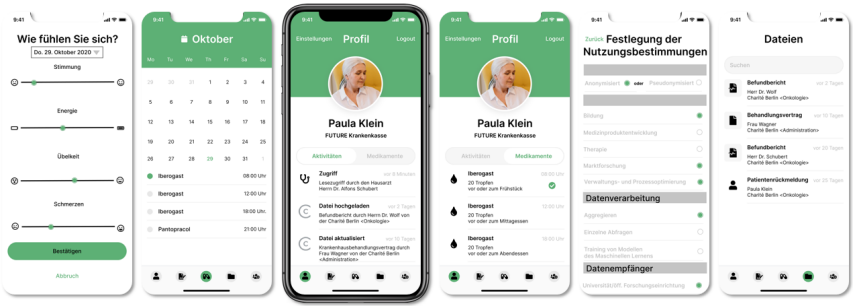

Abbildung A.1: Designentwürfe des mobilen Prototyps für Patient*innen im ersten Anwendungsfall [16].

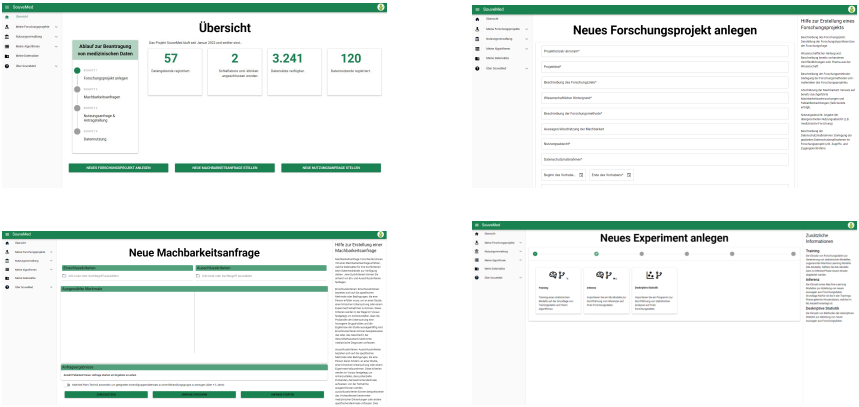

Abbildung A.2: Designs des webbasierten Prototypen für den zweiten Anwendungsfall aus Sicht der Datennutzenden [19].

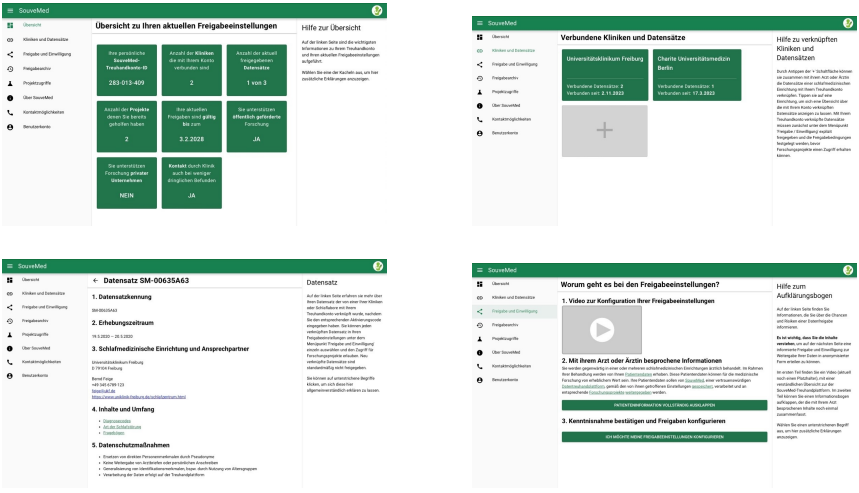

Abbildung A.3: Designs des webbasierten Prototypen für den zweiten Anwendungsfall aus Sicht der Datengebenden [19].

A.3 Taxonomien aus den Iterationen der Literaturrecherche

Tabelle A.2: Taxonomie bzgl. der Merkmale des Daten- und Identitätsmanagement, resultierend aus der ersten Iteration der Literaturrecherche.

		Datenmanagement														**Identitätsmanagement**				
		Blockchain-Typ					**Speicherort**			**Off-Chain Speicher**			**Blockchain Feature**							
It.	Referenz	Öffentliche	Private	Konsortium	Permissioned	Permissionless	On-Chain	Off-Chain	Hybrid	Dezentral	Zentral	Verteilt	Mehrere private BC	Anbindung an beliebte BC		Zentrale IdM	Föderiertes IdM	DTI	SSI	Benutzerzentrierte IdM
1	[207]			✓	✓			✓		✓								✓		
1	[183]	✓				✓			✓		✓									
1	[184]		✓	✓	✓		✓			✓			✓							
1	[200], App.1	✓				✓	✓													
1	[200], App.2	✓				✓		✓				✓						✓		
1	[208]			✓	✓			✓		✓								✓		
1	[178]	✓				✓		✓			✓							✓		
1	[180]		✓		✓			✓		✓								✓		
1	[182]			✓	✓			✓			✓					✓				
1	[202]	✓		✓	✓	✓		✓				✓		✓				✓		
1	[181]	✓		✓	✓	✓		✓		✓				✓				✓		
1	[175]		✓		✓			✓		✓								✓		
1	[177]		✓		✓			✓		✓								✓		
1	[179]		✓		✓			✓		✓								✓		
1	[192]		✓		✓			✓			✓					✓				
1	[173]		✓		✓			✓				✓								
1	[201]	✓				✓		✓												
1	[176]			✓	✓			✓			✓					✓				
1	[174]			✓	✓			✓			✓									

Tabelle A.3: Taxonomie bzgl. der Merkmale des Zugriffsmanagement sowie Datensicherheit, resultierend aus der ersten Iteration der Literaturrecherche.

		Zugriffsmanagement														**Sicherheitsmechanismen**								
		Autorität Allgemein			**Autorität F&E**			**Strategie**			**NDAC Logik**				**Zugriffskontrolle**					**Speicherung**			
It.	Referenz	System	Datensubjekt	Geteilt	System	Datensubjekt	Geteilt	DAC	Hybrid	NDAC	MAC	Rollenbasiert	Regelbasiert	Asym. Verschlüsselung	Sym. Verschlüsselung	Proxy Re-Encryption	Tokenisierung	CP-ABE	Sym. Verschlüsselung	Hybride Verschlüsselung	Asym. Verschlüsselung	Passwortschutz
1	[207]		✓						✓				✓	✓		✓				✓		
1	[183]		✓											✓						✓		
1	[184]		✓						✓				✓	✓							✓	
1	[200], App.1		✓					✓						✓					✓			
1	[200], App.2		✓					✓						✓					✓			
1	[208]			✓					✓				✓	✓		✓				✓		
1	[178]			✓						✓			✓					✓	✓			
1	[180]		✓			✓		✓											✓			
1	[182]		✓			✓		✓								✓					✓	
1	[202]		✓			✓		✓						✓		✓				✓		
1	[181]		✓			✓			✓				✓									
1	[175]		✓			✓				✓		✓							✓			
1	[177]		✓				✓	✓														
1	[179]		✓				✓	✓						✓			✓				✓	
1	[192]		✓					✓								✓					✓	
1	[173]		✓					✓													✓	
1	[201]	✓			✓					✓			✓			✓			✓			
1	[176]					✓		✓														
1	[174]		✓			✓												✓			✓	

Tabelle A.4: Taxonomie bzgl. der Merkmale des Daten- und Identitätsmanagement, resultierend aus der zweiten Iteration der Literaturrecherche.

		Datenmanagement														**Identitätsmanagement**				
		Blockchain-Typ					**Speicherort**			**Off-Chain Speicher**			**Blockchain Feature**							
It.	Referenz	Öffentliche	Private	Konsortium	Permissioned	Permissionless	On-Chain	Off-Chain	Hybrid	Dezentral	Zentral	Verteilt	Mehrere private BC	Anbindung an beliebte BC		Zentrale IdM	Föderiertes IdM	DTI	SSI	Benutzerzentrierte IdM
2	[212]		✓		✓			✓			✓									
2	[193]			✓	✓			✓			✓									
2	[204]	✓				✓		✓		✓								✓		
2	[185]		✓	✓	✓			✓			✓		✓					✓		
2	[186]	✓				✓		✓				✓		✓				✓		
2	[194]			✓	✓			✓				✓						✓		
2	[195]	✓				✓		✓				✓						✓		
2	[187]		✓		✓			✓				✓								
2	[188]	✓				✓		✓			✓					✓				
2	[189]	✓				✓		✓				✓				✓				
2	[213]			✓	✓			✓		✓										
2	[196]	✓				✓		✓				✓						✓		
2	[190]	✓				✓		✓				✓				✓				

Tabelle A.5: Taxonomie bzgl. der Merkmale des Zugriffsmanagement sowie Datensicherheit, resultierend aus der zweiten Iteration der Literaturrecherche.

