Ausbau des Probandenmanagementsystems der Universitätsmedizin Göttingen für standortübergreifende Kooperationen in der klinischen Forschung

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

„Doctor rerum naturalium“

der Georg-August-Universität Göttingen

im Promotionsprogramm in Computer Science (PCS)

der Georg-August University School of Science (GAUSS)

vorgelegt von

Johannes Pung

aus Trier

Göttingen, 2022

Betreuungsausschuss:

Prof. Dr. Otto Rienhoff
Institut für Medizinische Informatik, Universitätsmedizin Göttingen

Prof. Dr. Dieter Hogrefe
Institut für Informatik, Universität Göttingen

Prof. Dr. Dagmar Krefting
Institut für Medizinische Informatik, Universitätsmedizin Göttingen

Mitglieder der Prüfungskommission:

Referent:

Prof. Dr. Otto Rienhoff
Institut für Medizinische Informatik, Universitätsmedizin Göttingen

Korreferent:

Prof. Dr. Dieter Hogrefe
Institut für Informatik, Universität Göttingen

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Dagmar Krefting
Institut für Medizinische Informatik, Universitätsmedizin Göttingen

Priv.-Doz. Dr. Sara Y. Nußbeck
Zentrale Biobank UMG, Universitätsmedizin Göttingen

Prof. Dr. Ulrich Sax
Institut für Medizinische Informatik, Universitätsmedizin Göttingen

Prof. Dr. Philipp Wieder
Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

Tag der mündlichen Prüfung: 07.07.2022

Zusammenfassung

Klinische Forschung ist die Grundlage für den Fortschritt in der Medizin. Ihr wichtigstes Instrument sind klinische Studien, in denen an und mit Teilnehmenden geprüft wird, wie wirksam, verträglich und sicher neue Medikamente und Behandlungsmöglichkeiten sind. Bei vielen wissenschaftlichen Fragestellungen ist es sinnvoll, in standortübergreifenden Kooperationen gemeinsam zu forschen. Dies stellt hohe Anforderungen an die Organisation, die Kommunikation und die Arbeitsprozesse, aber auch an die eingesetzten IT-Systeme. So ist die standortübergreifende Verwaltung von Studienteilnehmenden mit einem Probandenmanagementsystem eine große Herausforderung. Besonders die zentrale Verarbeitung der sensiblen Daten der Studienteilnehmenden von unterschiedlichen Studienstandorten beeinflusst die Datenschutz- und IT-Sicherheitsanforderungen. Dabei ist auch das Zusammenspiel mit weiteren IT-Systemen relevant, weshalb die Aufgaben und Funktionen des Probandenmanagementsystems klar benannt und dessen Rolle in einer Forschungs-IT-Infrastruktur geklärt sein muss. In dieser Dissertation wurde der Ausbau des 2015 am Institut für Medizinische Informatik der Universitätsmedizin Göttingen entwickelten Probandenmanagementsystems für die Nutzung in standortübergreifenden Kooperationen in der klinischen Forschung untersucht.

In zwei systematischen Literaturrecherchen wurden Literaturquellen mit dem Ziel der Begriffs- und Funktionsbestimmung des Probandenmanagementsystems und der Einordnung des Systems in einer Forschungs-IT-Infrastruktur untersucht. Vier Hauptprozesse der Probandenverwaltung wurden identifiziert: Rekrutierung (1), Erfassung und Verwaltung der Einwilligungen (2), Pseudonymisierung (3) und Verwaltung administrativer Daten (4) der Studienteilnehmenden. Die Prozesse können durch IT-Systeme unterstützt werden, die untereinander über Schnittstellen interagieren und so in einer Forschungs-IT-Infrastruktur zusammenarbeiten. Dabei ist die Zusammensetzung der IT-Systeme abhängig von den Anforderungen des Forschungsprojektes. Das bestehende Göttinger Probandenmanagementsystem vereint die IT-Unterstützung der genannten Hauptprozesse ausgenommen der Rekrutierung und übernimmt damit zentrale Aufgaben und Funktionen in einer Forschungs-IT-Infrastruktur. Zur Erarbeitung eines Konzeptes zum Ausbau des Probandenmanagementsystems wurden die Anforderung an ein in standortübergreifenden Kooperationen nutzbaren Systems identifiziert. Das

Konzept gliedert sich angelehnt an die identifizierten Anforderungen in die Teile: Datentrennung bei der Probandenverwaltung (1), vernetze Probandenverwaltung (2), Zugangs- und Zugriffskontrolle (3) und Datenschutzkonzept (4). Das Konzept sieht die Anwendung von Grundprinzipien datenschutzgerechter Lösungen wie Datensparsamkeit und informationelle Gewaltenteilung vor. Um eine strikte Datentrennung bei der Probandenverwaltung zu gewährleisten, dürfen nur ausgewählte Pseudonyme in der Datenbank des Probandenmanagementsystems gespeichert werden. Das Anlegen neuer Pseudonyme sowie die Abfrage bestehender Pseudonyme wird über eine Schnittstelle zum Pseudonymisierungsdienst realisiert. Der sichere Zugriff auf Daten und Funktionen wird durch technische Datenschutzmaßnahmen zur Zugangs- und Zugriffskontrolle gewährleistet. Vorbereitend auf ein konkretes Forschungsprojekt wurden alle für ein Datenschutzkonzept relevanten Punkte zusammengetragen. Die konzeptionierten Maßnahmen wurden prototypisch implementiert, um so die Machbarkeit des Konzeptes zu bestätigen.

Standortübergreifende Kooperationen in der klinischen Forschung stellen hohe Anforderungen an ein Probandenmanagementsystem. Werden Datenschutz- und IT-Sicherheitsanforderungen nicht berücksichtigt und getroffene Maßnahmen in einem Datenschutzkonzept nicht ausreichend begründet, können nützliche IT-Systeme in der klinischen Forschung nicht eingesetzt werden. Das Konzept für die standortübergreifende Nutzung des Probandenmanagementsystems setzt an diesem Punkt an, berücksichtigt etablierte Datenschutzmodelle und setzt aktuelle IT-Sicherheitsempfehlungen um. So kann das Probandenmanagementsystem zur klinischen Forschung beitragen und Prozesse der Probandenverwaltung unterstützen. Damit werden Ressourcen in klinischen Forschungsprojekten effizient genutzt und neue Therapien können schneller in die Versorgung gelangen.

Abstract

Clinical research is key for medical progress. Clinical trials are the most important instrument of clinical research, which are used to test efficacy and safety of new drugs and new forms of treatment on human participants. Joint research in multi-site collaborations are of particular importance to enable significant research results. Multi-site research imposes high demands on organization, communication and work processes, but also on the IT systems used. The multi-site administration of study participants with a Study Participant Management System is a major challenge. In particular, the central processing of sensitive data of study participants from different study sites determines the data protection and data security requirements. Further, the interaction with other IT systems is highly relevant, which is why the tasks and functions of the Study Participant Management System must be clearly specified and its role in a research IT infrastructure clarified. In this dissertation, the extension of the Study Participant Management System developed in 2015 at the Department of Medical Informatics, University Medical Center Göttingen, for the use in multi-site collaborations in clinical research, was investigated.
Two systematic literature reviews were conducted with the goal to define the term Study Participant Management System and to specify its functions and how the system integrates into a research IT infrastructure. Four main participant-related management processes were identified: recruitment (1), obtaining and managing informed consent (2), pseudonymization (3), and management of administrative data (4) of study participants. These processes can be supported by IT systems that interact with each other via interfaces and thus collaborate in a research IT infrastructure. The composition of the IT systems depends on the requirements of the research project. The Study Participant Management System from Göttingen combines the computer-supported management for the main processes, except for recruitment, and thus undertakes central tasks and functions in a research IT infrastructure. To develop a concept for the extension of the Study Participant Management System, the requirements for a system that can be used in multi-site collaborations were identified. Based on the identified requirements, the concept was divided into the following sections: data separation in participant management (1), interconnected participant management (2), access control (3) and data protection concept (4). The concept for the system extension provides the implementation of basic principles

of data protection such as privacy by design and informational separation of powers. To ensure strict data separation in participant management, only selected pseudonyms may be stored in the database of the Study Participant Management System. Requests of new pseudonyms as well as the queries of existing pseudonyms are realized via an interface to the pseudonymization service. Secure access to data and functions is ensured by technical data protection measures for access control. In preparation for tangible research projects, all relevant aspects for a data protection concept were composed. The conceptualized measures were implemented prototypically to confirm the feasibility of the concept.

Multi-site collaborations in clinical research place high demands on a Study Participant Management System. If data protection and IT security requirements are not taken into account and measures taken are not sufficiently justified in a data protection concept, useful IT systems cannot be used in clinical research. The concept for the multi-site use of the Study Participant Management System addresses this issue, takes into account established data protection models and implements current IT security recommendations. Thereby, the Study Participant Management System can contribute to clinical research and support trial processes by managing participants' data. As a result, resources in clinical research projects are used efficiently and new therapies can be introduced into medical care faster.

Inhaltsverzeichnis

1 Einleitung **1**
1.1 Problemvorstellung . . . 1
1.2 Fragestellungen und Ziele dieser Arbeit . . . 5
1.3 Methodik und theoretische Vorüberlegungen . . . 6

2 Begriffsbestimmung und Einordnung des Probandenmanagementsystems im Forschungsumfeld **8**
2.1 Begriff des Probandenmanagementsystems . . . 9
2.1.1 Begriffsherkunft . . . 11
2.1.2 Begriffsdefinition . . . 14
2.2 Prozesse der Probandenverwaltung . . . 15
2.2.1 Rekrutierung von Studienteilnehmenden . . . 16
2.2.2 Erfassung und Verwaltung der Einwilligungen zur Studienteilnahme 17
2.2.3 Führen der Patienten-/Probandenliste und Pseudonymisierung . . . 18
2.2.4 Verwaltung administrativer Probandendaten . . . 19
2.3 IT-gestützte Probandenverwaltung in einer Forschungs-IT-Infrastruktur . . 20
2.3.1 Forschungsdatenerfassung . . . 21
2.3.2 Rekrutierungsunterstützung . . . 22
2.3.3 Einwilligungsverwaltung . . . 23
2.3.4 Pseudonymisierung . . . 24
2.3.5 Verwaltung administrativer Probandendaten . . . 25

3 Konzeption des standortübergreifenden Probandenmanagementsystems **27**
3.1 Methodik und Grundlagen . . . 28
3.1.1 Methodik: Design and Creation . . . 28
3.1.2 Rahmenbedingungen und Umfeld . . . 30
3.1.3 Datenschutzkonzept . . . 31
3.2 Anforderungen an eine standortübergreifende Probandenverwaltung . . . 32
3.3 Iteration 1: Datentrennung bei der Probandenverwaltung . . . 36
3.3.1 Problem und Zielsetzung . . . 36
3.3.2 Methodik . . . 38
3.3.3 Ergebnisse . . . 41
3.3.4 Diskussion . . . 50
3.3.5 Schlussfolgerung . . . 51

3.4 Iteration 2: Vernetzte Probandenverwaltung . . . 53
3.4.1 Problem und Zielsetzung . . . 53
3.4.2 Methodik . . . 54
3.4.3 Ergebnisse . . . 55
3.4.4 Diskussion . . . 61
3.4.5 Schlussfolgerung . . . 62
3.5 Iteration 3: Zugangs- und Zugriffskontrolle . . . 63
3.5.1 Problem und Zielsetzung . . . 63
3.5.2 Methodik . . . 65
3.5.3 Ergebnisse . . . 72
3.5.4 Diskussion . . . 88
3.5.5 Schlussfolgerung . . . 91
3.6 Iteration 4: Datenschutzkonzept . . . 92
3.6.1 Problem und Zielsetzung . . . 92
3.6.2 Methodik . . . 93
3.6.3 Ergebnisse . . . 94
3.6.4 Diskussion . . . 98
3.6.5 Schlussfolgerung . . . 99

4 Prototypische Implementierung des Probandenmanagementsystems 100
4.1 Grundlagen . . . 101
4.1.1 Software- und Systemarchitektur des Probandenmanagementsystems 101
4.1.2 Softwarestack des Probandenmanagementsystems . . . 105
4.1.3 Interoperabilität . . . 106
4.2 Datentrennung und vernetzte Probandenverwaltung . . . 107
4.2.1 Datenmodell . . . 107
4.2.2 Schnittstelle . . . 117
4.3 Zugangs- und Zugriffskontrolle . . . 124
4.3.1 Sichere Authentifizierung . . . 124
4.3.2 Verschlüsselte Verbindung . . . 130
4.3.3 Zugriffskontrolle . . . 132
4.3.4 Protokollierter Zugriff . . . 140

5 Diskussion und Ausblick 145
5.1 Diskussion . . . 145
5.2 Ausblick . . . 150

Literaturverzeichnis 153

Abkürzungsverzeichnis 162

Abbildungs- und Listingsverzeichnis 165

Anhang 169

Kapitel 1

Einleitung

1.1 Problemvorstellung

Klinische Forschung ist ein wichtiges Instrument der medizinischen Weiterentwicklung und des medizinischen Fortschritts. Im Rahmen von klinischen Studien werden Medikamente, Behandlungsformen, medizinische Interventionen oder Medizinprodukte auf ihre Wirksamkeit und Sicherheit überprüft. Die Durchführung klinischer Studien unterliegt dabei umfassenden ethischen und gesetzlichen Bestimmungen. Zu Beginn einer Studie ist es notwendig, Probandinnen und Probanden[1] zu rekrutieren und sie über die geplante Forschung und ihre Ziele aufzuklären. Zu den Aufgaben eines Studienzentrums gehört die Verwaltung von Kontaktdaten und der Einwilligungserklärung der Studienteilnehmenden zur Teilnahme an der Studie sowie die Terminplanung und Organisation von Untersuchungen im Studienzusammenhang. Die Verwaltung von Studienteilnehmenden ist somit Kernkomponente jedes klinischen Forschungsprojektes.
Die wissenschaftliche Qualität und Aussagekraft einer klinischen Studie wird wesentlich durch das Studiendesign bestimmt. Die Fallzahlplanung ist dabei ein entscheidender Teil. In medizinischen Forschungsvorhaben ist es oft notwendig, eine große Anzahl von

[1] Proband = Versuchs- oder Testperson. In der medizinischen Forschung ein gesunder Teilnehmer an einer klinischen Studie im Unterschied zu einem erkrankten Studienteilnehmer (Patient). Abseits der Definition jedoch weniger trennscharfe Aufteilung; häufig Proband als Oberbegriff für gesunde oder erkrankte Versuchs- oder Testperson. In dieser Arbeit wird im Folgenden vorrangig außerhalb feststehender Begriffe die geschlechtsneutrale Bezeichnung *Studienteilnehmende* verwendet.

Studienteilnehmenden zu rekrutieren, um signifikante Forschungsergebnisse erzielen zu können. Die Fallzahl ist jedoch häufig durch Faktoren wie die Dauer der Rekrutierungszeit, eine geringe Anzahl an Patientinnen und Patienten die den Einschlusskriterien der klinischen Studie genügen oder die Förderdauer der Studie begrenzt. Rekrutierungsprobleme und die überregionale Verteilung von Studienteilnehmenden erschweren die Durchführung von klinischen Studien. Um klinische Studien an geeigneten Kollektiven durchzuführen, ist es daher notwendig, in überregionalen und internationalen Forschungsverbünden zu forschen. Unter Verbundforschung wird die einrichtungs-, standort- und disziplinübergreifende Zusammenarbeit von Projektpartnern zur gemeinsamen Durchführung von Forschungsprojekten verstanden. Die Verbünde bestehen aus mehreren Einzelprojekten, die mit Blick auf eine gemeinsame übergeordnete Fragestellung zusammenarbeiten. Medizinische Verbundforschungsprojekte sind aufgrund ihrer Größe besonders auf eine effiziente Verwaltung ihrer Studienteilnehmenden angewiesen. Unterstützung der Prozesse zur Verwaltung von Studienteilnehmenden können IT-Systeme bieten.

Geeignete Systeme sind dafür Probandenmanagementsysteme, die in ihrem Funktionsumfang verschiedene Prozesse zur Verwaltung von Studienteilnehmenden abbilden. Am Institut für Medizinische Informatik der Universitätsmedizin Göttingen erfasste Jens Schwanke im Rahmen seiner Dissertation „Konzeption und Entwicklung eines Probandenmanagementsystems am Beispiel der Universitätsmedizin Göttingen“ Anforderungen an ein Probandenmanagementsystem und entwickelte einen funktionsfähigen Prototyp für Forschungsprojekte an der Universitätsmedizin Göttingen [Schwanke, 2015]. Da in einer vorangegangenen Analyse kein adäquates System zur Verwaltung von Studienteilnehmenden gefunden werden konnte, wurden die Anforderungen an ein solches System ermittelt [Schwanke u. a., 2013b]. Die durchgeführte Anforderungsanalyse umfasste sowohl Interviews mit Forschenden eines klinischen Forschungsprojektes als auch die Aufarbeitung relevanter Literaturquellen [vgl. Schwanke, 2015, S. 6, S. 8 f.]. Das daraufhin entwickelte und in einem Forschungsprojekt in Göttingen eingesetzte Probandenmanagementsystem unterstützt die Forschenden bei der Verwaltung der im Forschungsumfeld pseudonymisierten Identitäten der Studienteilnehmenden und bietet eine Prozessunterstützung bei der Aufgaben- und Terminplanung innerhalb einer klinischen Studie. Die Nutzung des Probandenmanagementsystems

in großen Verbundforschungsprojekten nennt Schwanke als Ansatzpunkt für zukünftige Untersuchungen [vgl. Schwanke, 2015, Kapitel 8.10 Ausblick, S. 97 f.]. Medizinische Verbundforschungsprojekte, die sich durch eine einrichtungs- und standortübergreifende Zusammenarbeit von Projektpartnern auszeichnen, sind eine Herausforderung für den Einsatz eines solchen Probandenmanagementsystems.

Aus IT-Sicht erfordert der Aufbau von Forschungsverbünden häufig eine zentrale Zusammenführung und Speicherung von Daten jenseits der einzelnen Ursprungsinstitutionen. Wo immer medizinischen Daten außerhalb der datenerhebenden Institution verarbeitet werden, müssen Datenschutzanforderungen in besonderem Maße bei der Konzeption einer Forschungs-IT-Infrastruktur und ihrer Systeme berücksichtigt werden. Besonders betroffen davon ist das Probandenmanagementsystem, dessen Kernaufgabe die Verwaltung der besonders schützenswerten und sensiblen Daten der Studienteilnehmenden ist. Im Normalfall ist die Kenntnis der identifizierenden Daten nur im direkten Behandlungs- oder Studienzusammenhang zulässig. Das bedeutet, dass identifizierende Daten die datenerhebende Institution im Allgemeinen nicht verlassen dürfen. Spezielle Datenschutzkonzepte mit besonderer Begründung und sorgfältig gewählten Schutzmaßnahmen können jedoch eine Ausnahme bilden. Der Ausbau des Probandenmanagementsystems für standortübergreifende Kooperationen in der klinischen Forschung erfordert eine dezidierte Betrachtung der aktuellen Datenschutzanforderungen, Prüfung möglicher Konzepte und das Ableiten neuer Anforderungen an das System.

Elementarer Teil des Datenschutzes ist die Datensicherheit. Damit die personenbezogenen Daten ausreichend geschützt sind, müssen entsprechende technische Maßnahmen für ihre Sicherheit ergriffen werden. Bisher abgeschottet durch den Betrieb im geschützten Netzwerksegment der Universitätsmedizin Göttingen, ist das Probandenmanagementsystem aus dem Internet nicht erreichbar und somit vor Zugriffen außerhalb der Klinik geschützt. Die Abschottung bedingt jedoch auch, dass kein zentrales Identitätsmanagement, wie es in Verbundforschungsprojekten häufig üblich ist, außerhalb der Klinik über eine Webschnittstelle angebunden werden kann. Zudem sind Zugänge für Systembenutzende außerhalb der eigenen Einrichtung, wie sie für Verbundforschungsprojekte mit mehreren Standorten benötigt werden, ebenfalls nicht möglich. Beim Einsatz des Probandenmanagementsystems in Verbundforschungsprojekten kann das System also nicht länger im abgeschotteten

Netzwerksegment der Universitätsmedizin Göttingen bleiben, da Schnittstellen zu anderen Systemen und Standorten geschaffen werden müssen. Daher ergeben sich neue Anforderungen an die Absicherung eines Systems, dass nun über das Internet zu einem möglichen Angriffsziel für Kriminelle wird. Doch auch innerhalb des Probandenmanagementsystems sind Sicherheitsanforderungen zu beachten, um Datenmissbrauch zu vermeiden. Es müssen Maßnahmen getroffen werden, die verhindern, dass Systembenutzende Zugriff auf Daten erhalten, die nicht für sie bestimmt sind. Mit dem Ausbau des Probandenmanagementsystems für standortübergreifende Kooperationen in der klinischen Forschung müssen verstärkt Datensicherheitsanforderungen berücksichtigt werden.

Der Ausbau eines bestehenden Systems bedingt auch dessen genauere Betrachtung. Der Begriff des Probandenmanagementsystems wurde zwar durch die Veröffentlichungen der Medizinischen Informatik in Göttingen geprägt [Schwanke u. a., 2012, 2013a,b; Stahmann u. a., 2014a,b], aber in der Literatur zu Themen wie *IT-Systeme in der klinischen Forschung* oder *Forschungs-IT-Infrastrukturen* nicht aufgegriffen. Im Fokus der Arbeit von Schwanke [2015] stand die Konzeption und Entwicklung eines Systems für ein konkretes Forschungsprojekt, sodass eine Schärfung des Begriffs nicht ausreichend stattgefunden hat. Um *Probandenmanagementsystem* als Fachbegriff in der wissenschaftlichen Diskussion nutzbar zu machen, ist jedoch eine Begriffsschärfung zwingend notwendig. Bleibt eine Begriffsdefinition aus, erschwert dies die Abgrenzung des Probandenmanagementsystems zu anderen in klinischen Forschungsprojekten eingesetzten Systemen und deren Funktionsumfang. Gerade in Verbundforschungsprojekten ergänzen sich unterschiedlich spezialisierte Systeme in ihren Funktionen, um in ihrer Gesamtheit die klinische Forschung zu unterstützen. Neben der Begriffsdefinition muss daher das Probandenmanagementsystem in einer Forschungs-IT-Infrastruktur verortet werden, um das Zusammenspiel mit weiteren Systemen zu beschreiben, Funktionen abzugrenzen und alle Anwendungsfälle zu betrachten.

1.2 Fragestellungen und Ziele dieser Arbeit

Die Verwaltung von Studienteilnehmenden ist eine wichtige Komponente eines klinischen Forschungsprojektes. Unterstützung bei den Prozessen zur Verwaltung von Studienteilnehmenden bietet ein Probandenmanagementsystem. Für seinen Einsatz in klinischen Forschungsverbünden müssen jedoch noch Herausforderungen überwunden werden, die sich aus dem gesteigerten Schutzbedarf bei einer zentralen Verarbeitung der sensiblen Daten von unterschiedlichen Studienstandorten ergeben. Bevor diese Herausforderungen angegangen werden, muss zunächst klargestellt werden, wie der Begriff des Probandenmanagementsystems zu verstehen ist. Ein einheitliches Verständnis von Begrifflichkeiten ist die Voraussetzung, um die Funktionen eines Probandenmanagementsystems zu erläutern und es in Beziehung zu anderen Systemen in einer Forschungs-IT-Infrastruktur zu setzen. Gleichzeitig dienen die Definition und die Einordnung des Probandenmanagementsystems der weiteren Arbeit als wichtige Grundlage. Daraus ergibt sich die erste Forschungsfrage dieser Arbeit:

Forschungsfrage 1

> Was versteht man unter einem Probandenmanagementsystem und wo ist es strukturell innerhalb einer medizinischen Forschungs-IT-Infrastruktur einzuordnen?

Darauf aufbauend können die neuen Anforderungen an ein Probandenmanagementsystem untersucht werden, das in klinischen Verbundforschungsprojekten zum Einsatz kommen soll. Diese Anforderungen unterscheiden sich vom bisherigen Konzept des in Göttingen entwickelten Probandenmanagementsystems für nur einen Studienstandort. Im Vordergrund stehen dabei weniger die Anforderungen der Systembenutzenden an Programmfeatures, denn diese sind nahezu identisch mit den bereits erfassten Anforderungen eines Probandenmanagementsystems für einen Studienstandort. Vielmehr steht die Konzeption eines multizentrisch zum Einsatz kommenden Systems im Vordergrund. Die in der Problemstellung beschriebenen Datenschutz- und Datensicherheitsanforderungen für den Betrieb eines standortübergreifenden Probandenmanagementsystems erfordern neue technische Konzepte. Insbesondere die Verwaltung der identifizierenden Daten von Studienteilnehmenden außerhalb eines Studienstandorts bedingt strenge Datenschutzanforderungen. Es ist zu prüfen, welche Bestimmungen und Regeln gelten. Darauf aufbauend muss ein

datenschutzkonformes, sicheres Probandenmanagementsystem konzipiert werden, das in klinischen Verbundforschungsprojekten zum Einsatz kommen kann. Diese Ziele zusammenfassend soll folgende zweite Forschungsfrage in dieser Arbeit untersucht werden:

Forschungsfrage 2

Wie kann ein standortübergreifendes Probandenmanagementsystem unter Berücksichtigung von Datenschutz- und IT-Sicherheitsanforderungen konzipiert werden?

Dieses Konzept bildet die Voraussetzung für den Ausbau des Probandenmanagementsystems für den standortübergreifenden Einsatz. Dabei muss untersucht werden, wie unter Berücksichtigung der bisher eingesetzten Technologien (Programmiersprache und Software-Frameworks) sowie der Softwarearchitektur das Konzept in einen Softwareentwurf überführt und die wichtigsten Funktionen prototypisch implementiert werden können. Die dritte Forschungsfrage dieser Arbeit lautet daher:

Forschungsfrage 3

Aufbauend auf dem entwickelten Konzept aus Forschungsfrage 2 soll ein standortübergreifendes Probandenmanagementsystem softwaretechnisch konzipiert und prototypisch implementiert werden.

1.3 Methodik und theoretische Vorüberlegungen

Die Beantwortung von Forschungsfrage 1 lässt sich in die Schritte *Begriffsbestimmung des Probandenmanagementsystems* und *Einordnung des Probandenmanagementsystems in eine Forschungs-IT-Infrastruktur* aufteilen. Zur terminologischen Klärung des Begriffs wurden zwei Formen der Literaturanalyse angewandt. Zum einen eine Literaturrecherche über ausgewählte Schlagworte in Suchmaschinen für wissenschaftliche Internet-Quellen, zum anderen eine strukturierte Literaturanalyse auf Basis bereits bekannter Veröffentlichungen des Instituts für Medizinische Informatik in Göttingen und Sichtung der darin zitierten Quellen. Ziel war es, einen möglichst umfassenden Überblick über die vorhandene Literatur zu erlangen, um die Begriffsdefinition des Probandenmanagementsystems zu ermöglichen. Der

Einordnung des Probandenmanagementsystems in eine Forschungs-IT-Infrastruktur ging ebenfalls eine Literaturrecherche voraus. In einer systematischen Literaturanalyse wurden Prozesse der Probandenverwaltung und die sie unterstützende IT-Systeme identifiziert. Darauf aufbauend konnte das Zusammenspiel der probandendatenverwaltenden Systeme in einer Forschungs-IT-Infrastruktur erläutert werden. Die Beantwortung von Forschungsfrage 1 diente der weiteren Arbeit für die Untersuchung des Ausbaus des Probandenmanagementsystems für standortübergreifende Kooperationen in der klinischen Forschung als wichtige Grundlage.

Die Antwort auf Forschungsfrage 2 wurde in Form eines Konzeptes für den Ausbau des Probandenmanagementsystems erarbeitet. Zur Konzeption wurden zunächst die Anforderungen für ein standortübergreifend nutzbares Probandenmanagementsystem erfasst. Dabei wurden Aspekte der Datentrennung bei der Probandenverwaltung, die Vernetzung der Probandenverwaltung, die Zugangs- und Zugriffskontrolle und das Datenschutzkonzept untersucht. In einer iterativen Vorgehensweise wurden anschließend Lösungen erarbeitet, wie die erfassten Anforderungen in einem standortübergreifend nutzbaren Probandenmanagementsystem umgesetzt werden können.

Zur Beantwortung von Forschungsfrage 3 wurde das Konzept des Probandenmanagementsystems in einem Softwareentwurf abgebildet und prototypisch implementiert. Berücksichtigt wurden dabei die bestehende Software- und Systemarchitektur des Probandenmanagementsystems sowie die bisher eingesetzten Software-Frameworks. Umgesetzt wurden die erarbeitenden Konzepte zur Sicherstellung einer datenschutzkonformen Datentrennung und vernetzten Probandenverwaltung sowie der Zugangs- und Zugriffskontrolle für das Probandenmanagementsystem. Die Machbarkeit des erarbeiteten Gesamtkonzeptes konnte mit der Implementierung bestätigt werden. Zur Veranschaulichung der Implementierung wurde ein Demonstrator des standortübergreifend nutzbaren Probandenmanagementsystems entwickelt.

Kapitel 2

Begriffsbestimmung und Einordnung des Probandenmanagementsystems im Forschungsumfeld

Dieses Kapitel widmet sich der Beantwortung der ersten Forschungsfrage dieser Arbeit, die lautet, den Begriffsinhalt des Probandenmanagementsystems zu bestimmen und das Probandenmanagementsystem in einer Forschungs-IT-Infrastruktur einzuordnen. Die Einordnung kann nur gelingen, wenn der Funktionsumfang des Probandenmanagementsystems klar beschrieben ist. Voraussetzung dafür ist ein einheitliches Verständnis von Begrifflichkeiten. Existieren eine Vielzahl von Vorstellungen zu den Funktionalitäten des Probandenmanagementsystems, kann dies zu Missverständnissen führen. Daher beschäftigt sich der erste Teil dieses Kapitels mit dem Begriff des Probandenmanagementsystems. Darauf aufbauend werden im zweiten Teil die Prozesse mit Probandenverwaltungsanteil in Forschungsprojekten betrachtet. Die Bestandsaufnahme der Prozesse, noch losgelöst von einer technischen Unterstützung, dient auch dem weiteren Verständnis der IT-gestützten Probandenverwaltung. Im dritten Teil dieses Kapitels werden IT-Systeme untersucht, die in klinischen Forschungsprojekten im Zusammenhang der Probandenverwaltung eingesetzt werden können. Die Betrachtung des Zusammenspiels der Systeme ermöglicht deren Einordnung in eine Forschungs-IT-Infrastruktur.

2.1 Begriff des Probandenmanagementsystems

Ohmann und Speer [2016] kategorisieren im IT-Report der TMF[1] zur Bestands- und Bedarfsanalyse von IT-Infrastrukturen für die patientenorientierte medizinische Forschung die in Abbildung 2.1 gezeigten Funktionsbereiche einer klinischen Studie.

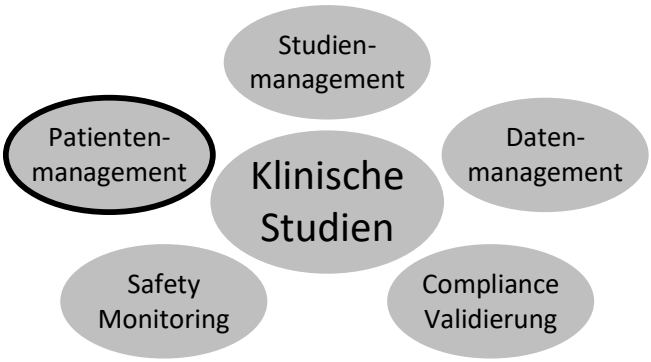

Abbildung 2.1: Funktionsbereiche für die Organisation und Durchführung klinischer Studien. Eigene Abbildung in Anlehnung an [Ohmann und Speer, 2016, S. 30].

Der Funktionsbereich des Patientenmanagements oder der im Zusammenhang von medizinischer Forschung synonym gebrauchten Begriffe des Probandenmanagements oder der Probandenverwaltung steht im Fokus dieses Unterkapitels. Die Formulierung einer Definition wird von Ohmann und Speer [2016] als ein Punkt des Handlungsbedarfs zur Weiterentwicklung von IT-Infrastrukturen für klinische Studien gesehen:

> „Mittelfristig gilt es somit die folgenden Bereiche zu stärken: [...] Definition einer einheitlichen Terminologie zu Probanden- und Einwilligungsverwaltungssystemen mit einer Abgrenzung des Funktionsumfangs und der Anforderungen.“ [Ohmann und Speer, 2016, Bewertung und Handlungsbedarf, S. 42]

Diese Forderung deckt sich mit dem Ziel der ersten Forschungsfrage dieser Arbeit den Begriff des Probandenmanagementsystems zu schärfen, um ihn als Fachbegriff in der wissenschaftlichen Diskussion nutzbar zu machen. Auf der Suche nach einer Begriffsdefinition wurde auf Basis einer einfachen Literaturrecherche der aktuelle Stand[2]

[1] TMF - Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V.: `https://www.tmf-ev.de/` [Stand: 01.03.2022]

[2] Zeitpunkt der Recherche: Frühjahr 2018

zum Thema *Probandenmanagementsystem* erfasst und analysiert. Als geeignete Suchbegriffe wurden die in Tabelle 2.1 gelisteten Schlagworte bestimmt. Die Tabelle umfasst die als relevant angesehenen deutschen und englischsprachigen Begriffe.

Probandenmanagementsystem	Probandenverwaltungssystem
participant management system	study participant management system
trial participant management system	clinical trial participant management system

Tabelle 2.1: Literatursuchbegriffe zum Thema *Probandenmanagementsystem*

Die ausgewählten Schlagworte wurden mittels der Suchmaschine Google Scholar[3] und der Suche von Web of Science[4] untersucht. Die Suche nach dem Schlagwort *Probandenverwaltungssystem* lieferte in beiden Suchmaschinen keine Treffer. Für das Schlagwort *Probandenmanagementsystem* lieferte Google Scholar ausschließlich Treffer im Zusammenhang mit dem am Institut für Medizinische Informatik in Göttingen entwickelten Probandenmanagementsystem. In englischsprachigen Veröffentlichungen der Autoren aus Göttingen wurde der Begriff des *study participant management system* verwendet. Dieser Begriff findet jedoch über die aus Göttingen bekannten Veröffentlichungen hinaus keine weitere Verwendung. Die Suche von Web of Science listet für alle in der Tabelle 2.1 gelisteten Schlagworte lediglich einen einzigen Treffer. In einem Artikel zur Risikoreduktion bei Herz-Kreislauf-Erkrankungen [Gordon u. a., 2002] wurde ein *computerized participant management system* verwendet, zu dem jedoch keine weiteren Details beschrieben wurden.

Es musste festgestellt werden, dass nach einer ersten Recherche keine allgemeingültige Definition für ein Probandenmanagementsystem aus vorhandener Literatur abgeleitet werden kann. Daher wurde versucht, sich dem Begriff des Probandenmanagementsystems mittels einer strukturierten Literaturanalyse zu nähern. Ziel ist hierbei, ausgehend von bereits bekannten Veröffentlichungen und der Sichtung der darin zitierten Quellen, einen möglichst umfassenden Überblick über die vorherrschende Literatur zu erlangen. Als Einstiegspunkt in die strukturierte Literaturanalyse dienten die oben erwähnten Veröffentlichungen der Medizinischen Informatik in Göttingen zum Probandenmanagementsystem. Auf Grundlage dieser ermittelten Literatur befasst

[3] Google Scholar: `https://scholar.google.de/` [Stand: 01.03.2022]

[4] Web of Science: `https://apps.webofknowledge.com/` [Stand: 01.03.2022]

sich das folgende Unterkapitel zur Begriffsherkunft mit der Entstehung des Begriffs *Probandenmanagementsystem*. Im Anschluss wird mit der gewonnenen Erkenntnis im Unterkapitel 2.1.2 (S. 14) eine Definition des Begriffs *Probandenmanagementsystem* vorgeschlagen.

2.1.1 Begriffsherkunft

Die Herkunft des Begriffs *Probandenmanagementsystem* lässt sich auf Veröffentlichungen des Instituts für Medizinische Informatik der Universitätsmedizin Göttingen zurückführen. Alle Publikationen entstanden im Zusammenhang der Arbeit der *Klinischen Forschergruppe*[5] *241* (KFO 241)[6]. Die KFO 241 führte ab dem Jahr 2012 eine longitudinale Studie zur Untersuchung von Psychosen durch [Reich-Erkelenz u. a., 2012]. Sie repräsentiert eine interdisziplinäre Forschungsgruppe aus Grundlagenforschenden und klinisch Forschenden, um die dem Krankheitsverlauf der Schizophrenie und Bipolaren Störung zugrunde liegenden Schlüsselfaktoren aufzudecken.
Auftakt der Entstehung des Begriffs *Probandenmanagementsystem* bildet das 2012 im Fachjournal Molecular Psychiatry erschienene Perspective Paper „Managing sensitive phenotypic data and biomaterial in large-scale collaborative psychiatric genetic research projects: practical considerations“ von Demiroglu u. a. [2012]. Es beschreibt die Bedeutung groß angelegter kooperativer Forschungsprojekte für eine zukünftige psychiatrische Genforschung und stellt Anforderungen an ein zentrales IT-Framework vor. Durch ein modulares IT-Toolbox-Konzept sollen die Forschenden bei ihrer Arbeit an einem Großforschungsprojekt unterstützt werden. Ein Vorläufer der vorgestellten IT-Toolbox wurde bereits zwei Jahre zuvor von Dangl u. a. [2010] in „The IT-Infrastructure of a Biobank for an Academic Medical Center“ zur Erforschung von individualisierten Rektumkarzinom-Therapien der *Klinischen Forschergruppe 179* (KFO 179) in Göttingen beschrieben. Die IT-Toolbox besteht dabei aus fünf Komponenten:

> „The IT toolbox [...] comprises five tools: (1) an identity and pseudonym tool, (2) an administrative tool, (3) a biomaterial tool, (4) a phenotype tool, and

[5] Klinische Forschungsgruppe (früher Klinische Forschergruppe), Förderform der Deutsche Forschungsgemeinschaft (DFG): `https://www.dfg.de/foerderung/programme/koordinierte_programme/klinische_forschungsgruppen/index.html` [Stand: 01.03.2022]

[6] Klinische Forschergruppe 241: `http://www.kfo241.de/` [Stand: 01.03.2022]

(5) a query tool.“ [Demiroglu u. a., 2012, S. 1181]

Alle fünf IT-Tools sollen dabei webbasiert sein, um den beteiligten Forschenden Zugriff von unterschiedlichen Standorten zu ermöglichen. Das Verwaltungstool, gedacht für die Verwaltung einer großen Anzahl an Probanden, wird dabei wie folgt beschrieben:

> „The administrative tool is tightly linked to the identity and pseudonym tool. This is essential as the administrative tool should allow for the management of identifying data and of follow-up dates for local participants. It should enable the presentation of family relations needed for family-based studies, and indicate the status of agreement of the informed consent signed by the participant.“ [Demiroglu u. a., 2012, S. 1181]

Demiroglu u. a. [2012] benennen damit Anforderungen an ein *Probanden-Verwaltungstool*, wie dem Management der identifizierenden Daten und der Termine für lokale Teilnehmende, der Darstellung der familiären Beziehungen und dem Status der Einverständniserklärung der Studienteilnehmenden. Das Verwaltungstool soll das Erstellen von Pseudonymen eines Studienteilnehmenden vereinfachen, indem für die verschiedenen IT-Tools eine einmalige Eingabe der identifizierenden Daten eines Studienteilnehmenden erfolgt. Außerdem erinnert das angedachte Verwaltungstool die Studienleiterin oder den Studienleiter daran, Studienteilnehmende einzuladen und ermöglicht eine Terminverwaltung der Follow-up-Termine [Demiroglu u. a., 2012]. In darauf folgenden Veröffentlichungen wurde stets Bezug auf das Verwaltungstool der IT-Toolbox genommen.

Erstmals Verwendung findet der Begriff *Probandenmanagementsystem* in einem Meeting Abstract der 57. GMDS[7]-Jahrestagung 2012 [Schwanke u. a., 2012]. Der Abstract beschreibt die Untersuchungen der Medizinischen Informatik in Göttingen ein Customer-Relationship-Management-System (CRM-System) auszuwählen, um es als Probandenmanagementsystem zu nutzen. Ein CRM-System ist eine Software, die vorrangig in Unternehmen zur Verwaltung von Kundenbeziehungen eingesetzt wird. Diese Systeme unterstützen die Speicherung von Informationen über Kundinnen und Kunden und Interessenten, können jedoch auch einen größeren Funktionsumfang abdecken,

[7] GMDS - Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V.: `https://www.gmds.de/` [Stand: 01.03.2022]

wie beispielsweise Funktionen für das Vertriebsmanagement. Die Untersuchung der Eignung von CRM-Systemen als Probandenmanagementsystem wurde unternommen, da in einer anfangs durchgeführten Marktanalyse keine den Anforderungen der Autorinnen und Autoren entsprechende Software zur Verwaltung von Probandendaten gefunden werden konnte [vgl. Schwanke u. a., 2012]. Auf Basis mehrerer in Göttingen durchgeführter Forschungsprojekte wurde eine Anforderungsanalyse für die Verwaltung von Studienteilnehmenden erhoben. Das Ergebnis der Untersuchung ergab, dass mehr als die Hälfte der Anforderungen durch ein CRM-System abgedeckt werden kann, die fehlenden weiteren Anforderungen jedoch sehr aufwendig umzusetzen sind [vgl. Schwanke u. a., 2012]. Daher wurde am Institut für Medizinische Informatik in Göttingen mit der Entwicklung eines Probandenmanagementsystems begonnen.