		Zugriffsmanagement															**Sicherheitsmechanismen**								
		Autorität Allgemein			**Autorität F&E**			**Strategie**			**NDAC Logik**			**Zugriffskontrolle**					**Speicherung**					
It.	Referenz	System	Datensubjekt	Geteilt	System	Datensubjekt	Geteilt	DAC	Hybrid	NDAC	MAC	Rollenbasiert	Regelbasiert	Asym. Verschlüsselung	Sym. Verschlüsselung	Proxy Re-Encryption	Tokenisierung	CP-ABE	Sym. Verschlüsselung	Hybride Verschlüsselung	Asym. Verschlüsselung	Passwortschutz
2	[194]			✓			✓			✓	✓					✓			✓			
2	[189]			✓						✓		✓										
2	[213]	✓						✓								✓			✓			
2	[196]		✓			✓				✓	✓		✓					✓	✓			
2	[188]		✓					✓						✓					✓			
2	[190]		✓					✓														
2	[204]		✓			✓				✓		✓										
2	[186]		✓							✓		✓	✓	✓			✓					✓
2	[195]		✓					✓								✓					✓	
2	[185]		✓					✓													✓	
2	[212]		✓					✓						✓					✓			
2	[193]		✓			✓				✓			✓					✓	✓			
2	[187]		✓					✓											✓			

Tabelle A.6: Taxonomie bzgl. der Merkmale des Daten- und Identitätsmanagement, resultierend aus der dritten Iteration der Literaturrecherche.

		Datenmanagement													**Identitätsmanagement**				
		Blockchain-Typ					**Speicherort**			**Off-Chain Speicher**			**Blockchain Feature**						
It.	Referenz	Öffentliche	Private	Konsortium	Permissioned	Permissionless	On-Chain	Off-Chain	Hybrid	Dezentral	Zentral	Verteilt	Mehrere private BC	Anbindung an beliebte BC	Zentrale IdM	Föderiertes IdM	DTI	SSI	Benutzerzentrierte IdM
3	[197]		✓		✓			✓			✓								
3	[203]		✓	✓				✓			✓								
3	[198]			✓	✓			✓				✓	✓					✓	
3	[199]			✓	✓				✓			✓			✓				
3	[206]			✓	✓			✓			✓				✓				
3	[191]		✓		✓			✓			✓				✓				

Tabelle A.7: Taxonomie bzgl. der Merkmale des Zugriffsmanagement sowie Datensicherheit, resultierend aus der dritten Iteration der Literaturrecherche.

		Zugriffsmanagement													**Sicherheitsmechanismen**									
		Autorität Allgemein			**Autorität F&E**			**Strategie**			**NDAC Logik**				**Zugriffskontrolle**					**Speicherung**				
It.	Referenz	System	Datensubjekt	Geteilt	System	Datensubjekt	Geteilt	DAC	Hybrid	NDAC	MAC	Rollenbasiert	Regelbasiert	Asym. Verschlüsselung	Sym. Verschlüsselung	Proxy Re-Encryption	Tokenisierung	CP-ABE	Sym. Verschlüsselung	Hybride Verschlüsselung	Asym. Verschlüsselung	Passwortschutz
3	[197]		✓					✓						✓							✓		
3	[203]				✓					✓		✓				✓					✓		
3	[198]			✓					✓			✓			✓				✓				
3	[199]			✓					✓			✓							✓				
3	[206]			✓				✓								✓					✓		
3	[191]		✓			✓			✓				✓	✓								✓	

Abbildungsverzeichnis

Tabellenverzeichnis

Eigene Veröffentlichungen

Journalartikel

[1] M. Schinle, C. Erler, M. Kaliciak, C. Milde, S. Stock, M. Gerdes, W. Stork, *et al.*, "Digital health apps in the context of dementia: Questionnaire study to assess the likelihood of use among physicians," *JMIR Formative Research*, vol. 6, no. 6, p. e35961, 2022.

[2] C. Erler, A.-M. Bauer, F. Gauger, and W. Stork, "Decision model to design trust-focused and blockchain-based health data management applications," *Blockchains*, vol. 2, no. 2, pp. 79–106, 2024.

Konferenzbeiträge und sonstige Veröffentlichungen

[3] S. Stock, C. Erler, and W. Stork, "Realistic simulation of progressive vision diseases in virtual reality," in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–2, 2018.

[4] S. Stock, C. Erler, W. Stork, G. Labuz, H. S. Son, R. Khoramnia, and G. U. Auffarth, "Suitability of virtual reality for vision simulation–a case study using glaucomatous visual fields," *Investigative Ophthalmology & Visual Science*, vol. 60, no. 9, pp. 2441–2441, 2019.

[5] M. Schinle, C. Erler, P. N. Andris, and W. Stork, “Integration, execution and monitoring of business processes with chaincode,” in *2020 2nd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS)*, pp. 63–70, IEEE, 2020.

[6] M. Schinle, C. Erler, and W. Stork, “Distributed Ledger Technology for the systematic Investigation and Reduction of Information Asymmetry in Collaborative Networks,” in *53rd Hawaii International Conference on System Sciences, HICSS*, pp. 1–10, 2020.

[7] M. Schinle, C. Erler, A. R. Vetter, and W. Stork, “How to disclose selective information from permissioned dlt-based traceability systems?,” in *2020 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS)*, pp. 153–158, IEEE, 2020.

[8] M. Schinle, C. Erler, S. Leenstra, S. Stock, M. Gerdes, and W. Stork, “A Decision Process Model for De-Identification Methods on the Example of Psychometric Data,” in *2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, pp. 1–6, IEEE, 2021.

[9] M. Schinle, C. Erler, and W. Stork, “Data Sovereignty in Data Donation Cycles - Requirements and Enabling Technologies for the Data-driven Development of Health Applications,” in *Proceedings of the 54th Hawaii International Conference on System Sciences*, pp. 3972–3981, 2021.

[10] M. Schinle, C. Erler, T. Schneider, J. Plewnia, and W. Stork, “Data-driven development of digital health applications on the example of dementia screening,” in *2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 1–6, IEEE, 2021.

[11] C. Erler, M. Schinle, M. Dietrich, and W. Stork, “Decision model to design a blockchain-based system for storing sensitive health data,” in *30th European Conference on Information Systems, ECIS 2022, Timisoara, Romania*, 2022.

[12] M. Schinle, C. Erler, M. Hess, and W. Stork, “Explainable artificial intelligence in ambulatory digital dementia screenings,” *Challenges of Trustable AI and Added-Value on Health*, p. 123, 2022.

[13] C. Erler, S. Hu, A. Danelski, W. Stork, A. Sunyaev, and M. Gersch, “Threat Modeling to Design a Decentralized Health Data Management Application,” in *International Conference on Information Technology & Systems*, pp. 443–455, Springer, 2023.

[14] A. Danelski, M. Gersch, and C. Erler, “Whitepaper zum Projekt"BloG^3-Blockchainbasiertes Gesundheitsdatenmanagement für gesamtheitliche Gesundheitsprofile". Delphi-Analyse: Szenarien und Geschäftsmodelle,” tech. rep., Discussion Paper, 2023.

[15] L. Schweickart, C. Erler, J. Juhl, C. Zimmermann, and W. Stork, “User-centered representation of data flows in mhealth applications,” in *2023 International Conference on Electrical, Communication and Computer Engineering (ICECCE)*, pp. 1–6, IEEE, 2023.

[16] C. Erler and C. Zimmermann, “Blockchain-basiertes Gesundheitsdatenmanagement für gesamtheitliche Gesundheitsprofile Blog3; Teilvorhaben: Abbildung von Informationskoordinationsprozessen in Blockchain-basierten Gesundheitsdatenmanagementsystemen: Abschlussbericht zum Verbundvorhaben,” 2023.

[17] C. Erler, S. Perret, G. Biri, and W. Stork, “Investigation of Current Translation Challenges and Barriers to the Use of Artificial Intelligence in the German Healthcare System,” in *57th Hawaii International Conference on System Sciences, HICSS*, pp. 3587–3595, 2024.

[18] R. Burmeister, C. Erler, F. Gauger, R. Dressle, and B. Feige, “Advancing Sleep Research through Dynamic Consent and Trustee-Based Medical Data Processing,” in *International Conference on Digital Society (ICDS 2024)*, 2024.

[19] C. Erler, R. Burmeister, and F. Gauger, “Vertrauenswürdiges Datentreuhandmodell zur souveränen Verwaltung und effektiven Nutzung von medizinischen Daten in der Schlafforschung; Teilvorhaben: Informationstechnologische Konzepte und Rahmenbedingungen zur Schaffunge eines nachhaltigen und souverän verwalteten Datentreuhandmodells: Abschlussbericht zum Verbundvorhaben,” 2024.

[20] G. Biri, A. Vasilache, T. Hu, M. A. N. Themistocli, S. Nitzsche, J. Juhl, C. Erler, S. Fuhrhop, W. Stork, and J. Becker, “Sleep Stage and Apnea Classification from Single-Lead ECG Using Artificial and Spiking Neural Networks,” in *2024 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pp. 79–84, 2024.

[21] C. Erler, G. Biri, T. Stein, and M. Bouras, “Digitale Identitäten im Gesundheitswesen am Beispiel von Patientenakte und Knochenmarkspenderregister,” in *Digitale Identitäten und Nachweise: Lösungsansätze für vertrauenswürdige Interaktionen zwischen Menschen, Unternehmen und Verwaltung* (J. Anke, M. Kubach, and J. Sürmeli, eds.), Springer Fachmedien Wiesbaden, 2025.

Betreute Abschlussarbeiten

[22] A.-M. Bauer, “Konzeption und Entwicklung des Identitäts- und Zugriffsmanagement für einen dezentralen Datentreuhänder,” Masterarbeit, KIT Karlsruher Institut für Technologie, 2023.

[23] A. Gnan, “De-Identifikation von Gesundheitsdaten in einem Datentreuhänder für die Schlafforschung,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2024.

[24] A. Vetter, “Entwicklung eines Blockchain-basierten Supply Chain Traceability Systems für nachhaltige Produkte in der Lebensmittelindustrie,” Masterarbeit, KIT Karlsruher Institut für Technologie, 2019.

[25] A. Wesner, “Identitätsmanagement für die Datenspende im Kontext des Digitale-Versorgung-Gesetzes,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2020.

[26] B. Liao, “Transparenz der Datenverarbeitung innerhalb einer Datentreuhänder-Plattform für Schlafstudien,” Masterarbeit, KIT Karlsruher Institut für Technologie, 2023.

[27] C. Retter, “Entwicklung eines intelligenten Disease-Management-Systems zur Unterstützung der Risikofaktorenbehandlung bei einer Demenz,” Masterarbeit, KIT Karlsruher Institut für Technologie, 2022.

[28] C. Staudenmaier, “Entwicklung von Daten-Integrationswerkzeugen für gesamtheitliche Gesundheitsprofile,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2020.

[29] E. Liem, “Konzeption eines Software-Ökosystems für sicherheitskritische Open-Source-Softwareentwicklung im hochregulierten Kontext,” Masterarbeit, KIT Karlsruher Institut für Technologie, 2020.

[30] J. Denzel, “Zugriffsverwaltung von Gesundheitsdaten mittels Distributed Ledger Technologie,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2020.

[31] J. Frey, “Konzeption und Entwicklung einer Pipeline zur Bewertung und Verbesserung der Datenqualität in der Schlafforschung im Rahmen eines Datentreuhänders,” Masterarbeit, KIT Karlsruher Institut für Technologie, 2024.

[32] J. Plewnia, “Privatsphäre-wahrende Analyseverfahren für medizinische Daten,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2020.

[33] L. A. Hüglin, “Digitale Einwilligungen zur souveränen Verwaltung und effektiven Nutzung von medizinischen Daten,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2022.

[34] L. Klassen, “Eine Sicherheits- und Performanzanalyse von Gesundheitsdatenmanagement-Anwendungen,” Masterarbeit, KIT Karlsruher Institut für Technologie, 2024.

[35] M. Dietrich, “Vergleich von On-Chain und Off-Chain Ansätzen zur Speicherung sensitiver Daten in einem Blockchain-basierten System,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2020.

[36] M. Lekesiz, “Konzeption und Entwicklung eines webbasierten Tools zur Prozessüberwachung in Blockchain-basierten Systemen für eine onkologische Behandlung,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2022.

[37] M. Neumahr, “Entwicklung eines Geschäftsmodells für einen Datentreuhänder im Rahmen eines KI-Reallabors,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2025.

[38] M. Stebner, “Analyse und Optimierung der Usability einer altersgerechten mobilen Applikation für die Erkennung und Verlaufskontrolle von Demenz,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2020.

[39] O. Özer, “DLT-basierte Abstimmungssysteme zur Priorisierung von Informationsbedarfen in kollaborativen Netzwerken,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2020.

[40] P. Andris, “Konzeption und Entwicklung eines Frameworks zur Integration modellierter Geschäftsprozesse in ein Blockchain-basiertes System,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2019.

[41] P. Jung, “Entwicklung eines Systems zur Integration von Datenspenden von demenziell erkrankten Personen in den Datenspendekreislauf des Gesundheitswesens,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2020.

[42] P. P. Worrach, “Künstliche Intelligenz in Kombination mit Distributed Ledger Technology als Mehrwertdienst für einen onkologischen Patienten,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2021.

[43] S. Leenstra, “Methoden zur De-Identifikation im Kontext der Datenspende,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2021.

[44] T. Georgiev, “Konzeption und Aufbau einer Datentreuhänderplattform für eine Gesundheitsdatenspende,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2022.

[45] T. Schneider, “Machine Learning Modelle zur Überwachung der psychischen Gesundheit - die Entwicklung eines neuartigen Demenztests,” Bachelorarbeit, KIT Karlsruher Institut für Technologie, 2020.

[46] Y. Jiang, “Aufbau einer Forschungsplattform für digitale Studienkonzepte im Kontext demenzieller Erkrankungen,” Masterarbeit, KIT Karlsruher Institut für Technologie, 2021.

Literaturverzeichnis

[47] L. Ismail, H. Materwala, A. P. Karduck, and A. Adem, "Requirements of health data management systems for biomedical care and research: Scoping review," *J. Med. Internet Res.*, vol. 22, p. e17508, July 2020.

[48] A. Blasimme, M. Fadda, M. Schneider, and E. Vayena, "Data sharing for precision medicine: Policy lessons and future directions," *Health Affairs*, vol. 37, no. 5, pp. 702–709, 2018.

[49] Europäische Kommission, "Questions and answers - EU Health: European Health Data Space (EHDS)." Online verfügbar unter `https://ec.europa.eu/commission/presscorner/detail/en/qanda_22_2712` (letzter Zugriff 27.03.2025).

[50] S. Pohlmann, A. Kunz, D. Ose, E. C. Winkler, A. Brandner, R. Poss-Doering, J. Szecsenyi, and M. Wensing, "Digitalizing health services by implementing a personal electronic health record in germany: Qualitative analysis of fundamental prerequisites from the perspective of selected experts," *J. Med. Internet Res.*, vol. 22, p. e15102, Jan. 2020.

[51] R. Thiel, L. Deimel, D. Schmidtmann, K. Piesche, T. Hüsing, J. Rennoch, V. Stroetmann, and K. Stroetmann, "#SmartHealthSystems Digitalisierungsstrategien im internationalen Vergleich," 2018. Online verfügbar unter `https://www.bertelsmann-stiftung.de/fileadmin/files/Projekte/Der_digitale_Patient/VV_SHS-Gesamtstudie_dt.pdf`(letzter Zugriff 09.03.2025).

[52] Bundesregierung, "Gesetz zur Modernisierung der gesetzlichen Krankenversicherung (GKV-Modernisierungsgesetz – GMG)." Online verfügbar

unter https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&start=//*%5b@attr_id=%27bgbl103s2190.pdf%27%5d#__bgbl__%2F%2F*%5B%40attr_id%3D%27bgbl103s2190.pdf%27%5D__1742812137425 (letzter Zugriff 24.03.2025).

[53] Bundesregierung, "Gesetz für sichere digitale Kommunikation und Anwendungen im Gesundheitswesen sowie zur Änderung weiterer Gesetze: E-Health-Gesetz." Online verfügbar unter https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl115s2408.pdf#__bgbl__%2F%2F*%5B%40attr_id%3D%27bgbl115s2408.pdf%27%5D__1742813162148 (letzter Zugriff 24.03.2025).

[54] Bundesministerium für Gesundheit, "Schnellere Termine, mehr Sprechstunden, bessere Angebote für gesetzlich Versicherte - Terminservice- und Versorgungsgesetz (TSVG)." Online verfügbar unter https://www.bundesgesundheitsministerium.de/terminservice-und-versorgungsgesetz.html (letzter Zugriff 25.03.2025).

[55] gematik GmbH, "Über uns," 2025. Online verfügbar unter https://www.gematik.de/ueber-uns/ (letzter Zugriff 25.02.2025).

[56] gematik GmbH, "Arena für digitale Medizin - Whitepaper Telematikinfrastruktur 2.0 für ein föderalistisch vernetztes Gesundheitssystem," 2025. Online verfügbar unter https://www.gematik.de/media/gematik/Medien/Telematikinfrastruktur/Dokumente/gematik_Whitepaper_Arena_digitale_Medizin_TI_2.0_Web.pdf (letzter Zugriff 03.03.2025).

[57] gematik GmbH, "Aktuelles | Pilotphase der ePA für alle erfolgreich gestartet." Online verfügbar unter https://www.bundesaerztekammer.de/themen/aerzte/digitalisierung/digitale-anwendungen/telematikinfrastruktur/epa (letzter Zugriff 27.03.2025).

[58] gematik GmbH, "Aktuelles | Pilotphase der ePA für alle erfolgreich gestartet." Online verfügbar unter https://www.gematik.de/newsroo

m/news-detail/aktuelles-pilotphase-der-epa-fuer-alle-erfolgreich-gestartet (letzter Zugriff 27.03.2025).