In einer weiteren Veröffentlichung folgten ein Jahr später detaillierte Ergebnisse der Untersuchung auf Eignung eines CRM-Systems als Probandenmanagementsystem [Schwanke u. a., 2013b]. Der englischsprachige Artikel verwendet als Übersetzung des Begriffs *Probandenmanagementsystem* den Begriff *study participant management system* und erläutert ausführlich die Schritte der Anforderungserhebung und der Bewertung der untersuchten CRM-Systeme. Dazu wurden in einem Workshop mit Forschenden der KFO 241 Anforderungen an das Probandenmanagementsystem erfasst. In einem Meeting Abstract der 58. GMDS-Jahrestagung 2013 [Schwanke u. a., 2013a] folgte die Beschreibung des Entwicklungsprozesses und der verwendeten Softwareframeworks zur Entwicklung eines Probandenmanagementsystems. 2014 erschien in einem weiteren Meeting Abstract [Stahmann u. a., 2014a] und einem Artikel [Stahmann u. a., 2014b] die Beschreibung eines Anwendungsszenarios für ein Probandenmanagementsystem zur Datenauswertung. Beide Veröffentlichungen beschreiben die standortbezogene Datenzusammenführung durch ein Probandenmanagementsystem und haben zum Ziel, ein Konzept zur Verringerung der Arbeitsbelastung und zur Vermeidung von Fehlern beim Verknüpfen der Daten aus den verschiedenen Studiendatenbanken mit einem Probandenmanagementsystem zu erstellen. Alle vorgestellten Artikel und Meeting Abstracts mündeten in der prototypischen Implementierung des „Göttinger-Probandenmanagementsystems“ im Rahmen der Dissertation von Jens Schwanke [Schwanke, 2015]. Die Dissertation trägt den Titel „Konzeption und Entwicklung eines Probandenmanagementsystems am Beispiel der Universitätsmedizin Göttingen“. In die Dissertation flossen außerdem die Erkenntnisse

einer Bachelorarbeit zum Thema „Erweiterung eines Probandenmanagementsystems zur Generierung von übergreifenden Forschungspseudonymen“ [Schwarzmaier, 2014] und einer Masterarbeit mit dem Thema „Abbildung und Verwaltung von Einwilligungserklärungen in einem Probandenmanagementsystem“ [Matoni, 2014] ein.
Betrachtet man diese aufeinander aufbauende Reihe an Veröffentlichungen, ist erkennbar, dass das Probandenmanagementsystem auf Initiative der Medizinischen Informatik in Göttingen in Verbindung mit den Anforderungen aus der klinischen Forschung entstand. Da ein CRM-System nicht alle gewünschten Anforderungen erfüllte, entschied man sich zur Entwicklung eines eigenen Probandenmanagementsystems, abgeleitet aus den zuvor gesammelten Anforderungen der KFO 241.
Zum Funktionsumfang des Göttinger Probandenmanagementsystems gehört die Erfassung von Probandenstammdaten wie Name, Geschlecht und Geburtsdatum, aber auch Kontaktdaten der Probandinnen und Probanden wie Adresse, Telefonnummer und E-Mail-Adresse. In einer Forschungsakte der Probandin oder des Probanden können Aufgaben- und Termine aufgenommen werden. Außerdem können Pseudonyme sowie der Status der Einwilligungsklärung erfasst werden. In verschiedenen tabellarischen Ansichten erhalten Systembenutzende eine Übersicht über die erfassten Daten.

2.1.2 Begriffsdefinition

Der Begriff des Probandenmanagementsystems wurde durch die Veröffentlichungen der Medizinischen Informatik in Göttingen eingeführt, nachdem mit der Entwicklung eines solchen Systems begonnen wurde. Sowohl den Begriff als auch den Funktionsumfang des Probandenmanagementsystems anhand eines einzelnen Softwareprodukts zu definieren, ist für den wissenschaftlichen Diskurs jedoch nicht zielführend, da Bestandteile anders benannter Lösungen, sowohl Softwareprodukte als auch Arbeitsprozesse, außer acht gelassen werden. Vergleichbare Systeme konnten jedoch nicht durch eine Literaturrecherche mit dem Probandenmanagementsystem ähnlichen Begriffen identifiziert werden. Daher musste sich der Beantwortung der Frage, was ein Probandenmanagementsystem ist, von Seiten der zum Probandenmanagement zählenden Prozesse genähert werden.
Der Begriff des Probandenmanagements oder synonym der Probandenverwaltung ist als Oberbegriff für verschiedene Prozesse zu sehen. Diese Prozesse der Probandenverwaltung

in der klinischen Forschung sind Gegenstand des nächsten Unterkapitels 2.2 (Prozesse der Probandenverwaltung, S. 15). Softwareprodukte, die zur Probandenverwaltung eingesetzt werden, müssen als eine variierende Zusammenstellung von Softwarekomponenten betrachtet werden, die IT-Unterstützung in allen Bereichen der klinischen Forschung liefern, in denen mit administrativen Probandendaten gearbeitet wird. Administrative Probandendaten sind dabei die Daten von Studienteilnehmenden, die im verwaltenden Zusammenhang eines Forschungsprojektes benötigt werden und abzugrenzen sind von den für die Forschung erhobenen Probandendaten. Die Komponenten, die zu einer IT-gestützten Probandenverwaltung in einer Forschungs-IT-Infrastruktur beitragen, werden in Unterkapitel 2.3 (IT-gestützte Probandenverwaltung in einer Forschungs-IT-Infrastruktur, S. 20) genauer vorgestellt. Aus den gewonnenen Erkenntnissen wurde die folgende Definition abgeleitet:

> „Ein Probandenmanagement- oder Probandenverwaltungssystem ist eine Softwarekategorie, die in der klinischen Forschung Anwendung findet, um Prozesse bei der Verwaltung von Studienteilnehmenden in klinischen Studien zu unterstützen.“

Zurückkommend auf die erste Forschungsfrage, was unter dem Begriff des Probandenmanagementsystems zu verstehen ist, konnte mit der gewählten Definition bereits eine Antwort gegeben werden. Die sich anschließenden Fragen, welche Prozesse zur Probandenverwaltung zu zählen sind und welche Software bei der Verwaltung unterstützt, werden in den folgenden Unterkapiteln beantwortet.

2.2 Prozesse der Probandenverwaltung

Probandenmanagement oder *Probandenverwaltung* sind als Oberbegriffe für verschiedene Prozesse in der klinischen Forschung zu sehen. Diese Prozesse sind Gegenstand des folgenden Unterkapitels und wurden in Form einer systematischen Literaturrecherche untersucht. Die Ergebnisse wurde in einer Übersichtsarbeit (engl. *Scoping Review*) mit dem Titel „Key components and IT assistance of participant management in clinical research: a scoping review“ im Journal JAMIA Open veröffentlicht (Anhang A, Pung und Rienhoff [2020]). Ziel des Scoping Reviews war es, Prozesse die sich mit dem Management von Studienteilnehmenden befassen zu identifizieren und zu beschreiben

und die IT-Unterstützung für diese Prozesse zu beleuchten. Dazu wurden mittels eines Literaturauswahlprozesses, dem PRISMA[8]-Statement [vgl. Moher u. a., 2009] folgend, wichtige Studienprozesse der Probandenverwaltung identifiziert: Rekrutierung (1), Erfassung und Verwaltung der Einwilligungen (2), Pseudonymisierung (3) und Verwaltung administrativer Daten (4) der Studienteilnehmenden. Abbildung 2.2 zeigt die Prozesse mit Probandenverwaltungsanteil im Verlauf einer klinischen Studie.

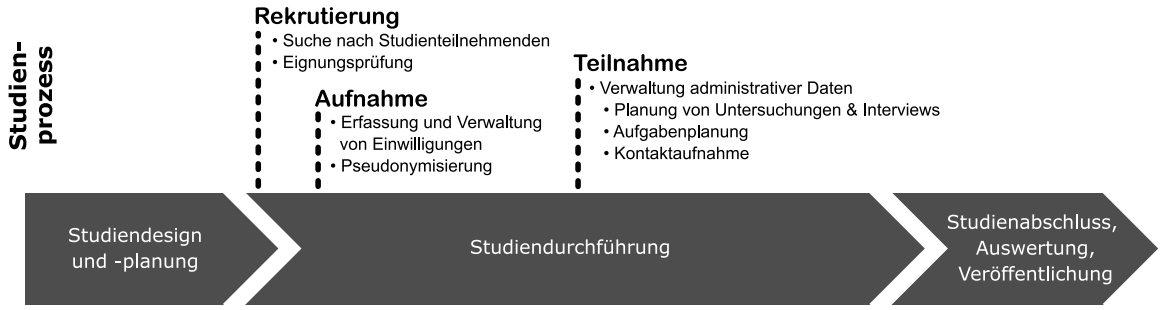

Abbildung 2.2: Verlauf einer klinischen Studie mit Prozessen der Probandenverwaltung. Während der Studiendurchführung werden Studienteilnehmende rekrutiert und bei Eignung aufgenommen. Zur Aufnahme gehört die Erfassung und Verwaltung der Einwilligungen zur Studienteilnahme sowie die Pseudonymisierung. Während der Teilnahme an der Studie sind verschiedene administrative Daten der Studienteilnehmenden zu verwalten. Abbildung abgeändert aus Pung und Rienhoff [2020, S. 452].

Jeder dieser Prozesse mit Probandenverwaltungsanteil wird auf Basis der in Pung und Rienhoff [2020] ermittelten Literatur in den folgenden Unterkapiteln kurz erläutert. Die Vorstellung der Prozesse dient als wichtige Grundlage für das Verständnis des später folgenden Unterkapitels zur IT-gestützten Probandenverwaltung in einer Forschungs-IT-Infrastruktur.

2.2.1 Rekrutierung von Studienteilnehmenden

Die Patienten-/Probandenrekrutierung beschreibt den Prozess der Suche nach geeigneten Teilnehmenden für klinische Studien. Potenzielle Studienteilnehmende werden vom behandelnden ärztlichen Personal oder vom Studienpersonal durch ihnen vorliegende Informationen aus einer papierbasierten oder elektronischen Gesundheitsakte während ihrer klinischen Behandlung identifiziert. Möglich sind jedoch auch andere Wege zur Rekrutierung, wie z. B. Werbematerialien, digitale Medien oder die Verwendung von

[8] PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)

Daten Dritter, wie etwa Patientenorganisationen oder Patientenregistern. Vor der formellen Aufnahme in eine klinische Studie durchlaufen Personen, die an einer Teilnahme interessiert sind, einen Screeningprozess. Die im Studienprotokoll definierten Ein- und Ausschlusskriterien, wie etwa Alter, Geschlecht, Art und Stadium einer Erkrankung beschreiben Merkmale, die von allen Teilnehmenden erfüllt werden müssen.

2.2.2 Erfassung und Verwaltung der Einwilligungen zur Studienteilnahme

Die informierte Einwilligung (engl. *Informed Consent*) ist für die medizinische Behandlung und die medizinische Wissenschaft von zentraler Bedeutung. Grundsätzlich ist die informierte Einwilligung im Zusammenhang mit einer medizinischen Behandlung die Zustimmung der Patientin oder des Patienten zu medizinische Maßnahmen. Darüber hinaus erfüllt die informierte Einwilligung eine Mehrfachfunktion. Sie ist wesentlicher Bestandteil datenschutzrechtlicher Anforderungen, wie sie in der europäischen Datenschutz-Grundverordnung (DSGVO)[9] und im Bundesdatenschutzgesetz (BDSG)[10] festgehalten sind, als auch zentrales Element ethischer Richtlinien der Medizin[11] und des Berufsrechts von Ärztinnen und Ärzten[12]. Die Einwilligungserklärungen müssen damit sowohl den Anforderungen des Datenschutzes als auch den Ethikrichtlinien ihrer wissenschaftlichen Disziplin genügen. Die Leitlinien der Good Clinical Practice (GCP) des International Council for Harmonisation[13] definieren im Rahmen von klinischen Studien den Informed Consent wie folgt:

---

[9] Verordnung (EU) 2016/679 (Datenschutz-Grundverordnung): `https://eur-lex.europa.eu/eli/reg/2016/679` [Stand: 01.03.2022]

[10] Bundesdatenschutzgesetz: `https://www.gesetze-im-internet.de/bdsg_2018/` [Stand: 01.03.2022]

[11] Bundesärztekammer, WMA Deklaration von Helsinki - Ethische Grundsätze für die medizinische Forschung am Menschen: `https://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/International/Deklaration_von_Helsinki_2013_20190905.pdf` [Stand: 01.03.2022]

[12] Bundesärztekammer, (Muster-)Berufsordnung-Ärzte: `https://www.bundesaerztekammer.de/recht/berufsrecht/muster-berufsordnung-aerzte/muster-berufsordnung/` [Stand: 01.03.2022]

[13] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use: `https://www.ich.org/home.html` [Stand: 01.03.2022]

> „A process by which a subject voluntarily confirms his or her willingness to participate in a particular trial, after having been informed of all aspects of the trial that are relevant to the subject's decision to participate." [Good Clinical Practice, 2016, S. 5]

Des Weiteren gehört die konkret für das Forschungsvorhaben formulierte Einwilligungserklärung zu den bei einer Ethikkommission zur Begutachtung einzureichenden Unterlagen, um ein Ethikvotum für ein medizinisches Forschungsprojekt zu erhalten. Die Kommission prüft, ob das Vorhaben ethisch und rechtlich vertretbar ist. Ein positives Ethikvotum ist die Voraussetzung für die Antragstellung und das Genehmigungsverfahren von klinischen Studien und wird ebenfalls von Förderinstitutionen verlangt.
Die Einwilligung wird durch ein schriftliches, unterschriebenes und datiertes Einwilligungsformular (engl. *Informed Consent Form*, ICF) dokumentiert. Die Verwaltung und Prüfung der Einwilligungsformulare sowie der Umgang mit einem eventuellen Widerruf einer Einwilligung gehören zu wichtigen Prozessen in klinischen Forschungsprojekten. Die Verwaltung der Einwilligungen obliegt in der Verantwortung der Studienleitung. Gemäß GCP muss das unterschriebene Einwilligungsformular in einem Prüfarztordner (engl. *Investigator Site File*, ISF) aufbewahrt werden [Good Clinical Practice, 2016].

2.2.3 Führen der Patienten-/Probandenliste und Pseudonymisierung

In einem klassischen Szenario einer klinischen Studie wird jedem Teilnehmenden bei Einschluss in die Studie ein Identifikationscode (Pseudonym) zugewiesen. Der Identifikationscode wird in einer vertraulichen Liste (häufig Patienten-/Probandenliste genannt, engl. *Subject Identification Code List*) zusammen mit weiteren Informationen notiert, die erforderlich sind, um einen Teilnehmer mit seinem Identifikationscode zu verknüpfen. Dieses Verfahren, die Pseudonymisierung, ist wichtiges Mittel, um über eine Trennung von identifizierenden Daten und Forschungsdaten eine *informationelle Gewaltenteilung* zu erreichen. Die *informationelle Gewaltenteilung* ist eines der Grundelemente, um Datenschutz in Forschungsprojekten zu gewährleisten. Grundsätzlich

sollen Daten nur derjenigen Stelle zur Verfügung stehen, die mit der Verarbeitung der Daten zu einem bestimmten Zweck betraut ist. Medizinische Daten, die in Prüfbögen (engl. *Case Report Form*, CRF) dokumentiert werden, werden ausschließlich mit dem Identifikationscode dokumentiert und können nicht direkt mit der Identität des Studienteilnehmenden in Verbindung gebracht werden. Nur mithilfe der Zuordnungsliste können die erfassten medizinischen Daten mit den Identifikationsinformationen verknüpft werden. Die Verknüpfung von medizinischen Daten und Identität ist nötig und in den GCP-Richtlinien [Good Clinical Practice, 2016] gefordert, um die Kontaktaufnahme mit den Teilnehmenden zu ermöglichen, falls während einer Studie medizinische Unregelmäßigkeiten auftreten. Die pseudonymisierte Datenverarbeitung erfolgt grundsätzlich einwilligungsbasiert, sodass die Studienteilnehmenden vor der Verarbeitung über die Nutzung ihrer Daten für die Forschung aufgeklärt und über die Risiken der Nutzung belehrt wurden.

2.2.4 Verwaltung administrativer Probandendaten

Ein enger Kontakt zu den Studienteilnehmenden während der gesamten Studienlaufzeit ist unerlässlich, um Studienabbrecher oder Studienausscheider zu vermeiden und somit die Qualität und Aussagekraft der Studie durch eine hohe Zahl an Studienteilnehmenden sicher zu stellen. Neben dem Ausschluss des Studienteilnehmenden aus einer Studie durch die Prüfleiterin oder den Prüfleiter, weil die im Prüfplan festgelegten Einschlusskriterien nicht mehr erfüllt werden, unterscheidet man, ob der Studienteilnehmende aktiv von sich aus die Teilnahme an der Studie abgebrochen hat (*Drop-out*) oder zur Nachbeobachtung nicht mehr zur Verfügung stand, z. B. weil der Studienteilnehmende unbekannt verzogen ist (*Lost to follow-up*). Die am häufigsten verwendeten Strategien, um *Drop-out* oder *Lost to follow-up* zu vermeiden, sind Erinnerungs-, Kontakt- und Terminplanungsmethoden.
Das Studienpersonal einer klinischen Studie ist für die Terminplanung der Studienbesuche zuständig und sollte, wenn möglich, die Studienteilnehmenden an ihre Termine erinnern. Die Kontaktinformationen der Studienteilnehmenden werden an jedem Studienstandort für den internen Gebrauch während der Studie sicher aufbewahrt. Benötigt werden alle Kontaktinformationen, die dem Forschungsteam helfen, den Teilnehmenden bei Bedarf zu kontaktieren. Die Anzahl der Studientermine hängt von der Art und der Dauer der Studie ab. Je nach Prüfplan können dies Labortests, Röntgenaufnahmen, CT-Scans, körperliche

Untersuchungen u. v. m. sein. Um *Lost to follow-up*-Fälle zu vermeiden ist es sinnvoll, bei jedem Besuch die Kontaktdaten des Studienteilnehmenden auf ihre Aktualität zu überprüfen. Nach Beendigung der Studie werden die Kontaktinformationen an einem sicheren Ort so lange aufbewahrt, wie es die Vorschriften der Studie erfordern.

2.3 IT-gestützte Probandenverwaltung in einer Forschungs-IT-Infrastruktur

Während die vorangegangenen Unterkapitel sich der Frage nach der Einordnung des Probandenmanagementsystems in die Forschungs-IT-Infrastruktur über die Erläuterung der zur Probandenverwaltung zählenden Prozesse näherten, fügt dieses Unterkapitel die gewonnenen Erkenntnisse mit der IT-Sicht auf die Probandenverwaltung zusammen. Einer Forschungsinfrastruktur sind die der Forschung dienenden wissenschaftlichen Infrastrukturen wie Anlagen (Geräte, Instrumente, ...), Ressourcen (Sammlungen, Archive, Bibliotheken, Datenbanken, ...), Einrichtungen und Dienstleistungen zuzurechnen [vgl. RfII, 2016b, S. A-14]. Die Forschungs-IT-Infrastruktur bildet dabei eine Untermenge, die informations- und kommunikationstechnische Infrastruktur wie Rechner, Rechennetze und Software beinhaltet.

Je nachdem, welcher Bereich der medizinischen Forschung betrachtet wird, unterscheidet sich die genutzte Forschungs-IT-Infrastruktur. Zwar kann für diese Arbeit der Fokus auf Forschungsbereiche mit hohem Probandenverwaltungsanteil wie klinische Studien und epidemiologische oder klinische Register eingeschränkt werden, dennoch würde es den Rahmen dieses Unterkapitels sprengen, die zahlreichen Varianten von Forschungs-IT-Infrastrukturen vorzustellen. Ein grundlegender Aufbau ist jedoch allen Ansätzen gemein, sodass anhand einer generischen Form einer medizinischen Forschungs-IT-Infrastruktur einzelne Komponenten vorgestellt werden können. Die Infrastrukturkomponenten, die zur Probandenverwaltung dienen, lassen sich aus den in Pung und Rienhoff [2020] identifizierten IT-Systemen bestimmen. Dazu zählen die Komponenten *Rekrutierungsunterstützung*, *Einwilligungsverwaltung*, *Pseudonymisierung* und *Verwaltung administrativer Probandendaten*. Diese unterstützen in medizinischen Forschungsprojekten die Erfassung und Verarbeitung der für Forschungszwecke benötigten Daten (Infrastrukturkomponente *Forschungsdatenerfassung*). Je nach Anforderungen

eines Forschungsprojektes lassen sich die verschiedenen Infrastrukturkomponenten kombinieren. Das Göttinger Probandenmanagementsystem kann als eine Anwendung betrachtet werden, die Funktionen der hier vorgestellten Infrastrukturkomponenten zur Probandenverwaltung in einem System vereint. Abbildung 2.3 zeigt die generischen Bestandteile einer Forschungs-IT-Infrastruktur für medizinische Forschungsprojekte mit Probandenverwaltungsanteil sowie der Infrastrukturkomponente zur Forschungsdatenerfassung.

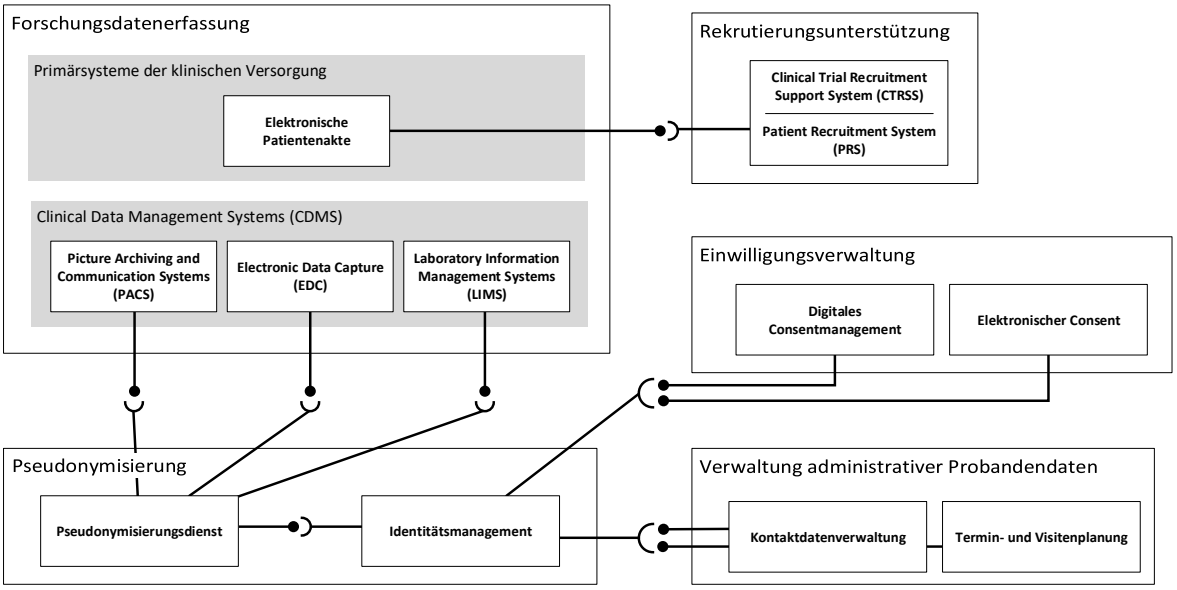

Abbildung 2.3: Generische Bestandteile einer Forschungs-IT-Infrastruktur für medizinische Forschungsprojekte. Mehrere Infrastrukturkomponenten zur Probandenverwaltung unterstützen die Erfassung von Forschungsdaten. Die IT-Systeme in den Infrastrukturkomponenten interagieren untereinander über Schnittstellen.

Die in Abbildung 2.3 gezeigten Infrastrukturkomponenten mit ihren IT-Systemen werden in den folgenden Unterkapiteln vorgestellt.

2.3.1 Forschungsdatenerfassung

Die Verwaltung von Studienteilnehmenden in Forschungsprojekten ist kein Selbstzweck, sondern ermöglicht die Erfassung und Verarbeitung der von den Studienteilnehmenden für Forschungszwecke benötigten Daten. Die Datenquellen können dabei vielfältig sein. Grundsätzlich wird zwischen Forschung an Primärdaten und Sekundärdaten in wissenschaftlichen Studien unterschieden. Unter Forschung an Primärdaten versteht

man die Durchführung von Studien mit Erhebung primärer, im direkten Bezug auf einen Untersuchungsgegenstand erhoben Studiendaten. Demgegenüber werden in der Forschung an Sekundärdaten bereits erhobene Daten und Ergebnisse durchgeführter Studien nachträglich analysiert [Röhrig u. a., 2009]. Für klinische Studien und Register werden im Regelfall Primärdaten erhoben, teilweise werden aber auch Sekundärdaten in die Auswertung einbezogen. Für die Erfassung und Verwaltung von Primärdaten werden *Clinical Data Management Systems* (CDMS) genutzt. Dazu gehören *Electronic Data Capture*-Systeme (EDC), *Laboratory Information Management Systems* (LIMS) oder *Picture Archiving and Communication Systems* (PACS). EDC-Systeme werden für die Datenerfassung mittels elektronischer Erhebungsbögen, sogenannter *Electronic Case Report Forms* (eCRF) eingesetzt und lösen dabei die papierbasierte Datenerfassung ab. Die Verwaltung von Biomaterialien zu Forschungszwecken erfolgt häufig über ein LIMS. Mit dem LIMS werden die Metadaten der Biomaterialien wie Zeit der Probenentnahme, Weiterverarbeitung im Transport, Labor- und Lagerbedingungen, Analyseergebnisse sowie Herausgabe und Verwendung erfasst. PACS sind Bilddatenmanagementsysteme zur Speicherung und Verwaltung von digitalen Bilddaten und ermöglichen den Zugriff auf die Bilder zur Auswertung. Natürlich finden sich in der Praxis auch weitere Werkzeuge für Teilaspekte des Datenmanagements, die in Kombination mit den genannten Softwaresystemen eingesetzt werden, sei es Tabellenkalkulationsprogramme der Bürosoftware (Office-Paket) oder selbst entwickelte Softwarewerkzeuge.

2.3.2 Rekrutierungsunterstützung

Ein häufig auftretendes Problem bei klinischen Studien ist die Erreichung des Rekrutierungsziels. Die Studien erreichen nicht die erforderlichen Rekrutierungszahlen oder rekrutieren nicht die erforderliche Anzahl von Studienteilnehmenden innerhalb der geplanten Rekrutierungszeit [McDonald u. a., 2006]. Zwischen 50% und 60% der randomisierten klinischen Studien erreichen nicht ihre ursprünglichen Rekrutierungsziele oder haben erhebliche Verzögerungen [Sully u. a., 2013; Walters u. a., 2017]. Eine schlechte Rekrutierung kann zu einer Verlängerung der Studien führen, die Kosten erhöht und durch eine mangelnde statistische Aussagekraft zu falschen Schlussfolgerungen führt. IT-Unterstützung bietet das Potenzial, die Rekrutierungsprozesse zu verbessern und das Rekrutierungsziel zu erreichen.

Clinical Trial Recruitment Support Systems (CTRSS) - oder auch *Patient Recruitment Systems* (PRS) genannt - wurden entwickelt, um den Patientenrekrutierungsprozess zu unterstützen, indem sie potenziell geeignete Studienteilnehmende vorschlagen. CTRSS verwenden elektronische Patientendaten, typischerweise aus elektronischen Gesundheitsakten, um die Eignung einer Patientin oder eines Patienten für die Teilnahme an einer klinischen Studie zu beurteilen. Das System benachrichtigt das Studienpersonal über die Eignung einer Person oder stellt ihnen eine Liste potenzieller Studienteilnehmender zur Verfügung. Es wird zwischen drei Arten von CTRSS unterschieden [Köpcke und Prokosch, 2014]: Systeme zur retrospektiven Identifizierung von Studienteilnehmenden auf der Grundlage vorhandener klinischer Daten (1), Systeme die ein klinisches Informationssystem auf definierte Ereignisse überwachen und bei Auftreten die Eignung einer Patientin oder eines Patienten prüfen (2) und eigenständige Systeme, in die Patienteninformationen eingegeben werden, um eine Eignungsprüfung durchzuführen (3).

2.3.3 Einwilligungsverwaltung

Bis auf wenige Sonderfälle erfolgt die Datenverarbeitung auf Basis einer informierten Einwilligung (vgl. Unterkapitel 2.2.2, S. 17). In den letzten Jahren nimmt verstärkt der Einsatz von elektronischen Methoden zur Einholung und Verwaltung der informierten Einwilligung zu [Soni u. a., 2017; Haussen u. a., 2017; Boutin u. a., 2016]. Elektronische Einwilligungen (*Electronic Informed Consent* (eIC) oder *E-consent* genannt) lassen sich in drei Stufen einteilen, die sich durch den Grad der Strukturierung des Dokumenteninhaltes unterscheiden [MITRE Corporation, 2014]. Als E-consent erster Stufe wird ein papierbasiertes Einwilligungsformular angesehen, das in eingescannter digitaler Form aufbewahrt wird. Diese Dokumente können zwar von IT-Systemen angezeigt, aber deren Inhalte nicht maschinell weiterverarbeitet werden. Ebenfalls ein papierbasiertes Einwilligungsformular wird der zweiten Stufe zugrunde gelegt, doch zusätzliche Metainformationen, bspw. Version der Einwilligung, Einwilligungsdatum und Einwilligungsentscheidungen des Studienteilnehmenden werden zu diesem Dokument erfasst. Dadurch steigt die Interoperabilität, denn durch die strukturierten Metainformationen werden die Informationen des Einwilligungsformulars für IT-Systeme interpretierbar. In der dritten Stufe von elektronischen Einwilligungen

wird nicht mehr papiergebunden eingewilligt, sondern alle benötigten Informationen strukturiert in elektronischen Formularen erfasst und digital unterschrieben.
Zunehmend wird von klinischen Studien berichtet, in denen das elektronische Management von Einwilligungen sowohl von den Studienteilnehmenden als auch von den Mitarbeitenden als effizienter und vereinfachender Ablauf wahrgenommen wird [Rothwell u. a., 2014; Vanaken, 2016]. E-Consent Management-Lösungen ermöglichen es, aufgrund der Prozessautomatisierung eine große Anzahl und Vielfalt von Forschungsprojekten in einem Krankenhaus zu verwalten [Bialke u. a., 2018]. Grady u. a. [2017] veröffentlichten im New England Journal of Medicine „The Changing Face of Informed Consent“ innovative Ansätze zur Verbesserung und Erweiterung des elektronischen Einwilligungsprozesses für Forschende und Studienteilnehmende. E-consent kann z. B. multimediale Informationen (Grafiken oder Videos) und interaktive Komponenten beinhalten, um das Verständnis der Studienteilnehmenden für die Studienzwecke zu verbessern, und konzentriert sich weniger auf die Unterzeichnung rechtlicher Dokumente.
Zwar wurden elektronische Einwilligungen von unabhängigen Ethikkommissionen (engl. *Institutional Review Boards*) in den Vereinigten Staaten von Amerika und einigen anderen Ländern akzeptiert, eine breite internationale Einführung ist jedoch eine Herausforderung, da jedes Land unterschiedliche gesetzliche Anforderungen an die Nutzung von elektronischen Einwilligungen hat.

2.3.4 Pseudonymisierung

Von der Erhebung über das Monitoring bis zur Auswertung und Nutzung der Forschungsdaten sind in der Regel identifizierende Informationen der Studienteilnehmenden wie Name und Adresse für die Forschung nicht erforderlich. Daher sind die Daten zu anonymisieren oder zu pseudonymisieren. Der im Unterkapitel 2.2.3 beschriebene Prozess der Pseudonymisierung ist ohne elektronische Unterstützung für kleinere Studien an einem Standort möglich, jedoch für größere Forschungsprojekte und steigenden technischen Anforderungen nicht mehr durchführbar. Um die Datenqualität zu erhöhen und das Datenmanagement in klinischen Studien zu verbessern, werden zunehmend elektronische Datenerfassungssysteme (engl. *Electronic Data Capture*, EDC) eingesetzt. Zusätzlich werden in großen Forschungsprojekten weitere Daten wie Bilddaten oder Daten zu Biomaterialproben in verschiedenen IT-Systemen erfasst. Die

Anbindung von Pseudonymisierungsdienst und Identitätsmanagement an diese Systeme ist daher nötig, um die zuvor erwähnte *informationelle Gewaltenteilung* zu gewährleisten (vgl. Unterkapitel 2.2.3). Gleichzeitig ermöglichen Pseudonymisierungsdienst und Identitätsmanagement regulatorische Anforderungen wie die Kontaktierung des Studienteilnehmenden in Notfällen oder Forschungsziele die das Zusammenführen beziehungsweise die Verknüpfung mit anderen klinischen Daten notwendig machen, gerecht zu werden. Pseudonymisierungsdienst und Identitätsmanagement können je nach Anwendungsfall zentral oder dezentral organisiert sein. In diesem Zusammenhang sind datenschutzkonforme Lösungen erforderlich, um sicherzustellen, dass (1) Probanden, die in mehreren Institutionen auftreten, standortübergreifend das gleiche Pseudonym erhalten und nicht als zwei verschiedene Identitäten behandelt werden (Identitätsmanagement), und dass (2) jedem Studienteilnehmenden pro angebundenem IT-System ein eindeutiges Pseudonym zugeordnet wird (Pseudonymisierungsdienst). Einen umfassenden Überblick über Methoden und Anwendungen für Pseudonymisierungsdienste geben Aamot u. a. [2013] sowie der Leitfaden zum Datenschutz in medizinischen Forschungsprojekten der TMF [Pommerening u. a., 2014]. Für Identitätsmanagement- und Pseudonymisierungsdienste ist grundsätzlich die Einschaltung eines Treuhänders (*Trusted Third Party*, TTP) empfehlenswert, der rechtlich unabhängig, räumlich und personell klar getrennt mit der Verwaltung der Identifizierungscodelisten beauftragt ist. Ausführliche Informationen über rechtliche Rahmenbedingungen, Entstehung, Einsatz und Zukunft von Datentreuhandmodellen in der medizinischen Forschung geben Buchner u. a. [2021] in ihrem Artikel „Das Modell der Datentreuhand in der medizinischen Forschung“.

2.3.5 Verwaltung administrativer Probandendaten

Die Erfassung von Metadaten über Studienteilnehmende und Forschungsprojekte sowie eine Aufgaben- und Terminplanung wurden als wichtige Funktionen für das Göttinger Probandenmanagementsystem angesehen [vgl. Schwanke, 2015, S. 37 f., S. 81]. Häufig verwendet das Studienpersonal für die Verwaltung administrativer Probandendaten keine für diesen Anwendungszweck zugeschnittene Software und behilft sich mit den Funktionen der allgemeinen Bürosoftware (Office-Paket). *Task and Workflow Management Systeme* für klinische Studien bieten Unterstützung bei

der Planung und Dokumentation von Aufgaben und verbundenen Arbeitsprozessen des Studienpersonals und leiten so durch den Verlauf einer klinischen Studie [Almeida u. a., 2019]. *Elektronische Besuchsplanungssysteme* für die klinische Forschung helfen, Studienbesuche mit Behandlungsbesuchen zu koordinieren. Unnötige oder redundante Besuche und Untersuchungen für die Studienteilnehmenden sowie ein hoher Verwaltungsaufwand für die beteiligten Einrichtungen können so vermieden werden [Weng u. a., 2013]. Die Verwaltung administrativer Probandendaten in Softwaresystemen ist oft keine eigenständige Anwendung. *Clinical Trial Management Systems* (CTMS) oder *Clinical Research Management Systems* (CRMS) unterstützen den gesamten klinischen Studienprozess. Zwar bieten CTMS je nach Produkt neben vielen weiteren Funktionen auch Probanden- und Patientenmanagement sowie Adress- und Kontaktdatenmanagement-Funktionen, diese sind jedoch in der akademischen Forschung aus Kostengründen nur selten anzutreffen [vgl. Ohmann und Speer, 2016, S. 34]. CTMS werden hauptsächlich in den kommerziellen Studien der Biotechnologie- und Pharmaindustrie eingesetzt [vgl. Haak u. a., 2017, S. 901].

Kapitel 3

Konzeption des standortübergreifenden Probandenmanagementsystems

In diesem Kapitel werden die Anforderungen an ein Probandenmanagementsystem untersucht, das in klinischen standortübergreifenden Forschungsprojekten zum Einsatz kommen soll. Ziel ist die Entwicklung eines Konzeptes, das in einem weiteren Schritt prototypisch implementiert werden soll. Zunächst werden im Unterkapitel *Methodik und Grundlagen* die für das Verständnis dieses Kapitels nötigen Informationen gegeben. Den Einstieg in das Konzept bildet eine Anforderungserhebung, auf dessen Basis im Hauptteil dieses Kapitels in mehreren Iterationen ein Konzept entwickelt wird. In vier Iterationen werden jeweils Problem, Zielsetzung und Methode vorgestellt, eine Lösung entwickelt und das Ergebnis anschließend diskutiert. Dies erfolgt angelehnt an die identifizierten Anforderungen für die Bereiche: Datentrennung bei der Probandenverwaltung, vernetze Probandenverwaltung, Zugangs- und Zugriffskontrolle und Datenschutzkonzept.

3.1 Methodik und Grundlagen

Dieses Unterkapitel legt die Grundlagen für den weiteren Verlauf des Kapitels, indem wichtige Begriffe und Methoden definiert und ausführlich erläutert werden. Zur Erarbeitung des Konzeptes wurde eine iterative Vorgehensweise gewählt, deren methodischer Ansatz im nächsten Unterkapitel vorgestellt wird. Darauf folgt die Beschreibung des Umfeldes, in dem das Probandenmanagementsystem zum Einsatz kommt. Des Weiteren werden Grundlagen zu Datenschutzkonzepten vorgestellt.

3.1.1 Methodik: Design and Creation

Das Vorgehen zur Erarbeitung des Konzeptes orientiert sich an der Forschungsstrategie *Design and Creation* nach Oates [2006]. *Design and Creation* ist ein iterativer Ansatz zur Entwicklung und Erforschung neuer IT-Artefakte. Zu den IT-Artefakten zählt Oates Methoden, Modelle, Komponenten und Systeme. Bei der Anwendung werden in mehreren Iterationen jeweils die Schritte *Awareness*, *Suggestion*, *Development*, *Evaluation* und *Conclusion* durchlaufen. In jeder Iteration wird jeweils mindestens ein IT-Artefakt erstellt [vgl. Oates, 2006, S. 108 ff.]. Nach Oates sind die IT-Artefakte, die am Ende jeder Iteration entstehen, neue Erkenntnisse, die in eine weitere Iteration eingehen. Dieses Vorgehen fand in zwei ermittelten Dissertationen Anwendung, die ähnlich dieser Arbeit die Konzeption einer Systemarchitektur durchführten und diese Methodik empfehlen [Heinze, 2015; Schreiweis, 2016]. Schreiweis [2016] orientierte sich an den Schritten von Oates und unterteilt gemäß einer wissenschaftlichen Gliederung in folgende Unterabschnitte:

1. Problem und Zielsetzung (nach Oates [2006] *Awareness*): In *Problem und Zielsetzung* wird das bestehende Problem analysiert und als konkret zu lösende Problemstellung formuliert.

2. Methodik (nach Oates [2006] *Suggestion*): *Suggestion* (dt. Anregung, Vorschlag, Vorstellung) beschreibt wie das Problem angegangen und ein Lösungsansatz entwickelt werden kann.

3. Ergebnisse (nach Oates [2006] *Development*): In diesem Teil werden die Lösungsansätze für die identifizierten Probleme detailliert ausgearbeitet und

ausführlich beschrieben. Es entsteht ein IT-Artefakt, durch welches das Problem gelöst werden soll.

4. Diskussion (nach Oates [2006] *Evaluation*): In der Evaluation wird das entwickelte IT-Artefakt auf den gewünschten Lösungseffekt und unerwartete Abweichungen vom gewünschten Lösungseffekt hin analysiert.

5. Schlussfolgerung (nach Oates [2006] *Conclusion*): Ergebnisse der durchlaufenen Schritte werden zusammengefasst und aufbereitet. Schlüsse die aus der gerade durchlaufenen Iteration resultieren, werden in der folgenden Iteration verwendet.

Abbildung 3.1 stellt den *Design and Creation*-Prozess für diese Arbeit dar.

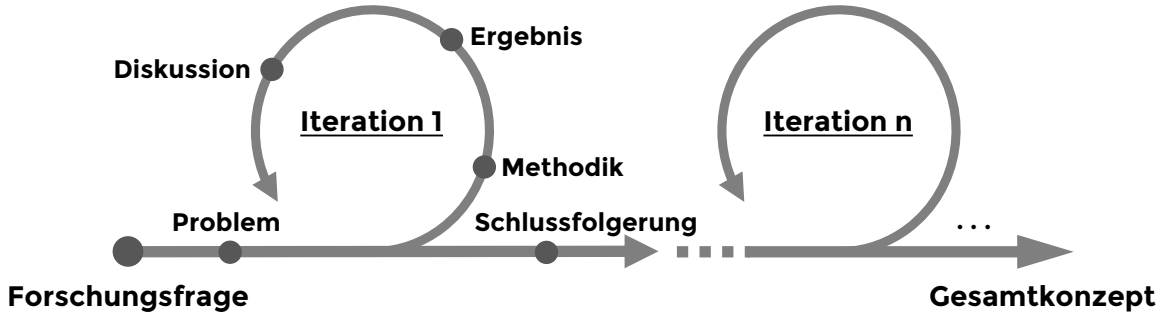

Abbildung 3.1: Darstellung des *Design and Creation*-Prozessmodells für diese Arbeit. Ausgehend von der zweiten Forschungsfrage dieser Dissertation wird durch eine Anforderungserhebung die Erarbeitung einer Lösung in Iterationen zerlegt. Unterteilt in Problem, Methodik, Ergebnis und Diskussion führen aufeinander aufbauend die gesammelten Schlüsse zu einem Gesamtkonzept.

Die Konzeption des standortübergreifenden Probandenmanagementsystems, gefordert in der zweiten Forschungsfrage dieser Arbeit, leitet sich aus zuvor beschriebenen Herausforderungen ab (vgl. Problemvorstellung, S. 1 und Fragestellungen und Ziele dieser Arbeit, S. 5). Der iterative Prozess von *Design and Creation* unterstützt solch eine problemorientierte Herangehensweise besonders, da durch das mehrmalige Durchlaufen der fünf Schritte das Problemverständnis schrittweise vertieft und ein Konzept verfeinert wird.

3.1.2 Rahmenbedingungen und Umfeld

Die Konzeption des standortübergreifenden Probandenmanagementsystems soll für den Anwendungsfall eines Verbundforschungsprojektes an der Universitätsmedizin Göttingen diskutiert werden. Die Diskussion mit Berücksichtigung der Rahmenbedingungen und des Umfeldes einer konkreten Einrichtung bietet den Vorteil, die Komplexität des Themas praxisnah erforschen zu können. Das noch nicht für standortübergreifende Kooperationen geeignete Probandenmanagementsystem von Schwanke [2015] wurde ebenfalls am Beispiel der Universitätsmedizin Göttingen konzeptioniert und entwickelt. Das System befindet sich seit Juni 2015 in drei Forschungsprojekten an der Universitätsmedizin Göttingen im produktiven Einsatz [vgl. Schwanke, 2015, S. 97]. Betrieben wird das Probandenmanagementsystem auf den Servern des Geschäftsbereichs Informationstechnologie, dem IT-Dienstleister für die Universitätsmedizin Göttingen. Die Netzwerkarchitektur der IT der Universitätsmedizin Göttingen orientiert sich an einem Schichtenmodell. Getrennt durch Firewalls bilden Netzsegmente unterschiedliche Sicherheitszonen für die Server und Clients der Universitätsmedizin. Im innersten Segment, dem sogenannten PatLAN[1], befinden sich alle Systeme, die ausschließlich für die Patientenversorgung genutzt werden. Das PatLAN ist getrennt von allen weiteren Segmenten und weist die höchste Schutzkategorie für besonders vertrauliche Patientendaten aus. Darüber liegt das sogenannte WissLAN[2], das Netzwerksegment, in dem IT-Systeme zur Forschung am Patienten untergebracht sind. Darin unterteilt gibt es weitere Zonen, in denen mit unterschiedlich schutzbedürftigen Daten gearbeitet wird. Die Zonen des WissLANs erlauben im jeweiligen Projekt- beziehungsweise Anwendungskontext einen Verbindungsaufbau zwischen den einzelnen Subnetzen und ins Internet. Aus dem Internet ist der Zugriff auf Server mit Web-Diensten nur in eine Pufferzone, der DMZ[3]-weltweit möglich.

---

[1] PatLAN: Bezeichnung eines Netzwerkes der Universitätsmedizin Göttingen. PatLAN steht für „Pat“ = Patientendaten und „LAN“ = Local Area Network

[2] WissLAN: Bezeichnung eines Netzwerkes der Universitätsmedizin Göttingen. WissLAN steht für „Wiss“ = Wissenschaft und „LAN“ = Local Area Network

[3] DMZ = Demilitarisierte Zone: Computernetz mit sicherheitstechnisch kontrollierten Zugriffsmöglichkeiten auf die daran angeschlossenen Server.

Ein vom Institut für Medizinische Informatik der Universitätsmedizin Göttingen bewährte Aufteilung der IT-Systeme für medizinische Forschungsprojekte in die Netzwerksegmente der Universitätsmedizin Göttingen ist in Abbildung 3.2 dargestellt.

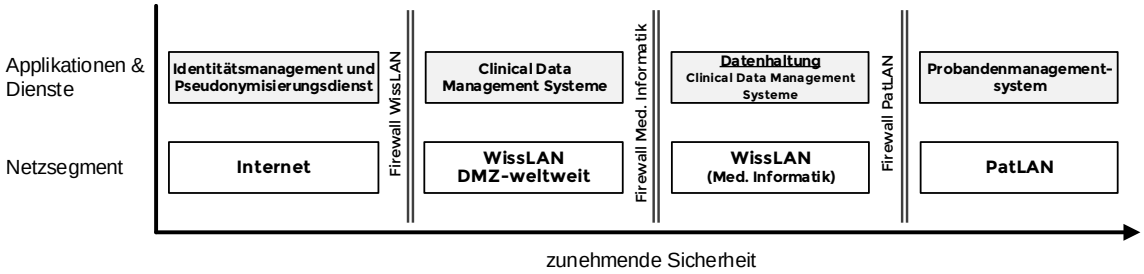

Abbildung 3.2: Darstellung der Netzwerksegmente und IT-Systeme für medizinische Forschungsprojekte am Institut für Medizinische Informatik der Universitätsmedizin Göttingen. Netzwerksegmentierung erhöht die IT-Sicherheit der verschiedenen zur medizinischen Forschung genutzten Applikationen und Dienste.

Das Probandenmanagementsystem befindet sich im PatLAN und ist damit abgeschottet von jeglichen Verbindungen in die Außenwelt [vgl. Demiroglu u. a., 2013, S. 13]. Innerhalb des WissLANs existiert ein dediziertes Netzwerksegment des Instituts für Medizinische Informatik. Dies ist durch eine Firewall vom restlichen Teil des WissLANs getrennt. In diesem Segment werden die Datenbanken der Clinical Data Management Systeme (CDMS) verwaltet. Die CDMS selber, die in Verbundforschungsprojekten aus dem Internet erreichbar sein müssen, sind in dem Segment DMZ-weltweit verortet und über Firewall-Regeln mit ihren Datenbanken verbunden. Die in den CDMS gespeicherten Daten haben einen Personenbezug über ein Pseudonym. Die Pseudonyme werden aus Datenschutzgründen von einem Pseudonymisierungsdienst außerhalb der Universitätsmedizin Göttingen erstellt, um eine Trennung zwischen Forschungsdaten und identifizierenden Daten der Studienteilnehmenden zu gewährleisten. Ein Treuhänder übernimmt die Verwaltung und Speicherung der identifizierenden Daten im Rahmen einer Funktionsübertragung.

3.1.3 Datenschutzkonzept

Ein Datenschutzkonzept wird in medizinischen Forschungsprojekten nötig, wenn für die Erfüllung der Projektziele personenbezogene Daten verwendet werden sollen. Die Pflicht ein Datenschutzkonzept zu verfassen, leitet sich aus der DSGVO indirekt ab. Art. 5 Abs. 2

DSGVO[4] beschreibt das grundlegende Prinzip der Rechenschaftspflicht. Darüber hinaus fordert Art. 24 DSGVO[5] von für die Verarbeitung Verantwortlichen, die Einhaltung des Datenschutzrechts sicherzustellen und zu belegen. Dazu zählt auch der Nachweis von geeigneten technisch-organisatorischen Maßnahmen (TOM), um einen ausreichenden Schutz der personenbezogenen Daten sicherzustellen. Im Datenschutzkonzept werden die für die datenschutzrechtliche Beurteilung notwendigen Informationen zur Erhebung, Verarbeitung und Nutzung personenbezogener Daten beschrieben. Außerdem wird die Rechtsgrundlage und Zweckbindung für die Erhebung, Verarbeitung und Nutzung der Daten dargelegt. Das Datenschutzkonzept ist die Grundlage für die im Datenschutzrecht vorgeschriebene Vorabprüfung des Vorhabens durch die zuständigen behördlichen oder betrieblichen Datenschutzbeauftragten. Außerdem ist das Datenschutzkonzept im Falle einer Prüfung durch die zuständige Aufsichtsbehörde ein wesentliches Dokument, das Auskunft über die Rechtmäßigkeit der Datenverarbeitung bei der Erhebung, Verarbeitung und Nutzung personenbezogener Daten gibt.

3.2 Anforderungen an eine standortübergreifende Probandenverwaltung

In klinischen Forschungsverbünden sind meist mehrere Einrichtungen an verschiedenen Standorten beteiligt. Die Anforderungen an ein standortübergreifend nutzbares Probandenmanagementsystem unterscheiden sich von den Anforderungen an ein System für nur einen Studienstandort. Im Vordergrund dieser Arbeit stand die Erhebung von Anforderungen, die für eine standortübergreifende Nutzung erfüllt sein müssen. Anforderungen an Programmfeatures sind dagegen nahezu identisch mit den bereits erfassten Anforderungen eines Systems für einen Studienstandort.
Da in der Literaturrecherche zum Begriff des Probandenmanagementsystems im zweiten Kapitel keine vergleichbaren Systeme ermittelt werden konnten (vgl. Unterkapitel 2.1, S. 9), entfällt das Ableiten von Anforderungen aus weiteren Systemen. Anforderung

[4] Art. 5 DSGVO Grundsätze für die Verarbeitung personenbezogener Daten: `https://dejure.org/gesetze/DSGVO/5.html` [Stand: 01.03.2022]

[5] Art. 24 DSGVO Verantwortung des für die Verarbeitung Verantwortlichen: `https://dejure.org/gesetze/DSGVO/24.html` [Stand: 01.03.2022]

leiten sich daher nur aus den zuvor beschriebenen identifizierten Herausforderungen ab (vgl. Problemvorstellung, S. 1 und Fragestellungen und Ziele dieser Arbeit, S. 5). Vier zentrale Anforderungen für den standortübergreifenden Einsatz des Probandenmanagementsystems wurden erarbeitet, um sich in den folgenden Unterkapiteln dem gesamten Konzept schrittweise zu nähern:

Anforderung 1: Ein vertraulicher Zugang und Zugriff auf die Probandendaten muss sichergestellt sein

Bei Probandendaten handelt es sich um Gesundheitsdaten, für die in Deutschland nach § 22 BDSG[6], spezifische Schutzmaßnahmen bei der Verarbeitung gelten. Hierzu gehört nach § 22 Abs. 2 Nr. 5 BDSG auch die "Beschränkung des Zugangs zu den personenbezogenen Daten innerhalb der verantwortlichen Stelle und von Auftragsverarbeitern". Vertraulichkeit und der Schutz vor Einsicht durch Unbefugte ist eines der grundlegenden Prinzipien des Datenschutzes im Rahmen von E-Health [vgl. Bauer u. a., 2018, S. 36, S. 41]. Unbefugte, also nicht mit der Datenverarbeitung betraute Dritte, dürfen keine Einsicht in die Datenbestände haben (*Need-to-know-Prinzip*). Dies sicherzustellen muss Teil der nach Art. 32 DSGVO[7] zu dokumentierenden technischen und organisatorischen Datenschutzmaßnahmen (TOM) sein. Technische Maßnahmen beziehen sich dabei auf den Datenverarbeitungsvorgang als solchen. Organisatorische Maßnahmen dagegen beziehen sich auf die Rahmenbedingungen des Datenverarbeitungsvorgangs. Die DSGVO benennt dabei jedoch keine konkreten Handlungsanweisungen, sondern beschreibt vielmehr Datenschutzziele.

Maßnahmen, um die genannten Datenschutzziele zu erreichen, sind eine Zutritts-, Zugangs- und Zugriffskontrolle zu den Daten. Hierbei konzentriert sich diese Arbeit auf die Maßnahmen, die durch die Softwareimplementierung des Probandenmanagementsystems beeinflusst werden. Technischen Maßnahmen, die im weitesten Sinne physisch umsetzbar sind, wie etwa die Zutrittskontrolle zu den Datenverarbeitungsanlagen, werden hier nicht berücksichtigt. Im Gegensatz

[6] § 22 BDSG Verarbeitung besonderer Kategorien personenbezogener Daten: `https://dejure.org/gesetze/BDSG/22.html` [Stand: 01.03.2022]

[7] Art. 32 DSGVO Sicherheit der Verarbeitung: `https://dejure.org/gesetze/DSGVO/32.html` [Stand: 01.03.2022]

zur Zugangskontrolle, die nur zwischen befugten und unbefugten Personen unterscheidet, regelt die Zugriffskontrolle, dass grundsätzlich befugte Nutzerinnen und Nutzer von Systemen nur auf Daten im Rahmen ihrer Befugnisse zugreifen können.
Zugangs- und Zugriffskontrolle ist also ein elementarer Bestandteil des Informationsschutzes und eine zwingende Datenschutzanforderung, um das Probandenmanagementsystem in Verbundforschungsprojekten betreiben zu können. Daher wird für diese Anforderung eine eigene Iteration (Iteration 3: Zugangs- und Zugriffskontrolle, S. 63) durchgeführt, in der eine geeignete Schutzmaßnahmen vorgestellt und diskutiert werden.

Anforderung 2: Eine Datentrennung (informationelle Gewaltenteilung) bei der Probandenverwaltung muss gegeben sein

Eine weitere Maßnahme, die getroffen wird um den Schutz der Probandendaten zu gewährleisten, ist die informationelle Gewaltenteilung durch Datentrennung (vgl. Unterkapitel 2.2.3, Führen der Patienten-/Probandenliste und Pseudonymisierung, S. 18). Anforderung 1 und Anforderung 2 sind eng verknüpft, da beide Anforderungen durch Anwendung des Need-to-know-Prinzips Maßnahmen zum Schutz vor Einsicht durch Unbefugte umsetzen. Das Prinzip der informationellen Gewaltenteilung auf das Probandenmanagementsystem anzuwenden, wird in der ersten Iteration (Iteration 1: Datentrennung bei der Probandenverwaltung, S. 36) der Konzeption untersucht, da das Gesamtkonzept des standortübergreifenden Probandenmanagementsystems erheblich von dieser Anforderung beeinflusst wird.

Anforderung 3: Das Probandenmanagementsystem muss sich in eine IT-Infrastruktur einbinden lassen und Schnittstellen zu Pseudonymisierungsdienst und Identitätsmanagement bereitstellen

Forschung in wissenschaftlichen Verbünden erfordert auch eine Vernetzung der IT-Systeme. Die Erreichung von Anforderung 2, eine Datentrennung sicherzustellen, erfordert, dass das Probandenmanagementsystem verteilte Systeme wie den Pseudonymisierungsdienst und das Identitätsmanagement durch Schnittstellen verbindet. Diese Interaktion des Probandenmanagementsystems mit weiteren Systemen wird in Iteration 2: Vernetzte Probandenverwaltung (Unterkapitel 3.4,

S. 53) erarbeitet.

Anforderung 4: Für den Betrieb des Probandenmanagementsystems muss ein Datenschutzkonzept vorhanden sein

Der Zweck eines Datenschutzkonzeptes wurde bereits im Unterkapitel 3.1.3 (Grundlagen Datenschutzkonzept, S. 31) erläutert. Möchte ein Forschungsprojekt ein Probandenmanagementsystem einsetzten, so muss auch das Probandenmanagementsystem in ihm beschrieben sein. Anforderungen an das Datenschutzkonzept, sowie das Vorgehen zur Entwicklung des Datenschutzkonzeptes wurden in einer weiteren Iteration (Iteration 4: Datenschutzkonzept, S. 92) erarbeitet. Da das Datenschutzkonzept abhängig von den vorherigen Anforderungen und deren Konzeption ist, kann das Datenschutzkonzept erst am Ende der gesamten Konzeption erstellt werden. Hierbei wird der Vorteil des iterativen Ansatzes besonders deutlich: Zum einen erfordert die Erstellung des Datenschutzkonzeptes eine systematische Vorgehensweise, zum anderen müssen die Erkenntnisse der vorherigen Iterationen in die Iteration des Datenschutzkonzeptes eingehen.

3.3 Iteration 1: Datentrennung bei der Probandenverwaltung

Eines der Ziele, das auf dem Weg zu einem Konzept für ein standortübergreifendes Probandenmanagementsystem erreicht werden muss, ist eine sinnvolle Trennung schützenswerter Daten. Eine Datentrennung im Sinne einer informationellen Gewaltenteilung bildet eine der Maßnahmen, um den Schutz der Probandendaten zu gewährleisten. Die Betrachtung weiterer Systeme rund um das Probandenmanagementsystem in einem Forschungsprojekt (vgl. Unterkapitel 2.3, IT-gestützte Probandenverwaltung in einer Forschungs-IT-Infrastruktur, S. 20) ist dazu nötig. Ein standortübergreifender Zugriff auf das Probandenmanagementsystem erfordert die bisherige Verortung des Systems in der IT-Infrastruktur zu überdenken. In dieser Iteration wird zunächst das Problem und die Zielsetzung der Datentrennung in Bezug auf das Probandenmanagementsystem beschrieben. Anschließend wird als Methodik ein Ansatz zur Lösung vorgestellt. Im Ergebnisteil wird eine IT-Infrastruktur skizziert, die sowohl eine sinnvolle Datentrennung als auch den Zugriff von verschiedenen Studienstandorten auf das Probandenmanagementsystem ermöglicht. Das Ergebnis im Spannungsfeld zwischen Datenschutzanforderungen und wünschenswerten Funktionen des Probandenmanagementsystems zu betrachten, ist Teil der anschließenden Diskussion. Abschließend werden die gewonnenen Erkenntnisse in der Schlussfolgerung kurz zusammengefasst und das weitere Vorgehen erläutert. Ergebnisse dieser Iteration wurden bei der 101. Sitzung der TMF AG Datenschutz[8] am 14. September 2021 als Vortrag mit dem Titel *Konzept zur datenschutzkonformen Kontaktdatenverwaltung in klinischen Forschungsprojekten* vorgestellt und diskutiert (Anhang B).

3.3.1 Problem und Zielsetzung

Im Unterkapitel 2.2.3 (Führen der Patienten-/Probandenliste und Pseudonymisierung, S. 18) wurde die Bedeutung der Pseudonymisierung für Forschungsprojekte bereits

[8] TMF - Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V., Arbeitsgruppe Datenschutz: `https://www.tmf-ev.de/Arbeitsgruppen_Foren/AGDS.aspx` [Stand: 01.03.2022]

erläutert. Forschende, die mit den medizinischen Daten der Studienteilnehmenden arbeiten, dokumentieren Forschungsdaten unter einem Pseudonym, um einen direkten Rückschluss auf die Identität des Teilnehmenden zu vermeiden. Die Liste mit den Zuordnungen von Pseudonymen zu Studienteilnehmenden wird vertraulich verwahrt. Unterkapitel 2.3.4 (Pseudonymisierung, S. 24) beschreibt wie Pseudonymisierungsdienst und Identitätsmanagement in Forschungsprojekten mit elektronischer Datenerfassung eingesetzt werden. Wichtig für das Problem der Datentrennung in standortübergreifenden Forschungsprojekten wird nun, wo Pseudonymisierungsdienst und Identitätsmanagement die Zuordnungslisten speichern und wo das Probandenmanagementsystem selbst betrieben wird.

Mit der Einschaltung eines Treuhänders, der für eine Datenzusammenführung (*Record Linkage*) notwendigen Pseudonyme verwaltet, wird eine informationelle Gewaltenteilung erreicht. Ein wichtiger Grundsatz, um personenbezogene Daten zu schützen. Der Treuhänder muss rechtlich unabhängig, räumlich und personell klar getrennt mit der Verwaltung der Zuordnungslisten beauftragt werden. So entsteht ein organisatorischer Schutz, da ein Personenbezug zu Forschungsdaten nur unter Beteiligung mehrerer Organisationen und der Einhaltung von Verfahrensregeln herzustellen ist.

Zu den Anforderungen, die Schwanke [2015] aus den Interviews zur Anforderungsanalyse ableitete, gehörte jedoch unter anderem die Speicherung von Pseudonymen im Probandenmanagementsystem (vgl. Ziel Nr. 14 [Schwanke, 2015, S. 26] und Anforderung Nr. 25 [Schwanke, 2015, S. 40]). Das System soll die Erfassung, Bearbeitung und Löschung von Pseudonymen für Studienteilnehmende ermöglichen. Das Probandenmanagementsystem speichert also sensible identifizierende Probandendaten (Name, Geburtsdatum, Geschlecht, sowie Adresse, Telefonnummer, E-Mail, ...) zusammen mit den Pseudonymen der Clinical Data Management Systeme (CDMS) ab. Ein streitbarer Punkt in Bezug auf die informationelle Gewaltenteilung, wenn der im Unterkapitel 3.1.2 (Rahmenbedingungen und Umfeld, S. 30) beschriebene Betrieb des Probandenmanagementsystems an der Universitätsmedizin Göttingen mit in die Betrachtung einbezogen wird. Sowohl das Probandenmanagementsystem als auch die CDMS liegen bei der administrativen Systembetreuung in der Verantwortung der IT der Universitätsmedizin Göttingen. Da das Probandenmanagementsystem sowohl Pseudonyme als auch identifizierende Probandendaten speichert, ist ein Record Linkage

ohne Hinzuziehen des Datentreuhänders möglich. Das eigentlich angestrebte Ziel, eine informationelle Gewaltenteilung durch Datentrennung zu erreichen, wird durch das Probandenmanagementsystem in der momentanen Umsetzung ausgehebelt. Eine IT-Administratorin oder ein IT-Administrator der Universitätsmedizin Göttingen oder eine weisungsbefugte Person (z. B. IT-Leiterin oder IT-Leiter) kann Zugang zu den im Probandenmanagementsystem gespeicherten identifizierenden Daten und ebenfalls Zugang zu den in den CDMS gespeicherten medizinischen Daten erhalten.
Neben dieser Herausforderung muss das Probandenmanagementsystem von den verschiedenen Studienstandorten, die sich an einer Studie beteiligen, erreichbar sein. Die in Unterkapitel 3.1.2 (Rahmenbedingungen und Umfeld, S. 30) beschriebene IT-Infrastruktur verhindert dies momentan. Das Probandenmanagementsystem befindet sich im PatLAN und ist damit abgeschottet von jeglichen Verbindungen in die Außenwelt (vgl. Unterkapitel 3.1.2, Abbildung 3.2, S. 31). Ziel dieser Iteration ist es, eine IT-Infrastruktur zu skizzieren, die eine Datentrennung im Sinne einer informationellen Gewaltenteilung als auch den Zugriff von verschiedenen Studienstandorten für das Probandenmanagementsystem ermöglicht.

3.3.2 Methodik

Die Herausforderung datenschutzkonforme IT-Infrastrukturen für Forschungsprojekte zu entwickeln ist nicht neu. In den Jahren 2001 bis 2003 haben sich Wissenschaftlerinnen und Wissenschaftler in der TMF zusammengeschlossen, um gemeinsam mit Vertretern der Datenschutzbehörden eine „Blaupause“ für die rechtskonforme Verarbeitung medizinischer Daten in unterschiedlichen Forschungskontexten zu entwickeln [vgl. Pommerening u. a., 2014, S. V]. Es entstanden die „Generischen Datenschutzkonzepte“ der TMF, die im Laufe der Jahre in mehreren Fassungen erschienen und aktuell den *Leitfaden zum Datenschutz in medizinischen Forschungsprojekten - Generische Lösungen der TMF 2.0* [Pommerening u. a., 2014] bilden. Der Leitfaden wurde zu einem Standardwerk für Datenschutz in der medizinischen Verbundforschung in Deutschland und zeigt mögliche Wege zur datenschutzkonformen Forschung auf. Einschränkend ist zu beachten, dass der Leitfaden vor Inkrafttreten der DSGVO erstellt wurde und Änderungen der aktuellen Gesetzeslage noch nicht berücksichtigt sind. Des Weiteren wird der Leitfaden zwar in einer Entschließung der Konferenz der Datenschutzbeauftragten zur Anwendung

empfohlen, die Vorschläge des Leitfadens sind aber rechtlich nicht verbindlich [vgl. Pommerening u. a., 2017, S. 3]. Der Aufbau der generischen Datenschutzkonzepte für medizinische Forschungsverbünde ist modular, um verschiedene Schwerpunkte zuzulassen. Abbildung 3.3 illustriert die Übersicht über die Module mit modulverbindender zentraler Komponente und den Verweisen auf die jeweiligen Kapitel des Datenschutzleitfadens.

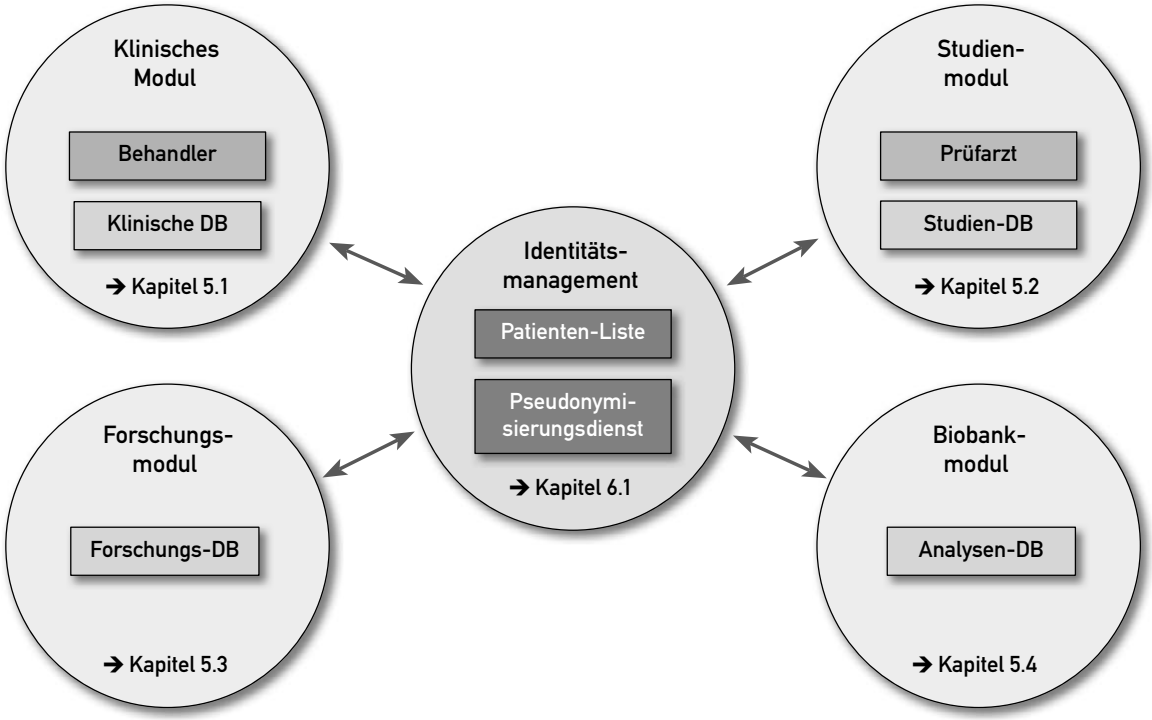

Abbildung 3.3: Übersicht über Module, Komponenten und Verweise auf Kapitel des *Leitfaden zum Datenschutz in medizinischen Forschungsprojekten - Generische Lösungen der TMF 2.0*. Der Kern jedes Moduls besteht aus einer Datenbank (DB), die ausschließlich Forschungsdaten aus dem Forschungskontext des Moduls enthält. Identitätsdaten werden in einer getrennten Patientenliste gehalten, Pseudonyme über den Pseudonymisierungsdienst verwaltet. Patientenliste und Pseudonymisierungsdienst bilden die zentrale modulverbindende Komponente des Identitätsmanagements. Abbildung aus [Pommerening u. a., 2014, S. 3].

Durch den modularen Aufbau lässt sich den Anforderungen von versorgungsnaher Forschung, klinischen Studien, epidemiologischen Projekten, Biobanken, Registern und Langzeitforschungsprojekten gerecht werden [vgl. Pommerening u. a., 2014, S. 6]. Die Module nutzen zur Wahrung der Persönlichkeitsrechte der Patientinnen und Patienten verschiedene Instrumentarien wie Patientenaufklärung und -einwilligung, Anonymisierung und Pseudonymisierung, informationelle Gewaltenteilung und Datentreuhänderschaft, organisatorische Maßnahmen und vertragliche Regelungen

sowie die sichere Gestaltung der IT-Infrastruktur [vgl. Pommerening u. a., 2014, S. 6]. Die Maßnahmen und Empfehlungen des TMF-Datenschutzleitfadens sind auch für die in dieser Arbeit konzeptionierte Forschungs-IT-Infrastruktur mit einem standortübergreifenden Probandenmanagementsystem relevant.

Der Datenschutzleitfaden sieht in seinen Infrastruktur-Komponenten kein Probandenmanagementsystem vor. Es bestehen jedoch Überschneidungen mit dem im Leitfaden vorkommenden Prozess des *Kontaktmanagements* [vgl. Pommerening u. a., 2014, S. 26, S. 172], einem der Anwendungsbereiche des Probandenmanagementsystems in einer klinischen Studie (vgl. Unterkapitel 2.2.4, S. 19). Das Kontaktmanagement wird im Leitfaden als Teil des Identitätsmanagements gesehen [vgl. Pommerening u. a., 2014, S. 26], jedoch nicht weiter als Infrastruktur-Komponente berücksichtigt.

> „Für das Kontaktmanagement könnte der Einsatz kommerzieller CRM-Software von Interesse sein; Erfahrungen hiermit liegen im TMF-Umfeld aber noch nicht vor.“ [Pommerening u. a., 2014, S. 172]

Auch wenn der TMF-Datenschutzleitfaden keine unmittelbar übernehmbare Lösung für das hier vorgestellte Problem liefert, ist er dennoch eine wertvolle Referenz für den Aufbau einer datenschutzkonformen Forschungs-IT-Infrastruktur mit einem Probandenmanagementsystem. Um eine Lösung für die im vorherigen Kapitel beschriebene Problemstellung zu entwickeln, wurden die Module des Datenschutzleitfadens sowie das Identitätsmanagement für Studienteilnehmende auf Vereinbarkeit mit dem Probandenmanagementsystem geprüft. Ziel war die Betrachtung der zugehörigen Daten und Datenflüsse, um einen Vorschlag zu erarbeiteten, wie das Probandenmanagementsystem in den modularen Aufbau des TMF-Leitfadens integriert werden kann. Eine Integration, die sich an den Grundprinzipien datenschutzgerechter Lösungen des TMF-Leitfadens orientiert, bildet damit gleichzeitig die Lösung für den ersten Teil der Problemstellung dieser Iteration. Zweiter Teil der Lösung ist es, die gewonnenen Erkenntnisse auf die Vereinbarkeit mit der IT-Infrastruktur, speziell der Netzwerksegmente der Universitätsmedizin Göttingen, zu prüfen und anzuwenden.

3.3.3 Ergebnisse

Die Anforderung für das Probandenmanagementsystem in klinischen Forschungsverbünden mit mehreren Studienstandorten eine Datentrennung im Sinne einer informationellen Gewaltenteilung zu erreichen, erfordert Prinzipien bereits etablierter Datenschutzkonzepte auf das Probandenmanagementsystem anzuwenden. Das Probandenmanagementsystem agiert in einer Forschungs-IT-Infrastruktur und sollte daher nach Möglichkeiten datenschutzkonforme Konzepte, wie sie beispielsweise im TMF-Datenschutzleitfaden beschrieben werden, nicht negativ beeinflussen. Es wurde daher eine Lösung erarbeitet, die das Probandenmanagementsystem als weitere Komponente im modularen Aufbau des Leitfadens ergänzt (siehe Abbildung 3.4).

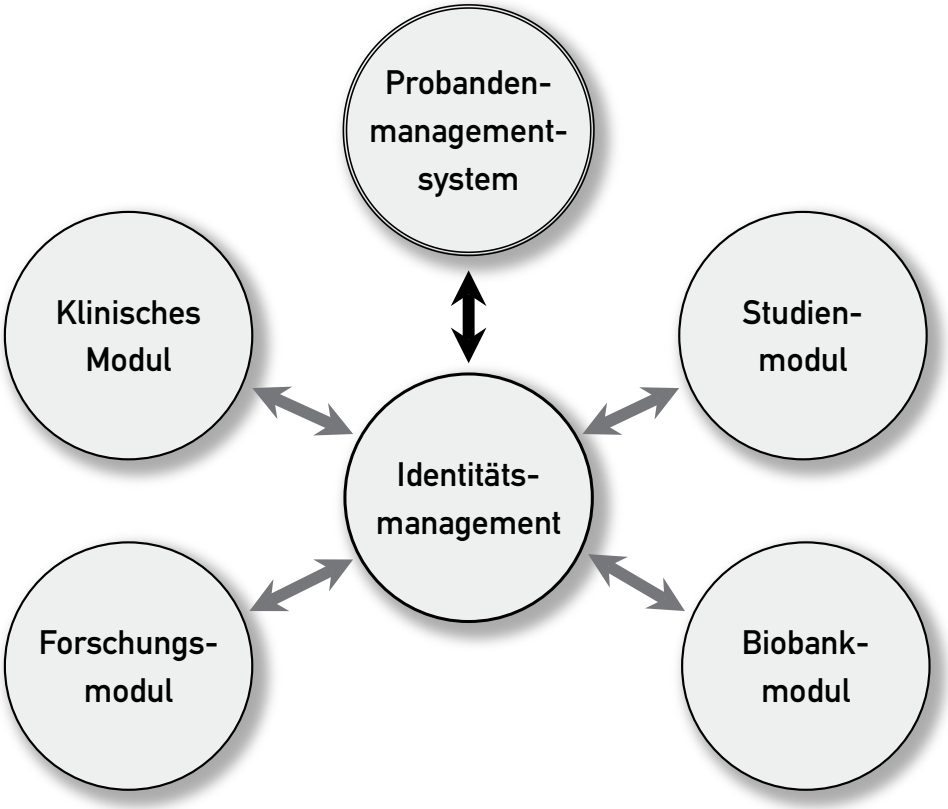

Abbildung 3.4: Übersicht über die Module des TMF-Datenschutzleitfadens, ergänzt um das Probandenmanagementsystem, das wie die Module mit dem Identitätsmanagement interagiert.

Die Module des TMF-Datenschutzleitfadens unterscheiden sich durch unterschiedliche Rahmenbedingungen und Vorgehensweisen und passen somit zu unterschiedlichen wissenschaftlichen Fragestellungen. Für eine eingehendere Beschreibung der Module sei

auf den TMF-Datenschutzleitfaden selbst verwiesen [Pommerening u. a., 2014]. Je nach Anforderung eines Forschungsprojektes lassen sich die Module beliebig kombinieren.
Die in dieser Arbeit beschriebenen Prozesse der Probandenverwaltung (Unterkapitel 2.2, Prozesse der Probandenverwaltung, S. 15) decken sich weitestgehend mit den Anwendungsfällen des Studienmoduls [vgl. Pommerening u. a., 2014, S. 76 ff.]. Eine Kombination von Studienmodul, zentralem Identitätsmanagement und Probandenmanagementsystem ist also eine praxisrelevante Zusammenstellung. Aber auch die Kombination mit weiteren Modulen bis hin zu einem Maximalmodell eines medizinischen Forschungsverbundes unter Einbeziehung aller Module ist denkbar. Durch deren Zusammenwirken ergeben sich komplexe Datenflüsse und Kommunikationsbeziehungen, die verschiedene Maßnahmen erfordern, um ein angemessenes Datenschutzniveau zu wahren. Das Identitätsmanagement hat im Zusammenspiel der verschiedenen Module eine zentrale Stellung. Für standortübergreifende Verbundforschungsprojekte ist fast immer eine Variante mit zentralem Identitätsmanagement notwendig [vgl. Pommerening u. a., 2014, S. 82]. Die grundsätzliche Aufgabe des Identitätsmanagements ist es, die Zuordnung von Pseudonymen zueinander und zu den Identitätsdaten herzustellen (vgl. auch Unterkapitel 2.3.4, Pseudonymisierung, S. 24). Die Trennung der verschiedenen Module und des Identitätsmanagements unterstützt eine strikte Datentrennung zwischen Identitätsdaten auf der einen Seite und medizinischen Daten auf der anderen Seite (informationelle Gewaltenteilung). Unterschiedliche Pseudonyme je Anwendungskontext erhöhen dabei die Wahrscheinlichkeit einer ungewollten Datenzusammenführung. An dieses Prinzip muss auch das ergänzte Probandenmanagementsystem anknüpfen und in Interaktion mit dem Identitätsmanagement treten (vgl. Abbildung 3.4, S. 41). Durch die Betrachtung der Daten und Datenflüsse der Komponenten Identitätsmanagement, Studienmodul und Probandenmanagementsystem wird die erarbeitet Lösung im Detail erläutert.

Daten und Datenflüsse des Identitätsmanagements

Zu den zentralen Aufgaben des Identitätsmanagements gehören das Führen von Listen mit personenidentifizierenden Daten und die Pseudonymisierung. Für die technische Umsetzung gibt es verschiedene Softwarewerkzeuge, auf die im Implementierungskapitel

(Unterkapitel 4.2.2, Implementierung Schnittstelle, S. 117) eingegangen wird. Grundprinzip der Pseudonymisierung und Aufgabe des Identitätsmanagements ist die Zuweisung einer eindeutigen Kennung zu einem Studienteilnehmenden anhand seiner personenidentifizierenden Daten. Der einfachste Fall ist eine einstufige Pseudonymisierung. Identifizierende Personenattribute, typischerweise Vorname, Nachname, Geburtsdatum und Geschlecht, sind Grundlage für die eindeutige Zuweisung eines Pseudonym erster Stufe, dem Patientenidentifikator (PID[9]). Bei der Zuweisung können verschiedene Verfahren (z. B. *Matching-Algorithmen*) verwendet werden, um die Qualität der Zuordnung von Person und PID zu gewährleisten. Es ist sicherzustellen, dass bei der Vergabe der PID Synonymfehler (eine Person hat mehrere PIDs) und Homonymfehler (zwei oder mehr Personen haben einen identischen PID) mit möglichst hoher Sicherheit vermieden werden, und zwar auch dann, wenn Personenattribute beispielsweise durch unterschiedliche Schreibweise voneinander abweichen [vgl. Pommerening u. a., 2014, S. 109]. Im TMF-Datenschutzleitfaden wird dieser Teil des Identitätsmanagements, also der Funktion zur Erzeugung und Verwaltung der notwendigen Pseudonyme erster Stufe sowie zur Speicherung der zugehörigen Identitätsdaten, der sogenannten „Patientenliste“ zugezählt. Abbildung 3.5 verdeutlicht die Funktionsweise der Patientenliste bei der Vergabe von PIDs.

[9] Abkürzung PID wird in diesem Zusammenhang teilweise auch für *Personen*identifikator verwendet.

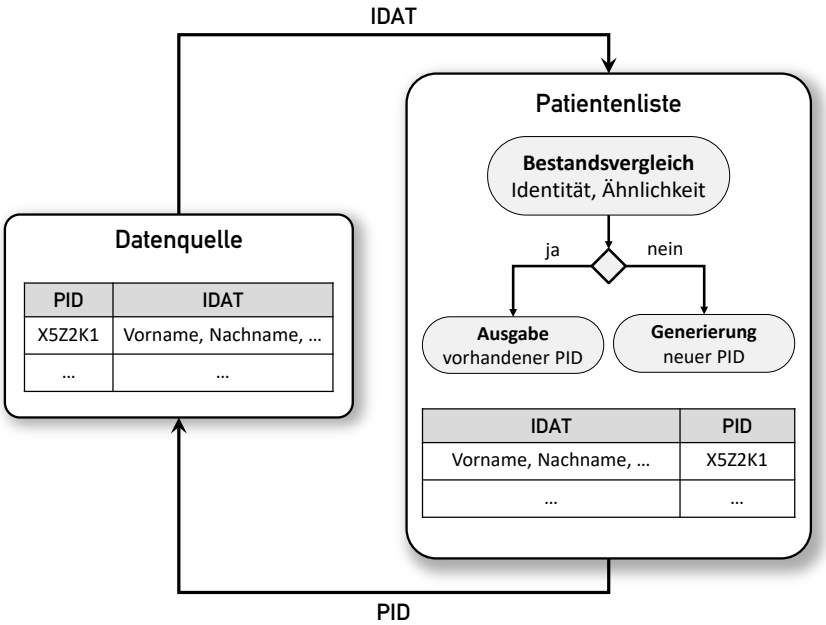

Abbildung 3.5: Funktionsweise der Patientenliste als Teil des Identitätsmanagements. Vergabe eines Patientenidentifikators (PID) anhand von identifizierenden Attributen (IDAT). Eigene Abbildung angelehnt an [Pommerening u. a., 2014, S. 112].