[59] gematik GmbH, "Aktuelles | Stellungnahme zum CCC-Vortrag zur ePA für alle." Online verfügbar unter https://www.gematik.de/newsroom/news-detail/aktuelles-stellungnahme-zum-ccc-vortrag-zur-epa-fuer-alle (letzter Zugriff 27.03.2025).

[60] IGES Institut, "Wissenschaftliche Evaluation des Produktivbetriebs der Anwendungen der Telematikinfrastruktur 2022." Online verfügbar unter https://www.gematik.de/media/gematik/Medien/Telematikinfrastruktur/TI-Atlas/IGES-Studie_Wissenschaftliche_Evaluation_des_Produktivbetriebs_der_Anwendungen_der_TI_2022.pdf (letzter Zugriff 28.03.2025).

[61] IGES Institut, "Wissenschaftliche Evaluation des Produktivbetriebs der Anwendungen der Telematikinfrastruktur 2023." Online verfügbar unter https://www.gematik.de/media/gematik/Medien/Telematikinfrastruktur/TI-Atlas/IGES-Studie_Wissenschaftliche_Evaluation_des_Produktivbetriebs_der_Anwendungen_der_TI_2023.pdf (letzter Zugriff 28.03.2025).

[62] IGES Institut, "Wissenschaftliche Evaluation des Produktivbetriebs der Anwendungen der Telematikinfrastruktur 2024." Online verfügbar unter https://www.gematik.de/media/gematik/Medien/Telematikinfrastruktur/TI-Atlas/Studienbericht_IGES_2024.pdf (letzter Zugriff 28.03.2025).

[63] gematik GmbH, "Aktuelles | Die bundesweite Einführung der ePA für alle startet am 29. April 2025," 2025. Online verfügbar unter https://www.gematik.de/newsroom/news-detail/aktuelles-die-bundesweite-einfuehrung-der-epa-fuer-alle-startet-am-29-april-2025 (letzter Zugriff 25.04.2025).

[64] gematik GmbH, “TI 2.0 - Unser Weg in die Zukunft.” Online verfügbar unter https://www.gematik.de/telematikinfrastruktur/ti-2-0 (letzter Zugriff 28.03.2025).

[65] gematik GmbH, “Die elektronische Patientenakte ab 2025: Basisinformationen zu Aufgaben, Pflichten und Zugriffsrechten.” Online verfügbar unter https://www.kbv.de/media/sp/PraxisInfoSpezial_ePA.pdf (letzter Zugriff 28.03.2025).

[66] M. Gersch, “Digitalisierung im Gesundheitswesen,” in *Handbuch Digitalisierung*, pp. 1016 – 1042, Vahlen, 2021.

[67] N. Kannengießer, S. Lins, T. Dehling, and A. Sunyaev, “What Does Not Fit Can be Made to Fit! Trade-Offs in Distributed Ledger Technology Designs,” in *Proceedings of the 52nd Hawaii International Conference on System Sciences* (T. Bui, ed.), Proceedings of the Annual Hawaii International Conference on System Sciences, Hawaii International Conference on System Sciences, Jan. 2019.

[68] H. Ghayvat, M. Sharma, P. Gope, and P. K. Sharma, “Sharif: Solid pod-based secured healthcare information storage and exchange solution in internet of things,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 8, pp. 5609–5618, 2022.

[69] K. Peffers, T. Tuunanen, M. Rothenberger, and S. Chatterjee, “A Design Science Research Methodology for Information Systems Research,” *JMIS*, 2007.

[70] G. Jonitz, K. Piwernetz, and E. A. M. Neugebauer, *Das deutsche Gesundheitssystem – ein Überblick*, pp. 63–75. Springer Fachmedien Wiesbaden, 2024.

[71] T. Latal, T. Hinze, N. Roeder, and D. Franz, “Aufbau des selbstverwalteten gesundheitswesens in deutschland,” *Zeitschrift für Herz-,Thorax- und Gefäßchirurgie*, vol. 31, 08 2016.

[72] GKV-Spitzenverband, “Aufgaben und Ziele,” 2025. Online verfügbar unter https://www.gkv-spitzenverband.de/gkv_spitzenverband/der_verband/aufgaben_und_ziele/aufgaben_und_ziele.jsp (letzter Zugriff 25.02.2025).

[73] gematik GmbH, “Die Telematikinfrastruktur,” 2025. Online verfügbar unter https://www.gematik.de/telematikinfrastruktur/ (letzter Zugriff 25.02.2025).

[74] gematik GmbH, “Whitepaper Datenschutz und Informationssicherheit in der Telematikinfrastruktur – So schützt die TI Gesundheitsdaten,” 2021. Online verfügbar unter https://fachportal.gematik.de/schnelleinstieg/downloadcenter/datenschutz-und-informationssicherheit-in-der-ti (letzter Zugriff 25.02.2025).

[75] gematik GmbH, “Die Struktur der gematik,” 2025. Online verfügbar unter https://www.gematik.de/ueber-uns/struktur (letzter Zugriff 24.02.2025).

[76] T. Kriedel, “Digitale Praxis: Nutzen für Patienten und Ärzte,” 2019. Online verfügbar unter https://www.kvbawue.de/api-file-fetcher?fid=3235 (letzter Zugriff 03.03.2025).

[77] C. Fitte, P. Meier, A. Behne, D. Miftari, and F. Teuteberg, “Die elektronische gesundheitsakte als vernetzungsinstrument im internet of health,” in *INFORMATIK 2019: 50 Jahre Gesellschaft für Informatik – Informatik für Gesellschaft*, pp. 111–124, Bonn: Gesellschaft für Informatik e.V., 2019.

[78] Bundesamt für Sicherheit in der Informationstechnik, “Wie funktioniert ein Virtual Private Network (VPN)?,” 2025. Online verfügbar unter https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Cyber-Sicherheitsempfehlungen/Router-WLAN-VPN/Virtual-Private-Networks-VPN/virtual-private-networks-vpn_node.html (letzter Zugriff 12.03.2025).

[79] gematik GmbH, “WANDA - die Weiteren Anwendungen für den Datenaustausch in der Telematikinfrastruktur,” 2025. Online verfügbar unter https://fachportal.gematik.de/anwendungen/weitere-anwendungen (letzter Zugriff 14.05.2025).

[80] Bundesministerium für Gesundheit, “Lauterbach: Elektronische Patientenakte ab Ende 2024 für alle verbindlich ,” 2024. Online verfügbar unter https://www.bundesgesundheitsministerium.de/presse/interviews/fas-030324-elektronische-patientenakte.html (letzter Zugriff 03.03.2025).

[81] gematik GmbH, “ePA für alle,” 2025. Online verfügbar unter https://www.gematik.de/anwendungen/epa-fuer-alle (letzter Zugriff 03.03.2025).

[82] Forschungsdatenzentrum-Gesundheit, “Gesundheitsdaten,” 2025. Online verfügbar unter https://www.forschungsdatenzentrum-gesundheit.de/gesundheitsdaten-am-fdz-gesundheit (letzter Zugriff 25.02.2025).

[83] S. C. Semler, M. Boeker, R. Eils, D. Krefting, M. Loeffler, J. Bussmann, F. Wissing, and H.-U. Prokosch, “Die Medizininformatik-Initiative im Überblick – Aufbau einer Gesundheitsforschungsdateninfrastruktur in Deutschland,” *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, vol. 67, pp. 616–628, 2024.

[84] Deutsche Forschungsdatenportal für Gesundheit, “Forschen für Gesundheit,” 2025. Online verfügbar unter https://forschen-fuer-gesundheit.de/ (letzter Zugriff 04.03.2025).

[85] Deutsche Forschungsdatenportal für Gesundheit, “Prozesse der Antragstellung und Datennutzung über das Forschungsdatenportal,” 2025. Online verfügbar unter https://forschen-fuer-gesundheit.de/daten-und-bioproben/prozesse-der-antragstellung-und-datennutzung-in-der-mii/ (letzter Zugriff 04.03.2025).

[86] Deutsche Forschungsdatenportal für Gesundheit, “Registrierung,” 2025. Online verfügbar unter https://forschen-fuer-gesundheit.de/registrierung/ (letzter Zugriff 04.03.2025).

[87] Deutsche Forschungsdatenportal für Gesundheit, “Daten und Bioproben finden,” 2025. Online verfügbar unter https://forschen-fuer-gesundheit.de/daten-und-bioproben/daten-finden/ (letzter Zugriff 04.03.2025).

[88] Deutsche Forschungsdatenportal für Gesundheit, “Daten und Bioproben für ein Forschungsprojekt beantragen,” 2025. Online verfügbar unter https://forschen-fuer-gesundheit.de/daten-und-bioproben/daten-und-bioproben-fur-ein-forschungsprojekt-beantragen/ (letzter Zugriff 04.03.2025).

[89] Deutsche Forschungsdatenportal für Gesundheit, “Daten und Proben analysieren,” 2025. Online verfügbar unter https://forschen-fuer-gesundheit.de/daten-und-bioproben/daten-und-proben-analysieren/ (letzter Zugriff 04.03.2025).

[90] Medizininformatik-Initiative, “Digitale FortschrittsHubs Gesundheit ,” 2025. Online verfügbar unter https://www.medizininformatik-initiative.de/de/use-cases-und-projekte/digitale-fortschrittshubs-gesundheit (letzter Zugriff 04.03.2025).

[91] Medizininformatik-Initiative, “Mustertext zur Patienteneinwilligung,” 2021. Online verfügbar unter https://www.medizininformatik-initiative.de/de/mustertext-zur-patienteneinwilligung (letzter Zugriff 04.03.2025).

[92] Medizininformatik-Initiative, “Der Kerndatensatz der Medizininformatik-Initiative,” 2021. Online verfügbar unter https://www.medizininformatik-initiative.de/de/der-kerndatensatz-der-medizininformatik-initiative (letzter Zugriff 04.03.2025).

[93] Nationale Forschungsdateninfrastruktur (NFDI) e.V., "NFDI: Daten als gemeinsames Gut für exzellente Forschung, organisiert durch die Wissenschaft in Deutschland.," 2025. Online verfügbar unter https://www.nfdi.de/ (letzter Zugriff 04.03.2025).

[94] N. Hartl, E. Wössner, and Y. Sure-Vetter, "Nationale Forschungsdateninfrastruktur (NFDI)," *Informatik Spektrum*, vol. 44, pp. 1–4, 10 2021.

[95] D. Strech, S. Graf von Kielmansegg, S. Zenker, M. Krawczak, and S. C. Semler, "Gutachten Datenspende - Bundesgesundheitsministerium," Mar 2020.

[96] Universitätsmedizin Greifswald, "Core Units," 2025. Online verfügbar unter https://www.medizin.uni-greifswald.de/de/forschung/forschungsservice/core-units/ (letzter Zugriff 04.03.2025).

[97] Unabhängige Treuhandstelle der Universitätsmedizin Greifswald, "Die Datentreuhänder," 2025. Online verfügbar unter https://www.ths-greifswald.de/ueber-uns/ (letzter Zugriff 04.03.2025).

[98] NAKO e.V., "Die NAKO Gesundheitsstudie - Deutschlands größte Langzeitstudie zur Erforschung von Volkskrankheiten," 2025. Online verfügbar unter https://nako.de/ (letzter Zugriff 04.03.2025).

[99] Bundesdruckerei GmbH, "Datentreuhänder - Datentreuhänder-Plattform mit Vertrauensstellendienst on demand," 2025. Online verfügbar unter https://www.bundesdruckerei-gmbh.de/de/loesungen/datentreuhaender (letzter Zugriff 04.03.2025).

[100] Bundesdruckerei GmbH, "DSGVO-konforme Nutzung und Verknüpfung sensibler Daten," 2025. Online verfügbar unter https://www.bundesdruckerei-gmbh.de/files/dokumente/pdf/datenblatt-datentreuhaender.pdf (letzter Zugriff 04.03.2025).

[101] Forschungspraxennetz Baden-Württemberg, "Forschungspraxennetz Baden-Württemberg," 2025. Online verfügbar unter

https://www.forschungspraxennetz-bw.de/(letzter Zugriff 04.03.2025).

[102] P. Schmutz, A. Krauss, S. Dörflinger, A. Becker, R. Koch, A. Polanc, E. Feil, C. Salm, K. Scheeser, F. Peters-Klimm, and C. Thies, "Fopranet-bw: an infrastructure for clinical studies in practice-based research networks in the german health system," in *Digital health and informatics innovations for sustainable health care systems : Proceedings of MIE 2024, Athens, 25-29 August 2024 (Studies in health technology and informatics, volume 316)* (J. Mantas, A. Hasman, and G. Demiris, eds.), pp. 190 – 194, 2024.

[103] Europäische Kommission, "A European strategy for data," 2024. Online verfügbar unter https://digital-strategy.ec.europa.eu/en/policies/strategy-data)(letzter Zugriff 04.03.2025).

[104] Europäische Kommission, "Recommendation on a European Electronic Health Record exchange format," 2019. Online verfügbar unter https://digital-strategy.ec.europa.eu/en/library/recommendation-european-electronic-health-record-exchange-format(letzter Zugriff 04.03.2025).

[105] Europäische Kommission, "European Health Data Space Regulation (EHDS)," 2025. Online verfügbar unter https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en(letzter Zugriff 04.03.2025).

[106] Gaia-X European Association for Data and Cloud AISBL, "About Gaia-X," 2025. Online verfügbar unter https://gaia-x.eu/about/(letzter Zugriff 04.03.2025).

[107] Bundesministerium für Wirtschaft und Klimaschutz, "Das Gaia-X Ökosystem - Souveräne Dateninfrastruktur für Europa," 2025. https://www.bmwk.de/Redaktion/DE/Dossier/gaia-x.html(letzter Zugriff 09.03.2025).

[108] International Data Spaces e. V., “Our mission: Creating the future of the global, digital economy,” 2025. Online verfügbar unter https://internationaldataspaces.org/(letzter Zugriff 04.03.2025).

[109] B. Otto, A. Rubina, A. Eitel, A. Teuscher, A. M. Schleimer, C. Lange, D. Stingl, E. Loukipoudis, G. Brost, G. Boege, H. Pettenpohl, J. Langkau, J. Gelhaar, K. Mitani, M. Hupperz, M. Huber, N. Jahnke, R. Brandstädter, S. Wessel, and S. Bader, “GAIA-X and IDS,” 2021.

[110] B. Otto, S. Steinbuss, A. Teuscher, S. Lohmann, S. Bader, P. Birnstil, M. Böhmer, G. Brost, J. Cirullies, A. Eitel, T. Ernst, S. Geisler, J. Gelhaar, R. Gude, C. Haas, M. Huber, C. Jung, J. Jürjens, C. Lange, D. Lis, C. Mader, N. Menz, R. Nagel, F. Patzer, H. Pettenpohl, J. Pullmann, C. Quix, D. Schulz, J. Schütte, and et al., “Reference Architecture Model. Version 3.0,” 2019. Online verfügbar unter https://publica.fraunhofer.de/handle/publica/299836(letzter Zugriff 09.03.2025).

[111] Estonian Business and Innovation Agency, “X-Road – interoperability services,” 2025. Online verfügbar unter https://e-estonia.com/solutions/x-road-interoperability-services/x-road/(letzter Zugriff 09.03.2025).

[112] Estonian Business and Innovation Agency, “e-Health,” 2025. Online verfügbar unter https://e-estonia.com/solutions/e-health/e-health-records/(letzter Zugriff 09.03.2025).

[113] Estonian Business and Innovation Agency, “Cyber security,” 2025. Online verfügbar unter https://e-estonia.com/solutions/cyber-security/ksi-blockchain/(letzter Zugriff 09.03.2025).

[114] Findata – Social and Health Data Permit Authority, “Finnish Social and Health Data Permit Authority Findata,” 2025. Online verfügbar unter https://findata.fi/en/(letzter Zugriff 07.03.2025).

[115] Findata – Social and Health Data Permit Authority, “The range of Kanta Services materials for which secondary use permits are available is expanding,” 2025. Online verfügbar unter https://findata.fi/en/news/the-range-of-kanta-services-materials-for-which-secondary-use-permits-are-available-is-expanding/(letzter Zugriff 07.03.2025).

[116] Australian Digital Health Agency, “My Health Record,” 2025. Online verfügbar unter https://www.digitalhealth.gov.au/healthcare-providers/initiatives-and-programs/my-health-record(letzter Zugriff 07.03.2025).

[117] V. Strotbaum, M. Pobiruchin, B. Schreiweis, M. Wiesner, and B. Strahwald, “Your data is gold – data donation for better healthcare?,” *it - Information Technology*, vol. 61, pp. 219 – 229, 2019.

[118] The Department of Health and Aged Care, “Use of My Health Record data,” 2024. Online verfügbar unter https://www.health.gov.au/topics/health-technologies-and-digital-health/what-we-do/use-of-my-health-record-data(letzter Zugriff 07.03.2025).

[119] Healthcare Information and Management Systems Society, Inc. (HIMSS), “Interoperability in Healthcare,” 2025. Online verfügbar unter https://www.himss.org/resources/interoperability-healthcare(letzter Zugriff 12.03.2025).

[120] M. L. Braunstein, “Healthcare in the Age of Interoperability: The Promise of Fast Healthcare Interoperability Resources,” *IEEE Pulse*, vol. 9, no. 6, pp. 24–27, 2018.

[121] Bundesdruckerei GmbH, “Telematikinfrastruktur: sicherer Datenaustausch im Gesundheitswesen.” Online verfügbar unter https://www.bundesdruckerei.de/de/innovation-hub/telematikinfrastruktur (letzter Zugriff 24.03.2025).