Die einstufige Pseudonymisierung ist in der Regel angemessen, wenn die Pseudonyme (hier PID) innerhalb einer Einrichtung bleiben. Komplexere Forschungsprojekte benötigen zur Erhöhung des Datenschutzniveaus meist ein mehrstufiges Pseudonymisierungsverfahren. Ein Pseudonym erster Stufe (z. B. PID, andere Identifier sind aber auch möglich) verweist direkt auf die Personendaten, ein Pseudonym höherer Stufe verweist dagegen jeweils auf das Pseudonym der vorherigen Stufe. Durch die mehrstufige Pseudonymisierung kann eine Entkoppelung verschiedener Anwendungskontexte (z. B. Fragebogendaten, Bilddaten, Biomaterialien, etc.) erreicht werden (siehe Abbildung 3.6). Die Studienteilnehmenden werden in verschiedenen Bereichen des Forschungsverbundes durch unterschiedliche Pseudonyme repräsentiert.

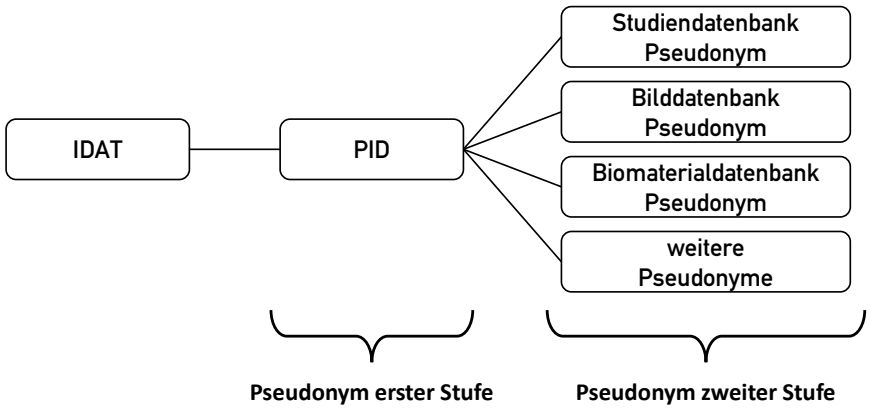

Abbildung 3.6: Zusammenhang verschiedener Identifikatoren bei einer mehrstufigen Pseudonymisierung. Personendaten oder identifizierende Stammdaten (IDAT) werden einem Pseudonym erster Stufe, dem Patientenidentifikator (PID) zugeordnet. Aus dem PID werden jeweils für einen Anwendungskontext weitere Pseudonyme zweiter Stufe abgeleitet. Bei getrennter Datenhaltung steigt dadurch für unautorisierte Personen der Aufwand zur Re-Identifizierung.

Die Erzeugung von Pseudonymen zweiter Stufe kann durch einen von der Patientenliste getrennten Pseudonymisierungsdienst erbracht werden. Die Pseudonyme können entweder mittels eines kryptografischen Verfahrens aus dem Pseudonym erster Stufe transformiert werden oder sie können unabhängig vom Pseudonym erster Stufe neu generiert und in einer kontextspezifischen Zuordnungstabelle mit dem Pseudonym erster Stufe verknüpft werden. Hier sind unterschiedliche Umsetzungen möglich, die zum Forschungsprojekt passen müssen [vgl. Pommerening u. a., 2014, S. 178]. Bei größeren Forschungsverbünden mit Daten aus verschiedenen Quellen ist es sinnvoll, die Patientenliste und Pseudonymisierung über eine zentrale, unabhängige und vertrauenswürdige Stelle betreiben zulassen (vgl. Unterkapitel 2.3.4, Pseudonymisierung: Prinzip Treuhänder, S. 24).

Die bisher vorgestellten Daten und Datenflüsse des Identitätsmanagements ändern sich nicht grundlegend durch die Ergänzung des Probandenmanagementsystems. Da das Probandenmanagementsystem in Verbundforschungsprojekten zum Einsatz kommen soll, ist die Verwendung eines zentral organisierten Identitätsmanagements mit mehrstufiger Pseudonymisierung für unterschiedliche Anwendungskontexte geboten. Wichtig für die Vorstellung der weiteren Daten und Datenflüsse des Studienmoduls und des Probandenmanagementsystems ist die Abgrenzung der für das Identitätsmanagement benötigten Daten gegenüber den Daten des Probandenmanagementsystems. Ein

Identitätsmanagement, wie es im Zusammenspiel mit dem Probandenmanagementsystem betrieben wird, muss sich ausschließlich auf seine Kernfunktionen beschränken. Für das hier vorgestellte Konzept zählen dazu die eindeutige Zuweisung einer PID anhand von identifizierenden Daten und deren Aufbewahrung, sowie der Generierung von weiteren Pseudonymen und die Speicherung der Zuordnungstabellen. Das bedeutet, dem Identitätsmanagement werden ausschließlich die identifizierenden Stammdaten übergeben, die für die eindeutige Zuordnung einer Person benötigt werden. Im Sinne der Datensparsamkeit werden keine weiteren Personendaten oder identifizierende Stammdaten gespeichert. Diese Forderung weicht von Ansätzen ab, in denen Kontaktdaten und weitere Kontextinformationen der Studienteilnehmenden dem Identitätsmanagement zugezählt werden. Dies wird im TMF-Datenschutzleitfaden unter Prozessen des Kontaktmanagements (Pflege der Kontaktdaten) als optionaler Teil des Identitätsmanagements gesehen [vgl. Pommerening u. a., 2014, S. 26, S. 114]. Ein Treuhänder wird so in die Lage versetzt, Patienten anhand der Adressdaten direkt anzusprechen [vgl. Pommerening u. a., 2014, S. 151]. Des Weiteren wird der zentralen Patientenliste die Möglichkeit eingeräumt, neben identifizierenden Daten auch organisatorische Daten, die Bezug zu Einrichtung, Arzt oder Prüfer erlauben, zu speichern [vgl. Pommerening u. a., 2014, S. 82 f., S. 110]. Bei Verwendung eines Probandenmanagementsystems werden diese Funktionen jedoch nicht dem Identitätsmanagement zugeschrieben, sondern dem Probandenmanagementsystem. Das Probandenmanagementsystem übernimmt die Verwaltung aller Metadaten der Studienteilnehmenden über die Pseudonymisierung hinaus. Diese werden im Abschnitt *Daten und Datenflüsse des Probandenmanagementsystems* genauer beschrieben.

Daten und Datenflüsse des Studienmoduls

Das Studienmodul dient der Durchführung klinischer Forschungsprojekte wie klinischer Studien. Im einfachsten Fall wird für das Studienmodul eine Studiendatenbank und eine Patientenliste benötigt. Für standortübergreifende Verbundforschungsprojekte, die in dieser Arbeit betrachtet werden, ist jedoch eine Variante mit zentralem Identitätsmanagement notwendig. Zweck der Trennung von Studiendatenbank und Identitätsmanagement ist es, eine strikte Datentrennung zwischen Identitätsdaten (IDAT) und medizinischen Daten (MDAT) zu schaffen. Die behandelnde Einrichtung

dokumentiert alle studienrelevanten Daten als MDAT zum Studienteilnehmenden zusammen mit einem Pseudonym und speichert diese in einer Studiendatenbank. Zur Dokumentation der MDAT werden Clinical Data Management Systeme (CDMS) eingesetzt (vgl. Unterkapitel 2.3, IT-gestützte Probandenverwaltung in einer Forschungs-IT-Infrastruktur, S. 20). Die Datenflüsse zwischen Studie, CDMS und Identitätsmanagement sind in Abbildung 3.7 erläutert. Im Studienkontext arbeitet das Personal mit dem CDMS, in dem angemeldete Systembenutzende neue Studienteilnehmende nach Eingabe der Studienteilnehmenden-IDAT anlegen können. Die IDAT werden nicht im CDMS gespeichert, sondern an das Identitätsmanagement durchgeleitet. Nach der Generierung und Zuweisung von PID und anwendungskontextbezogenem Pseudonym (PSN) wird das PSN an das CDMS übergeben, unter dem Systembenutzende die MDAT dokumentieren.

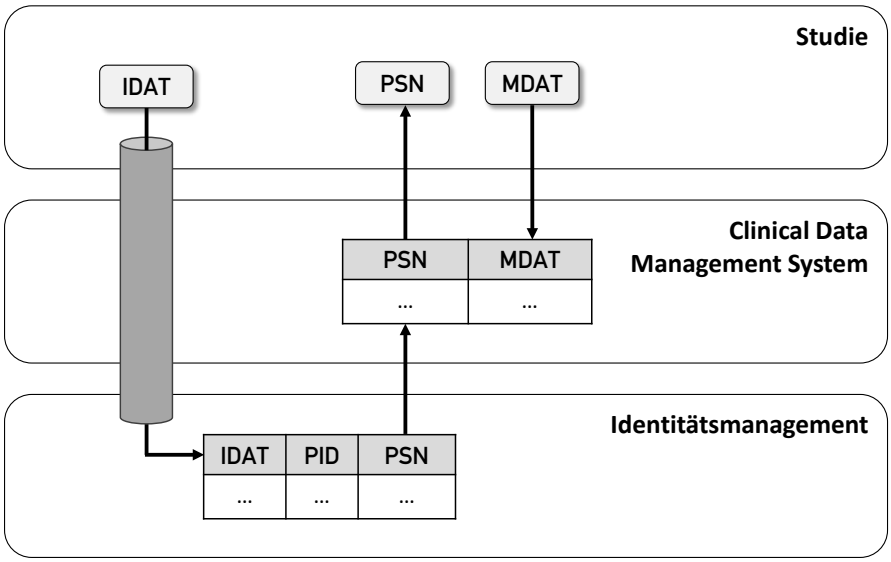

Abbildung 3.7: Darstellung der Datenflüsse von identifizierenden Stammdaten (IDAT), Pseudonym (PSN) und medizinischen Daten (MDAT) zwischen Studienkontext, Clinical Data Management System und Identitätsmanagement.

Eine spätere Zusammenführung von Daten verschiedener CDMS ist möglich, da im zentralen Identitätsmanagement die unterschiedlichen Pseudonyme eines Patienten zusammen mit einer einheitlichen PID gespeichert werden. Das Probandenmanagementsystem übt auf diese Datenflüsse keinen Einfluss aus.

Daten und Datenflüsse des Probandenmanagementsystems

Das Probandenmanagementsystem speichert in seiner Datenbank Probandenstammdaten wie Name, Geschlecht und Geburtsdatum, aber auch Kontaktdaten der Studienteilnehmenden wie Adresse, Telefonnummer und E-Mail-Adresse. Hinzu kommen Daten zum Studienverlauf der Teilnehmenden wie Studienstatus, Studiengruppe, Rekrutierungsstandort, Rekrutierungsdatum, Einwilligungsstatus, Kontaktereignisse, Termine, Aufgaben und Anmerkungen. Diese Daten werden für die im Unterkapitel 2.2 (S. 15) beschriebenen Prozesse der Probandenverwaltung benötigt. Da hier hochsensible Identifikationsdaten verwaltet werden, besteht ein besonders hoher Anspruch auf Vertraulichkeit, auf dessen Einhaltung in Iteration 3: Zugangs- und Zugriffskontrolle (Unterkapitel 3.5, S. 63) eingegangen wird.

Für eine datenschutzkonforme standortübergreifende Probandenverwaltung dürfen die Pseudonyme der CDMS nicht im Probandenmanagementsystem gespeichert werden (vgl. Problemstellung Iteration 1, 36). Diese müssen, wie im Abschnitt zu den Daten und Datenflüssen des Identitätsmanagements beschrieben, von einem externen Identitätsmanagement erstellt und gespeichert werden. Dennoch wird für das Probandenmanagementsystem eine Verknüpfungsmöglichkeit zu den im Identitätsmanagement gespeicherten Daten benötigt. Das Probandenmanagementsystem hat zum Ziel, dem Studienpersonal, das zum datenschutzgerechten Umgang mit dem IDAT-Datensatz verpflichtet ist, die Arbeitsprozesse in einer Studie zu erleichtern. Dazu gehört auch, Pseudonyme der CDMS für einzelne Studienteilnehmende abfragen zu können. Diese Funktion steht jedoch im Konflikt mit dem Speichern der Pseudonyme der CDMS in der Datenbank des Probandenmanagementsystems. Ein Kompromiss war daher, das Probandenmanagementsystem in die Lage zu versetzen, berechtigten Systembenutzenden (siehe Iteration 3: Zugangs- und Zugriffskontrolle, S. 63) die Pseudonymabfrage zu ermöglichen, diese jedoch nicht in der Datenbank des Probandenmanagementsystems zu speichern. Die Pseudonymabfrage über eine Schnittstelle zum Identitätsmanagement zu ermöglichen wird in Iteration 2: Vernetzte Probandenverwaltung (S. 53) betrachtet. Abbildung 3.8 zeigt exemplarisch die Aufteilung der Pseudonyme in einem Forschungsprojekt mit einem CDMS und einem Probandenmanagementsystem.

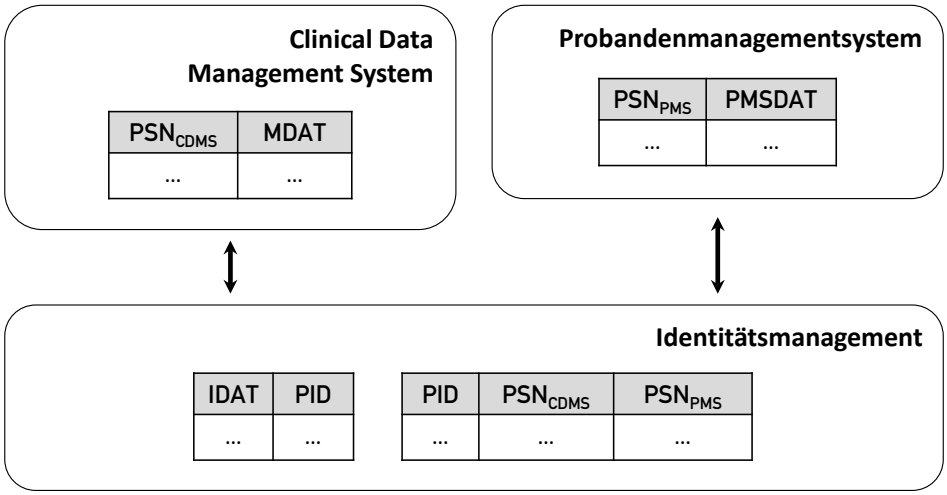

Abbildung 3.8: Übersicht über Daten und Datenflüsse des Probandenmanagementsystems (PMS) mit einem Clinical Data Management Systeme (CDMS) und dem Identitätsmanagement. Die medizinischen Daten (MDAT) im Forschungskontext werden mit dem Pseudonym PSN_{CDMS} gespeichert. Die Daten des Probandenmanagementsystems (PMSDAT) werden mit dem Pseudonym PSN_{PMS} gespeichert. Das Identitätsmanagement erstellt mithilfe übergebender IDAT (Prinzip siehe Abbildung 3.5, S. 44) die Pseudonyme und verwahrt die Zuordnung der Pseudonyme untereinander.

Mit Hinzunahme des Probandenmanagementsystems in eine IT-Infrastruktur zu CDMS und Identitätsmanagement wird das Probandenmanagementsystempseudonym (PSN_{PMS}) als ein weiteres Pseudonym eingeführt. Das Probandenmanagementsystempseudonym dient als Identifier zur Kommunikation mit dem Identitätsmanagement und zur pseudonymen Anzeige von Datensätzen, für die keine identifizierenden Daten benötigt werden. Der Zugewinn an Datenschutz liegt also zum einen in der Unterscheidung verschiedener Klassen von Pseudonymen und zum anderen darin, welcher Personenkreis von ihnen Kenntnis erhält.

Zweiter Teil der Konzeption dieser Iteration ist die Überlegung, in welchem Teil einer Netzwerkarchitektur der Universitätsmedizin Göttingen das Probandenmanagementsystem betrieben wird (vgl. Unterkapitel 3.1.2, Abbildung 3.2). Die Anforderungen das Probandenmanagementsystem von verschiedenen Studienstandorten erreichen zu können und es mit dem Identitätsmanagement zu vernetzen, erfordern eine Verbindung zum Internet. Grundsätzlich entsteht dadurch ein höheres Risiko für die Informationssicherheit, auf das in Iteration 3: Zugangs- und Zugriffskontrolle (Unterkapitel 3.5, S. 63) ausführlich eingegangen wird. Dennoch wird unter den geplanten Sicherheitsvorkehrungen ein Betrieb

des Probandenmanagementsystems mit Verbindung zur Außenwelt möglich. Das Probandenmanagementsystem wird nicht mehr im PatLAN sondern im WissLAN betrieben (siehe Abbildung 3.9).

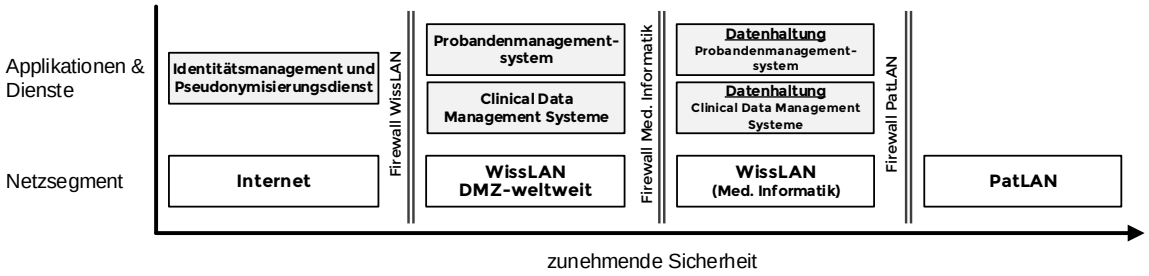

Abbildung 3.9: Darstellung der Netzwerksegmente und ihren IT-Systemen für medizinische Forschungsprojekte am Institut für Medizinische Informatik der Universitätsmedizin Göttingen. Erweiterung der Abbildung 3.2 (S. 31).

Dadurch wird ein Zugriff von Verbundpartnern auf das Probandenmanagementsystem möglich. Außerdem wird ein geregelter Zugriff auf ein externes Identitätsmanagement möglich, sodass das Probandenmanagementsystem Pseudonyme abrufen kann. Die Option, das Probandenmanagementsystem innerhalb eines Netzwerks zusammen mit den CDMS zu betreiben, ergibt sich erst durch die im vorherigen Abschnitt vorgestellte Datentrennung. Das Probandenmanagementsystem speichert die Zuordnungslisten der Pseudonymen nicht mehr in der eigenen Datenbank. Eine IT-Administratorin oder ein IT-Administrator der Universitätsmedizin Göttingen, der Zugang zur Datenbank des Probandenmanagementsystems und Zugang zu den Daten der CDMS hat, kann somit beide Datenquellen nicht mehr über die Pseudonyme verknüpfen.

3.3.4 Diskussion

Das Probandenmanagementsystem hat zum Ziel, administrative Probandendaten zu verwalten und so die Prozesse der Probandenverwaltung zu unterstützen. Dazu müssen sensible identifizierende Probandendaten gespeichert werden. Demgegenüber steht das Ziel eines treuhänderisch verwalteten Identitätsmanagements, Zuordnungen von Pseudonymen zu Studienteilnehmenden zu verwalten und die dazu benötigten identifizierenden Daten der Studienteilnehmenden in einem Forschungsprojekt exklusiv sicher zu verwahren. Diese beiden Ziele stehen in einem gewissen Spannungsverhältnis, eine einfache Datentrennung im Sinne einer informationellen Gewaltenteilung zu

erreichen. Streitbarer Punkt ist, wem die Verantwortung zukommt, identifizierende Probandendaten in einem Forschungsprojekt zu speichern. In dem hier vorgestellten Konzept wird die Verantwortung aufgeteilt und die informationelle Gewaltenteilung nach dem Verwendungszweck der Daten unterschieden. Das Probandenmanagementsystem speichert die administrativen Probandendaten, ohne die eine Probandenverwaltung nicht existieren kann. Das Identitätsmanagement wird in dem hier vorgestellten Konzept auf die Verwahrung ausschließlich der identifizierenden Probandendaten reduziert, die es für seine Kernaufgabe, die eindeutige Zuweisung von Pseudonymen, benötigt. Treuhänderisch verwaltet werden die Zuordnungslisten zwischen den Pseudonymen. Diese Zuordnungslisten werden nur herausgegeben, wenn eine Datenzusammenführung geplant ist. Aus diesem Grund darf das Probandenmanagementsystem nicht mehr zusätzlich zu seinen administrativen Probandendaten die Pseudonyme der Zuordnungslisten speichern. Das gewählte Konzept passt sich somit in die im TMF-Datenschutzleitfaden beschrieben Trennung von Daten ein.

Abstimmungsbedarf besteht mit den zuständigen Datenschutzbeauftragten eines konkreten Forschungsprojektes zu der konzeptionierten Funktion, ausgewählte Pseudonyme der CDMS über eine Schnittstelle temporär abfragen zu dürfen. Eine gezielte Pseudonymabfrage ist bei der Probandenverwaltung eine sinnvolle Funktion, birgt jedoch auch die Gefahr des Missbrauchs bei massenhafter Ausführung. Zur Minimierung des Risikos werden in Iteration 3: Zugangs- und Zugriffskontrolle (Unterkapitel 3.5, S. 63) geeignete Maßnahmen vorgestellt, beispielsweise zur Steuerung der Berechtigungen und dem Protokollieren getätigter Pseudonymabfragen.

Das Probandenmanagementsystem sollte, um in Verbundforschungsprojekten eingesetzt werden zu können, über das Internet von unterschiedlichen Studienstandorten erreichbar sein. Auch hier sind geeignete Maßnahmen zur Informationssicherheit in Iteration 3 erarbeitet worden. Dennoch muss sich in konkreten Forschungsprojekten zeigen, ob Kliniken eines Forschungsverbundes einer Speicherung der sensiblen Daten ihrer Studienteilnehmenden außerhalb ihrer eigenen Einrichtung zustimmen würden.

3.3.5 Schlussfolgerung

Das Ergebnis dieser ersten Iteration ist ein Konzept, welches ermöglicht, das Probandenmanagementsystem unter datenschutzrechtlichen Anforderungen wie der

informationellen Gewaltenteilung als Bestandteil einer Forschungs-IT-Infrastruktur zu betreiben. Neben den sich aus dem Konzept ergebenen technischen Änderungen am Probandenmanagementsystem im weiteren Verlauf dieser Arbeit ist die Einordnung in bereits etablierte Datenschutzkonzepte eine Bereicherung bei der Weiterentwicklung zukünftiger Forschungs-IT-Infrastrukturen. Voraussetzung die Datentrennung praktikabel umzusetzen ist jedoch eine Interaktion des Probandenmanagementsystems mit dem Identitätsmanagement, die in der folgenden Iteration geplant werden muss.

3.4 Iteration 2: Vernetzte Probandenverwaltung

Ein standortübergreifendes Probandenmanagementsystem ist Bestandteil einer Forschungs-IT-Infrastruktur und nur im Zusammenspiel mit weiteren Systemen nutzbar. Das in Iteration 1 vorgestellte Konzept verlangt eine über die Studiendauer fortführende Interaktion mit dem Identitätsmanagement. Die Betrachtung der notwendigen softwareseitigen Datenschnittstellen zum Austausch von Befehlen und Daten zwischen beiden Systemen ist Thema dieses Unterkapitels. Auch diese Iteration gliedert sich in 5 Schritte. Zunächst werden im folgenden Abschnitt das Problem und die Zielsetzung in Bezug auf die Schnittstellen des Probandenmanagementsystems erläutert. Der Methodikteil erläutert anschließend, wie das Problem angegangen wurde. Der Abschnitt *Ergebnisse* beschreibt die Schnittstellen des Probandenmanagementsystems mit ihren Anforderungen und Datenflüssen als Vorarbeit für einen Softwareentwurf. Den Abschluss der Iteration bilden eine Diskussion und eine Schlussfolgerung über die gewonnenen Erkenntnisse.

3.4.1 Problem und Zielsetzung

Iteration 1 machte deutlich, das standortübergreifende Verbundforschungsprojekte zur datenschutzkonformen Datentrennung ein zentrales Identitätsmanagement benötigen, das über Software-Schnittstellen mit anderen Systemen kommuniziert. Im vorgestellten Konzept wurde auch die notwendige Interaktion zwischen Identitätsmanagement und Probandenmanagementsystem begründet. Die Untersuchung der softwareseitigen Datenschnittstellen zwischen beiden Softwaresystemen ist ein weiterer notwendiger Schritt für die Konzeption des standortübergreifenden Probandenmanagementsystems.
Bereits Schwanke [2015] erfasste in seiner Arbeit die Anforderung für den Abruf von Pseudonymen aus externen Systemen [vgl. Schwanke, 2015, Ziel Nr. 15, S. 26].

> „Externe Systeme sind in diesem Zusammenhang Softwaresysteme zur Pseudonymisierung von Probanden in medizinischen Verbundforschungsprojekten. In diesen Projekten wird typischerweise ein zentrales System betrieben, welches die Generierung von Pseudonymen unter Angabe von Probandenstammdaten ermöglicht [41, 47].“ [Schwanke,

2015, S. 89]

Sein Architekturkonzept sieht daher bereits Schnittstellen zu externen Systemen vor, worauf bei der weiteren Konzeption und Implementierung in dieser Arbeit aufgebaut werden kann. Dennoch ist die bloße Nennung der benötigten Schnittstelle für diese Arbeit nicht ausreichend, sowohl um eine gute Interoperabilität, also die Fähigkeit eines Systems mit anderen Systemen zusammenzuarbeiten, für das Probandenmanagementsystem zu begründen, als auch für die prototypische Implementierung hinreichend Informationen bereitzustellen. Schwanke [2015] erläuterte den Prozess der Pseudonymisierung nicht detaillierter:

> „Die Stammdaten des Probanden werden entsprechend der Schnittstelle des externen Systems transformiert und nachfolgend an dieses übermittelt.“ [Schwanke, 2015, S. 79]

Des Weiteren ist ein entscheidender Unterschied zum in dieser Arbeit benötigten Konzept, dass [Schwanke, 2015] vorsieht, Pseudonyme externer Systeme über das Probandenmanagementsystem anzulegen und zu speichern, was für die Lösung dieser Arbeit ausgeschlossen wird (vgl. Unterkapitel 3.3.3, Ergebnisse, S. 41).
Ziel dieser Iteration ist es, die Schnittstellen des Probandenmanagementsystems mit ihren Anforderungen und Datenflüssen zu beschreiben. Grundlage bildet das in Iteration 1 vorgestellte Konzept, das mit dem Fokus auf die benötigten Schnittstellen verfeinert werden soll.

3.4.2 Methodik

Um die Anwendungsfälle der Schnittstelle des Probandenmanagementsystems mit dem Identitätsmanagement zu beschreiben, wurde die in Iteration 1 skizzierten Datenflüsse herangezogen. Zur grafischen Visualisierung der Prozesse wurden Anwendungsfall-, Aktivitäts- und Sequenzdiagramme nach den Vorgaben der Unified Modeling Language (UML) erstellt. Die Anwendungsfälle abstrahieren von einer konkreten technischen Lösung, da diese ihm Rahmen der prototypische Implementierung in Kapitel 4 (S. 100) detaillierter betrachtetet wird.

3.4.3 Ergebnisse

Zwei Anwendungsfälle des Probandenmanagementsystems, die eine Interaktion mit der Schnittstelle zum Identitätsmanagement verlangen, wurden aus dem bisherigen Konzept abgeleitet. Die Anwendungsfälle dienen der Konzeption der Schnittstelle und teilen sich in die Schritte *Pseudonymzuweisung* und *Pseudonymabfrage* (siehe Abbildung 3.10):

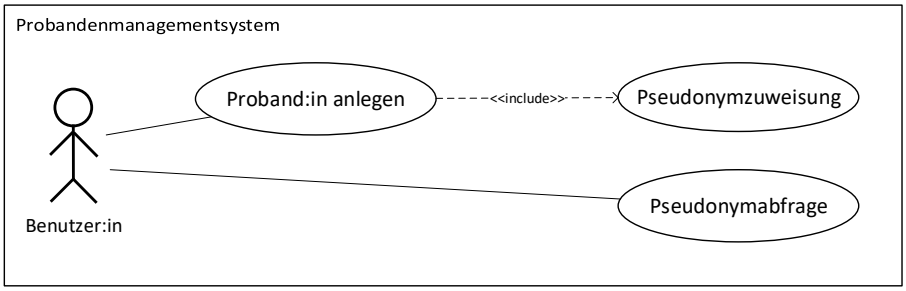

Abbildung 3.10: Anwendungsfalldiagramm für die Interaktion mit der Schnittstelle des Identitätsmanagements eines Systembenutzenden über das Probandenmanagementsystem. 1.) Anwendungsfall: Anlegen eines neuen Studienteilnehmenden im Probandenmanagementsystem, der den Anwendungsfall der Pseudonymzuweisung beinhaltet (Include-Beziehung). 2.) Anwendungsfall: Abfrage eines Pseudonyms eines Studienteilnehmenden.

1. Pseudonymzuweisung:

Das Anlegen eines neuen Studienteilnehmenden im Probandenmanagementsystem kann durch die Eingabe seiner Stammdaten erfolgen [vgl. Schwanke, 2015, Anforderung A-F-15, S. 38]. Eine Duplikaterkennung stellt sicher, dass Studienteilnehmende nicht mehrfach im Probandenmanagementsystem erfasst werden [vgl. Schwanke, 2015, Anforderung A-F-16, S. 38]. Auch eine Datenübernahme aus weiteren Systemen, wie einem Krankenhausinformationssystem, wurde von Schwanke geplant, jedoch aufgrund der Komplexität nicht implementiert [vgl. Schwanke, 2015, S. 68]. Für die Konzeption der Schnittstelle in dieser Arbeit spielt die Datenübernahme keine Rolle, entscheidend ist, das identifizierende Probandenstammdaten vorliegen. Erfasste Probandenstammdaten sind Vorname, Nachname, Titel, Namensvorsatz, Namenszusatz, Geschlecht, Geburtsdatum, Geburtsname und Geburtsort. Bei der Anlage in der Benutzeroberfläche des Probandenmanagementsystems einzugebende Pflichtfelder sind Vorname, Nachname, Geschlecht und Geburtsdatum. Studienteilnehmende werden im Probandenmanagementsystem Forschungsprojekten über Forschungsakten zugeordnet.

Abbildung 3.11 verdeutlicht in einem Aktivitätsdiagramm in UML-Notation die Schritte der lokalen Anlage eines Studienteilnehmenden im Probandenmanagementsystem mit Eingabeprüfung der Pflichtfelder, Prüfung auf Duplikate und der Zuordnung eines Studienteilnehmenden zu einem Forschungsprojekt.

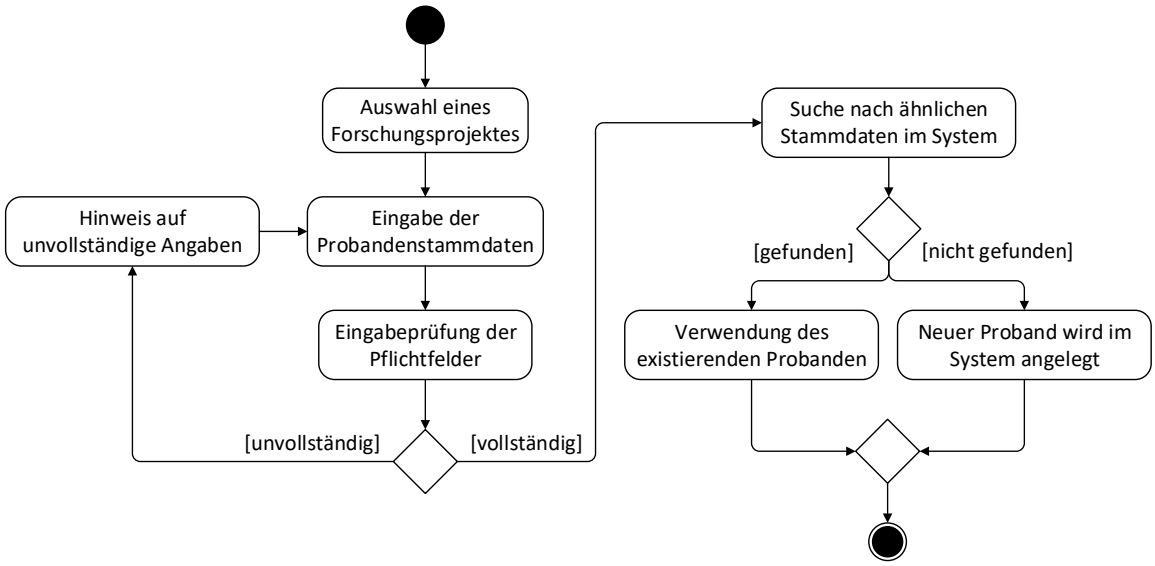

Abbildung 3.11: Aktivitätsdiagramm der Schritte zur Anlage eines Studienteilnehmenden im Probandenmanagementsystem.

Nach diesen Schritten ist der Studienteilnehmende im Probandenmanagementsystem angelegt, jedoch noch mit keinem Pseudonym verknüpft. Für die Pseudonymzuweisung sind folgende Punkte für die spätere Implementierung der Schnittstelle relevant:

(a) Konfiguration des Identitätsmanagements:
Das Probandenmanagementsystem muss Systembenutzenden ermöglichen, pro Forschungsprojekt den Dienst eines Identitätsmanagements zu konfigurieren. Dies umfasst zum einen technische Parameter wie beispielsweise Adresse und Zugangsdaten zu einem Webservice und zum anderen Kontextinformationen zu den durch das Identitätsmanagement verwalteten Pseudonymen. Für den in diesem Abschnitt beschriebenen Schritt des Anlegens eines Studienteilnehmenden sind dies beispielsweise die Information, dass ein Pseudonym vom Typ *Probandenmanagementsystempseudonym* mit der Kennung PSN_{PMS} für einen Studienteilnehmenden nach Übergabe einer Liste von Identifikatoren abgerufen werden soll. Diese Informationen müssen ebenfalls auf Seite des

Identitätsmanagements konfiguriert werden, um das Zusammenspiel beider Systeme zu ermöglichen.

(b) Auswahl der Identifier einer Personenidentität:
Das Identitätsmanagement hat zur Aufgabe einen Studienteilnehmenden anhand seiner personenidentifizierenden Daten eine eindeutige Kennung zuzuweisen, aus der weitere Pseudonyme abgeleitet werden können. Diese personenidentifizierenden Daten müssen dem Identitätsmanagement vom Probandenmanagementsystem über die Schnittstelle übergeben werden und daher im Probandenmanagementsystem konfigurierbar sein. Typischerweise sind die obengenannten Pflichtfelder des Probandenmanagementsystems die Grundlage für die im Identitätsmanagement eingesetzten Verfahren zur Zusammenführung von Datensätzen. Denkbar ist auch die Erfassung weiterer Identifikatoren im Probandenmanagementsystem, die durch ein weiteres System vergeben wurden, wie beispielsweise einer ID aus einem Krankenhausinformationssystem. Der Identifier identifiziert dabei die Personenidentität eindeutig in diesem System. Die Konfiguration der Pflichtfelder und der an die Schnittstelle übergebenen Werte im Probandenmanagementsystem hängt dabei von den Anforderungen des jeweiligen Forschungsprojektes ab.

(c) Abfragezeitpunkt:
Nach der lokalen Anlage des Studienteilnehmenden im Probandenmanagementsystem muss ein geeigneter Zeitpunkt gefunden werden, um das Pseudonym abzufragen. Denkbar ist ein automatischer Abruf nach der lokalen Anlage des Studienteilnehmenden im System oder ein manuell durch Systembenutzenden ausgelösten Abruf des Pseudonyms.

(d) Sicherheit:
Die Schnittstelle zwischen Probandenmanagementsystem und Identitätsmanagement überträgt sensible Probandendaten und muss dementsprechend vor Manipulation und unberechtigtem Mitlesen geschützt werden. Die konkreten Schutzmaßnahmen hängen dabei vom gewählten Identitätsmanagement ab. Das Probandenmanagementsystem muss diese Vorgaben umsetzen.

(e) Verfügbarkeit und Performanz:
Performanz-Kenngrößen an Durchsatz und Bearbeitungszeit sowie Verfügbarkeitsangaben in einer Performanz-Spezifikation als Anforderung an die Schnittstelle des Probandenmanagementsystems festzulegen, ist für diese Arbeit nicht nötig. Die geplanten Pseudonymabfragen des Probandenmanagementsystems sind wenig komplex und zeitlich unkritisch. Forschungsprojekte mit besonderen Anforderungen an Verfügbarkeit und Performanz müssen spezielle Anforderungen abhängig von ihrem Betriebskonzept prüfen.

Die Interaktion zwischen Probandenmanagementsystem und Identitätsmanagement zur Pseudonymabfrage ist vereinfacht, ohne auf technische Details einer späteren Implementierung einzugehen, in Abbildung 3.12 dargestellt. Nach erfolgreicher Anmeldung des Probandenmanagementsystems beim Identitätsmanagement wird ein Probandenmanagementsystempseudonym angefordert. Dazu wird eine Liste mit Identifikatoren der Studienteilnehmenden übergeben, mit denen das Identitätsmanagement ein eindeutiges Pseudonym generieren und an das Probandenmanagementsystem übermitteln kann.

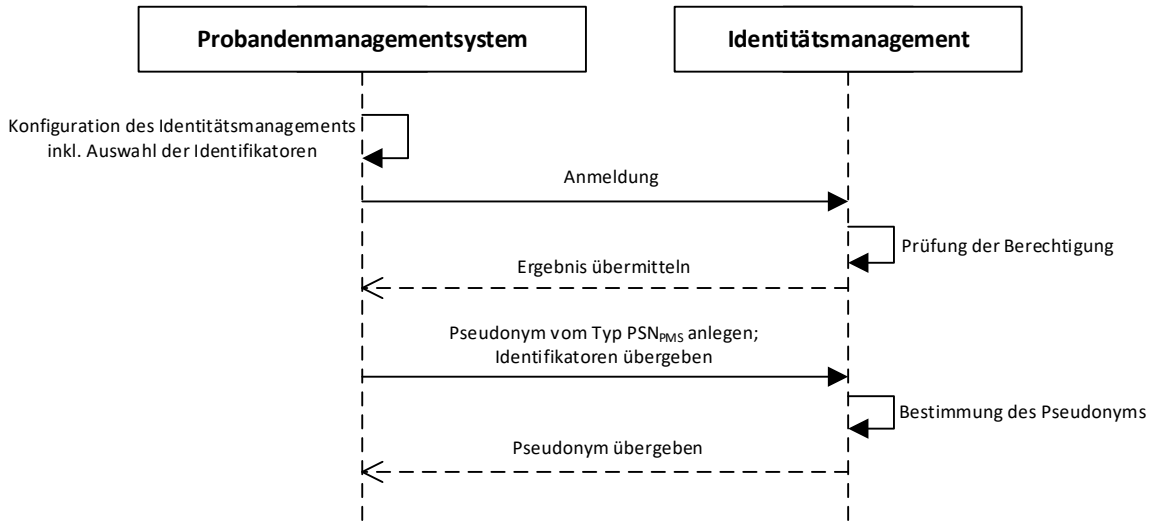

Abbildung 3.12: Sequenzdiagramm zur Interaktion zwischen Probandenmanagementsystem und Identitätsmanagement bei der Pseudonymabfrage.

Das erstellte Probandenmanagementsystempseudonym wird anschließend in der Datenbank des Probandenmanagementsystems verknüpft mit dem Datensatz des Studienteilnehmenden gespeichert.

2. Pseudonymabfrage:

Iteration 1 zeigte, dass in einem Verbundforschungsprojekt mit Clinical Data Management Systemen (CDMS) medizinische Forschungsdaten unter Pseudonymen verwaltet werden. Ähnlich dem Probandenmanagementsystem sind auch die CDMS über eine Schnittstelle mit dem Identitätsmanagement verbunden. Identitätsdaten werden in diesen Systemen zum Abruf eines Pseudonyms eingegeben, jedoch die Identitätsdaten niemals gespeichert (vgl. Ergebnisse Iteration 1, Daten und Datenflüsse des Studienmoduls, S. 46). Möchte das Studienpersonal Datensätze eines Studienteilnehmenden im CDMS betrachten oder bearbeiten, so ist dies nur unter Kenntnis des Pseudonyms möglich. Daraus ergibt sich für das Probandenmanagementsystem der Anwendungsfall der Pseudonymabfrage: berechtigtes Personal darf über das Probandenmanagementsystem die Pseudonyme der CDMS abfragen. Die Pseudonyme der CDMS dürfen jedoch nicht in der Datenbank des Probandenmanagementsystems gespeichert werden (vgl. Ergebnisse Iteration 1, Daten und Datenflüsse des Probandenmanagementsystems, S. 48). Das Konzept für die Pseudonymabfrage aus Sicht des Probandenmanagementsystems über eine geeignete Schnittstelle muss die folgenden Punkte berücksichtigen:

(a) Konfiguration des Identitätsmanagements:
Die in Abschnitt *Pseudonymzuweisung* beschriebene Konfiguration des Identitätsmanagements muss um Angaben abzufragender Pseudonyme der CDMS erweitert werden.