[122] I. Martenstein and A. Wienke, “Das neue E-Health-Gesetz Was kommt auf Kliniken und niedergelassene Ärzte zu?,” *Laryngo Rhino Otologie*, vol. 95, pp. 417 – 418, 2016.

[123] Bundesministerium für Gesundheit, “Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz - DVG).” Online verfügbar unter https://www.bundesgesundheitsministerium.de/digitale-versorgung-gesetz.html (letzter Zugriff 25.03.2025).

[124] Bundesministerium für Gesundheit, “Verordnung zur Neufassung der Datentransparenzverordnung und zur Änderung der Datentransparenz-Gebührenverordnung.” Online verfügbar unter https://www.bundesgesundheitsministerium.de/service/gesetze-und-verordnungen/guv-19-lp/vo-datentransparenzverordnung.html (letzter Zugriff 26.03.2025).

[125] Bundesministerium für Gesundheit, “Gesetz zum Schutz elektronischer Patientendaten in der Telematikinfrastruktur (Patientendaten-Schutz-Gesetz – PDSG).” Online verfügbar unter https://www.bundesgesundheitsministerium.de/patientendaten-schutz-gesetz.html (letzter Zugriff 26.03.2025).

[126] Bundesministerium für Gesundheit, “Gesetz zur Beschleunigung der Digitalisierung des Gesundheitswesens (Digital-Gesetz – DigiG).” Online verfügbar unter https://www.bundesgesundheitsministerium.de/ministerium/gesetze-und-verordnungen/guv-20-lp/digig.html (letzter Zugriff 25.03.2025).

[127] Bundesministerium für Gesundheit, “Gesundheitsdatennutzungsgesetz (GDNG).” Online verfügbar unter https://www.bundesgesundheitsministerium.de/ministerium/gesetze-und-verordnungen/guv-20-lp/gesundheitsdatennutzungsgesetz.html (letzter Zugriff 25.03.2025).

[128] Europäische Kommission, “Daten-Governance-Gesetz erläutert.” Online verfügbar unter https://digital-strategy.ec.europa.eu/de/policies/data-governance-act-explained (letzter Zugriff 25.03.2025).

[129] Europäisches Parlament, “Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung).” Online verfügbar unter https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32016R0679&from=DE (letzter Zugriff 25.03.2025).

[130] P. Schaar, *Anonymisieren und Pseudonymisieren als Möglichkeit der Forschung mit sensiblen, personenbezogenen Forschungsdaten*, pp. 95–100. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014.

[131] Bundesministerium des Innern und für Heimat, “Bundesdatenschutzgesetz.” Online verfügbar unter https://www.bmi.bund.de/DE/themen/verfassung/datenschutz/bundesdatenschutzgesetz/bundesdatenschutzgesetz-node.html (letzter Zugriff 25.03.2025).

[132] B. Juraschko and V. Saur, K. G., *Praxishandbuch Recht für Bibliotheken und Informationseinrichtungen /*. Berlin ;: De Gruyter Saur,, 2., völlig überarbeitete auflage ed., [2020].

[133] I. Budin-Ljøsne, H. J. Teare, J. Kaye, S. Beck, H. B. Bentzen, L. Caenazzo, C. Collett, F. D’Abramo, H. Felzmann, T. Finlay, M. K. Javaid, E. Jones, V. Katić, A. Simpson, and D. Mascalzoni, “Dynamic consent: a potential solution to some of the challenges of modern biomedical research,” *BMC Medical Ethics*, vol. 18, no. 1, p. 4, 2017.

[134] Bundesministerium für Bildung und Forschung, “Richtlinie zur Förderung von Projekten zur Skalierung und Akzeptanzsteigerung von intersektoralen Datentreuhandmodellen in der Praxis,” 2023. Online verfügbar unter https://www.bmbf.de/SharedDocs/Bekanntmachungen/DE/2023/

10/2023-10-13-Bekanntmachung-Datentreuhandmodelle.html(letzter Zugriff 10.03.2025).

[135] Lindner, Maximilian and Straub, Sebastian, "Datentreuhänderschaft - Status Quo und Entwicklungsperspektiven," 2023. Online verfügbar unter https://www.iit-berlin.de/wp-content/uploads/2023/02/SDW_Datentreuhand.pdf(letzter Zugriff 10.03.2025).

[136] D. Feth, B. Rauch, D. Krohmer, J. von Albedyll, and K. B. Villela, "Datentreuhänder – Begriffliche Einordnung und Definition (Teil 1)," 2022. Online verfügbar unter https://www.iese.fraunhofer.de/blog/datentreuhaender-definition/(letzter Zugriff 10.03.2025).

[137] A. Blankertz, P. v. Braunmühl, P. Kuzev, F. Richter, H. Richter, and M. Schallbruch, "Datentreuhandmodelle-Themenpapier," 2020.

[138] Software Engineering Institute, "What Is Your Definition of Software Architecture?." Carnegie Mellon University, 2017. Online verfügbar unter http://www.sei.cmu.edu/architecture/definitions.html (letzter Zugriff 03.02.2025).

[139] M. Fowler, "Growing an Architecture," 2004. Online verfügbar unter https://www.martinfowler.com/articles/designDead.html#GrowingAnArchitecture (letzter Zugriff 03.02.2025).

[140] ISO/IEC/IEEE 42010:2022, "Software, Systems and Enterprise—Architecture Description," 2022. Online verfügbar unter https://www.iso.org/standard/74393.html (letzter Zugriff 03.02.2025).

[141] G. Booch, J. Rumbaugh, and I. Jacobson, "Unified Modeling Language User Guide, The (2nd Edition) (Addison-Wesley Object Technology Series)," *J. Database Manag.*, vol. 10, 01 1999.

[142] N. Rozanski and E. Woods, *Software Systems Architecture: Working with Stakeholders Using Viewpoints and Perspectives*. Addison-Wesley, 2012.

[143] S. Toth, *Vorgehensmuster für Softwarearchitektur*. München: Hanser, 3., aktualisierte und erweiterte auflage ed., [2019].

[144] L. Bass, P. Clements, and R. Kazman, *Software Architecture in Practice*. Addison-Wesley Professional, 3rd ed., 2012.

[145] S. Brown, "The C4 model for visualising software architecture." Online verfügbar unter `https://c4model.com/` (letzter Zugriff 07.02.2025).

[146] ISO/IEC/IEEE 42010:2022, "Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Product quality model (Edition 2, 2023)," 2023. Online verfügbar unter `https://www.iso.org/standard/78176.html` (letzter Zugriff 04.02.2025).

[147] G. Starke and C. H. Verlag, *Effektive Softwarearchitekturen - Ein praktischer Leitfaden*. Hanser eLibrary, München :: Hanser,, 10., überarbeitete auflage ed., [2024].

[148] K. Dittmann, K. Dirbanis, and T. Meier, *Project Management (IPMA®): Study Guide for Level D and Basic Certificate (GPM)*. Haufe Fachbuch, Haufe Group, 2021.

[149] B. Betzwieser, S. Franzbonenkamp, T. Riasanow, M. Böhm, H. Kienegger, and H. Krcmar, "A Decision Model for the Implementation of Blockchain Solutions," in *Americas Conference on Information Systems*, 2019.

[150] X. Xu, H. D. Bandara, Q. Lu, I. Weber, L. Bass, and L. Zhu, "A Decision Model for Choosing Patterns in Blockchain-Based Applications," in *2021 IEEE 18th International Conference on Software Architecture (ICSA)*, IEEE, Mar. 2021.

[151] The Object Management Group, "About the Business Process Model and Notation Specification Version 2.0.2." Online verfügbar unter `https://www.omg.org/spec/BPMN` (letzter Zugriff 12.02.2025).

[152] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, *et al.*, "The fair guiding principles for scientific data management and stewardship," *Scientific data*, vol. 3, no. 1, pp. 1–9, 2016.

[153] N. El Ioini and C. Pahl, "A Review of Distributed Ledger Technologies," in *On the Move to Meaningful Internet Systems. OTM 2018 Conferences* (H. Panetto, C. Debruyne, H. A. Proper, C. A. Ardagna, D. Roman, and R. Meersman, eds.), (Cham), pp. 277–288, Springer International Publishing, 2018.

[154] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," May 2009.

[155] X. Xu, I. Weber, and M. Staples, *Architecture for Blockchain Applications*. Cham, Switzerland: Springer Nature, 1 ed., Mar. 2019.

[156] M. Castro and B. Liskov, "Practical Byzantine fault tolerance," in *Proceedings of the Third Symposium on Operating Systems Design and Implementation*, p. 173–186, USENIX Association, 1999.

[157] The Linux Foundation, "Linux Foundation Decentralized Trust Launches with 17 Projects, 100+ Founding Members." Online verfügbar unter `https://www.linuxfoundation.org/press/linux-foundation-decentralized-trust-launches-with-17-projects-100-founding-members` (letzter Zugriff 13.03.2025).

[158] The Linux Foundation, "Hyperledger Aries." Online verfügbar unter `https://github.com/hyperledger/aries` (letzter Zugriff 13.03.2025).