(b) Auswahl der Identifier einer Personenidentität:
Der Identifier, der zur Abfrage der Pseudonyme beim Identitätsmanagement verwendet wird, ist das Probandenmanagementsystempseudonym. Da das Probandenmanagementsystempseudonym bei der Anlage des Studienteilnehmenden im Probandenmanagementsystem auch im Identitätsmanagement mit einer eindeutigen Probandenidentität verknüpft wurde, eignet sich dieses Pseudonym als Schlüssel zur Abfrage weiterer Pseudonyme.

(c) Abfragezeitpunkt:
Die Pseudonymabfrage wird durch berechtigte Systembenutzende des Probandenmanagementsystems gezielt ausgelöst.

(d) Sicherheit:
Analog zu dem im Abschnitt *Sicherheit bei der Pseudonymzuweisung* beschriebenen Vorgehen muss das Probandenmanagementsystem die Sicherheitsmaßnahmen der Schnittstellen des Identitätsmanagements umsetzten.

(e) Verfügbarkeit und Performanz:
Spezielle Anforderung an Verfügbarkeit und Performanz sind für die Schnittstelle dieser Arbeit nicht nötig. Es gelten die gleichen Anforderungen, wie sie im Abschnitt *Pseudonymzuweisung* beschrieben wurden.

Abbildung 3.13 zeigt den Ablauf der Pseudonymzuweisung eines CDMS und des Probandenmanagementsystems sowie der Pseudonymabfrage des Probandenmanagementsystems.

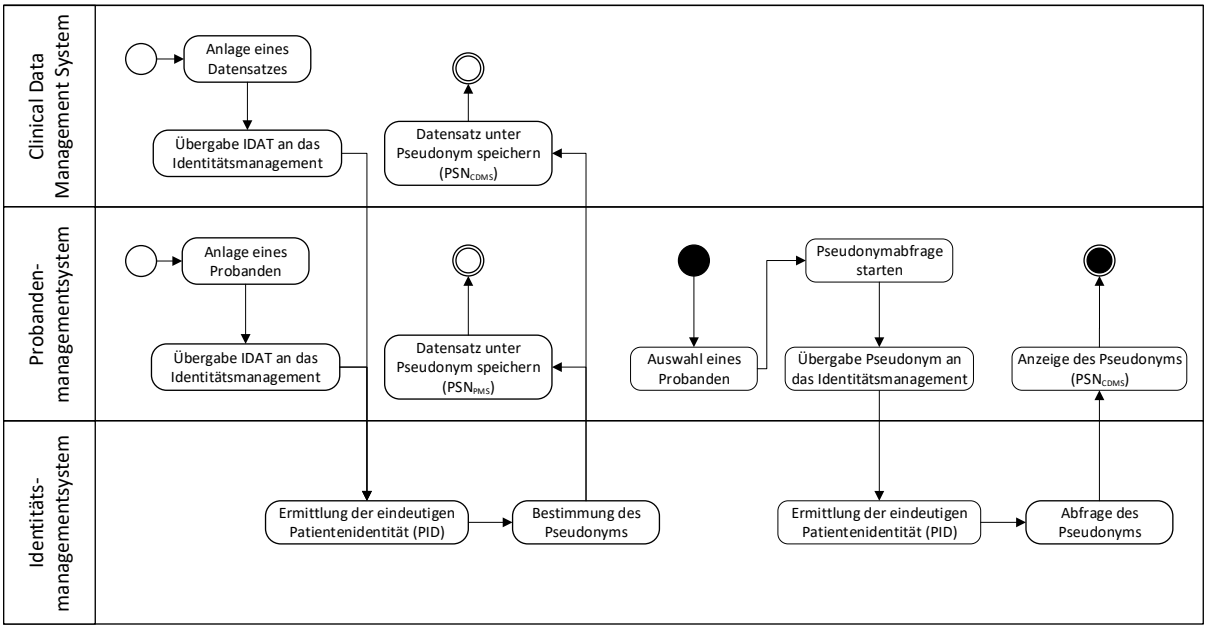

Abbildung 3.13: Übersicht über die Datenflüsse von Pseudonymzuweisung und Pseudonymabfrage zwischen Clinical Data Management System und Probandenmanagementsystem.

Der gesamte Prozess von *Pseudonymzuweisung* und *Pseudonymabfrage* zwischen Clinical Data Management System und Probandenmanagementsystem gliedert sich in drei Teile. Die erste Interaktion mit dem Identitätsmanagement über die Schnittstelle erfolgt entweder durch das Probandenmanagementsystem bei Anlage eines Studienteilnehmenden

oder durch das CDMS bei Anlage eines Datensatzes. Beide Systeme übermitteln, um ein eigenes Pseudonym zum Speichern ihrer Daten zu erhalten, probandenidentifizierende Daten an das Identitätsmanagement. Das Identitätsmanagement ermittelt für diese identifizierenden Daten die eindeutige Probandenidentität. In einem zweiten Schritt wird dann für das anfragende System ein Pseudonym generiert und zurück übermittelt. Der schwarze Punkt in Abbildung 3.13 ist der Startknoten für den dritten Teil des Prozesses, der Pseudonymabfrage. Im Probandenmanagementsystem wählt der Systembenutzende einen Probandendatensatz aus, für den er das Pseudonym des CDMS erfahren möchte. Als Identifikator des Studienteilnehmenden wird dem Identitätsmanagement über die Schnittstelle das Probandenmanagementsystempseudonym übergeben. Das Probandenmanagementsystempseudonym ist dem Probandenmanagementsystem seit Anlage des Studienteilnehmenden im System durch den Teilschritt der Pseudonymzuweisung bekannt. Das Identitätsmanagement ist in der Lage, zwischen den verschiedenen Systemen zu verknüpfen und so das CDMS-Pseudonym an das Probandenmanagementsystem zu übermitteln. Es ist möglich, dass kein CDMS-Pseudonym abgerufen werden kann, wenn noch kein Datensatz im CDMS angelegt wurde und somit kein Pseudonym generiert wurde.

3.4.4 Diskussion

Das Ergebnis dieser Iteration wirkt trivial, lediglich zwei Anwendungsfälle zur Pseudonymzuweisung und Pseudonymabfrage wurden als Anforderung an die Schnittstelle beschrieben. Dennoch sind diese Funktionen unverzichtbar, um die in Iteration 1 eingeführte Datentrennung bei der Probandenverwaltung möglich zu machen. Ohne diese würde die Konzeption des standortübergreifenden Probandenmanagementsystems an den genannten Datenschutzanforderungen scheitern. Hinsichtlich der Umsetzung der Anwendungsfälle wird im Kapitel 4 (Prototypische Implementierung des Probandenmanagementsystems, S. 100) die detaillierte Untersuchung der für Forschungsprojekte verfügbaren Identitätsmanagement / Pseudonymisierungsdienst-Lösungen nötig. In dieser Iteration wurden lediglich die Anwendungsfälle mit ihren Anforderungen beschrieben und nicht die Machbarkeit untersucht. Dabei könnte die Pseudonymabfrage zur größeren Hürde bei der Umsetzung werden, da diese Funktion nicht zu den klassischen Anwendungsfällen eines

Identitätsmanagements gehört. Die Pseudonymzuweisung hingegen ist die Kernaufgabe eines Identitätsmanagements, IDAT werden entgegengenommen und ein Pseudonym zurückgeliefert. Ob das aufrufende System ein Probandenmanagementsystem oder ein Clinical Data Management System ist, spielt keine Rolle.

3.4.5 Schlussfolgerung

In dieser zweiten Iteration der Konzeption wurde das Ziel angegangen, das Probandenmanagementsystem über eine Schnittstelle mit dem Identitätsmanagement zu verknüpfen (vgl. Unterkapitel 3.2, Anforderungen an eine standortübergreifende Probandenverwaltung: Anforderung 3, S. 32). Zwei Anwendungsfälle beschrieben die Interaktion mit der Schnittstelle aus Sicht des Probandenmanagementsystems. Der iterative Ansatz dieser Konzeptionsphase unterstützte dabei, die Erkenntnisse aus Iteration 1 um Aspekte der Schnittstellenanforderungen zu verfeinern. Die Beschreibung der Abläufe und der Anforderungen an die Software-Schnittstellen in dieser Iteration ist ein wichtiger Schritt in Hinblick auf die Implementierung.

3.5 Iteration 3: Zugangs- und Zugriffskontrolle

Zur Lösung von Anforderung 1: *Ein vertraulicher Zugang und Zugriff auf die Probandendaten muss sichergestellt sein* wird in dieser Iteration ein Konzept erarbeitet. Bei Probandendaten handelt es sich um besonders schützenswerte Daten, für die das Probandenmanagementsystem Schutzmaßnahmen implementieren muss. In den vorangegangenen Iterationen wurde ein Konzept entworfen, dass durch einen verteilten Zugriff auf die Daten den Einsatz des Probandenmanagementsystems in standortübergreifenden Forschungsprojekten ermöglicht. Durch den verteilten Zugriff steigen jedoch auch die Datenschutz- und Sicherheitsanforderungen an das System. Teil dieser dritten Iteration ist es, die Probleme und Herausforderungen für das Probandenmanagementsystem in Hinblick auf eine sichere Zugangs- und Zugriffskontrolle zu erarbeiten und Anforderungen zur Zielerreichung zu definieren. Anschließend werden die Möglichkeiten zur Erfüllung der Anforderungen als Methodik untersucht. Im Ergebnisteil wird dann ein Konzept für jede Anforderung vorgestellt. Die Konzeptdiskussion zur Eignung der Maßnahmen zur Zugangs- und Zugriffskontrolle für das Probandenmanagementsystem bildet den Abschluss der Iteration. Teile der Ergebnisse dieser Iteration wurden bei der 65. GMDS[10]-Jahrestagung 2020 als Vortrag mit dem Titel *Access control policy for a study participant management system for multi-site clinical research* [Pung u. a., 2021] vorgestellt (Anhang C).

3.5.1 Problem und Zielsetzung

Das Probandenmanagementsystem, wie es durch Schwanke [2015] konzeptioniert und umgesetzt wurde, erlaubt eine Zugangskontrolle über eine Anmeldung im System mit Benutzername und Passwort [vgl. Schwanke, 2015, S. 36]. Die Sicherheit beruht bis zu einem gewissen Grad auf der Selbstverantwortung und Gewissenhaftigkeit der Anwender (Wahl eines starken Passwortes, sichere Aufbewahrung des Passwortes und Verzicht eines automatischen Logins durch ein gespeichertes Passwort). Eine erweiterte Zugangskontrolle durch die Überarbeitung der Sicherheitsmechanismen des Anmeldeverfahrens des Probandenmanagementsystems war daher nötig. Daraus ergab sich *Anforderung A1: Sichere Authentifizierung.*

[10] Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) e.V.

Bei der Übertragung von Informationen über Kommunikationsnetze besteht die Gefahr, dass Informationen unbefugt abgehört oder manipuliert werden. Um die Vertraulichkeit und die Integrität der Informationen sicherzustellen, müssen Maßnahmen zur Absicherung der Übertragung ergriffen werden. Daraus ergab sich *Anforderung A2: Verschlüsselte Verbindung*.

Eine Zugriffskontrolle auf die Probandendaten erlaubte das ursprünglich konzipierte Probandenmanagementsystem nur auf Forschungsprojektebene. Das bedeutet, dass Systembenutzende, die Zugang zu einem Forschungsprojekt im Probandenmanagementsystem haben, alle Probandendaten des Forschungsprojektes einsehen können. Zur Wahrung des beschriebenen Need-to-know-Prinzips (vgl. Unterkapitel 3.2, S. 32) dürfen Forschende jedoch nur direkten Zugriff auf Probandendaten der ihnen berechtigten Studien und der ihnen berechtigten Standorte erhalten. Daher muss für ein standortübergreifend nutzbares Probandenmanagementsystem ein Berechtigungskonzept erstellt werden, aus dem eindeutig abzulesen ist, welche Benutzergruppen es gibt und wer die damit verbundenen Rechte bezüglich des Datenzugriffs hat. Das Berechtigungskonzept ist von hoher Bedeutung, um technische Maßnahmen eines Zugriffskontrollmodells implementieren zu können, als auch um die Berechtigungen in einem Datenschutzkonzept darzulegen. Daraus ergab sich *Anforderung A3: Berechtigungskonzept*.

Für den Betrieb des Probandenmanagementsystems sind technische Maßnahmen zu ergreifen, um die Einhaltung des Berechtigungskonzeptes sicherzustellen. Daher muss das Probandenmanagementsystem ein geeignetes Zugriffskontrollmodell implementieren. Daraus ergab sich *Anforderung A4: Zugriffskontrollmodell*.

Eine weitere Maßnahme, um einen vertraulichen Zugriff sicherzustellen, ist die personenbezogene Protokollierung von Zugriffen auf Datensätze im System. Diese Protokollierung ist durch § 22 Abs. 2 Nr. 2 BDSG[11] im deutschen Datenschutzrecht geregelt: "Maßnahmen, die gewährleisten, dass nachträglich überprüft und festgestellt werden kann, ob und von wem personenbezogene Daten eingegeben, verändert oder entfernt worden sind". Das Probandenmanagementsystem muss daher eine Möglichkeit bieten, den Zugriff auf die Datenbestände zu protokollieren. Daraus ergab sich

[11] § 22 BDSG Verarbeitung besonderer Kategorien personenbezogener Daten: `https://dejure.org/gesetze/BDSG/22.html` [Stand: 01.03.2022]

Anforderung A5: Protokollierter Zugriff.

3.5.2 Methodik

Anforderung A1: Sichere Authentifizierung

Im deutschen Sprachgebrauch häufig synonym verwendet, unterscheidet die Fachsprache die Begriffe *Authentisierung*[12] für das Vorlegen eines Nachweises einer bestimmten Identität und *Authentifizierung*[12] für die Prüfung der Echtheit der behaupteten Identität. Bei der *Autorisierung*[12] werden für eine erfolgreich authentifizierte Person die ihr auf einem System eingeräumten Rechte freigeschaltet. Im Beispiel des Probandenmanagementsystems behauptet der Systembenutzende zuerst seine Zugangsberechtigung, indem er einen Benutzernamen eingibt. Zusätzlich authentisiert er sich, indem er sein Passwort angibt. Das Probandenmanagementsystem identifiziert den Systembenutzenden anhand dieser Angaben und führt anschließend die Authentifizierung durch, also die Verifizierung der erbrachten Behauptung über die Authentizität. Dieser Vorgang ist die einfachste Form der Zugangskontrolle, wie sie häufig in Softwaresystemen eingesetzt wird. Wie in der Problemstellung beschrieben, sollte jedoch ein System, das sensible Probandendaten von mehreren Studienstandorten beinhaltet, sicherere Verfahren zur Authentifizierung umsetzen. Um Maßnahmen zur Verbesserung der Authentifizierung des Probandenmanagementsystems abzuleiten, wurde die aktuelle *Handlungsempfehlung sichere Authentifizierung* [LfDNI, 2020] der Landesbeauftragten für den Datenschutz in Niedersachsen untersucht.

Anforderung A2: Verschlüsselte Verbindung

Besonderer Bedarf eine verschlüsselte Verbindung zu nutzen besteht immer dann, wenn sensible Daten übermittelt werden, wie es beim Probandenmanagementsystem der Fall ist. Zur Auswahl eines geeigneten Verfahrens wurden die aktuellen technischen Richtlinien des Bundesamts für Sicherheit in der Informationstechnik (BSI) verwendet

[12] Glossar der Cyber-Sicherheit - Bundesamt für Sicherheit in der Informationstechnik: `https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Glossar-der-Cyber-Sicherheit/Functions/glossar.html` [Stand: 01.03.2022]

[BSI, 2020b,c]. Diese geben Empfehlungen zu kryptographischen Verfahren sowie grundsätzliche Hinweise für die Verwendung von Verschlüsselungsprotokollen. Für das Probandenmanagementsystem wurde anhand der Empfehlungen ein Konzept abgeleitet.

Anforderung A3: Berechtigungskonzept

Beim Thema *Berechtigung* geht es darum, Ressourcen vor unbefugter Nutzung und Missbrauch zu schützen. Das können physische Objekte wie Räume, Fahrzeuge oder Geräte sein oder auch gespeicherte Daten oder Softwarefunktionen [vgl. Tsolkas und Schmidt, 2017, S. 1]. Ressourcen werden in Inhaltsressourcen und Funktionsressourcen unterschieden, je nachdem was durch die Ressource bereitgestellt wird [vgl. Tsolkas und Schmidt, 2017, S. 1 f.]. Für das Probandenmanagementsystem ist eine Inhaltsressource beispielsweise die Kontaktadresse eines Studienteilnehmenden, eine Funktionsressource beispielsweise die Funktion diese Kontaktadresse zu drucken. Eine Ressource kann auch gleichzeitig Funktions- und Inhaltsressource sein, wie beispielsweise das Auflösen eines Pseudonyms für einen Studienteilnehmenden von einem externen Pseudonymisierungsdienst. Das elementarste Element für die Strukturierung von Berechtigungen ist die Berechtigung selbst, die aus zwei Komponenten besteht: Der zu berechtigenden Ressource und der zu berechtigenden Operation für diese Ressource [vgl. Tsolkas und Schmidt, 2017, S. 3 f.].

Eine effektive Zugangs- und Zugriffskontrolle kann nur gewährleistet werden, wenn klar definiert wurde, welche Berechtigungen für die Systembenutzenden auf die Daten und Funktionen eines IT-Systems gelten. In einem Berechtigungskonzept werden diese Regeln beschrieben [vgl. ISDSG, 2021]. Ein Berechtigungskonzept wird für das Probandenmanagementsystem benötigt, um die Zugriffsregeln schriftlich im Datenschutzkonzept vorzuweisen, als auch um ein geeignetes Zugriffskontrollmodell auswählen und implementieren zu können (Anforderung A4).

Bei der Erstellung des Berechtigungskonzeptes wird in der Regel das Ziel angepeilt, alle Informationen im Unternehmen über Nutzer, ihre Aufgaben und Rollen, über Geräte und Anwendungen, sowohl lokal als auch im Netzwerk und in der Cloud, zusammenzustellen. In weiteren Schritten werden dann aus den Informationen digitalen Identitäten abgebildet, z. B. Nutzer, Geräte und Anwendungen, denen dann Zugriffsrechte zugewiesen werden [vgl. Schonschek, 2019]. Berechtigungskonzepte müssen vollständig, umfassend und aktuell

sein. Solch ein vollständiges Konzept konnte im Rahmen dieser Arbeit jedoch nicht erstellt werden, da dafür ein konkretes laufendes Forschungsprojekt fehlt. Dennoch reicht es für die obengenannten Ziele dieser Arbeit aus, typische Systembenutzende des Probandenmanagementsystems und ihre Berechtigungen zu betrachten. Damit lassen sich beispielsweise alle Berechtigungen zusammenfassen, die sich für eine Mitarbeitendengruppe in einer klinischen Studie aus den dortigen Arbeitsprozessen ergeben. Auf Basis der in Unterkapitel 2.2 (Prozesse der Probandenverwaltung, S. 15) gesammelten Informationen und Literatur sowie der Sichtung der durch Schwanke [2015] durchgeführten Stakeholderinterviews wurde für das Probandenmanagementsystem die Basis für ein Berechtigungskonzept definiert.

Anforderung A4: Zugriffskontrollmodell

Aufbauend auf dem Berechtigungskonzept musste für das Probandenmanagementsystem ein Zugriffskontrollkonzept entwickelt werden, indem ein zu den Anforderungen passendes Zugriffskontrollmodell ausgewählt wurde. Zugriffskontrollmodelle sind Entwurfsmuster (engl. design patterns) zur Zugriffskontrolle und beschreiben, auf welche Art und Weise Systembenutzende Zugriff auf IT-Systeme erhalten. Entwurfsmuster sind bewährte Lösungsschablonen für wiederkehrende Entwurfsprobleme in der Softwarearchitektur und -entwicklung. Vier Zugriffskontrollmodelle, die häufig Anwendung finden, sind:

- Discretionary Access Control (DAC)
  Die Entscheidung, ob auf eine Ressource zugegriffen werden darf, wird auf der Basis der Identität der Benutzerin oder des Benutzers getroffen. Benutzerinnen und Benutzer können als Eigentümerin oder Eigentümer einer Ressource den Zugriff an andere Benutzerinnen und Benutzer übertragen und legen so die Zugriffsrechte für eine Ressource fest.

- Mandatory Access Control (MAC)
  Die Entscheidungen über Zugriffsberechtigungen werden nicht nur auf der Basis der Identität der Benutzerin oder des Benutzers und der Ressource auf die zugegriffen werden soll gefällt, sondern aufgrund von Regeln und zusätzlicher Informationen.

- Role Based Access Control (RBAC)
  Bei der rollenbasierten Zugriffskontrolle werden Benutzerinnen und Benutzern

Rollen zugeordnet. Eine Rolle im Sinne des Berechtigungskonzeptes beschreibt eine Funktion bezogen auf ein Aufgabengebiet, einen Verantwortungsbereich oder ein Tätigkeitsspektrum, die Rolleninhabende ausüben. Durch die Rollenzuordnung erhält die Benutzerin oder der Benutzer die den Rollen zugeordneten Zugriffsrechte.

- Attribute Based Access Control (ABAC)
  Der Zugriff der Benutzerin oder des Benutzers auf eine Ressource wird anhand von Attributen der Benutzerin oder des Benutzers, der Ressource, der Systemumgebung sowie auf diese Attribute angewendeten Regeln gesteuert.

Für alle Modelle wurden Entstehung, Grundlagen und Funktion recherchiert und stark zusammengefasst, bevor auf deren Basis die Eignung der Modelle für das Probandenmanagementsystem im Ergebnissteil bewertet wurde:

▸ Discretionary Access Control (DAC)

Das DAC-Modell wurde Ende der 1960er Jahren entwickelt [vgl. Kashmar u. a., 2019, S. 894]. Es basiert auf dem Eigentümer-Prinzip, d.h. eine Benutzerin oder ein Benutzer als Eigentümerin oder Eigentümer einer Ressource legt die Zugriffsrechte auf diese fest [vgl. Eckert, 2008, S. 232]. Das Modell wird *discretionary* (dt. Ermessenspielraum) genannt, da eine Benutzerin oder ein Benutzer mit entsprechenden Rechten anderen Benutzerinnen und Benutzern nach eigenem Ermessen Zugriffsrechte einräumen kann. Dateiberechtigungen in Linux- und Windows-Betriebssystemen sind Beispiele für die Implementierungen von Varianten des DAC-Modells. Eine Form der Umsetzung von DAC ist eine Zugriffsmatrix wie in Tabelle 3.1 gezeigt.

	Datei 1	Datei 2	Datei 3
Benutzer:in A	own, read, write		own, read, write
Benutzer:in B	read	own, read, write	read, write

Tabelle 3.1: Beispiel einer Zugriffskontrollmatrix wie sie mit dem Discretionary Access Control Zugriffskontrollmodell umgesetzt werden kann. Benutzer:in A, die/der Eigentümer:in der Dateien 1 und 3 ist, hat Benutzer:in B berechtigt Datei 1 zu lesen sowie Datei 3 zu lesen und zu bearbeiten. Benutzer:in B, die/der Eigentümer:in der Datei 2 ist, gewährt Benutzer:in A keine Rechte an dieser Datei.

▸ Mandatory Access Control (MAC)

In den 1970er Jahren wurde das MAC-Zugriffskontrollmodell vorgestellt [vgl. Kashmar

u. a., 2019, S. 894]. Im Gegensatz zu DAC baut die Zugriffskontrolle von MAC auf die von einer zentralen Autorität festgelegten Regeln auf [Samarati und de Vimercati, 2001]. Bei MAC können Benutzende Rechte nicht selbst definieren. Jeder Ressource ist eine Sicherheitsstufe zugeordnet, bspw. *vertraulich* oder *geheim*. Jeder Benutzende innerhalb eines Systems erhält eine entsprechende Sicherheitsstufe. Aufgrund der restriktiven Strategie wird MAC häufig im Regierungs- und Militärumfeld eingesetzt.

Ressource	Sicherheitsstufe
Mitarbeiterliste	vertraulich
Finanzbericht	geheim

Benutzende	Sicherheitsstufe
Benutzer:in A, Benutzer:in B	vertraulich
Benutzer:in C	geheim

Tabelle 3.2: Beispiel für eine Rechtevergaben nach dem MAC-Zugriffskontrollmodell. Die Ressource Mitarbeiterliste ist als *vertraulich*, der Finanzbericht als *geheim* definiert worden. Benutzer:in A und B dürfen vertrauliche Ressourcen, Benutzer:in C geheime Ressourcen einsehen.

▸ Role Based Access Control (RBAC)

Das RBAC-Modell wurde 1992 von Ferraiolo und Kuhn [1992] beschrieben, im Jahr 2000 in ein einheitliches RBAC-Modell von Sandhu u. a. [1996] integriert, 2001 als Standard dem National Institute of Standards and Technology (NIST) vorgestellt [Ferraiolo u. a., 2001], und 2004 als ANSI[13]-Norm 359-2004 verabschiedet. Die rollenbasierte Zugriffskontrolle verwirklicht die Umsetzung der Rechtevergabe anhand von Rollen. Benutzende bekommen über ihre zugewiesenen Rollen ihre Zugriffsrechte zugeteilt. Abbildung 3.14 verdeutlicht das Grundprinzip der rollenbasierten Zugriffskontrolle.

[13] American National Standards Institute (ANSI)

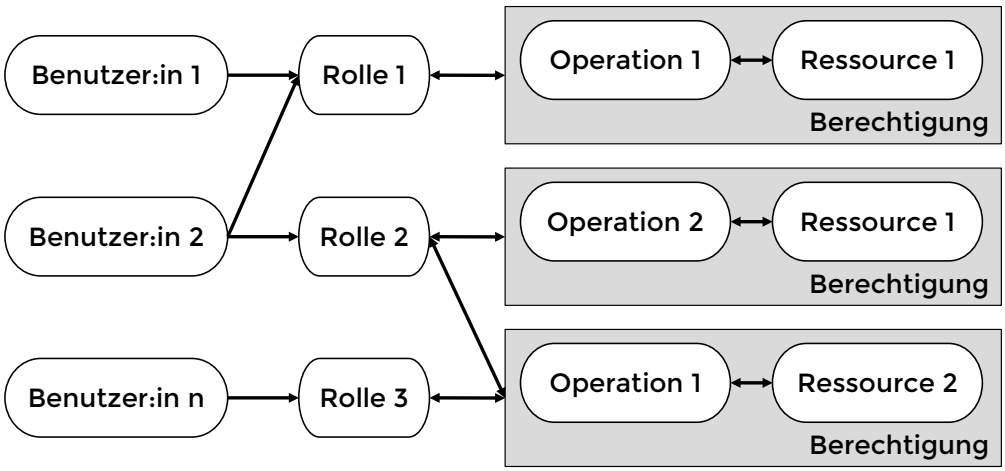

Abbildung 3.14: Grundprinzip der rollenbasierten Zugriffskontrolle. Benutzende können ein oder mehrere Rollen einnehmen, denen ein oder mehrere Berechtigungen zugewiesen werden. Eine Berechtigung besteht aus einer Operation auf eine Ressource. Eigene Abbildung mit abgeänderter Terminologie angelehnt an [Hagenhoff, 2008, S. 295].

Die Verwaltung der Zugriffsrechte anhand von Rollen ermöglicht die Entkopplung der Berechtigungen von den Identitäten [vgl. Tsolkas und Schmidt, 2017, S. 12]. Des Weiteren lassen sich mit Rollenhierarchien Organisationsstrukturen und Vererbungen von Zugriffsrechten abbilden [vgl. Seufert, 2002, S. 4]. Für detaillierte Informationen sei auf die Webseite der NIST[14] mit Hintergrundinformationen und Implementierungsbeispielen zur rollenbasierten Zugriffskontrolle verwiesen.

▸ Attribute Based Access Control (ABAC)

Bei DAC, MAC und RBAC basiert die Zugriffskontrolle auf der Identität eines Benutzenden, der eine Berechtigung auf eine Ressource besitzt entweder direkt oder indirekt über Rollen oder Gruppen, die diesem Benutzenden zugewiesen sind. Eine Alternative besteht darin, Benutzenden Rechte basierend auf Attributen der Benutzenden, Attributen der Ressource oder Umgebungsbedingungen zu gewähren oder zu verweigern. Dieser Ansatz wird als attributbasierte Zugriffskontrolle bezeichnet [Priebe u. a., 2005; Hu u. a., 2014]. Abbildung 3.15 verdeutlicht das Grundprinzip der attributbasierten Zugriffskontrolle.

[14] Role Based Access Control, National Institute of Standards and Technology: `https://csrc.nist.gov/projects/role-based-access-control` [Stand: 01.03.2022]

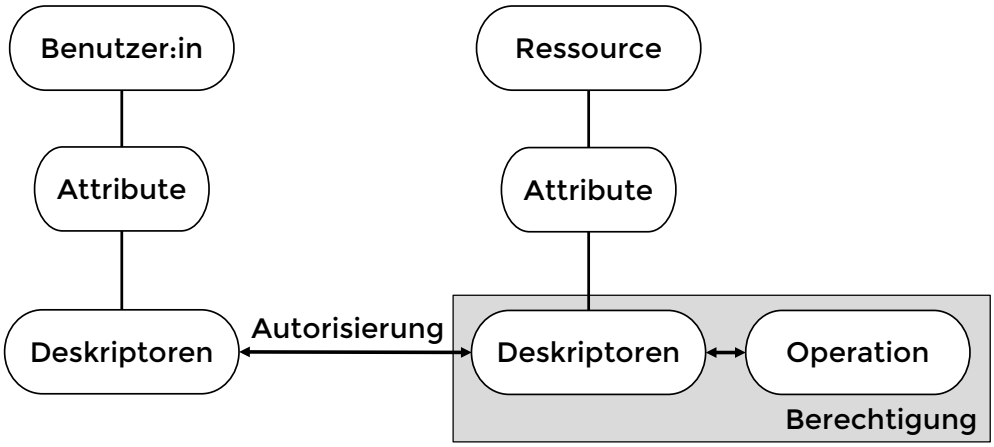

Abbildung 3.15: Grundprinzip der attributbasierten Zugriffskontrolle. Benutzende sowie Ressourcen besitzen eine Menge von Attributen. Deskriptoren bestehen aus einer Menge von Attributen die Bedingungen beinhalten (z. B. Zugehörigkeit zu Organisationseinheit X und Aufgabe = Projektleiter). Durch Prüfung dieser Bedingungen, wird eine Berechtigung auf eine Operation gewährt oder verweigert. Eigene Abbildung mit abgeänderter Terminologie angelehnt an [Hagenhoff, 2008, S. 296].

Für ausführliche Informationen sei auf die Webseite der NIST[15] verwiesen, auf der auch der *Guide to Attribute Based Access Control (ABAC) Definition and Considerations* [Hu u. a., 2014] zu finden ist, in dem ABAC näher beschrieben wird.

Anforderung A5: Protokollierter Zugriff

Sowohl die Datenschutzgesetze der Länder, des Bundes als auch die europäische Datenschutz-Grundverordnung enthalten Regelungen, nach denen Protokollierung zum Nachweis der ordnungsgemäßen Verarbeitung personenbezogener Daten verwendet werden darf und sollte. Hat eine Benutzerin oder ein Benutzer einen Zugriff oder eine Aktion durchgeführt, ermöglicht die Protokollierung, dass im Nachhinein die Urheberschaft des Zugriffs oder der Aktion eindeutig der entsprechenden Person zuordenbar ist. Protokollierung bildet dann eine wichtige Maßnahme, um *Verbindlichkeit* zu gewährleisten, einem der grundlegenden Schutzziele bei der Sicherheit von Informationssystemen [vgl. Eckert, 2008, S. 6, S. 11]. In der Regel enthalten Protokolle personenbeziehbare Daten, weshalb bei jeder Protokollierung datenschutzrechtliche Vorgaben zu beachten sind. Um diese Vorgaben für das Probandenmanagementsystem

[15] Attribute Based Access Control, National Institute of Standards and Technology: `https://csrc.nist.gov/projects/attribute-based-access-control` [Stand: 01.03.2022]

zu berücksichtigen, wurde die *Praxishilfe zur Protokollierung und zur Erstellung von Protokollierungskonzepten im Gesundheitswesen* [Ammon u. a., 2020] herangezogen. Die Praxishilfe entstand aus einer Zusammenarbeit der *Arbeitsgruppe Datenschutz & IT-Sicherheit* des bvitg[16], der Arbeitsgruppe *Datenschutz und IT-Sicherheit im Gesundheitswesen* der GMDS[17] und IHE Deutschland[18]. Im Jahr 2020 verfasst, eignet sich die Praxishilfe aufgrund ihrer Aktualität hervorragend als Referenz um Maßnahmen der Protokollierung für das Probandenmanagementsystem zu planen. Weiterhin empfiehlt sich die Betrachtung des *Baustein Protokollierung* [BSI, 2020a] aus dem *BSI IT-Grundschutz*[19].

3.5.3 Ergebnisse

Anforderung A1: Sichere Authentifizierung

Aus der *Handlungsempfehlung sichere Authentifizierung* der Landesbeauftragten für den Datenschutz in Niedersachsen wurden die folgenden Empfehlung ausgewählt, aus denen Maßnahmen zur Verbesserung der Authentifizierung des Probandenmanagementsystems abgeleitet wurden:

1. „Entscheidend für die Auswahl einer geeigneten Authentifizierung ist das Anwendungs-Szenario und die Betrachtung des darin enthaltenen Risikos z. B. zur Anmeldung an einem lokalen System (z. B. im LAN) oder zur Nutzung externer Accounts (z. B. Mail oder Webseiten), [...]“ [vgl. LfDNI, 2020, S. 5]

   Da weitere Entscheidungen von dieser Empfehlung abhängen, wurde dieser Punkt der Handlungsempfehlung zuerst bedacht. Zunächst musste das Risiko beurteilt werden, ob die Anmeldung im System aus einem eingeschränkten oder unbeschränkten Netzwerk erfolgt. Das Probandenmanagementsystem muss, um in Verbundforschungsvorhaben zum Einsatz kommen zu können, von verschiedenen Studienstandorten und damit aus dem Internet erreichbar sein (vgl. Iteration 1: Datentrennung bei der Probandenverwaltung, S. 36). Ein weltweiter Zugriff auf

[16] Bundesverband Gesundheits-IT - bvitg e. V.

[17] Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) e.V.

[18] IHE Deutschland e.V.; IHE (Integrating the Healthcare Enterprise)

[19] IT-Grundschutz, Bundesamt für Sicherheit in der Informationstechnik (BSI): `www.bsi.bund.de/grundschutz` [Stand: 01.03.2022]

den isolierten Netzwerkbereich des Probandenmanagementsystems kann jedoch vermieden werden, indem der Zugriff auf autorisierte IP-Adressen eingeschränkt wird. Da in Verbundforschungsprojekten die beteiligten Projektpartner bekannt sind, können IP-Adressen oder IP-Adressräume der teilnehmenden Einrichtungen erfasst werden. Die Kommunikation zwischen Probandenmanagementsystem und externen Systemen wird dann durch die Beschränkung von IP-Adressen oder IP-Adressräume abgesichert. Nur autorisierten IP-Adressen ist grundlegend ein Verbindungsaufbau zum Probandenmanagementsystem möglich. Der Zugriff auf den Server des Probandenmanagementsystems kann durch die Maßnahme der IP-Zugriffsbeschränkung als ein zugangseingeschränktes Netzwerk betrachtet werden.

2. „In gesicherten Bereichen kann somit selbst ein Passwort, das lediglich acht Zeichen umfasst und aus Groß- und Kleinbuchstaben sowie Ziffern zusammengesetzt ist, einen hinreichenden Schutz sicherstellen. [...] Ganz anders stellt sich die Situation in ungesicherten Bereichen dar, die einen Brute-Force-Angriff nicht wirksam ausschließen; hier kommt es darauf an, Passworte so komplex und lang zu gestalten, dass ein erfolgreiches Durchprobieren aller möglichen Kombinationen jeden sinnvollen Zeitrahmen sprengt. [...]“ [vgl. LfDNI, 2020, S. 6]

Dieser Empfehlung folgend, wurde für das Probandenmanagementsystem die Maßnahme abgeleitet, die Passwortwahl der Benutzenden hinsichtlich der Passwortstärke bei der Vergabe zu überprüfen. Das Probandenmanagementsystem muss bei der erstmaligen Anmeldung und Auswahl des benutzereigenen Passwortes prüfen, ob das gewählte Passwort mindestens 8 Zeichen lang, Groß- und Kleinbuchstaben sowie Ziffern enthält.

3. „Unabhängig von der Passwortstärke muss, um eine höhere Sicherheit zu erreichen, zu einem Passwort ein gesalteter Hash gebildet werden und nur dieser darf auf der prüfenden Seite gespeichert werden.“ [vgl. LfDNI, 2020, S. 5]

Zusätzlich zur Einrichtung der Passwortrichtlinien müssen die Passwörter der Benutzenden des Probandenmanagementsystems sicher gespeichert werden. Die Speicherung muss mit einer geeigneten kryptografischen Hashfunktion erfolgen.

4. „Benutzer erhalten je nach eingesetztem Authentifizierungsverfahren eine oder mehrere Faktoren (Benutzerkennung, persönliches Passwort, Token, Pin-Nummer, EC-Karte...), um sich so gegenüber dem System als berechtigt ausweisen zu können.“ [vgl. LfDNI, 2020, S. 1]

Eine Verbesserung der Sicherheit des Identitätsnachweises (Authentisierung) eines Nutzers lässt sich mittels der Kombination zweier unterschiedlicher und insbesondere unabhängiger Komponenten (Faktoren) erreichen. Das Bundesamt für Sicherheit in der Informationstechnik (BSI) definiert in ihrem IT-Grundschutz-Kompendium eine starke Authentisierung wie folgt:

> „Starke Authentisierung bezeichnet die Kombination von zwei oder mehr Authentisierungstechniken, wie Passwort plus Transaktionsnummern (Einmalpasswörter) oder plus Chipkarte. Daher wird dies auch häufig als Zwei- oder Mehr-Faktor-Authentisierung bezeichnet.“ [vgl. IT-Grundschutz-Kompendium, BSI, 2021, S. 8]

Sowohl die Landesbeauftragte für den Datenschutz in Niedersachsen in der *Handlungsempfehlung sichere Authentifizierung* [LfDNI, 2020] als auch das BSI im IT-Grundschutz-Kompendium [BSI, 2021] raten dazu, für sicherheitskritische Anwendungen die Zwei-Faktor-Authentisierung zu verwenden. Bei der Multi-Faktor-Authentisierung (MFA), zu der die Zwei-Faktor-Authentisierung zählt, werden üblicherweise drei Kategorien von Faktoren unterschieden [vgl. LfDNI, 2020]:

- Nachweis der Kenntnis einer Information (Wissensfaktor): z. B. Passwort, Persönliche Identifikationsnummer (PIN), Transaktionsnummer (TAN), Einmalkennwort (engl. *one-time password*, OTP)
- Verwendung eines Besitztums (Besitzfaktor): z. B. physischer Schlüssel, Digitales Zertifikat
- Gegenwart der Benutzerin oder des Benutzers selbst (Biometriefaktor): z. B. Fingerabdruck

Grundsätzlich erhöht ein zweiter Faktor immer die Sicherheit, es kommt jedoch auch auf die Art der Umsetzung und Verwendung des zweiten Faktors an. Je nach Anwendungsgebiet führt die Wahl der Authentisierungsmethoden zu verschiedenen Vor- und Nachteilen des angestrebten Sicherheitsbedarfs sowie der Praktikabilität im

Alltag der Benutzenden. Gängige Verfahren zur Zwei-Faktor-Authentisierung sind die Verwendung von TAN/OTP-Systemen oder die Verwendung kryptographischer Token (vgl. Webseite des BSI, Zwei-Faktor-Authentisierung: Was sind gängige Systeme zur Zwei-Faktor-Authentisierung?[20]). Bei TAN/OTP-Systemen wird ein Einmalkennwort als zweiter Faktor verwendet. Kryptographische Token enthalten einen geheimen kryptographischen Schlüssel, der in einem Softwarezertifikat oder in Hardware auf einer Chipkarte oder einem speziellen USB-Stick gespeichert, zur Authentisierung genutzt werden kann.

Durch Nutzung einer Zwei-Faktor-Authentisierung soll für das Probandenmanagementsystem die Sicherheit des Anmeldeverfahrens erhöht werden. Als erster Authentifizierungsfaktor wird weiterhin die Kombination aus Benutzername und Passwort verwendet (Wissensfaktor). Als zweiter Authentifizierungsfaktor vor der Eingabe des Benutzernamens und des Passworts wird der Besitz eines Softwarezertifikates vorausgesetzt (Besitzfaktor). Das Probandenmanagementsystem wird nur noch Verbindungen zu einem Rechner erlauben, auf dem ein für den Benutzenden individuelles Softwarezertifikat (Clientzertifikat) installiert wurde.

5. „Dem authentifizierten Benutzer wird der Zugang zum System, zur Anwendung oder zu Teilen der Anwendung entsprechend den vergebenen Rechten eröffnet.“ [vgl. LfDNI, 2020, S. 1]

   Diese wichtige Empfehlung soll ausführlich in der *Anforderung A3: Berechtigungskonzept* und der *Anforderung A4: Zugriffskontrollmodel* betrachtet werden.