[159] The Linux Foundation, "Introduction to Hyperledger Indy." Online verfügbar unter `https://github.com/hyperledger-archives/education/blob/master/LFS171x/docs/introduction-to-hyperledger-indy.md` (letzter Zugriff 13.03.2025).

[160] P.-L. Aublin, S. B. Mokhtar, and V. Quéma, "RBFT: Redundant Byzantine Fault Tolerance," in *2013 IEEE 33rd International Conference on Distributed Computing Systems*, pp. 297–306, 2013.

[161] M. A. Bouras, Q. Lu, F. Zhang, Y. Wan, T. Zhang, and H. Ning, "Distributed Ledger Technology for eHealth Identity Privacy: State of The Art and Future Perspective," *Sensors (Basel, Switzerland)*, vol. 20, no. 2, 2020.

[162] A. Jøsang and S. Pope, "User Centric Identity Management," 2005.

[163] B. Rawal, G. Manogaran, and A. Peter, "Cybersecurity and identity access management," 01 2023.

[164] Bundesverband der Deutschen Industrie e.V., "Anonymisierung personenbezogener Daten: Ein branchenübergreifender Praxisleitfaden für Industrieunternehmen," 2020. Online verfügbar unter `https://bdi.eu/publikation/news/anonymisierung-personenbezogener-daten/`(letzter Zugriff 12.03.2025).

[165] C. Wegener, T. Milde, and W. Dolle, *Informationssicherheits-Management*. Springer Vieweg Berlin, Heidelberg, 01 2016.

[166] C. Paar and J. Pelzl, *Understanding Cryptography: A Textbook for Students and Practitioners*. Springer Publishing Company, Incorporated, 1st ed., 2009.

[167] W. Xiong and R. Lagerström, "Threat modeling – A systematic literature review," *Comput. Secur.*, vol. 84, p. 53–69, 2019.

[168] A. Shostack, *Threat modeling: Designing for Security*. Wiley, 2014.

[169] N. Shevchenko, T. A. Chick, P. O'Riordan, T. P. Scanlon, and C. Woody, "Threat Modeling: A Summary of Available Methods," 2018.

[170] M. Deng, K. Wuyts, R. Scandariato, B. Preneel, and W. Joosen, "A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements," *Requirements Engineering*, vol. 16, pp. 3–32, 2011.

[171] T. UcedaVélez, "Threat Modeling w/PASTA: Risk Centric Threat Modeling Case Studies," tech. rep., 2017. Technical Report, Open Web Application Security Project (OWASP).

[172] K. Tuma, G. Çalikli, and R. Scandariato, "Threat analysis of software systems: A systematic literature review," *J. Syst. Softw.*, vol. 144, pp. 275–294, 2018.

[173] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, "Blockchain for Secure EHRs Sharing of Mobile Cloud Based E-Health Systems," *IEEE Access*, vol. 7, pp. 66792–66806, 2019.

[174] J. Liu, X. Li, L. Ye, H. Zhang, X. Du, and M. Guizani, "BPDS: A blockchain based Privacy-Preserving data sharing for electronic medical records," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2018.

[175] Z. Xiao, Z. Li, Y. Liu, L. Feng, W. Zhang, T. Lertwuthikarn, and R. S. Mong Goh, "EMRShare: A Cross-Organizational medical data sharing and management framework using permissioned blockchain," in *2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS)*, pp. 998–1003, 2018.

[176] A. Theodouli, S. Arakliotis, K. Moschou, K. Votis, and D. Tzovaras, "On the design of a Blockchain-Based system to facilitate healthcare data sharing," in *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, pp. 1374–1379, 2018.

[177] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, "MedRec: Using blockchain for medical data access and permission management," in *2016 2nd International Conference on Open and Big Data (OBD)*, pp. 25–30, 2016.

[178] E. Zaghloul, T. Li, and J. Ren, "Security and Privacy of Electronic Health Records: Decentralized and Hierarchical Data Sharing using Smart Contracts," in *2019 International Conference on Computing, Networking and Communications (ICNC)*, pp. 375–379, 2019.

[179] P. Zhang, J. White, D. C. Schmidt, G. Lenz, and S. T. Rosenbloom, "FHIRChain: Applying blockchain to securely and scalably share clinical data." July 2018.

[180] M. Hanley and H. Tewari, "Managing Lifetime Healthcare Data on the Blockchain," in *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pp. 246–251, 2018.

[181] Edward Y. Chang, Shih-Wei Liao, Chun-Ting Liu, Wei-Chen Lin, Pin-Wei Liao, Wei-Kang Fu, Chung-Huan Mei, and Emily J. Chang, "DeepLinQ: Distributed Multi-Layer ledgers for Privacy-Preserving data sharing," *2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, pp. 173–178, 2018.

[182] Y. Wang, A. Zhang, P. Zhang, and H. Wang, "Cloud-Assisted EHR Sharing With Security and Privacy Preservation via Consortium Blockchain," *IEEE Access*, vol. 7, pp. 136704–136719, 2019.

[183] H. Li, L. Zhu, M. Shen, F. Gao, X. Tao, and S. Liu, "Blockchain-Based Data Preservation System for Medical Data," *Journal of medical systems*, vol. 42, no. 8, p. 141, 2018.

[184] A. Zhang and X. Lin, "Towards Secure and Privacy-Preserving Data Sharing in e-Health Systems via Consortium Blockchain," *Journal of medical systems*, vol. 42, no. 8, p. 140, 2018.

[185] Y. Wang and M. He, "CPDS: A cross-blockchain based privacy-preserving data sharing for electronic health records," in *2021 IEEE 6th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*, IEEE, Apr. 2021.

[186] J. G. L. A. Jayasinghe, K. G. S. Shiranthaka, T. Kavith, M. H. D. V. Jayasinghe, K. Y. Abeywardena, and K. Yapa, "Blockchain-based secure

environment for electronic health records," in *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, IEEE, Oct. 2022.

[187] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, "A cooperative architecture of data offloading and sharing for smart healthcare with blockchain," in *2021 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, IEEE, May 2021.

[188] I. Boumezbeur and K. Zarour, "Blockchain-Based Electronic Health Records Sharing Scheme with Data Privacy Verifiable," *Applied Medical Informatics*, vol. 43, no. 4, pp. 124–135, 2021.

[189] A. Gupta, R. Rodrigues, A. Tripathi, R. Coutinho, and J. Gomes, "Blockchain for EHR: an off-chain based approach," in *2022 IEEE Region 10 Symposium (TENSYMP)*, IEEE, July 2022.

[190] S. Sabu, H. M. Ramalingam, M. Vishaka, H. R. Swapna, and S. Hegde, "Implementation of a secure and privacy-aware E-Health record and IoT data sharing using blockchain," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 429–433, 2021.

[191] R. K. Lomotey, S. Kumi, and R. Deters, "Data Trusts as a Service: Providing a platform for multi-party data sharing," *International Journal of Information Management Data Insights*, vol. 2, no. 1, p. 100075, 2022.

[192] T. T. Thwin and S. Vasupongayya, "Blockchain Based Secret-Data Sharing Model for Personal Health Record System," in *2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA)*, pp. 196–201, 2018.

[193] L. Zhang, T. Zhang, Q. Wu, Y. Mu, and F. Rezaeibagha, "Secure Decentralized Attribute-Based Sharing of Personal Health Records With Blockchain," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 12482–12496, 2022.

[194] S. Lee, J. Kim, Y. Kwon, T. Kim, and S. Cho, "Privacy Preservation in Patient Information Exchange Systems Based on Blockchain: System Design

Study," *Journal of Medical Internet Research*, vol. 24, no. 3, pp. N.PAG–N.PAG, 2022.

[195] R. Zou, X. Lv, and J. Zhao, "SPChain: Blockchain-based medical data sharing and privacy-preserving ehealth system," *Information Processing & Management*, vol. 58, no. 4, p. 102604, 2021.

[196] E. Zaghloul, T. Li, M. W. Mutka, and J. Ren, "*d*d-MABE: Distributed multilevel Attribute-Based EMR management and applications," *IEEE Transactions on Services Computing*, vol. 15, no. 3, pp. 1592–1605, 2022.

[197] H.-A. Lee, H.-H. Kung, J. G. Udayasankaran, B. Kijsanayotin, A. B. Marcelo, L. R. Chao, C.-Y. Hsu, and A. B Marcelo, "An Architecture and Management Platform for Blockchain-Based Personal Health Record Exchange: Development and Usability Study," *Journal of Medical Internet Research*, vol. 22, no. 6, pp. N.PAG–N.PAG, 2020.

[198] F. Zhao, J. Yu, and B. Yan, "Towards cross-chain access control model for medical data sharing," *Procedia Computer Science*, vol. 202, pp. 330–335, 2022.

[199] L. Li, Z. Yue, and G. Wu, "Electronic medical record sharing system based on hyperledger fabric and InterPlanetary file system," in *2021 The 5th International Conference on Compute and Data Analysis*, (New York, NY, USA), ACM, Feb. 2021.