Zusammenfassend wurden folgenden Maßnahmen für Umsetzung der *Anforderung 1: Sichere Authentifizierung* konzeptioniert:

- Zugriffsbeschränkung auf bestimmte IP-Adressen
- Zwei-Faktor-Authentisierung über Benutzername/Passwort und Clientzertifikat

[20] Zwei-Faktor-Authentisierung, Bundesamt für Sicherheit in der Informationstechnik (BSI): `https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Cyber-Sicherheitsempfehlungen/Accountschutz/Zwei-Faktor-Authentisierung/zwei-faktor-authentisierung_node.html` [Stand: 01.03.2022]

- Überprüfung der Passwortstärke
- Sichere Speicherung der Passwörter

Anforderung A2: Verschlüsselte Verbindung

Das Probandenmanagementsystem muss als Webanwendung das Übertragungsprotokoll HTTP (Hypertext Transfer Protocol) beziehungsweise HTTPS (Hypertext Transfer Protocol Secure) nutzen. Bei HTTPS wird die Verbindung durch die Protokolle SSL/TLS verschlüsselt und gesichert. SSL (Secure Sockets Layer) beziehungsweise dem heute verwendeten Nachfolger TLS (Transport Layer Security), sind Verschlüsselungsprotokolle zur sicheren Datenübertragung im Internet. Das SSL-Protokoll existiert in den Versionen 1.0, 2.0 und 3.0, wobei die Version 1.0 nicht veröffentlicht wurde. TLS 1.0 ist eine direkte Weiterentwicklung von SSL 3.0. Die verschiedenen Versionen sind von der Internet Engineering Task Force (IETF) spezifiziert und im Internet verfügbar[21]. Zur Auswahl eines sicheren Protokolls für das Probandenmanagementsystem wurde die *Technische Richtlinie TR-02102-2 Kryptographische Verfahren* [BSI, 2020b,c] des BSI herangezogen. Momentan werden die Varianten TLS 1.2 und TLS 1.3 als ausreichend sicher eingestuft. Alle darunter liegenden Versionen werden nicht mehr empfohlen [vgl. BSI, 2020c, S. 7]. Für das Probandenmanagementsystem werden die als unsicher angesehenen Versionen von SSL/TLS deaktiviert.
Damit eine Verbindung zwischen Server und Webbrowser per HTTPS aufgebaut werden kann, spielen digitale Zertifikate, so genannte SSL-Server-Zertifikate, eine bedeutende Rolle. SSL-Server-Zertifikate haben eine ähnliche Aufgabe wie die in der vorherigen Anforderung beschriebenen Clientzertifikate zur sicheren Authentifizierung der Systembenutzenden. Der Unterschied zwischen den Zertifikatsarten besteht darin, dass Clientzertifikate den Client beziehungsweise die Einzelperson identifizieren und mit Server-Zertifikaten der Betreiber der Website authentifiziert wird. Ein Webbrowser stellt eine Verbindung mit dem Server her und validiert die Authentizität des SSL-Serverzertifikats. Dies ermöglicht dem Webbrowser zu überprüfen, ob die Website

[21] TLS 1.0: `https://tools.ietf.org/html/rfc2246`
TLS 1.1: `https://tools.ietf.org/html/rfc4346`
TLS 1.2: `https://tools.ietf.org/html/rfc5246`
TLS 1.3: `https://tools.ietf.org/html/rfc8446`

die ist, für die sie sich ausgibt sowie seine Interaktion mit der Webseite zu verschlüsseln. Das SSL-Server-Zertifikat selbst ist dabei ein digitales Dokument, das alle benötigten Informationen zur Prüfung beinhaltet und im Idealfall von einer allgemein anerkannten Zertifizierungsstelle ausgegeben wird. Für das Probandenmanagementsystem, dessen Sever im universitären Umfeld betrieben wird, kann das SSL-Server-Zertifikat durch den DFN-PKI ausgestellt werden. DFN-PKI ist der Public Key Infrastructure (PKI) Dienst des Deutschen Forschungsnetzes (DFN)[22], als Zertifizierungsstelle (engl. Certificate Authority, CA) für Mitgliedseinrichtungen des DFN-Vereins. Das Zertifikat für das Probandenmanagementsystem wird entsprechend der Anleitung für Zertifikatanträge der DFN-PKI[23] und der Zertifizierungsrichtlinie der DFN-PKI[24] erstellt, beantragt und verwaltet. Das Zertifikat muss entsprechend der Zertifizierungsrichtlinie mit einem mindestens 2048 Bit langem RSA[25]-Schlüssel erstellt werden. Für die Nutzung des Verschlüsselungsverfahrens RSA empfiehlt das BSI in der *Technische Richtlinie Kryptographische Verfahren* gegenwärtig ebenfalls die Verwendung von Schlüsseln mit einer Länge von mindestens 2048 Bit [vgl. BSI, 2020c, S. 15]. Um einen RSA-Schlüssel, der über das Jahr 2022 hinaus gemeinhin als hinreichend sicher angesehen wird, zu generieren [vgl. BSI, 2020c, S. 15], wird für das Probandenmanagementsystem eine Schlüssellänge von 4096 Bit verwendet.

Um das Probandenmanagementsystem ausschließlich über eine verschlüsselte Verbindung zu betreiben, wird eine Weiterleitung von HTTP auf HTTPS eingerichtet sowie unverschlüsselter Datenverkehr ausgeschlossen. HTTP Strict Transport Security (HSTS) ist ein Mechanismus, der es Webservern ermöglicht zu deklarieren, dass auf sie nur über eine sichere Verbindung (HTTPS) zugegriffen werden kann. Der Mechanismus wird in der technische Spezifikation RFC 6797[26] beschrieben.

---

[22] Verein zur Förderung eines Deutschen Forschungsnetzes e. V.: `https://www.dfn.de/` [Stand: 01.03.2022]

[23] DFN-PKI: Anleitung Webschnittstelle Nutzer und Administratoren: `https://www.pki.dfn.de/fileadmin/PKI/Anleitung_DFN-Test-PKI.pdf` [Stand: 01.03.2022]

[24] DFN-PKI: Zertifizierungsrichtlinie: `https://www.pki.dfn.de/fileadmin/PKI/DFN-PKI_CP.pdf` [Stand: 01.03.2022]

[25] RSA (Rivest-Shamir-Adleman) ist ein asymmetrisches kryptographisches Verfahren.

[26] Request for Comments (RFC) 6797: HTTP Strict Transport Security (HSTS): `https://tools.ietf.org/html/rfc6797`

Anforderung A3: Berechtigungskonzept

Ein Berechtigungskonzept muss abbilden, welche Zugriffe für Benutzende eines Systems in Hinblick auf den Datenschutz erlaubt sind und welche nicht. Klinische Forschungsprojekte sind komplex und viele verschiedene Personen sind daran beteiligt, das Projekt von Anfang bis zum Ende durchzuführen. Erfreulicherweise lassen sich die Prozesse der Probandenverwaltung (vgl. Unterkapitel 2.2, S. 15) und damit der Benutzerkreis des Probandenmanagementsystems eingrenzen. Studienärztinnen und -ärzte, Studienassistentinnen und -assistenten, Wissenschaftliche Mitarbeitende und IT-Mitarbeitende wurden als Stakeholder bei der Entwicklung des Probandenmanagementsystems bereits identifiziert (vgl. [Schwanke u. a., 2013b, S. 342] und [Schwanke, 2015, S. 9]). Darüber hinaus ist für die Entwicklung des Berechtigungskonzeptes die den Berufsgruppen zugeteilten Berechtigungen in Bezug auf die im System enthaltenen Daten und Funktionen wichtig. Grundsätzlich wird bei der Erstellung eines Berechtigungskonzeptes eine Grundmenge von Berechtigungen definiert, die anschließend Benutzerinnen und Benutzer zugewiesen werden können. Bei der Auswahl der wichtigsten Berechtigungen für das Konzept des standortübergreifend nutzbaren Probandenmanagementsystems mussten zwei Aspekte besonders berücksichtigt werden: Zum einen die Betrachtung der bereits bestehenden Funktionen und der im Probandenmanagementsystems dokumentierten Daten, wie sie ursprünglich durch Schwanke implementiert wurden. Zum anderen die Trennung in verschiedenen Studien und mehrere Standorte, um das Probandenmanagementsystems in multizentrischen Studien einsetzen zu können.

Schwanke [2015] erfasste die Anforderung an ein leicht konfigurierbares Rechte- und Rollenmanagement, das den Zugriff auf das System und dessen Funktionen regelt [vgl. Schwanke, 2015, Anforderung A-F-3, S. 36]. Das Rechte- und Rollenmanagement ist für jedes Projekt einzeln konfigurierbar [vgl. Schwanke, 2015, Anforderung A-F-6, S. 37]. Diese Anforderungen wurden in der prototypischen Implementierung des Probandenmanagementsystems umgesetzt. Die Zuordnung von Systembenutzenden zu Forschungsprojekten erfolgt in der Benutzerverwaltung. Innerhalb dieser können Administratorinnen und Administratoren des Systems Benutzerinnen und Benutzer anlegen und diesen Rollen zuordnen [vgl. Schwanke, 2015, S. 59]. Innerhalb der Projektverwaltung wird die Zuordnung von Rechten zu projektspezifischen Rollen

vorgenommen [vgl. Schwanke, 2015, S. 62]. Die vom Probandenmanagementsystem vorgegebenen Rechte werden der jeweiligen Rolle zugeordnet [vgl. Schwanke, 2015, S. 64]. Aus dem der Arbeit beiliegenden Quellcode konnte eine Liste der vorgegebenen Rechte entnommen werden, eine Dokumentation der Rechte liegt nicht vor (siehe Listing 3.1).

```
CREATE_RECORD_WITH_LOCAL_PARTICIPANT, CREATE_RECORD_WITH_HIS_PARTICIPANT,
CREATE_RECORD_WITH_EXISTING_PARTICIPANT, SHOW_RECORD, EDIT_RECORD,
EDIT_PARTICIPANT, MODULE_DE_IDENTIFICATION, MODULE_CONTACT_PERSON,
MODULE_TASK, MODULE_APPOINTMENT, MODULE_CONSENT, MODULE_CONTACT_EVENT
```

Listing 3.1: Liste der durch den Programmierer des Probandenmanagementsystems vorgegebenen Rechte zur Zugriffssteuerung der Systembenutzenden des Probandenmanagementsystems.

Mit der Umsetzung des Rechte- und Rollenmanagements wurden wichtige Grundlagen für den Betrieb des Probandenmanagementsystems geschaffen. Jedoch sollte im Kontext eines probandenverwaltenden Systems, und damit der Verwaltung von besonders schützenswerten Daten, Vorgaben für Berechtigungen gemacht werden, die in das Berechtigungskonzept eingehen und sich am Need-to-know-Prinzip orientieren. Zu groß ist die Gefahr, dass die in Listing 3.1 aufgeführten Rechte zur Zugriffssteuerung auf schützenswerte Daten nicht ausreichen oder undifferenziert vergeben werden. Für das Probandenmanagementsystem wurden vier Kategorien von Daten bestimmt, die sich im Schutzstatus unterscheiden:

Probandenstammdaten: Name, Geschlecht, Geburtsdatum, Geburtsort

Adressdaten: Straße, Postleitzahl, Stadt, Land, Telefonnummern, E-Mail-Adressen

Pseudonyme: Kennzeichen über die ein Personenbezug zu weiteren Systemen hergestellt werden kann.

Forschungsaktendaten: Studienstatus, Studiengruppe, Rekrutierungsstandort, Rekrutierungsdatum, Einwilligungsstatus, Kontaktereignisse, Termine, Aufgaben, Anmerkungen

Mit Hilfe der Datenkategorien wurden in einem konservativen Datenschutzszenario zwei für das Probandenmanagementsystem wichtige Benutzerrollen definiert:

Benutzerrolle 1: Studienkoordinator:in

Studienkoordinierende können eine Übersicht aller Studienteilnehmenden ihrer Studie einsehen und bekommen beispielsweise den Status der Studienteilnehmenden innerhalb der Studie angezeigt (vgl. Datenkategorie Forschungsaktendaten). Klarnamen und Kontaktdaten der Studienteilnehmenden, sowie die Pseudonyme der Forschungssysteme sind für sie, um eine Übersicht über die in einer Studie eingeschlossenen Studienteilnehmenden zu erhalten, nicht nötig. Das Probandenmanagementsystempseudonym (PSN_{PMS}) kann zur Kommunikation mit den Studienmitarbeitenden genutzt werden, die ihrerseits darüber einen Personenbezug herstellen können. Abbildung 3.16 zeigt einen Entwurf (Mockup) einer Übersichtstabelle der Studienteilnehmenden für einen Studienkoordinierenden. Probandenstammdaten sind zur Veranschaulichung geschwärzt, auch wenn in einer späteren Implementierung diese Spalten aus Platzgründen weggelassen werden.

PSN $_{PMS}$	Nachname	Vorname	Geburtsdatum	Studienstatus	Einschlussdatum
ab12zy	[redacted]	[redacted]	[redacted]	Eingeschlossen	02.02.2018
cd34xw	[redacted]	[redacted]	[redacted]	Eingeschlossen	12.02.2018
ef56vu	[redacted]	[redacted]	[redacted]	Eingeschlossen	22.02.2018

Abbildung 3.16: Mockup einer Übersicht für Studienkoordinierende im Probandenmanagementsystem. Die Liste zeigt Studienteilnehmende mit ihrem Studienstatus, jedoch keine Stammdaten der Studienteilnehmenden.

Studienkoordinierende sind zur Ausführung von Reports berechtigt, die keine Klarnamen der Studienteilnehmenden enthalten. Zugriff auf Adressdaten erhalten Studienkoordinierende nicht. Feingranularer berechtigt werden kann der Zugriff auf weitere Daten und Funktionen der Forschungsakte der Studienteilnehmenden, wie z. B. das Schreiben von Anmerkungen durch den Studienkoordinator. Diese zusätzlichen Zugriffe sind den Prozessen im Projekt anzupassen, sollten jedoch streng nach dem Need-to-know-Prinzip vergeben werden. Unterteilt in den Datenkategorien des Probandenmanagementsystems ergibt sich für Studienkoordinierende die folgende Tabelle zur Zugriffsberechtigung:

Probandenstammdaten	Adressdaten	Pseudonyme (externer Systeme)	Forschungsaktendaten
✖	✖	✖	✔

Benutzerrolle 2: Studienmitarbeiter:in

Zu den Studienmitarbeitenden, die direkt mit den Studienteilnehmenden arbeiten, gehören je nach Art des Forschungsprojektes Studienärztinnen und Studienärzte, Studienassistentinnen und Studienassistenten und Wissenschaftliche Mitarbeitende. Ihnen wird der Zugriff auf die Probandenstammdaten, Adressdaten, Pseudonyme und die weiteren Forschungsaktendaten gewährt. Abbildung 3.17 zeigt einen Mockup einer Übersichtstabelle der Studienteilnehmenden für einen Studienmitarbeitenden.

PSN $_{PMS}$	Nachname	Vorname	Geburtsdatum	Studienstatus	Einschlussdatum
ab12zy	Nachname1	Vorname1	02.02.2001	Eingeschlossen	02.02.2018
cd34xw	Nachname2	Vorname2	12.02.2002	Eingeschlossen	12.02.2018
ef56vu	Nachname3	Vorname3	22.12.2000	Eingeschlossen	22.02.2018

Abbildung 3.17: Mockup einer Übersicht von Studienteilnehmenden für Studienmitarbeitende im Probandenmanagementsystem. Studienmitarbeitende dürfen Stammdaten der Studienteilnehmenden und Daten der Forschungsakten einsehen, da sie für ihre Arbeit auf diese Informationen angewiesen sind.

Studienmitarbeitende müssen mit den Stammdaten der Studienteilnehmenden arbeiten. Je nach Projekt kann der Zugriff auf weitere Pseudonyme aus anderen Forschungssystemen gewährt werden. Für die Benutzerrollen der Studienmitarbeitenden kann eine Vielzahl von granularen Berechtigungen vergeben werden (vgl. auch Listing 3.1). Beispielsweise muss es nicht allen Mitarbeitenden erlaubt sein Pseudonyme von Systemen aufzulösen, mit denen sie nicht arbeiten oder es ist ihnen nicht erlaubt, neue Probandendatensätze anzulegen, weil sie nur Datensätze pflegen. Unterteilt in den Datenkategorien des Probandenmanagementsystems ergibt sich für Studienmitarbeitende die folgende Tabelle zur Zugriffsberechtigung:

Probandenstammdaten	Adressdaten	Pseudonyme (externer Systeme)	Forschungsaktendaten
✔	✔	✔	✔

Neben den unterschiedlichen Schutzstatus der Datenkategorien im Probandenmanagementsystem hat auf dieses Konzept der Studienstandort einen großen Einfluss. Es musste bedacht werden, wie die sichere Verwaltung von Probandendaten aus verschiedenen Studienstandorten im Probandenmanagementsystem umgesetzt

werden kann. Abbildung 3.18 und Abbildung 3.19 zeigen an einem Beispiel, wie die Zugriffskontrolle im Probandenmanagementsystem durch Schwanke [2015] umgesetzt wurde.

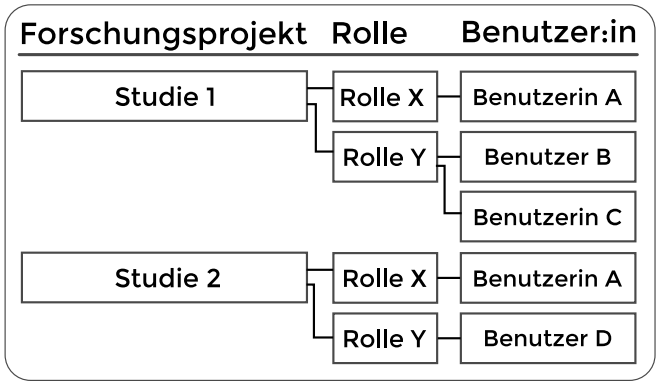

Abbildung 3.18: Beispiel für die Struktur des Rechte- und Rollenmanagements im Probandenmanagementsystem für einen Standort. Forschungsprojekten, hier zwei Studien, sind jeweils zwei studienspezifische Rollen zugewiesen. Den Rollen wiederum sind die Benutzerinnen und Benutzer des Probandenmanagementsystems zugewiesen.

Für Forschungsprojekte lassen sich im Probandenmanagementsystem unterschiedliche Rollen definieren. Rollen werden für jedes Forschungsprojekt einzeln definiert. Den Rollen werden anschließend die Benutzenden des Probandenmanagementsystems zugewiesen (siehe Abbildung 3.18). Den Benutzenden des Probandenmanagementsystems werden damit die einer Rolle hinterlegten Zugriffsrechte (vgl. Listing 3.1) zugeteilt. Zugriffsrechte sind immer einer Rolle zugewiesen (siehe Abbildung 3.19).

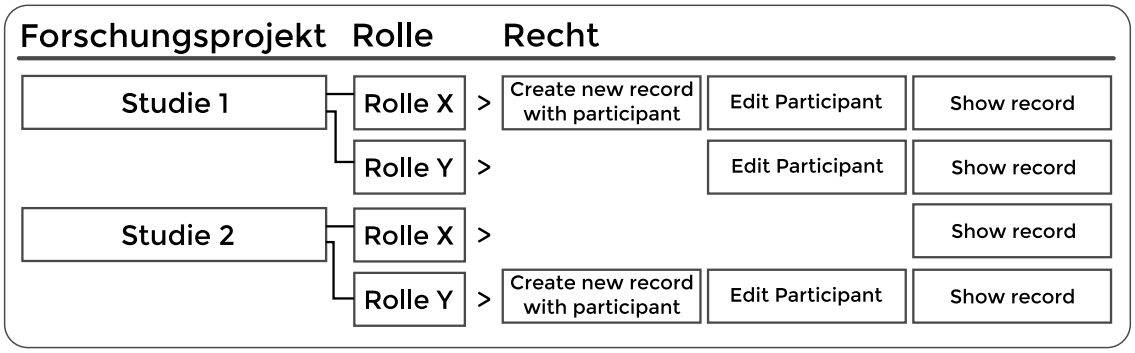

Abbildung 3.19: Beispiel für die Struktur des Rechte- und Rollenmanagements im Probandenmanagementsystem für einen Standort. Definierte Zugriffsrechte des Probandenmanagementsystems werden studienspezifische Rollen eines Forschungsprojektes zugewiesen.

Eine Unterscheidung welche Gruppe von Studienteilnehmenden oder welche Daten von Studienteilnehmenden innerhalb einer Studie eingesehen dürfen, ist nicht möglich.

Dieser Schwäche wurde schon in einem ersten Schritt mit der Unterscheidung der Datenkategorien begegnet. In einem zweiten Schritt wurde das Berechtigungskonzept um eine Standortkomponente erweitert. Im Sinne der Berechtigungssteuerung ist der Standort als Attribut für die Zugriffsteuerung zu verwenden. Jeder Benutzende des Probandenmanagementsystems erhält eine Zuweisung zu mindestens einem Standort. Abbildung 3.20 verdeutlicht im Vergleich zu den vorherigen Abbildungen den Unterschied zu einer Probandenmanagementsystem für nur eine Institution.

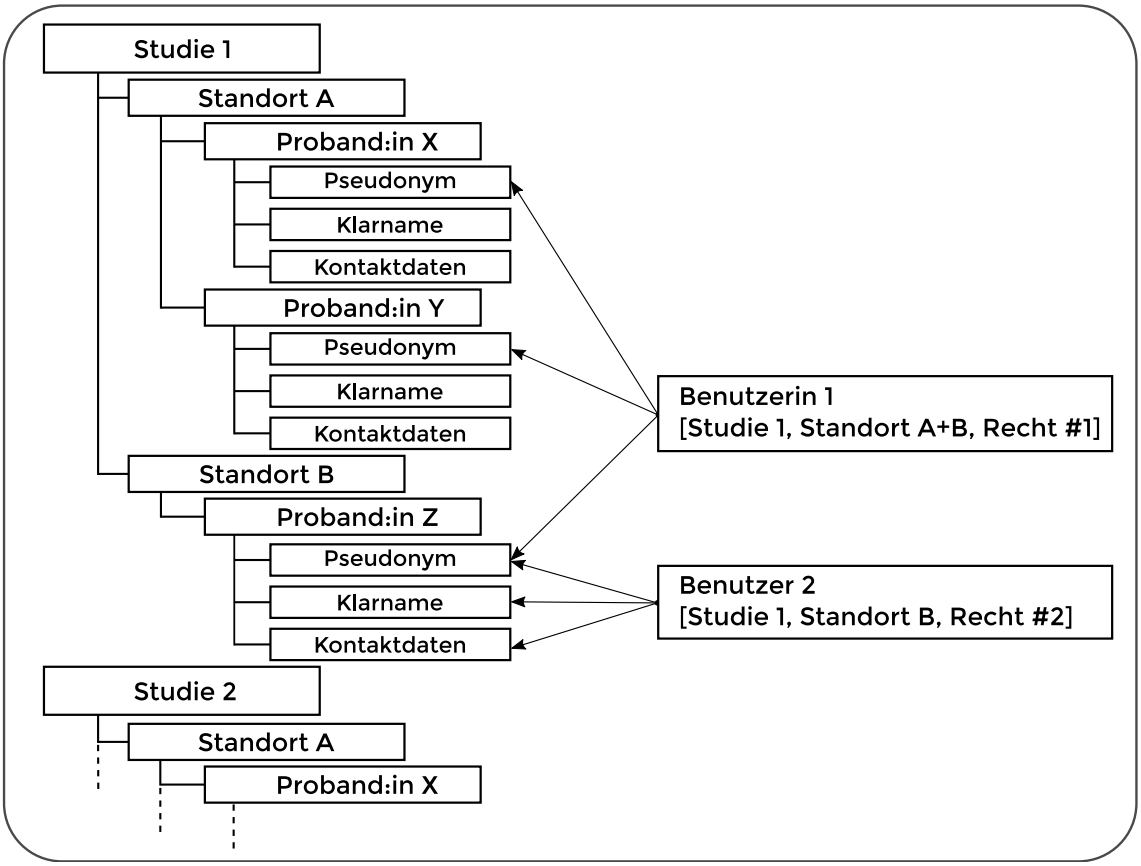

Abbildung 3.20: Beispiel zur Struktur und Rechteverwaltung in einem Probandenmanagementsystem mit Studien an mehreren Standorten.

Eine Studie kann an unterschiedlichen Standorten durchgeführt werden und muss diese Unterscheidung auch in den Benutzerberechtigungen erlauben. Die in Abbildung 3.20 gezeigte Benutzerin 1, beispielsweise eine Studienkoordinatorin, darf alle Studienteilnehmenden der Studie einsehen, ohne Klarnamen und Kontaktdaten der Studienteilnehmenden einsehen zu müssen. Benutzer 2, beispielsweise ein Datenmanager der Studie am Standort B, pflegt die Kontaktdaten der Studienteilnehmenden und

muss diese einsehen und bearbeiten können. Er sollte jedoch keinen Zugriff auf Studienteilnehmende des Studienstandorts A erhalten.

Anforderung A4: Zugriffskontrollmodell

Das Zugriffskontrollmodell muss die im Berechtigungskonzept definierten Berechtigungen für ein System umsetzen. Die wichtigsten Zugriffskontrollmodelle *Discretionary Access Control (DAC)*, *Mandatory Access Control (MAC)*, *Role Based Access Control (RBAC)* und *Attribute Based Access Control (ABAC)* wurden bereits im Methodikteil dieser Anforderung vorgestellt. Diese wurden im Folgenden zur Eignung im Probandenmanagementsystem bewertet.
Primäres Ziel der Berechtigungssteuerung ist es, einzelne Berechtigungen in größere Einheiten zusammenzufassen, um sie damit schneller, übersichtlicher und eleganter zuordnen zu können [vgl. Tsolkas und Schmidt, 2017, S. 1]. Zur Erreichung dieses Ziels, welches auch für das Probandenmanagementsystem als wichtig angesehen wird, zeigen zwei Zugriffskontrollmodelle Schwächen. Das DAC-Modell, welches zwar einfach und intuitiv implementierbar ist, erfordert, dass die Zugriffsberechtigungen für jeden Nutzer explizit modelliert werden müssen. Auch das MAC-Modell, obwohl es ein sicheres Zugriffssystem ist, erfordert eine sehr sorgfältige Planung und kontinuierliche Überprüfung der zugewiesenen Rechte – sowohl für Ressourcen als auch für Benutzer. DAC und MAC haben also einen hohen Verwaltungsaufwand im Betrieb. Diesem Nachteil kann durch das Zusammenfassen von Benutzern zu Benutzerrollen begegnet werden. Ein Vorgehen, das bereits in der vorherigen Anforderung mit dem Ergebnis des grundlegenden Berechtigungskonzeptes angewendet wurde. Verschiedene Rechte zum Schutz unterschiedlicher Datenkategorien wurden in zwei Benutzerrollen für das Probandenmanagementsystem zusammengefasst. Durch die Entkopplung der Berechtigungen von den Identitäten wurde RBAC als rollenbasierte Zugriffskontrolle in Forschung und Praxis zu einem häufig verwendeten Standard. Abbildung 3.21 zeigt diesen Vorteil für das Probandenmanagementsystem grafisch.

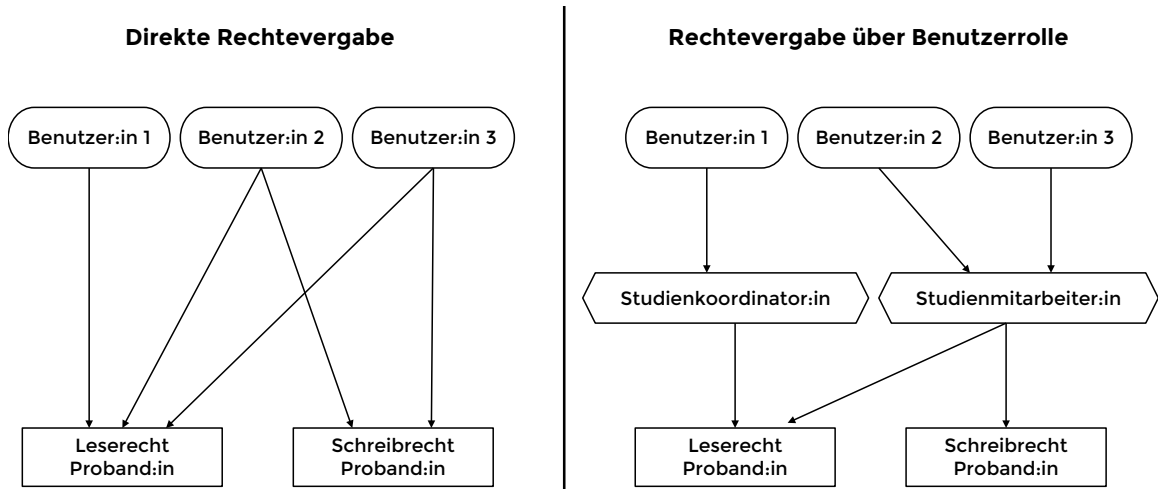

Abbildung 3.21: Vergleich von Vergabekonzepten für Rechte im Probandenmanagementsystem: Direkte Rechtevergabe gegenüber einer Rechtevergabe über Benutzerrollen. Die Komplexität der Rechtevergabe reduziert sich durch die Einführung von Rollen, denen die Benutzenden zugeordnet sind. Im Gegensatz zur direkten Rechtevergabe sind Rechte nicht Benutzenden, sondern Rollen zugeordnet.

Durch die Abkehr von einer auf einzelne Benutzerinnen und Benutzer bezogenen Vergabe von Zugriffsrechten sind RBAC-Modelle besonders für IT-Systeme geeignet, die durch eine sich ändernde Menge an Benutzerinnen und Benutzern und durch eine über lange Zeiträume gleichbleibende Menge an Aufgaben charakterisiert sind [vgl. Eckert, 2008, S. 266].

Das grundlegende Berechtigungskonzept des Probandenmanagementsystems sieht neben Rechten, die an Benutzerrollen geknüpft sind, auch eine Standortberechtigung vor. Die Standortberechtigung ausschließlich über Rollen abzubilden, wäre möglich, führt aber zu einer unüberschaubaren Menge an Rollen. Angenommen für eine Studie mit minimalen Anforderungen gäbe es lediglich die Rollen Studienkoordinator:innen und Studienmitarbeiter:innen. Die Studie wird an 10 Studienstandorten durchgeführt. Für die Abbildung der Rechte im Probandenmanagementsystem nach diesem Rollenmodell würden 20 Rollen nötig werden, nämlich pro Standort je die Rollen Studienkoordinatoren und Studienmitarbeiter. Der Vorteil von RBAC, die Anzahl der Zuordnungen zwischen den Identitäten und den Ressourcen zu reduzieren, wird wieder zunichtegemacht. An dieser Stelle kommt die attributbasierte Zugriffskontrolle (ABAC) ins Spiel, die ebenfalls in die Entscheidung über das anzuwendende Zugriffskontrollmodell einbezogen wurde. Bei ABAC werden Berechtigungen von einer bestimmten Ausprägung einer Eigenschaft abhängig gemacht. So wäre die Berechtigung zur Einsicht von Studienteilnehmenden eines

Studienstandorts vom Attribut *Studienstandort* des Benutzenden abhängig. ABAC allein bringt den Nachteil mit sich, dass viele berechtigungssteuernde Attribute die Anzahl der benötigten Regeln erhöhen und damit der Berechtigungsumfang des Benutzenden unübersichtlich wird. Die Lösung für diese Anforderung liegt also in der Kombination der Modelle RBAC und ABAC. Durch eine Kombination der beiden Modelle lässt sich der Aufwand für die Berechtigungssteuerung senken. Tsolkas und Schmidt [2017] sprechen bei dieser Variante einer Kombination der Modelle von einem *Rollendominierten Ansatz (RBAC-A role centric)*. Dieser Ansatz entspricht dem klassischen RBAC-Modell, nur dass die Berechtigungen dynamisch (d. h. während der aktuellen Sitzung) aufgrund von Attributwerten noch weiter eingeschränkt werden können.

Berechtigungszuordnung: $User \rightarrow Role \rightarrow Attr1 \ldots n \rightarrow permissions$

Im oben genannten Beispiel werden statt insgesamt 20 verschiedenen Rollen nur noch die beiden Rollen Studienkoordinator:in und Studienmitarbeiter:in und ein Attribut *Studienstandort* benötigt.
Damit konnte ein Zugriffskontrollmodell für das Probandenmanagementsystem ausgewählt werden, mit dem sich die im Berechtigungskonzept definierten Berechtigungen umsetzen lassen.

Anforderung A5: Protokollierter Zugriff

Die Praxishilfe zur Protokollierung und zur Erstellung von Protokollierungskonzepten im Gesundheitswesen [Ammon u. a., 2020] bietet im ersten Teil eine Einführung in die Thematik *Protokollierung*, während der zweite Teil Aufbau und Struktur eines Protokollierungskonzeptes beinhaltet. Das Protokollierungskonzept ist eine Dokumentation, in der unter anderem Art und Umfang, Verwendungszweck, Rechtsgrundlage sowie Zugriffsberechtigung der Protokollierung festgehalten werden [vgl. Ammon u. a., 2020, S. 10]. Wie bei *Anforderung A3: Berechtigungskonzept* muss auch das Protokollierungskonzept vollständig, umfassend und aktuell sein, konnte jedoch im Rahmen dieser Arbeit nicht erstellt werden, da dafür ein konkretes laufendes Forschungsprojekt fehlt. Für die weitere Konzeption des Probandenmanagementsystems ist aber ausreichend, die zentralen Inhalte des Protokollierungskonzeptes für das Probandenmanagementsystem zu benennen.

Grundlegend unterliegen alle Protokolldaten einer strikten Zweckbindung, so auch für das Probandenmanagementsystem. Protokolldaten im Probandenmanagementsystem werden angelegt zum Zweck von:

1. Stichprobenartige Datenschutzkontrolle bezüglich Ordnungsmäßigkeit der Verarbeitung
2. Prüfung des Sachverhalts bei einer vorliegenden beziehungsweise bei einem Verdacht auf eine vorliegende Verletzung des Schutzes personenbezogener Daten
3. Gewährleistung der Sicherheit der Verarbeitung

Rechtsgrundlage für die Protokollierung bildet Art. 32 DSGVO[27] in Verbindung mit § 22 Abs. 2 BDSG[28]. Für das Probandenmanagementsystem wird eine ereignisbezogene Protokollierung zur Erreichung der dargestellten Zwecke als angemessen angesehen. Protokolliert werden der angemeldete Benutzer, der Zeitpunkt einer Aktion sowie die durchgeführte Aktion selbst. Dies erfolgt für Probandenstammdaten, Adressdaten und Pseudonyme (vgl. Datenkategorien des Probandenmanagementsystem, definiert in *Anforderung A3: Berechtigungskonzept*, S. 79). Eine gesetzliche Vorgabe für die Speicherdauer von Protokollen existiert nicht, Maßstab für Speicherdauer der Protokolldaten ist die *Erforderlichkeit der Aufgabenerfüllung* [vgl. Ammon u. a., 2020, S. 18]. Daher sollte mit dem Datenschutzbeauftragten eines Forschungsprojektes in dem das Probandenmanagementsystem genutzt wird, ein für die Datenschutzkontrolle angemessener Zeitraum festgelegt werden. Die Benutzenden des Probandenmanagementsystems müssen über die Anlässe ihrer Protokollierung, ihre Einsichtsmöglichkeit in die Protokolle und die datenschutzkonforme Verarbeitung, Nutzung und Löschung der Protokolldaten unterrichtet werden. Die Auswertung der Protokollierung muss dem engen Kreis der dafür Berechtigten ermöglichen, festzustellen, wer wann welche personenbezogenen Daten in welcher Weise verarbeitet oder genutzt hat.

---

[27] Art. 32 DSGVO Sicherheit der Verarbeitung: `https://dejure.org/gesetze/DSGVO/32.html` [Stand: 01.03.2022]

[28] § 22 BDSG Verarbeitung besonderer Kategorien personenbezogener Daten: `https://dejure.org/gesetze/BDSG/22.html` [Stand: 01.03.2022]

3.5.4 Diskussion

Schon lange hat sich die Erkenntnis durchgesetzt, dass eine hundertprozentige IT-Sicherheit praktisch nicht zu erreichen ist, schon gar nicht zu akzeptablen Kosten und mit leistbarem Aufwand. Es geht also nicht um absolute Sicherheit, sondern um die Verringerung von Risiken. In dieser Iteration wurde ein Konzept erarbeitet, das eine angemessene Zugangs- und Zugriffskontrolle zu einem Probandenmanagementsystem, das von mehreren Standorten genutzt wird, ermöglicht. Für fünf Anforderungen wurden Maßnahmen erarbeitet, die in die Implementierung eingehen werden. Diese Ergebnisse wurden im Folgenden reflektiert und bewertet.

Anforderung A1: Sichere Authentifizierung

Eine sichere Authentifizierung wird gewährleistet, indem Zugriffe auf das Probandenmanagementsystem auf bestimmte IP-Adressen eingeschränkt werden, eine Zwei-Faktor-Authentifizierung eingesetzt wird, Passwörter auf ihre Güte geprüft werden und diese Passwörter sicher gespeichert werden. Damit setzt das Probandenmanagementsystem Handlungsempfehlungen zur sicheren Authentifizierung der Landesbeauftragten für den Datenschutz in Niedersachsen [LfDNI, 2020] um. Mit Umsetzung der Empfehlungen soll jedoch keine Aussage über die absolute Sicherheit der Authentifizierung zum Probandenmanagementsystem getroffen werden. Weitere Maßnahmen zur Steigerung der IT-Sicherheit können ergänzt werden. Dennoch setzt das Konzept zur *Anforderung A1: Sichere Authentifizierung* Maßnahmen um, die die Wahrscheinlichkeit für den Zugriff durch Unberechtigte erheblich reduzieren. So sinkt durch die Zwei-Faktor-Authentifizierung das Risiko einer erfolgreichen Attacke auf das System, weil das reine Knacken des Passworts nicht mehr ausreicht, um Zugriff zu erhalten. Der Angreifer müsste auch Zugriff auf das physische Gerät erlangen, das mit dem Benutzerkonto verknüpft ist. Jeder zusätzliche Authentifizierungs-Faktor macht ein System also sicherer, wenn die einzelnen Faktoren unabhängig voneinander sind. Demgegenüber steht der Aufwand, der von administrativer Seite beim Betrieb des Probandenmanagementsystems geleistet werden muss. Für jede Benutzerin und jeden Benutzer muss ein Arbeitsrechner zur Verfügung stehen, der über eine feste IP-Adresse verfügt, die dem Server des Probandenmanagementsystems bekannt gemacht

werden muss, um den Zugriff zu ermöglichen. Des Weiteren muss auf diesem Rechner ein Clientzertifikat installiert werden, es entsteht also ein höherer Verwaltungs- und Beratungsaufwand für den Betrieb des Probandenmanagementsystems. Da in klinischen Verbundforschungsprojekten jedoch die Anzahl der Benutzenden eines Probandenmanagementsystems überschaubar ist, wurde im Sinne einer Abwägung zwischen Aufwands und Sicherheitszugewinn die Nutzung von Clientzertifikaten für geeignet befunden.

Anforderung A2: Verschlüsselte Verbindung

Eine sichere Verbindung ist für die Nutzung des Probandenmanagementsystems unabdingbar. Die vorgestellte Lösung sieht ein RSA/4096-Bit Serverzertifikat, signiert durch den DFN-PKI, vor, mit einer über HSTS erzwungenen HTTPS Verbindung über TLS 1.2 oder TLS 1.3. Die Wahl der Protokolle und Cipher-Suites[29] wurde aufbauend auf den technischen Richtlinie des BSI getroffen. Dabei wurden die Richtlinien nur für die Entscheidung zur Wahl der Verschlüsselungsprotokolle und Cipher-Suites herangezogen, auch wenn diese weitaus mehr Informationen und Empfehlungen für Experten enthält. Die Richtlinie weist zu Beginn explizit aus, keine Empfehlungen für konkrete Anwendungen zu geben. Bei der Auswahl der Verschlüsselungsprotokolle konnte für das Probandenmanagementsystem rigoros auf TLS 1.2 oder TLS 1.3 eingeschränkt werden, da auf eine Abwärtskompatibilität, wie es für gewöhnliche Webseiten relevant ist, nicht geachtet werden muss. Die in den Richtlinien ausgesprochenen Empfehlungen müssen jährlich überprüft und bei Bedarf angepasst werden, um aktuelle Sicherheitsempfehlungen zu berücksichtigen. Die Umsetzung des hier vorgestellten Konzeptes soll, nachdem es im Kapitel 4 (Prototypische Implementierung des Probandenmanagementsystems, S. 100) umgesetzt wurde, validiert werden. Ein frei verfügbarer Online-Service zur Überprüfung der Verschlüsselung von Webservern ist z. B. SSL Report von Qualys SSL Labs[30].

Anforderung A3: Berechtigungskonzept

Mit dem erarbeiteten Berechtigungskonzept sind die wichtigsten Grundlagen für ein datenschutzkonformes Probandenmanagementsystem im multizentrischen Einsatz

[29] Eine Cipher-Suite ist eine Sammlung kryptographischer Verfahren zur Verschlüsselung.

[30] SSL Server Test, Qualys, Inc.: `https://www.ssllabs.com/ssltest/` [Stand: 01.03.2022]

festgelegt worden. Die Trennung der wesentlichen Funktionsrollen und das Modell der standortverteilten Berechtigung ermöglichten in Anforderung *A4 Zugriffskontrollmodell* ein geeignetes Modell zu konzeptionieren. Das Berechtigungskonzept hilft im Datenschutzkonzept die Einhaltung des Need-to-know-Prinzips zu belegen und der Nachweispflicht (vgl. Art. 5 Abs. 2 DSGVO[31]) nachzukommen. Ein vollständiges Berechtigungskonzeptdokument geht zwar darüber hinaus, ist jedoch ohne ein reales Forschungsprojekt mit seinen spezifischen Anforderungen nicht im Rahmen dieser Arbeit umzusetzen. Es ist die Unmenge an Möglichkeiten, Berechtigungen zu definieren und Benutzern zuzuteilen, welche die Thematik komplex macht.

Anforderung A4: Zugriffskontrollmodell

Zugriffskontrollmodelle beziehungsweise Zugriffskontrollstrategien sind bereits seit Jahrzehnten Gegenstand der Forschung. Es gibt zahlreiche Modelle, von denen nur die in der Literatur als bedeutend angesehenen in dieser Arbeit betrachtet wurden. Es wurde ersichtlich, dass alle Modelle, unabhängig davon, ob sie benutzer-, regel-, rollen- oder attributorientiert sind, Vorteile und Nachteile besitzen. Umsetzbar für das Probandenmanagementsystem sind alle Modelle, doch erwies sich ein rollendominiertes attributnutzendes Zugriffskontrollmodell als besonders geeignet, um die Anforderungen des Berechtigungskonzeptes zu erfüllen. Für das Konzept dieser Arbeit ist eine feine Ausarbeitung aller möglichen Rechte und Rollen nicht möglich, daher eignete sich der sehr flexible und erweiterungsfähig RBAC-Ansatz, mit einer attributbasierten Zugriffskontrollstrategie kombiniert, besonders. Erweiterbarkeit und eine übersichtliche Rechtevergabe sind wichtige Eigenschaften für das Konzept des standortübergreifenden Probandenmanagementsystems für die klinische Forschung.

Anforderung A5: Protokollierter Zugriff

Protokollierung unterstützt den Nachweis der ordnungsgemäßen Verarbeitung personenbezogener Daten und dient dem Zweck der Aufrechterhaltung von Datenschutz und Datensicherheit. Die vorgestellte Lösung ergänzt das Konzept für das Probandenmanagementsystem, indem Protokolldaten zweckgebunden, vollständig

[31] Art. 5 DSGVO Grundsätze für die Verarbeitung personenbezogener Daten: `https://dejure.org/gesetze/DSGVO/5.html` [Stand: 01.03.2022]

und datensparsam eingerichtet werden. Die Protokolle helfen Aktivitäten bei der Verarbeitung personenbezogener Daten so transparent zu machen, dass die Ordnungsmäßigkeit beziehungsweise ein Verstoß gegen die Ordnungsmäßigkeit einer Verarbeitung personenbezogener Daten nachweisbar ist.

3.5.5 Schlussfolgerung

In dieser Iteration entstand ein Konzept, um einen vertraulichen Zugang und Zugriff auf die Probandendaten in einem standortübergreifend nutzbaren Probandenmanagementsystem sicherzustellen. Fünf Anforderungen wurden bearbeitet um dieses Ziel zu erreichen:

1. Sichere Authentifizierung
2. Verschlüsselte Verbindung
3. Berechtigungskonzept
4. Zugriffskontrollmodell
5. Protokollierter Zugriff

Das entstandene Konzept bildete die Grundlage für eine prototypische Implementierung. Des Weiteren sind die Ergebnisse dieser Iteration eine wichtige Voraussetzung für das Datenschutzkonzept, das in der nächsten Iteration erarbeitet wurde.

3.6 Iteration 4: Datenschutzkonzept

Eine Grundvoraussetzung für ein rechtskonform umgesetztes Forschungsprojekt mit personenbezogenen Daten von Studienteilnehmenden ist ein umfassendes Datenschutzkonzept (vgl. Unterkapitel 3.1.3, Grundlagen Datenschutzkonzept, S. 31). Das Probandenmanagementsystem als Teil einer IT-Infrastruktur, wie in den vorhergehenden Iterationen beschrieben, muss in einem Datenschutzkonzept beschrieben sein. Die Erfüllung von Anforderung 4 für die Umsetzung der verbundübergreifenden Probandenverwaltung, welche ein Datenschutzkonzept für das Probandenmanagementsystem fordert (vgl. Unterkapitel 3.2, Anforderungen an eine standortübergreifende Probandenverwaltung, S. 32), ist Inhalt dieser letzten Iteration des Gesamtkonzeptes. Einleitend in die Iteration wird Problem und Zielsetzung in Bezug auf das benötigte Datenschutzkonzept beschrieben. Die Auswahl von Hilfsmitteln zur Erstellung des Datenschutzkonzeptes wie Leitfäden, Vorlagen und Checklisten wird anschließend im Abschnitt *Methodik* vorgestellt. Der Abschnitt *Ergebnisse* beschreibt die für Forschungsprojekte mit Probandenmanagementsystem relevanten Teile eines Datenschutzkonzeptes. Dieses Ergebnis wird im Anschluss diskutiert und eine Schlussfolgerung daraus gezogen.

3.6.1 Problem und Zielsetzung

Für den Ausbau des Probandenmanagementsystems wird besonders ein transparentes und leicht verständliches Datenschutzkonzept wichtig, da bei standortübergreifenden Kooperationen verschiedene Ethikkommissionen und Datenschutzbeauftragte an der Genehmigung des Forschungsvorhabens beteiligt sind. Zwar müsste auch für ein Forschungsprojekt, in dem das durch Schwanke [2015] entwickelte Probandenmanagementsystem zum Einsatz kommt, ein Datenschutzkonzept vorhanden sein, der Betrieb beziehungsweise die Anwendung des Systems stand jedoch nicht im Fokus seiner Arbeit und eine passende Vorlage für ein Datenschutzkonzept wurde nicht erarbeitet. Die Anforderungen an ein standortübergreifend nutzbares Probandenmanagementsystem (vgl. Unterkapitel 3.2, Anforderungen an eine standortübergreifende Probandenverwaltung, S. 32), zeigten deutlich auf, dass eine große Herausforderung in der sicheren und datenschutzkonformen Verwaltung der

Studienteilnehmenden liegt. Die Bestandteile zur Lösung dieser Herausforderung, erarbeitet in den vorangegangenen Iterationen, müssen nun in einem plausiblen Datenschutzkonzept münden.
Ziel dieser Iteration ist, alle relevanten Informationen zusammenzutragen, um in Forschungsprojekten, die das Probandenmanagementsystem einsetzen, ein Datenschutzkonzept erstellen zu können. Da Datenschutzkonzepte immer für Forschungsprojekt und nicht für einzelne Systeme erstellt werden, wird das Ergebnis dieser Iteration Teil eines späteren umfassenden und vollständigen Datenschutzkonzeptes für ein konkretes Forschungsprojekt sein. Die vollständige Erstellung eines umfassenden Datenschutzkonzeptes im Rahmen dieser Dissertation ist nicht möglich, da forschungsprojektspezifische Angaben noch nicht bekannt sind. Dennoch ist diese Vorarbeit sehr wichtig, um die datenschutzkonforme Einbettung des verbundübergreifenden Probandenmanagementsystems in eine IT-Infrastruktur in einem Datenschutzkonzept aufzuzeigen.

3.6.2 Methodik

Für die Erstellung eines Datenschutzkonzeptes ist die Verwendung von Leitfäden, Vorlagen und Checklisten hilfreich. Für medizinische Forschungsprojekte eignet sich besonders die *Checkliste zur Erstellung eines Datenschutzkonzeptes* der TMF [Antony u. a., 2017]. Die Checkliste vereint Erkenntnisse aus verschiedenen Leitfäden, Empfehlungen und Mustervorlagen vorangehender Jahre zum Thema Datenschutzkonzepte. Ziel der Checkliste ist es, den Erstellern von Datenschutzkonzepten Fragen in themengegliederter Form vorzugeben, die in einem Datenschutzkonzept beantwortet werden sollten.
In dieser Iteration wurde die Checkliste als Gliederungsvorlage für die datenschutzkonzeptrelevanten Themen des Probandenmanagementsystems verwendet. Die Checkliste teilt relevante Themen für ein Datenschutzkonzept in acht Teile (1. Darstellung des Forschungsvorhabens, 2. Organisatorische Struktur, 3. Datenschutzrelevante Rahmenbedingungen, 4. Grundlagen zum Schutz der Rechte und Freiheiten der betroffenen Personen, 5. Rechte der betroffenen Personen, 6. Organisatorische Maßnahmen, 7. Technische Maßnahmen und 8. Vergleich mit dem TMF-Datenschutzleitfaden) auf, anhand dieser auf das Probandenmanagementsystem

eingegangen wird. Bei der späteren Erstellung eines vollständigen Datenschutzkonzeptes für ein konkretes Forschungsprojekt ist zu empfehlen, die Checkliste mit ihren Fragen vollständig durchzugehen. Zudem sind zusätzliche Dokumente zum Datenschutzkonzept zu erstellen und dem Konzept beizulegen, auf die in dieser Dissertation aufgrund des Umfangs und der noch unbekannten projektspezifischen Informationen nicht eingegangen werden kann. Dazu gehören unter anderem ein IT-Sicherheitskonzept, eine Datenschutz-Folgenabschätzung, eine Patienten-/Probandeninformation zum Forschungsvorhaben mit Einwilligungserklärung, weitere Richtlinien und Standardarbeitsanweisungen sowie Kooperationsvereinbarungen und Verträge.

3.6.3 Ergebnisse

In acht Unterkapiteln, die sich an der Gliederung der Checkliste zur Erstellung eines Datenschutzkonzeptes orientieren, wurden jeweils die für ein Datenschutzkonzept relevanten Punkte für den Einsatz des Probandenmanagementsystems beschrieben.

I. Darstellung des Forschungsvorhabens

Die Darstellung des Forschungsvorhabens leitet in das Datenschutzkonzept ein. Beschrieben werden das Projektziel und das zur Erreichung geplante Vorgehen einschließlich der wissenschaftlichen Methodik. Darzulegen ist auch der Beitrag der Forschung zum Fortschritt in der Wissenschaft und dem erwarteten Nutzen für die betroffenen Personen.

Dieser Teil des Datenschutzkonzeptes sollte bezogen auf das Probandenmanagementsystem auch dessen Nutzen und Erfordernis indirekt begründen. Die Verwaltung sensibler Probandendaten in einem elektronischen Verwaltungssystem wird unabhängig von den in dieser Arbeit beschriebenen Datenschutz- und Sicherheitsvorkehrungen erwartungsgemäß kritisch gesehen. Zurecht muss daher ausreichend begründet werden, welche sensiblen Probandendaten erhoben werden und bei welchen Verarbeitungsprozessen ein expliziter Personenbezug unvermeidbar ist. Ein Forschungsvorhaben mit kurzer Laufzeit, wenigen Probanden, wenigen Probandenkontakten und keiner geplanten Kontaktierung der Studienteilnehmenden im Verlauf des Vorhabens wird den Einsatz eines elektronischen Probandenverwaltungssystems in Abwägung zu einer sicher verwahrten Liste auf Papier

nicht rechtfertigen.

II. Organisatorische Struktur

Dieser Abschnitt beschreibt die Personen und Organisationen, die an einem Forschungsprojekt beteiligt sind, sowie ihre Verantwortungsbereiche und Beziehungen zueinander.
In Bezug auf das Probandenmanagementsystem sind die für die Verarbeitung personenbezogener Daten Verantwortlichen sowie ihre organisatorischen Abhängigkeiten besonders relevant. Iteration 1 (Unterkapitel 3.3, S. 36) zeigte, dass zur Wahrung der informationellen Gewaltenteilung Forschungs-IT-Infrastruktur verteilt sein muss. Die verteilte Datenhaltung bedingt also die Einbindung von verschiedenen Kooperationspartnern, die einzeln beschrieben werden müssen. Ein Organigramm hilft dabei, alle Beteiligten zu überblicken. Wie in Iteration 1 dargestellt, sind die Datenkategorien und Datenflüsse in diesem Abschnitt zu nennen, um die Funktionen beziehungsweise die Aufgaben im Rahmen der Datenverarbeitung des Forschungsvorhabens aufzuzeigen. In einer Datenverteilungsmatrix sollte übersichtlich dargestellt werden, welcher Partner welche Informationen speichern und einsehen können. Auch sollte dargelegt werden, wie die Zusammenarbeit im Forschungsvorhaben vertraglich geregelt ist.

III. Datenschutzrelevante Rahmenbedingungen

Die datenschutzrelevanten Rahmenbedingungen bilden den Hauptteil des Datenschutzkonzeptes. In ihm werden die zur Durchführung des Forschungsprojektes vorgesehenen Prozesse mit ihren organisatorischen und rechtlichen Rahmenbedingungen vorgestellt. Diese umfassen die geplanten Anwendungsfälle mit Beschreibung der Daten, Datenflüsse und Datenverwendung. Anfangend bei der Erhebung über die Verarbeitung und Nutzung bis zur Löschung der Daten sollte dieser Teil alle datenschutzrelevanten Inhalte erläutern.
In diesem Teil des Datenschutzkonzeptes werden auch Details zur Verarbeitung der Daten des Probandenmanagementsystems beschrieben. Für das Probandenmanagementsystem muss anhand seiner Anwendungsfällen die Verarbeitung erläutert werden. Auch die Interaktion zwischen dem Probandenmanagementsystem und dem Identitätsmanagement,

wie in vorangehenden Iterationen beschrieben, ist Teil dieses Abschnitts.

IV. Grundlagen zum Schutz der Rechte und Freiheiten der betroffenen Personen

Dieser Teil des Datenschutzkonzeptes beschreibt grundsätzliche Überlegungen zum Datenschutz, die sich aus der Datenschutzgesetzgebung ergeben. Zu klären ist, auf welcher gesetzlichen Grundlage die Erhebung und Verarbeitung der Probandendaten erfolgt. Grundsätzlich ist die Bereitstellung personenbezogener Daten durch die Studienteilnehmenden freiwillig, sie bedarf einer Einwilligung und ist zweckgebunden. Form der Einwilligungserklärung, Anforderungen von Patientenorganisationen oder Ethikkommissionen, die Einbindung der betriebliche beziehungsweise behördliche Datenschutzbeauftragten sowie eine Abschätzung der Risiken sind unter anderem in diesem Abschnitt zu nennen.
Alle hier zu berücksichtigenden Punkte hängen stark von den Rahmenbedingungen des Forschungsprojektes ab. Das Probandenmanagementsystem muss entsprechend in die Überlegungen einbezogen werden.

V. Rechte der betroffenen Personen

Die DSGVO beinhaltet eine Reihe von Betroffenenrechten, die den Studienteilnehmenden im Rahmen der Verarbeitung ihrer Daten zustehen und die Verantwortliche eines Forschungsprojektes sicherstellen müssen. Neben Transparenz-, Informations- und Auskunftspflichten gehören dazu die Rechte auf Berichtigung beziehungsweise Vervollständigung falscher beziehungsweise unvollständiger Daten, die Rechte auf Löschung oder Einschränkung der Verarbeitung sowie ein allgemeines Widerspruchsrecht. Das Probandenmanagementsystem ist von den dort zu beschreibenden Maßnahmen teilweise betroffen. So sollte geregelt sein, wie die Auskunft über gespeicherten Probandendaten, eine Einschränkung der Verarbeitung oder die Löschung der Daten im Probandenmanagementsystem durchzuführen ist.

VI. Organisatorische Maßnahmen

Unter den organisatorischen Maßnahmen werden Verantwortlichkeiten und damit verbundenen Verpflichtungen und Regelungen für datenschutzrelevante Prozesse im

Forschungsprojekt aufgeführt. Eine Vielzahl von Maßnahmen kann zum Datenschutz beitragen. Unter anderem gehören dazu interne Regelungen wie Richtlinien und Dienstanweisungen für Beschäftigte, Zugangs- und Zugriffsregelungen oder weitere personelle Maßnahmen wie Schulungen im Rahmen des Forschungsvorhabens. Ebenfalls werden Gremien und Funktionen von Personen beschrieben (z. B. Datenschutzausschuss, IT-Sicherheitsbeauftragte etc.), als auch wie und durch wen die technischen und organisatorischen Sicherheitsmaßnahmen auf Einhaltung kontrolliert werden können.
Das Probandenmanagementsystem unterstützt die Umsetzung von organisatorischen Maßnahmen wie der Zugangs- und Zugriffskontrolle (Unterkapitel 3.5, S. 63). Das Berechtigungskonzept sowie der protokollierte Zugriff wie in Iteration 3 beschrieben sind an dieser Stelle ins Datenschutzkonzept zu übernehmen.

VII. Technische Maßnahmen

Die zur Einhaltung der Datenschutzanforderungen notwendigen technischen Maßnahmen werden in diesem Teil des Datenschutzkonzeptes beschrieben. Vorgestellt wird die Forschungs-IT-Infrastruktur mit ihren Systemkomponenten, eingesetzter Software, betriebene Server und ihre Standorte. Die IT-Infrastruktur muss in einem Umfang beschrieben werden, wie sie für die Identifizierung der nötigen Datenschutz- und IT-Sicherheitsmaßnahmen notwendig ist. Relevante Punkte sind unter anderem Authentifizierung und Berechtigungen, Absicherung des Netzwerks, eingesetzte Verschlüsselungstechnik, Backup-Strategie und Ausfallschutz. Technische Details können in einem separaten IT-Sicherheitskonzept beschrieben werden, auf das das Datenschutzkonzept verweist und das beigelegt werden muss.
In Iteration 3 (Unterkapitel 3.5, S. 63) wurde ein Konzept erarbeitet, um einen vertraulichen Zugang und Zugriff auf die Probandendaten sicherzustellen. Die konzeptionierten Maßnahmen, die prototypisch implementiert werden sollen, sind technische Maßnahmen, die in diesem Abschnitt des Datenschutzkonzeptes sowie dem IT-Sicherheitskonzept eingefügt werden müssen. Zusätzlich gibt Iteration 1 (Unterkapitel 3.3, S. 36) die Grundstruktur einer Forschungs-IT-Infrastruktur vor, die zu einer sinnvollen Datentrennung bei der Probandenverwaltung nötig wird.

VIII. Vergleich mit dem TMF-Datenschutzleitfaden

Dieser letzte Abschnitt im Datenschutzkonzept erleichtert die Beurteilung des Datenschutzkonzeptes durch einen Vergleich mit den Empfehlungen des TMF-Datenschutzleitfadens. Datenschutzkonzepte die sich am Leitfaden orientieren haben häufig eine höhere Akzeptanz unter Datenschutzbeauftragten, da sich der Leitfaden in generischer Form aus etablierten Datenschutzmodellen in medizinischen Forschungsprojekten zusammensetzt.
Für das Problem der Datentrennung beim Einsatz eines Probandenmanagementsystems (vgl. Unterkapitel 3.3.1, Iteration 1: Problem und Zielsetzung, S. 36), enthält der TMF-Datenschutzleitfaden jedoch keine unmittelbar verwendbare Empfehlung. Zur Lösung wurde daher in Iteration 1 eine Erweiterung zur Trennung von Daten beschrieben, die sich in das Grundmodell des TMF-Datenschutzleitfadens einpasst (vgl. Unterkapitel 3.3.2, Iteration 1: Methodik, S. 38). Im Datenschutzkonzept muss diese Erweiterung der Modelle des TMF-Datenschutzleitfaden zusammengefasst und begründet werden.

3.6.4 Diskussion

In medizinischen Verbundforschungsprojekten stellt ein schlüssiges Datenschutzkonzept oft eine Herausforderung dar. Komplizierte Forschungsfragen, viele beteiligte Forschungspartner sowie die benötigte Forschungs-IT-Infrastruktur, die im Datenschutzkonzept berücksichtig werden müssen, tragen zur Komplexität bei. Zwar existieren Mustervorlagen, Leitfäden und Checklisten, letztendlich ist ein Datenschutzkonzept jedoch genauso individuell wie das Forschungsprojekt selbst. Dabei kann ein Datenschutzkonzept niemals statisch sein, es muss entsprechend der technischen Weiterentwicklung, den rechtlichen Rahmenbedingungen und auch dem datenschutzbetreffenden aktuellen Diskurs angepasst sein. Auch der Umfang eines Datenschutzkonzeptes hängt von der Komplexität des Forschungsprojektes ab. Vor diesem Hintergrund wurde in dieser Iteration der Grundstein gelegt, um die Erstellung eines individuellen Datenschutzkonzeptes für den Einsatz des Probandenmanagementsystems zu erleichtern. Die gewählte Methodik orientiert an der Checkliste für Datenschutzkonzepte der TMF, verweist auf entsprechende Unterkapitel

der konzeptuellen Lösung der vorangehenden Iterationen und versucht bewusst nicht ein vollständiges Datenschutzkonzept eines konstruierten Forschungsprojektes zu antizipieren. Zwar ist die Nutzung von TMF-Datenschutzleitfaden und TMF-Checkliste für Datenschutzkonzepte keine Bedingung für die Akzeptanz der hier vorstellten Lösung, die Verwendung jedoch konsequent, da sich so das neu entwickelte Konzept für das Probandenmanagementsystem in etablierte Modelle eingliedert und einen Vergleich zulässt. In diesem Zusammenhang ist die Erwartung, dass die Nähe zum TMF-Datenschutzleitfaden sowohl das Verständnis des Probandenmanagementsystems in einer Forschungs-IT-Infrastruktur steigert, als auch zur Akzeptanz des Datenschutzkonzeptes bei Datenschutzbeauftragten und Aufsichtsbehörden beiträgt. Somit kann die Hürde für den Einsatz eines Probandenmanagementsystems in Forschungsprojekten aufgrund von Datenschutzbedenken mit den hier vorgestellten Punkten gesenkt werden.

3.6.5 Schlussfolgerung

Mit den vorgestellten Überlegungen zum Datenschutzkonzept für das Probandenmanagementsystem endet das Konzeptionskapitel. Die vorangehenden Iterationen mit ihren Ergebnissen zur Datentrennung, Vernetzung und Zugangs- und Zugriffskontrolle lieferten die Inhalte, die zur Erstellung eines schlüssigen Datenschutzkonzeptes benötigt werden. Die vorliegende Gliederung eines Datenschutzkonzeptes mit ihren Hinweisen ist ein wichtiger Teil der Konzeption des standortübergreifenden Probandenmanagementsystems, da es den praktischen Einsatz des Probandenmanagementsystems in einem konkreten Forschungsprojekt vorbereitet.

Kapitel 4

Prototypische Implementierung des standortübergreifenden Probandenmanagementsystems

Im vierten Teil dieser Arbeit wird die technische Weiterentwicklung des Probandenmanagementsystems prototypisch umgesetzt. Zuvor wurde in vier Iterationen ein umfassendes Konzept erarbeitet, um das Probandenmanagementsystem in standortübergreifenden Forschungsprojekten einsetzen zu können. Für die Teile des Konzeptes, für die eine technische Umsetzung möglich ist, wird im Folgenden ihre Implementierung vorgestellt. Dies umfasst die Datentrennung und vernetzte Probandenverwaltung sowie die Umsetzung der Zugangs- und Zugriffskontrolle für das Probandenmanagementsystem. Dazu benötigte wichtige Begriffe und Konzepte werden zuvor in einem Grundlagenkapitel vorgestellt. Dieses Implementierungskapitel dient der Überprüfung der Machbarkeit des Konzeptes. Zur Nachvollziehbarkeit der Implementierung des standortübergreifend nutzbaren Probandenmanagementsystems wurde ein Demonstrator[1] erstellt.

[1] Demonstrator des standortübergreifend nutzbaren Probandenmanagementsystems: `https://gitlab.gwdg.de/pung/demonstrator-probandenmanagementsystem`

4.1 Grundlagen

In diesem Unterkapitel werden einige Grundlagen für die Implementierung des standortübergreifenden Probandenmanagementsystems vorgestellt und wichtige Begriffe eingeführt. Software- und Systemarchitektur sind zwei Begriffe, deren Erläuterung benötigt wird, um die Komponenten des Probandenmanagementsystems und deren Zusammenspiel verständlich zu machen. Darauf aufbauend wird erläutert, was unter dem Begriff Softwarestack zu verstehen ist und wie der Softwarestack des Probandenmanagementsystems aussieht. Abschließend wird kurz der Begriff der Interoperabilität vorgestellt.

4.1.1 Software- und Systemarchitektur des Probandenmanagementsystems

Der Begriff der Architektur (aus dem lateinischen / altgriechischen: Baukunst) ist in der Informatik mit einer Vielzahl an verschiedenen Definitionen belegt [Hellige, 2004; Vogel u. a., 2009]. Der Architekturbegriff findet dabei in unterschiedlichen Bereichen Anwendung: Unter anderem die Systemarchitektur, als das Zusammenspiel eines IT-Systems mit anderen, die Prozessorarchitektur, als die Funktionsweise der Kernkomponenten eines Prozessors, oder die Softwarearchitektur, als die Funktionsweise von Computerprogrammen, Daten und deren Zusammenspiel. Für diese Arbeit relevant sind vor allem die Begriffe der Softwarearchitektur sowie der Systemarchitektur. Zahlreiche Personen haben sich an der Definition dieser beiden Begriffe versucht. Der gemeinsame Nenner dieser Definitionsversuche besteht darin, dass *Softwarearchitektur* die Komponenten eines Softwaresystems und ihr Zusammenspiel innerhalb des Softwaresystems beschreibt. *Systemarchitektur* ist der breitere Begriff, der auf Systeme und die in ihnen zusammenwirkenden Komponenten ausgerichtet ist.
Für eine detaillierte Terminologiediskussion des Begriffs *Softwarearchitektur* sei auf eine Sammlung an Definitionen der Carnegie Mellon University verwiesen [SEI, 2010]. Die Softwarearchitektur beschreibt noch nicht den detaillierten Softwareentwurf, vielmehr die Zusammenhänge zwischen den Anforderungen und dem zu konstruierenden System, möglichst mit einer Begründung für die Entwurfsentscheidungen [Hasselbring, 2006].

Der Entwurf einer Softwarearchitektur ist also der Erstellungsprozess einer Grobstruktur eines Softwaresystems. Dabei dienen funktionale und nichtfunktionale Anforderungen sowie technische und organisatorische Einflussfaktoren als Eingabe [vgl. Posch u. a., 2011, S. 95 ff.]. Eine funktionale Anforderung legt fest, was eine Software leisten muss. Nichtfunktionale Anforderungen hingegen sind meist unspezifischer und beschreiben, wie eine Software eine Leistung erbringen soll. Nichtfunktionale Anforderungen werden auch als Randbedingungen und Qualitätseigenschaften verstanden [vgl. Eigner u. a., 2012, S. 52].

Ein weiterer wichtiger Begriff im Zusammenhang mit der Softwarearchitektur ist der des Architekturmusters. Architekturmuster (engl. *architectural pattern* oder *architectual style*) sind abstrakte Beschreibungen empfohlener Vorgehensweisen zum Aufbau von Architekturen sowie deren Einsatzgebieten [vgl. Sommerville, 2018, S. 200].

Auf diesen Begriffserläuterungen aufbauend kann nun die Softwarearchitektur und die Systemarchitektur des Probandenmanagementsystems betrachtet werden. Die Softwarearchitektur des Probandenmanagementsystems orientiert sich an einem Architekturmuster. Schwanke [2015] wählte für das Probandenmanagementsystem das Architekturmuster *Ports and Adapters* [vgl. Schwanke, 2015, S. 15]. *Ports and Adapters*, das auch hexagonale Architektur genannt wird, wurde von dem amerikanischen Informatiker Alistair Cockburn entwickelt [Cockburn, 2005]. Schwanke [2015] stützte seine Entscheidung auf die Empfehlung von Vernon [2013] und Starke [2015], die die Verwendung dieses Architekturmusters zur Strukturierung der Fachdomäne empfehlen [vgl. Schwanke, 2015, S. 15]. Abbildung 4.1 zeigt die Terminologie und den Aufbau der *Ports* und der *Adapter*.

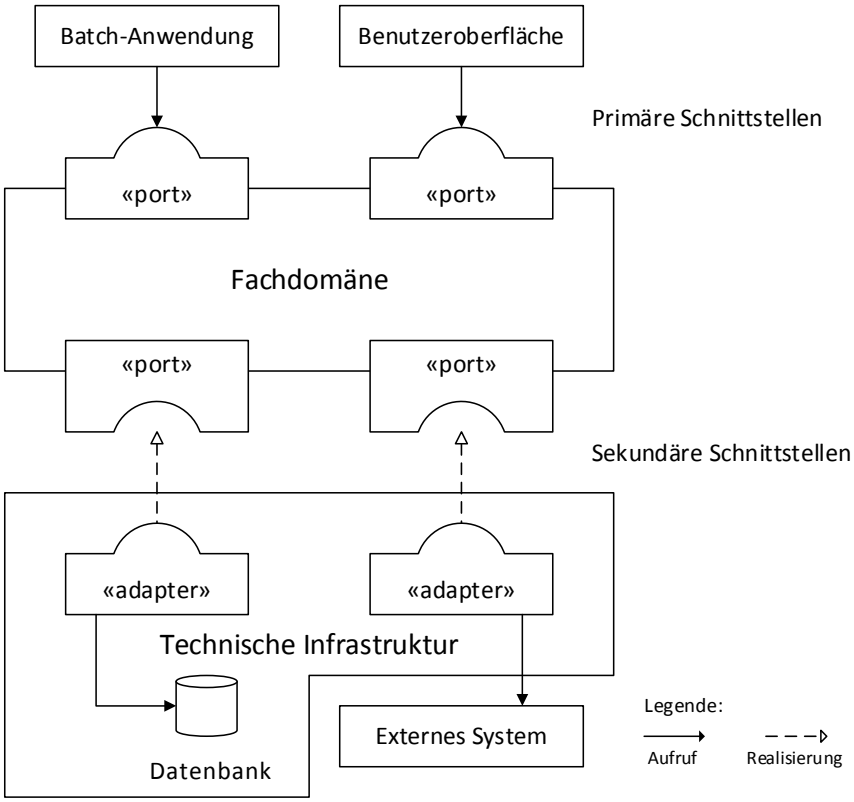

Abbildung 4.1: Darstellung des Architekturmusters *Ports and Adapters* nach Starke [2015] aus [Schwanke, 2015, S. 16].

Durch die Ports, die die primäre und sekundäre Schnittstellen bilden, wird die Fachdomäne gekapselt. Ziel der Kapselung ist die Isolierung der Fachdomäne einerseits von der Benutzeroberfläche und andererseits von technischer Infrastruktur wie der Datenbank. Primäre Schnittstellen werden von der Fachdomäne zur direkten Verwendung bereitgestellt. Sie bilden die Grundlage zur Implementierung einer Batch-Anwendung oder der Benutzeroberfläche. Sekundäre Schnittstellen müssen durch geeignete Adapter implementiert werden, z. B. für den Zugriff auf die Datenbank oder auf externe Systeme [vgl. Schwanke, 2015, S. 16].

Die Systemarchitektur des Probandenmanagementsystems, wie sie auch von Schwanke [2015] verwendet wurde, ist in Abbildung 4.2 dargestellt.

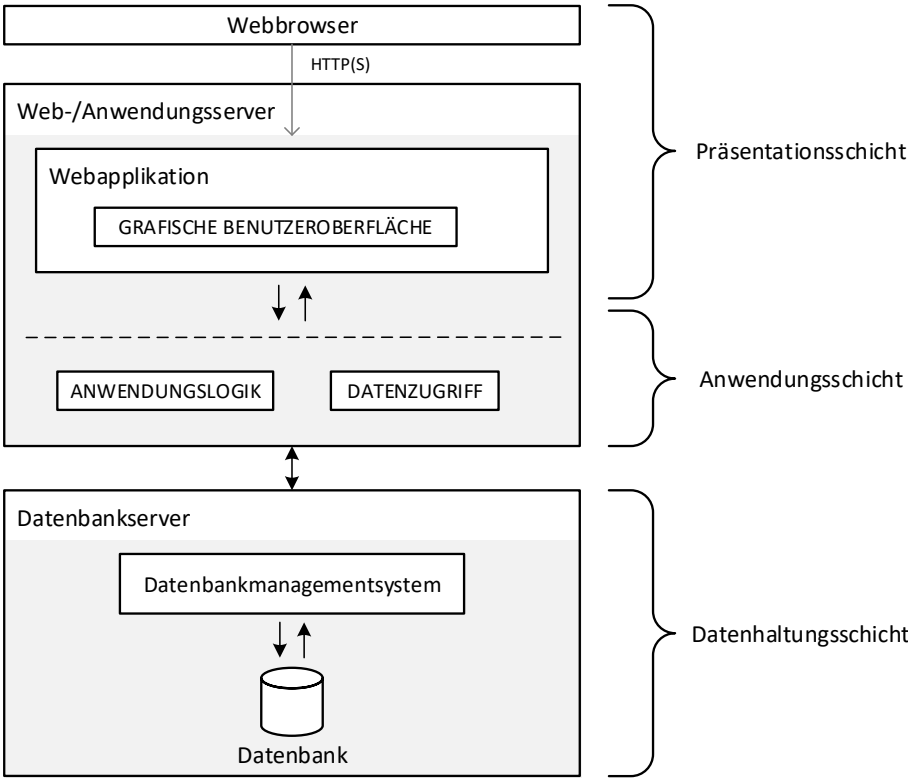

Abbildung 4.2: Übersicht über die Systemarchitektur des Probandenmanagementsystems. Das Probandenmanagementsystem setzt eine dreischichtige Architektur mit Präsentations-, Applikations- und Datenbankschicht um.

Das Probandenmanagementsystem ist als Webanwendung konzipiert. Die Anwendung läuft auf einem Webserver und Systembenutzende greifen auf die Benutzeroberfläche über einen Webbrowser zu. Die Systemarchitektur des Probandenmanagementsystems orientiert sich an der häufig angewandten Drei-Schichten-Architektur (englisch *three tier architecture*). Eine Präsentationsschicht (*client tier* oder *presentation tier*), auch als Front-End bezeichnet, ist für die Benutzerschnittstelle verantwortlich. Die zweite Schicht (*application-server tier* oder *middle tier*) vereint die Anwendungslogik, die dritte Schicht, die Datenhaltungsschicht (*data-server tier*) ist verantwortlich für das Speichern und Laden von Daten. Ein Datenbankserver beinhaltet das Datenbankmanagementsystem, also die Software die die Datenbank verwaltet, sowie die Datenbank als tatsächliche Ansammlung von Daten selbst.

4.1.2 Softwarestack des Probandenmanagementsystems

Als Softwarestack oder Solution-Stack wird eine Reihe von aufeinander aufbauenden Softwarekomponenten, die gemeinsam eine Plattform zur Ausführung einer gemeinsamen Anwendung bilden, bezeichnet. Ein Software-Stack kann aus einzelnen Komponenten, zum Beispiel Betriebssystem, Web-Server-Technologie und Laufzeitumgebung, Programmiersprachen und -werkzeugen, Anwendungen für die Datenpersistenz sowie Frameworks, Bibliotheken und Schnittstellen bestehen [vgl. Geißler und Ostler, 2019].
Verwendung im Probandenmanagementsystem findet die Programmiersprache Java[2], eine Kombination der Frameworks Spring[3] und Vaadin[4], ein Webserver mit Servlet-Container, z.B. Apache Tomcat[5], sowie ein relationales Datenbankmanagementsystem, wie z.B. MariaDB[6]. Des Weiteren kommt das Build-Management-Tool Maven[7] sowie einige weitere Softwarebibliotheken, unter anderem die JPA[8]-Referenzimplementierung Hibernate[9], zum Einsatz. Abbildung 4.3 zeigt eine Übersicht über den eingesetzten Softwarestack des Probandenmanagementsystems.

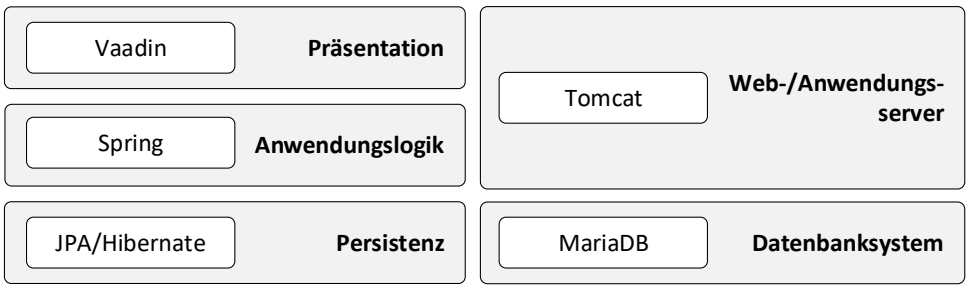

Abbildung 4.3: Übersicht über die Präsentations-, Anwendungslogik- und Persistenz-Schicht des Probandenmanagementsystems und den eingesetzten Softwarestack.

Das Spring Framework bietet ein umfassendes Programmier- und Konfigurationsmodell für Java-basierte Anwendungen. Ergänzt wird es durch Vaadin als Webframework

2 Java: `https://www.oracle.com/java/` [Stand: 01.03.2022]
3 Spring: `https://spring.io/projects/spring-framework` [Stand: 01.03.2022]
4 Vaadin: `https://vaadin.com/framework/` [Stand: 01.03.2022]
5 Apache Tomcat: `https://tomcat.apache.org/` [Stand: 01.03.2022]
6 MariaDB: `https://mariadb.com/` [Stand: 01.03.2022]
7 Apache Maven: `https://maven.apache.org/` [Stand: 01.03.2022]
8 Java Persistence API (JPA): `https://oracle.com/java/technologies/persistence-jsp.html` [Stand: 01.03.2022]
9 Hibernate: `https://hibernate.org/` [Stand: 01.03.2022]

für Internetanwendungen mit vielfältigen Interaktionsmöglichkeiten. JPA ist eine Programmierschnittstelle für Datenbankzugriffe und objektrelationales Mapping, dessen Referenzimplementierung Hibernate für die objektrelationale Abbildung der Persistenz-Schicht des Probandenmanagementsystems verwendet wird. Als Webserver kommt der Open-Source-Webserver und Servlet-Container Apache Tomcat zum Einsatz. MariaDB wird als Datenbankmanagementsystem zur Datenhaltung eingesetzt. Apache Maven ist ein Build-Management-Tool mit dem die Erstellung von Java-Programmen standardisiert verwaltet und durchgeführt werden kann.

4.1.3 Interoperabilität

Als Interoperabilität (von lateinisch opus „Arbeit“ und inter „zwischen“) bezeichnet man die Fähigkeit zur Zusammenarbeit von verschiedenen Systemen, Techniken oder Organisationen. Der Begriff zielt auf den pragmatisch entscheidenden Punkt ab, dass Geräte oder Systeme mit (ggf. noch unbekannten) Fremdgeräten oder Fremdsystemen kommunizieren/zusammenarbeiten können [vgl. RfII, 2016a, S. 15]. Es werden allgemein zwei Arten von Interoperabilität unterschieden: die syntaktische und die semantische Interoperabilität [Thun und Dewenter, 2017]. In Zusammenhang mit Software spricht man vor allem dann von syntaktischer Interoperabilität, wenn mehrere Programme dasselbe Dateiformat verwenden; im Zusammenhang mit verteilten Systemen, wenn alle Systemkomponenten dieselben Datenformate oder dasselbe Kommunikationsprotokoll verwenden [Wunder und Grosche, 2009; Gregorczyk, 2014]. Unter semantischer Interoperabilität ist zu verstehen, dass Anwendungssysteme nicht nur Daten austauschen können (syntaktische Interoperabilität), sondern die ausgetauschten Daten auch identisch interpretieren [Thun und Dewenter, 2017]. Die syntaktische Interoperabilität ist somit eine wichtige Voraussetzung für die semantische Interoperabilität.

4.2 Datentrennung und vernetzte Probandenverwaltung

Dieser zweite Teil des Implementierungskapitels beschreibt die Implementierung der in Iteration 1 (Unterkapitel 3.3, S. 36) und Iteration 2 (Unterkapitel 3.4, S. 53) erarbeiteten Konzepte zur Datentrennung und zur vernetzten Probandenverwaltung im Probandenmanagementsystem. Beide Konzepte sind über ihre zugrundeliegende Problemstellung, an welchem Ort identifizierende Probandendaten gespeichert werden und wie sie ausgetauscht werden können, eng miteinander verknüpft. Die konzeptionierte Datentrennung und vernetzte Probandenverwaltung erforderten Anpassungen am Datenmodell und Implementierungsarbeiten an den Schnittstellen des Probandenmanagementsystems für eine standortübergreifende Nutzbarkeit. Die Umsetzung baut auf der Implementierung des Probandenmanagementsystems von Schwanke [2015] und dessen Programmcode auf.