[200] D. Hawig, C. Zhou, S. Fuhrhop, A. S. Fialho, and N. Ramachandran, "Designing a Distributed Ledger Technology System for Interoperable and General Data Protection Regulation-Compliant Health Data Exchange: A Use Case in Blood Glucose Data," *J Med Internet Res*, vol. 21, no. 6, p. e13665, 2019.

[201] X. Zheng, R. R. Mukkamala, R. Vatrapu, and J. Ordieres-Mere, "Blockchain-based Personal Health Data Sharing System Using Cloud Storage," in *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 1–6, 2018.

[202] T. Zhou, X. Li, and H. Zhao, "Med-PPPHIS: Blockchain-Based Personal Healthcare Information System for National Physique Monitoring and Scientific Exercise Guiding," *Journal of medical systems*, vol. 43, no. 9, p. 305, 2019.

[203] H. Huang, P. Zhu, F. Xiao, X. Sun, and Q. Huang, "A blockchain-based scheme for privacy-preserving and secure sharing of medical data," *Computers & Security*, vol. 99, p. 102010, 2020.

[204] C. Hu, C. Li, G. Zhang, Z. Lei, M. Shah, Y. Zhang, C. Xing, J. Jiang, and R. Bao, "CrowdMed-II: a blockchain-based framework for efficient consent management in health data sharing," *World Wide Web*, vol. 25, no. 3, pp. 1489–1515, 2022.

[205] R. Wang, W.-T. Tsai, J. He, C. Liu, Q. Li, and E. Deng, *A Medical Data Sharing Platform Based On Permissioned Blockchains*, pp. 12–16.

[206] Q. Qin, B. Jin, and Y. Liu, "A Secure Storage and Sharing Scheme of Stroke Electronic Medical Records Based on Consortium Blockchain," *BioMed Research International*, pp. 1–14, 2021.

[207] E.-Y. Daraghmi, Y.-A. Daraghmi, and S.-M. Yuan, "MedChain: A Design of Blockchain-Based System for Medical Records Access and Permissions Management," *IEEE Access*, vol. 7, pp. 164595–164613, 2019.

[208] Gaby G. Dagher, Jordan Mohler, Matea Milojkovic, and Praneeth Babu Marella, "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology," *Sustainable Cities and Society*, vol. 39, pp. 283–297, 2018.

[209] H. Jin, Y. Luo, P. Li, and J. Mathew, "A Review of Secure and Privacy-Preserving Medical Data Sharing," *IEEE Access*, vol. 7, pp. 61656–61669, 2019.

[210] K. Wüst and A. Gervais, "Do you need a blockchain?," in *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)*, pp. 45–54, 2018.

[211] B. Schneier, *Angewandte Kryptographie - Protokolle, Algorithmen und Sourcecode in C: der Klassiker*. Pearson Education, 2006.

[212] Y.-L. Lee, H.-A. Lee, C.-Y. Hsu, H.-H. Kung, and H.-W. Chiu, "SEMRES - A Triple Security Protected Blockchain Based Medical Record Exchange Structure," *Computer Methods and Programs in Biomedicine*, vol. 215, p. 106595, 2022.

[213] G. Lin, H. Wang, J. Wan, L. Zhang, and J. Huang, "A blockchain-based fine-grained data sharing scheme for e-healthcare system," *J. Syst. Arch.*, vol. 132, p. 102731, Nov. 2022.

[214] W. A. Al-Hamdani, *Cryptography Based Access Control in Healthcare Web Systems*, pp. 66–79.

[215] D. Ferraiolo, D. Kuhn, and R. Chandramouli, *Role-based Access Control*. Artech House computer security series, Artech House, 2003.

[216] S. Rouhani and R. Deters, "Blockchain based access control systems: State of the art and challenges," in *IEEE/WIC/ACM International Conference on Web Intelligence*, WI '19, (New York, NY, USA), p. 423–428, Association for Computing Machinery, 2019.

[217] J. Hughes and E. Maler, "Security assertion markup language (saml) v2. 0 technical overview," *OASIS SSTC Working Draft sstc-saml-tech-overview-2.0-draft-08*, vol. 13, p. 12, 2005.

[218] R. L. Morgan, S. Cantor, S. Carmody, W. Hoehn, and K. Klingenstein, "Federated security: The shibboleth approach," *EDUCAUSE quarterly*, vol. 27, no. 4, pp. 12–17, 2004.

[219] R. C. Nickerson, U. Varshney, and J. Muntermann, "A method for taxonomy development and its application in information systems," *European Journal of Information Systems*, vol. 22, no. 3, pp. 336–359, 2013.

[220] K. DeSalvo, "Connecting health and care for the nation: a shared nation-wide interoperability roadmap Draft Version 1.0," *Office of the National Coordinator for Health Information Technology*, 2015.

[221] T. McGhin, K.-K. R. Choo, C. Z. Liu, and D. He, "Blockchain in healthcare applications: Research challenges and opportunities," *Journal of Network and Computer Applications*, vol. 135, pp. 62–75, 2019.

[222] X. Xu, I. Weber, M. Staples, L. Zhu, J. Bosch, L. Bass, C. Pautasso, and P. Rimba, "A Taxonomy of Blockchain-Based Systems for Architecture Design," in *2017 IEEE International Conference on Software Architecture (ICSA)*, pp. 243–252, 2017.

[223] C. Pahl, N. E. Ioini, and S. Helmer, "A Decision Framework for Blockchain Platforms for IoT and Edge Computing," in *International Conference on Internet of Things, Big Data and Security*, 2018.

[224] T. Koens and E. Poll, "What Blockchain Alternative Do You Need?," in *DPM/CBT@ESORICS*, 2018.

[225] S. Meunier, "When do you need blockchain? Decision models.," 2019. Online verfügbar unter `https://medium.com/@sbmeunier/when-do-you-need-blockchain-decision-models-a5c40e7c9ba1` (letzter Zugriff 12.02.2025).

[226] S. K. Lo, X. Xu, Y. K. Chiam, and Q. Lu, "Evaluating Suitability of Applying Blockchain," in *2017 22nd International Conference on Engineering of Complex Computer Systems (ICECCS)*, pp. 158–161, 2017.

[227] S. Hu, M. Schmidt-Kraepelin, S. Thiebes, and A. Sunyaev, "Mapping Distributed Ledger Technology Characteristics to Use Cases in Healthcare: A Structured Literature Review," *ACM Trans. Comput. Healthcare*, jul 2024. Just Accepted.

[228] P. Coley, "MoSCoW Prioritisation." Online verfügbar unter `https://www.coleyconsulting.co.uk/moscow.htm` (letzter Zugriff 08.11.2024).

[229] gematik GmbH, "Übergreifende Spezifikation - Performance und Mengengerüst TI-Plattform." Online verfügbar unter `https://gemspec.gematik.de/downloads/gemSpec/gemSpec_Perf/gemSpec_Perf_V2.29.0.pdf` (letzter Zugriff 11.04.2025).

[230] J. Hunker and C. W. Probst, “Insiders and Insider Threats - An Overview of Definitions and Mitigation Techniques,” *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, vol. 2, pp. 4–27, 2011.

[231] M. Saad, J. Spaulding, L. L. Njilla, C. A. Kamhoua, S. S. Shetty, D. Nyang, and A. Mohaisen, “Exploring the Attack Surface of Blockchain: A Systematic Overview,” *ArXiv*, vol. abs/1904.03487, 2019.

[232] Y. Liu, Q. Lu, H. young Paik, X. Xu, S. Chen, and L. Zhu, “Design Pattern as a Service for Blockchain-Based Self-Sovereign Identity,” *IEEE Software*, vol. 37, pp. 30–36, 2020.

[233] FZI Forschungszentrum Informatik, “Aktuell.” Online verfügbar unter `https://www.blog3.de/blog/` (letzter Zugriff 11.04.2025).

[234] K. Cuhls, “Die Delphi-Methode–eine Einführung,” *Delphi-Verfahren in den Sozial-und Gesundheitswissenschaften: Konzept, Varianten und Anwendungsbeispiele*, pp. 3–31, 2019.

[235] A. Bangor, P. T. Kortum, and J. T. Miller, “Determining what individual SUS scores mean: adding an adjective rating scale,” *Journal of Usability Studies archive*, vol. 4, pp. 114–123, 2009.

[236] G. Richter, C. Borzikowsky, B. F. Hoyer, M. Laudes, and M. Krawczak, “Secondary research use of personal medical data: patient attitudes towards data donation,” *BMC medical ethics*, vol. 22, no. 1, p. 164, 2021.

[237] T. Arlinghaus, K. Kus, P. Kajüter, and F. Teuteberg, “Designing Data Trustees: Status quo and Perspectives for Business Models,” *HMD Praxis der Wirtschaftsinformatik*, vol. 58, pp. 565–579, 2021.

[238] M. C. Lacity, “Addressing Key Challenges to Making Enterprise Blockchain Applications a Reality,” *MIS Q. Executive*, vol. 17, p. 3, 2018.