4.2.1 Datenmodell

Eine grundlegende Änderung zum Konzept von Schwanke [2015] ist es, keine externen Pseudonyme der Clinical Data Management Systeme mehr zu speichern (vgl. Unterkapitel 3.3.3, Iteration 1: Daten und Datenflüsse des Probandenmanagementsystems, S. 48). Stattdessen wird eine temporäre Pseudonymabfrage über eine Schnittstelle zum Identitätsmanagement eingerichtet. Die Implementierung der Schnittstelle des Probandenmanagementsystems zum Identitätsmanagement wird im Anschluss an dieses Unterkapitel in Unterkapitel 4.2.2 vorgestellt. Zuvor muss jedoch das Datenmodell des Probandenmanagementsystems vorbereitet werden, um verschiedene Metadaten zur Konfiguration der Schnittstelle aufnehmen zu können. Des Weiteren müssen die über die Schnittstelle angefragten Pseudonyme im Probandenmanagementsystem verarbeitet werden. Der geänderte Umgang mit den Pseudonymen der Studienteilnehmen erfordert Anpassung am Datenmodell der Probandenidentifikatoren. Die Anpassung des Datenmodells teilt sich somit in zwei Teile: der Anpassung des Datenmodells für Identitätsmanagementmetadaten und der Anpassung des Datenmodells für Probandenidentifikatoren.

Datenmodell für Identitätsmanagementmetadaten

Das von Schwanke entwickelte *Domänenmodell des Bereichs Pseudonymisierung* [vgl. Schwanke, 2015, S. 49] bietet eine gute Grundlage zur Weiterentwicklung des Datenmodells, wie es für das Konzept der standortübergreifenden Probandenverwaltung nötig wird. Abbildung 4.4 zeigt einen Auszug des Domänenmodells von Schwanke [2015]. Pseudonyme (Entität *Identifier*) beziehen sich auf ein oder mehrere Forschungssysteme (Entität *ResearchSystem*) wie Clinical Data Management Systeme, die sich wiederum auf ein Forschungsprojekt (Entität *Project*) beziehen. Pseudonyme können durch externe Systeme (Entität *ExternalSystem*) wie einem Identitätsmanagement generiert werden und sind damit vom Typ der Entität *ExternalIdentifier*. Externe Systeme sind, wie auch Forschungssysteme, einem Forschungsprojekt zugewiesen.

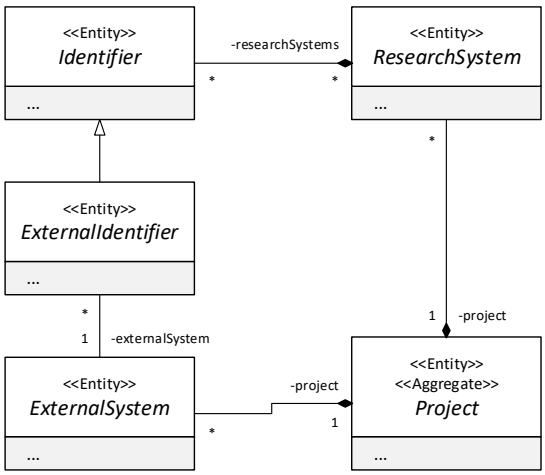

Abbildung 4.4: Zusammengefasstes Domänenmodell der Bereiche *Kommunikation* [vgl. Schwanke, 2015, Abbildung 22, S. 48] und *Pseudonymisierung* [vgl. Schwanke, 2015, Abbildung 23, S. 49] zur Darstellung der Beziehungen zwischen *Identifier*, *ExternalIdentifier*, *ExternalSystem*, *ResearchSystem* und *Project*.

Das Probandenmanagementsystem soll jedoch nicht mehr zwangsläufig alle Pseudonyme speichern, sondern auch eine temporäre Pseudonymabfrage über eine Schnittstelle zum Identitätsmanagement ermöglichen. Dieser Anwendungsfall war für das von Schwanke [2015] geplante System, das Pseudonyme uneingeschränkt speichert, nicht relevant. Daher setzte das bisherige Datenmodell lediglich voraus, dass die Existenz eines Pseudonyms über die Entität *Identifier* abgebildet wird. In diese Logik soll für temporär abzurufende Pseudonyme nicht eingegriffen werden, doch lässt die Abhängigkeit der Entität *ExternalSystem* von der Entität *Project* (vgl. Abbildung 4.4) nicht zu,

abzubilden, welche Forschungssysteme welche externen Identitätsmanagementsysteme nutzen, wenn keine Entität eines *Identifier* existiert. Externe Systeme müssen daher auch mit Forschungssystemen und nicht mehr nur mit Forschungsprojekten in Beziehung gebracht werden können. Zur Darstellung der Zusammenhänge zwischen Entitäten und Beziehungen eignet sich die Chen-Notation, eine grafische Notation für Entity-Relationship-Modelle, die nach dem Informatiker Peter Chen benannt wurde, der sie zusammen mit dem Entity-Relationship-Modell 1976 zur Darstellung von Datenmodellen einführte [Chen, 1976]. Abbildung 4.5 zeigt als Entity-Relationship-Modell in Chen-Notation die benötigten Entitäten und ihre Beziehungen zueinander, um alle benötigten Identitätsmanagementmetadaten des Probandenmanagementsystems abbilden zu können.

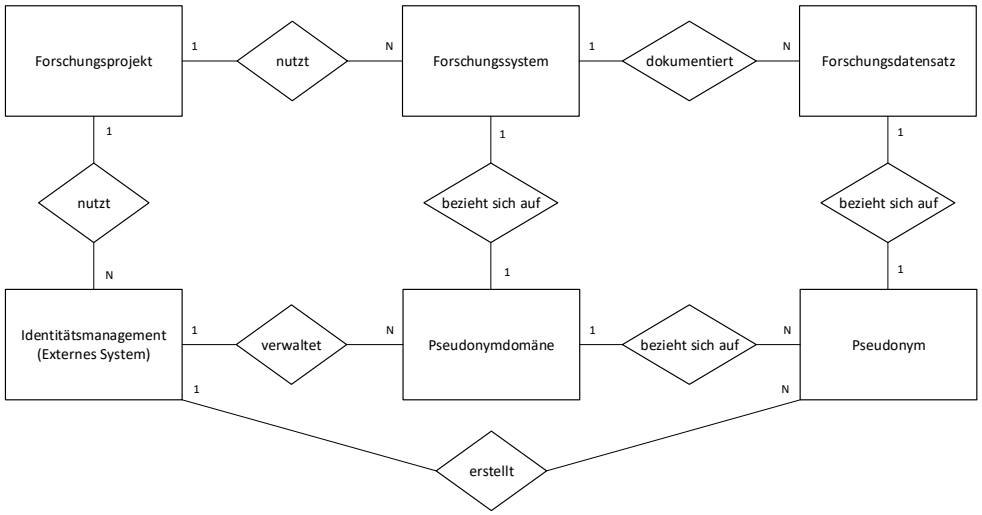

Abbildung 4.5: Vereinfachtes Entity-Relationship-Modell in Chen-Notation zur Darstellung der Entitäten Forschungsprojekt, -system, -datensatz, Identitätsmanagement, Pseudonymdomäne und Pseudonym und ihre Beziehungen zueinander.

Ein Forschungsprojekt nutzt ein oder mehrere Forschungssysteme, in denen medizinische Forschungsdaten dokumentiert werden. Die Dokumentation der Forschungsdaten erfolgt pseudonymisiert, sodass in einem Forschungsprojekt ein oder auch mehrere Identitätsmanagementsysteme zur Pseudonymisierung eingesetzt werden. Grundsätzlich beziehen sich Pseudonyme immer auf eine bestimmte Domäne, in der sie Gültigkeit besitzen, in der Regel innerhalb eines Forschungsprojektes für ein bestimmtes Forschungssystem, z. B. eines Clinical Data Management Systems. Die Pseudonymdomäne (Entität *IdentifierDomain*) wurde daher in das Datenmodell des

Probandenmanagementsystems aufgenommen (siehe Abbildung 4.6).

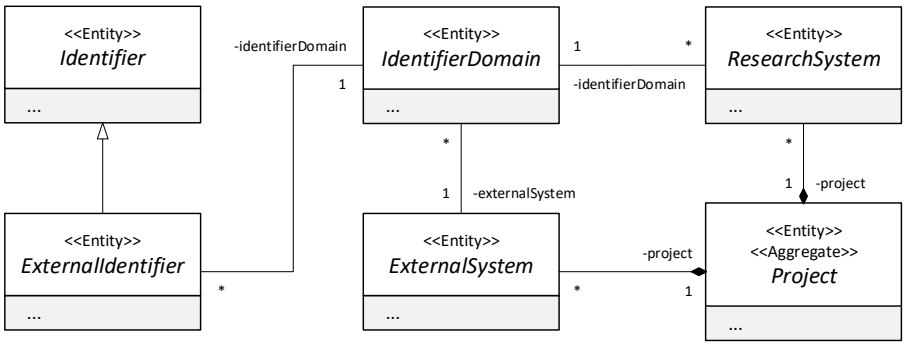

Abbildung 4.6: UML Klassendiagramm aus Abbildung 4.4 erweitert um die Entität *IdentifierDomain* und ihre Beziehungen zu den Entitäten *Identifier*, *ExternalIdentifier*, *ExternalSystem*, *ResearchSystem* und *Project*.

Die Entität *IdentifierDomain* bildet nun das verbindende Element zwischen den Entitäten *ResearchSystem*, *ExternalSystem* und *ExternalIdentifier*. In Abbildung 4.7 ist das zugehörige relationale Tabellendiagramm dargestellt.

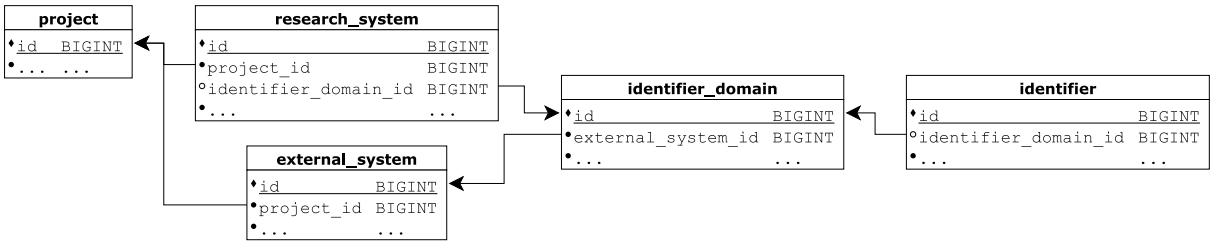

Abbildung 4.7: Auszug des Tabellendiagramms des Probandenmanagementsystems (vereinfacht auf Schlüsselspalten) mit den Tabellen die Daten und Beziehungen zwischen Forschungsprojekt, -system, externem Identitätsmanagement, Pseudonymdomäne und Pseudonym abbilden.

Forschungsprojekte werden über die Tabelle *project*, Forschungssysteme über die Tabelle *research_system* und externe Systeme über die Tabelle *external_system* repräsentiert. Die Pseudonymdomänen werden über die Tabelle *identifier_domain* und Pseudonyme über die Tabelle *identifier* abgebildet.

Die gesamte Konfiguration des Probandenmanagementsystems erfolgt forschungsprojektbezogen. Das bedeutet, eine Administratorin oder ein Administrator erfasst alle relevanten Informationen über das Forschungsprojekt im System und konfiguriert so die nötigen Einstellungen auf Ebene eines Forschungsprojektes für die Benutzung [vgl. Schwanke, 2015, S. 28]. Das angepasste Datenmodell erlaubt nun die benötigten Identitätsmanagementmetadaten für ein Forschungsprojekt im

System abzubilden. In der erweiterten Benutzeroberfläche zur Konfiguration des Probandenmanagementsystems kann eine Administratorin oder ein Administrator externe Systeme für ein Forschungsprojekt anlegen und Informationen zu diesen hinterlegen (siehe Abbildung 4.8).

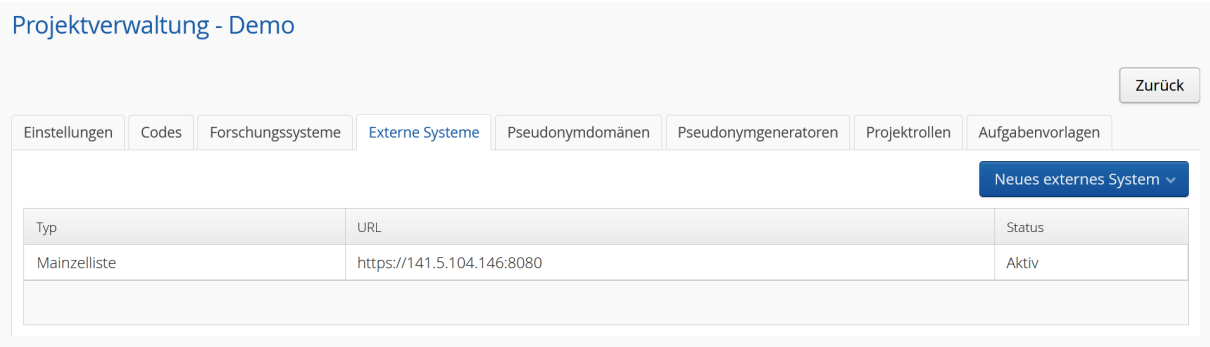

Abbildung 4.8: Screenshot der Benutzeroberfläche des Probandenmanagementsystems zur Konfiguration von externen Systemen.

Relevante Datenfelder, die ein externes Identitätsmanagement beschreiben, sind Name des Systems, URL und Port der Webschnittstelle sowie der Typ der Schnittstelle, der für die konkrete Implementierung dieser Schnittstelle relevant ist. Des Weiteren lassen sich Forschungssysteme und Pseudonymdomänen anlegen und konfigurieren. Die Pseudonymdomänen werden einem Forschungssystem zugewiesen (siehe Abbildung 4.9).

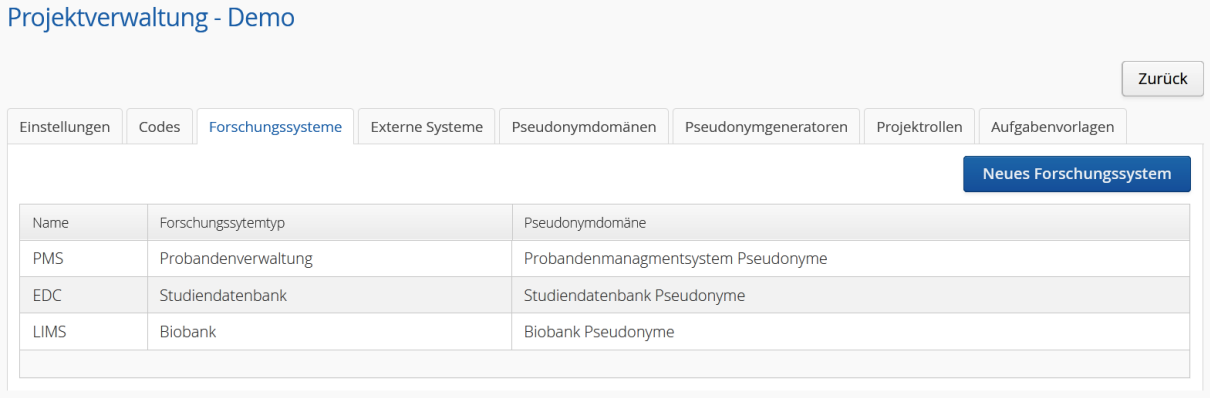

Abbildung 4.9: Screenshot der Benutzeroberfläche des Probandenmanagementsystems zur Konfiguration von Forschungssystemen. Konfiguriert werden der Name des Forschungssystems, der Forschungssystemtyp und eine zugehörige Pseudonymdomäne.

In diesem Abschnitt zum Datenmodell für Identitätsmanagementmetadaten lag der Fokus auf den neu geschaffenen Relationen zur Konfiguration der Schnittstellen zum Identitätsmanagement. Auf die genaue Abbildung der Pseudonyme wird im nächsten Unterkapitel zum Datenmodell für Probandenidentitäten genauer eingegangen.

Datenmodell für Probandenidentifikatoren

Mittelpunkt des Datenmodells des Probandenmanagementsystems bildet der Kerndatensatz eines Studienteilnehmenden mit seiner Forschungsakte in einem Forschungsprojekt. Die Beziehungen zwischen Probandin oder Proband, Forschungsakte und Forschungsprojekt lässt sich im Rahmen der Datenmodellierung wie folgt beschreiben: Eine Probandin oder ein Proband hat eine oder mehrere Forschungsakten, die jeweils einem Forschungsprojekt zugewiesen sind (siehe Abbildung 4.10).

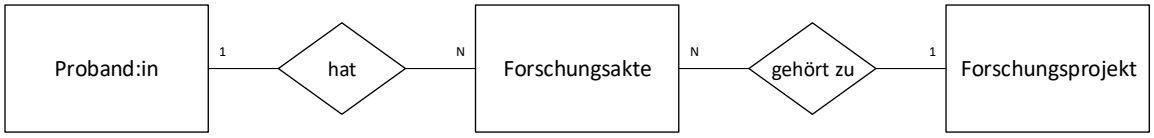

Abbildung 4.10: Vereinfachtes Entity-Relationship-Modell in Chen-Notation zur Darstellung der Beziehungen zwischen Proband:in, Forschungsakte und Forschungsprojekt.

Übertragen auf ein relationales Datenbankschema sind Daten und Beziehungen über die Tabellen *participant* (Probandin / Proband), *record* (Forschungsakte) und *project* (Forschungsprojekt) beschrieben (siehe Abbildung 4.11).

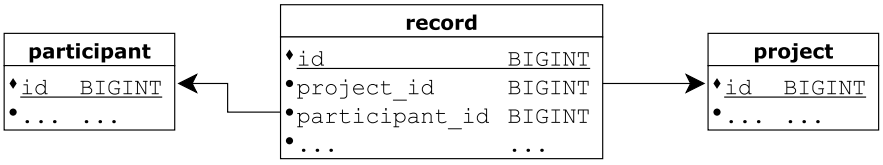

Abbildung 4.11: Auszug des Tabellendiagramms des Probandenmanagementsystems (vereinfacht) mit den Tabellen *participant*, *record* und *project* mit Beziehungen über Primär- und Fremdschlüssel.

Für jede Tabelle stellt die Spalte *id* den Primärschlüssel dar, der zur eindeutigen Identifizierung des Datensatzes verwendet wird. Der Wert eines Primärschlüssels muss in einer Tabelle einmalig sein, da er jeden Datensatz eindeutig kennzeichnet. Er bildet den Schlüssel, der in Folge für die Abbildung der Relationen verwendet wird. Die Werte dieses Schlüssels werden in referenzierenden Tabellen als Fremdschlüssel verwendet. Umgesetzt im Probandenmanagementsystem über das Hibernate-Framework als Java Persistence API (JPA)-Implementierung, zeigt Listing 4.1 einen Auszug des Programmcodes der Klasse *Participant* und Listing 4.2 der Klasse *AbstractPersistable*.

Die Klasse *AbstractPersistable* vererbt allen ableitenden Klassen, hier im Beispiel der Klasse und Entität *Participant*, eine Implementierung für eindeutige Schlüssel im Hibernate/JPA-Kontext (siehe Abbildung 4.12).

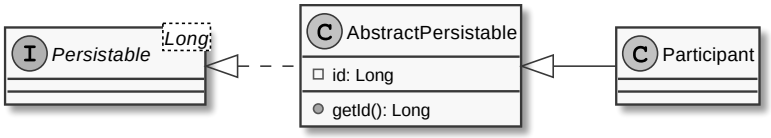

Abbildung 4.12: Klassendiagramm der Klassen *Persistable*, *AbstractPersistable* und *Participant* ihr Verhältnis zueinander.

```
@Entity
@Table(name = "participant")
public class Participant extends AbstractPersistable {

    @NotNull
    @Size(min = 1, max = 255)
    @Column(name = "first_name", nullable = false, length = 255)
    private String firstName;
...
```

Listing 4.1: Programmcode-Auszug der Klasse *Participant*, die eine Participant-Entität beschreibt und im Datenmodell die Tabelle *participant* erzeugt. Die Klasse *Participant* erbt von der Klasse *AbstractPersistable*. Exemplarisch mitdargestellt das Pflichtfeld eines Probandendatensatzes *firstName* mit seinen Metadaten.

```
1 @MappedSuperclass
2 public class AbstractPersistable implements Persistable<Long> {
3
4     @Id
5     @GeneratedValue(strategy = GenerationType.IDENTITY)
6     private Long id;
7 ...
```

Listing 4.2: Auszug des Programmcodes der Klasse *AbstractPersistable*, die das Hibernate Interface *Persistable* implementiert und über die Variable *id* allen ableitenden Klassen einen Primärschlüssel erzeugt.

Zeile 4 in Listing 4.2 beschreibt über die @Id-Annotation die Eigenschaft eines Schlüsselwertes für die Java-Variable *id* vom Java-Datentyp *Long*. Zeile 5 im

gleichen Listing bewirkt über die @GeneratedValue-Annotation mit der Strategie *GenerationType.IDENTITY*, dass diese Schlüsselwertspalte über die Datenbank mit eindeutigen, automatisch generierten Werten befüllt wird. Der Wert der Spalte *id* der Tabelle *participant* wird nach diesem Prinzip für die eindeutige Kennzeichnung eines Studienteilnehmenden genutzt und bildet so den ersten Identifikator einer Probandin oder eines Probanden im Probandenmanagementsystem. Dieser Identifikator ist nach dem Strukturierungsprinzip einer Software-Schichtenarchitektur ein technischer Schlüssel der Datenbank-Schicht, der von der Anwendungslogik zu trennen ist. Die ID eines Studienteilnehmenden in der Tabelle *participant* im Probandenmanagementsystem bleibt damit der Benutzerin oder des Benutzers des Systems verborgen. Identifikatoren eines Studienteilnehmenden, die zur Anzeige dienen, sind die Pseudonyme des Studienteilnehmenden.

Wie bereits im vorangehenden Unterkapitel beschrieben, beziehen sich Pseudonyme immer auf eine bestimmte Domäne, in der sie eindeutig sind und damit Gültigkeit besitzen. Für das Probandenmanagementsystem bildet ein bestimmtes Forschungssystem innerhalb eines Forschungsprojektes eine Domäne. Im Datenmodell des Probandenmanagementsystems ist ein Pseudonym an eine Forschungsakte gebunden, da die Forschungsakte nach Schwanke [2015, S. 47] die Teilnahme eines Studienteilnehmenden an einem Forschungsprojekt repräsentiert. Eine Probandin oder ein Proband besitzt also innerhalb eines Forschungsprojektes und seiner Forschungsakte für eine bestimmte Pseudonymdomäne, die ein Forschungssystem repräsentiert, verschiedene Pseudonyme. Abbildung 4.13 verdeutlicht diese Beziehungen der verschiedenen Entitäten in einem Entity-Relationship-Modell.

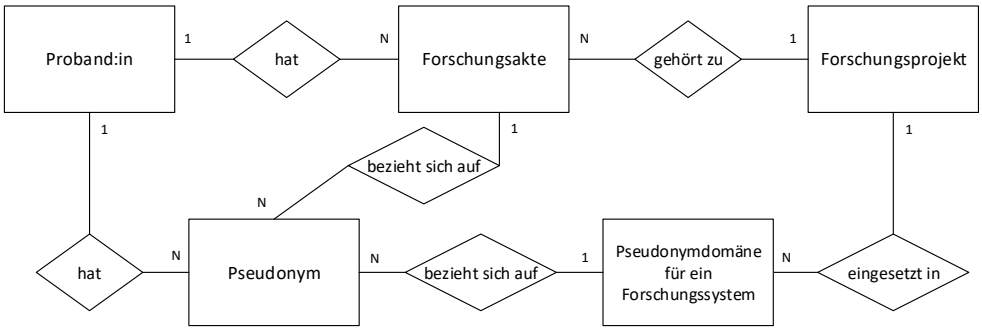

Abbildung 4.13: Vereinfachtes Entity-Relationship-Modell in Chen-Notation zur Darstellung der Beziehungen zwischen Proband:in, Pseudonym, Forschungsakte, -system und -projekt.

Zur Repräsentation der Pseudonyme erweitern die Tabellen *identifier*, *research_system*, *external_system* und *identifier_domain* das in Abbildung 4.11 dargestellte relationale Datenbankschema des Probandenmanagementsystems, dargestellt in Abbildung 4.14.

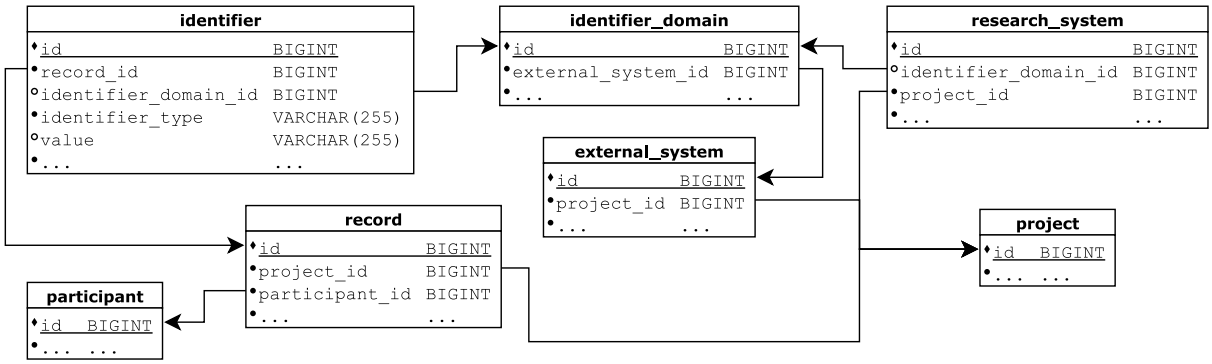

Abbildung 4.14: Auszug des Tabellendiagramms des Probandenmanagementsystems (vereinfacht) mit den Tabellen *identifier*, *identifier_domain*, *research_system*, *external_system*, *participant*, *record* und *project*, die Daten und Beziehungen zwischen Studienteilnehmenden, Pseudonymen, Forschungsakten, -systemen und -projekten abbilden.

Die Erweiterung des Datenmodells des Probandenmanagementsystems wurde so gewählt, dass die Kompatibilität mit in standortübergreifenden Forschungsprojekten nicht benötigen Funktionen des Probandenmanagementsystems, wie beispielsweise manuell generierten Pseudonymen, gewahrt bleibt. Die Namensvergabe neuer Tabellen und Spalten wurde an die bisher verwendete Namenskonvention angepasst.

Eine wichtige Erweiterung des bisherigen Datenmodells, die bereits im vorherigen Unterkapitel zum Datenmodell für Identitätsmanagementmetadaten angesprochen wurde, sind Pseudonymdomänen. Konfigurierbar über die neue Benutzeroberfläche des Probandenmanagementsystems können Administratorinnen oder Administratoren Metainformationen zu den im Forschungsprojekt verwendeten Pseudonymen hinterlegen. Datenfelder, die erfasst werden können, sind der Name des Pseudonyms, eine Beschreibung, ein Kürzel und die Information, ob ein Pseudonym im Probandenmanagementsystem gespeichert werden darf oder nicht. Die Unterscheidung von Pseudonymen in temporäre Abrufpseudonyme oder für das Probandenmanagementsystem persistierbare Pseudonyme ist eine wichtige Neuerung, um die zentrale Anforderung einer starken Datentrennung (vgl. Unterkapitel 3.3.3, S. 41) umzusetzen. In Iteration 2 wurde die Unterscheidung der Pseudonyme in Bezug auf die benötigte Schnittstelle zum Identitätsmanagement mit den Anwendungsfällen zur

Pseudonymzuweisung und zur Pseudonymabfrage (vgl. Abbildung 3.10, S. 55) erweitert. Für den Anwendungsfall der Pseudonymzuweisung beim Einsatz des Probandenmanagementsystems in standortübergreifenden Forschungsprojekten ist das persistierbare Probandenmanagementsystempseudonym (PSN_{PMS}) vorgesehen (vgl. Abbildung 3.8, S. 49). Dieses Pseudonym darf in der Datenbank des Probandenmanagementsystems gespeichert werden und erhält einen Eintrag in der Tabelle *identifier*. Das Pseudonym kann aus einer beliebigen Zahlen und Zeichenkette bestehen und wird in der Spalte *value* gespeichert (vgl. Abbildung 4.14). Zur Pseudonymzuweisung muss über die Schnittstelle zum Identitätsmanagement eine Liste von identifizierenden Datenfeldern des Studienteilnehmenden übergeben werden (vgl. Unterkapitel 4.2.2, Schnittstelle, S. 117). Um die Auswahl der Übergabeparameter zu speichern, wurde das Datenmodell des Probandenmanagementsystems erweitert und die Konfigurationsmöglichkeit in der Benutzeroberfläche des Probandenmanagementsystems geschaffen (siehe Abbildung 4.15).

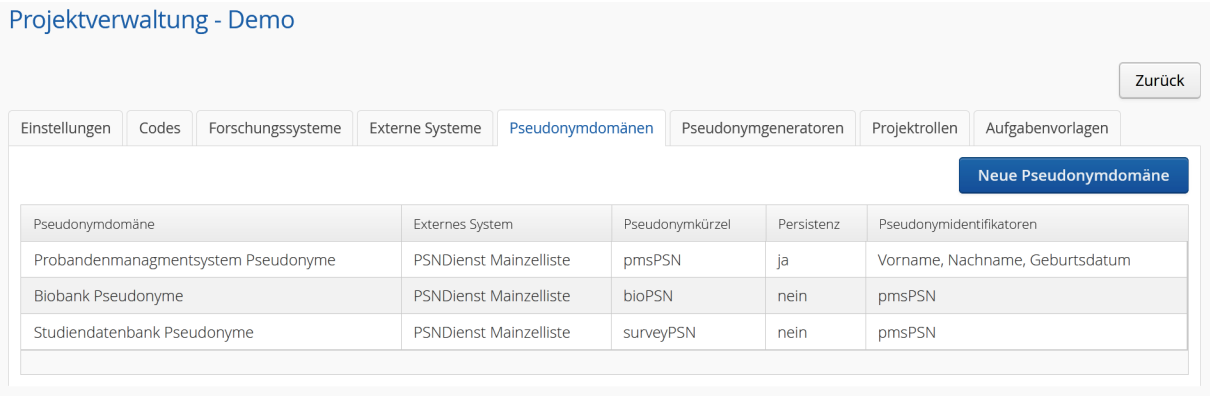

Abbildung 4.15: Screenshot der Benutzeroberfläche des Probandenmanagementsystems zur Konfiguration von Pseudonymdomänen. Konfiguriert werden der Name der Pseudonymdomäne, das zugehörige externe System, ein Pseudonymkürzel, die Information ob das Pseudonym persistiert werden darf sowie die Pseudonymidentifikatoren.

Abbildung 4.16 zeigt die Erweiterung des Datenmodells um die Tabellen *identifier_domain*, *domain_interface_mapping* und *interface_identifier*. Mit ihrer Hilfe lassen sich jeder Pseudonymdomäne eine beliebige Anzahl an Übergabeparameter zuweisen.

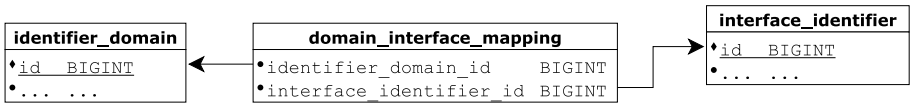

Abbildung 4.16: Auszug des Tabellendiagramms des Probandenmanagementsystems (vereinfacht) mit den Tabellen die Daten und Beziehungen zur Verwaltung von Pseudonymen abbilden.

Eine ähnliche Konfigurationsmöglichkeit besteht auch für den zweiten Anwendungsfall der Pseudonymabfrage. Auch hier wurde die Konfiguration in der Benutzeroberfläche geschaffen anzugeben, anhand welcher Parameter ein Abrufpseudonym angefragt wird. Im einfachsten Fall werden weitere Pseudonyme mit dem Übergabeparameter des PSN_{PMS} abgefragt (siehe Abbildung 4.15).
Mit den vorgestellten Anpassungen des Datenmodells wurde ein wichtiger Teil der im Konzeptionskapitel erarbeiteten Lösung implementiert. Als dritter Teil folgt die Implementierung der Interaktion mit der Schnittstelle externer Identitätsmanagementsysteme im nächsten Kapitel.

4.2.2 Schnittstelle

Iteration 2 beschrieb Anwendungsfälle, die hinsichtlich der Schnittstelle des standortübergreifenden Probandenmanagementsystems benötigt werden, abstrahierte dabei jedoch gezielt von einer technischen Lösung (vgl. Unterkapitel 3.4.2, Iteration 2, S. 54). Die nun folgende technische Betrachtung dient der Überprüfung der Machbarkeit der beiden Anwendungsfälle der Pseudonymzuweisung und -abfrage. Sie umfasst sowohl die prototypische Implementierung der Schnittstelle seitens des Probandenmanagementsystems als auch die Betrachtung der Gegenseite, dem Identitätsmanagement. Vorgaben bezüglich der Schnittstellenbeschreibung macht dabei die eingesetzte Software für das Identitätsmanagement; das Probandenmanagementsystem muss diese umsetzen. In Forschungsprojekten des Instituts für Medizinische Informatik der Universitätsmedizin Göttingen häufig für Identitätsmanagement und Pseudonymisierungsdienst eingesetzte Anwendungen sind die Mainzelliste[10],

[10] Mainzelliste: `https://www.unimedizin-mainz.de/imbei/informatik/ag-verbundforschung/mainzelliste.html` und `https://www.mainzelliste.de/` [Stand: 01.03.2022]

ihr Vorgänger PID-Generator[11] oder eine Kombination der Anwendungen E-PIX[12] und gPAS[13]. Alle Werkzeuge erlauben die Erzeugung von Patientenidentifikatoren aus identifizierenden Patientendaten, die dank fehlertoleranter Record-Linkage-Funktionalität auch bei schlechter Qualität identifizierender Patientendaten funktioniert (vgl. auch Unterkapitel 3.3.3, Iteration 1: Daten und Datenflüsse des Identitätsmanagements, S. 42). Die Mainzelliste ist ein webbasierter Identitätsmanagement-/Pseudonymisierungsdienst. Die 2013 vorgestellte Open-Source-Software stellt ihre Funktionen über eine REST-Schnittstelle bereit [Lablans u. a., 2015]. Der im Jahr 2005 vorgestellte PID-Generator ist ebenfalls ein webbasierter Identitätsmanagement-/Pseudonymisierungsdienst, der seine Funktionen über eine SOAP-Schnittstelle bereitstellt [Faldum und Pommerening, 2005]. Der PID-Generator wird aktuell nicht mehr weiterentwickelt, da es modernere Anwendungen mit gleicher Funktionalität gibt. E-PIX und gPAS wurden von 2012 bis 2015 im Rahmen des DFG-geförderten Projektes MOSAIC entwickelt [Bialke u. a., 2015a]. Die Funktionalität des Identitätsmanagements erbringt E-PIX, das Pseudonymmanagement wird mittels gPAS ergänzt. Die webbasierte Open-Source-Software stellt ihre Funktionen über eine SOAP-Schnittstelle bereit. Auch nach dem Projektende von MOSAIC werden die Anwendungen durch die Universitätsmedizin Greifswald weiterentwickelt. Projekte, bei denen die Treuhandstelle der Universitätsmedizin Greifswald beteiligt ist, verwenden als Schnittstelle zu weiteren Systemen ein E-PIX und gPAS übergeordneten Workflowmanager, den sogenannten Treuhandstellen-Dispatcher [vgl. Bialke u. a., 2015b]. Wünschenswert wäre eine softwareunabhängige, standardisierte Schnittstelle für Identitätsmanagement- und Pseudonymisierungsdienste, sodass das Probandenmanagementsystem nur eine Schnittstelle für seine Anwendungsfälle der Pseudonymisierung implementieren müsste und generell die Interoperabilität (siehe Unterkapitel 4.1.3, Grundlagen: Interoperabilität, S. 106) verschiedener die Schnittstelle nutzender Systeme gesteigert würde. Leider bieten alle hier genannten

[11] PID-Generator: `https://www.toolpool-gesundheitsforschung.de/produkte/pid-generator/` [Stand: 01.03.2022]

[12] E-PIX - Enterprise Identifier Cross-Referencing: `https://www.ths-greifswald.de/forscher/e-pix/` [Stand: 01.03.2022]

[13] gPAS - Generic Pseudonym Administration Service: `https://www.ths-greifswald.de/forscher/gpas/` [Stand: 01.03.2022]

Dienste unterschiedliche Schnittstellen an, vermutlich begründet in der zeitlichen Entstehung und dem unterschiedlichen Funktionsumfang der Dienste. Da für das Probandenmanagementsystem prinzipiell unterschiedliche Identitätsmanagement- und Pseudonymisierungsdienste die Pseudonymisierung übernehmen können, wählte Schwanke [2015], der bereits Schnittstellen zu externen Systemen in seinem Architekturkonzept vorsah [Schwanke, 2015, S. 89], einen erweiterbaren Ansatz, der für die Implementierung in dieser Arbeit genutzt wurde. Schwankes Architekturkonzept orientiert sich am Architekturmuster *Ports and Adapters* (vgl. Unterkapitel 4.1.1, Software- und Systemarchitektur des Probandenmanagementsystems, S. 101), indem sekundäre Schnittstellen durch geeignete Adapter implementiert werden [vgl. Schwanke, 2015, S. 16]. Die sekundäre Schnittstelle für externe Systeme, den Identitätsmanagement- und Pseudonymisierungsdiensten, wird über das Interface *ExternalSystemProvider* abgebildet [vgl. Schwanke, 2015, S. 51 f.]. Dieses Interface müssen die unterschiedlichen Adapter implementieren, die jeweils eine Implementierung einer Schnittstelle zu einem Identitätsmanagement- und Pseudonymisierungsdienst repräsentieren. Abweichend vom bisherigen Architekturkonzept von Schwanke [2015] (vgl. Unterkapitel 3.4.1, Iteration 2, Problem und Zielsetzung, S. 53) wurden für diese Arbeit die Anwendungsfälle zur Pseudonymzuweisung und –abfrage wie in Iteration 2 konzeptioniert, ergänzt. Abbildung 4.17 zeigt das vereinfachte Klassendiagramm zur Verdeutlichung des Adapterprinzips für Schnittstellen zu externen Identitätsmanagement- und Pseudonymisierungsdiensten.

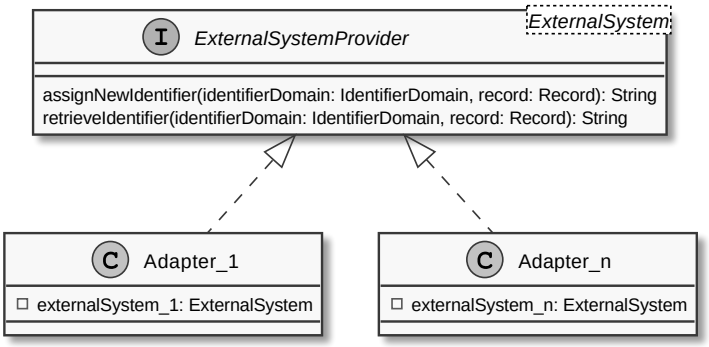

Abbildung 4.17: Vereinfachtes Klassendiagramm zur Darstellung der sekundären Schnittstelle für externe Systeme über das Interface *ExternalSystemProvider* und eine beliebige Anzahl an Adaptern, die das Interface implementieren.

Die Methode *assignNewIdentifier(...)* deckt den Anwendungsfall der Pseudonymzuweisung, die Methode *retrieveIdentifier(...)* den Anwendungsfall der Pseudonymabfrage ab. Wichtiger Übergabeparameter an beide Methoden ist ein Objekt der Klasse *IdentifierDomain*, über das die bei der Konfiguration des Probandenmanagementsystems hinterlegten Metadaten zu den Pseudonymen verfügbar gemacht werden (vgl. Unterkapitel 4.2.1, Datenmodell, S. 107). Des Weiteren wird ein Objekt der Klasse *Record* übergeben, über das auf die Daten der Forschungsakte des Studienteilnehmenden zugegriffen wird, für welche ein Pseudonym zugewiesen oder abgefragt werden soll. Die Implementierung eines Adapters mit den im Probandenmanagementsystem benötigten Funktionen erfolgte exemplarisch für die Mainzelliste.

Implementierung des Mainzelliste Adapters

Der grundsätzliche Kommunikationsablauf mit der Schnittstelle der Mainzelliste ist in Abbildung 4.18 dargestellt.

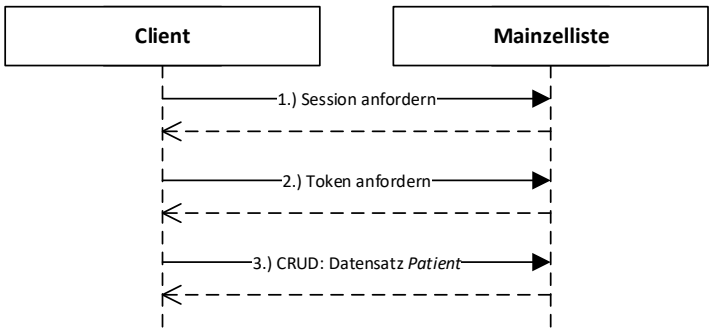

Abbildung 4.18: Sequenzdiagramm zum grundsätzlichen Kommunikationsablauf zwischen einem Client und der Schnittstelle der Mainzelliste. Die Schritte umfassen die Abfrage einer Session, eines Tokens sowie anschließende Create, Read, Update und Delete (CRUD) – Operationen auf den Datensatz *Patient*[14].

Dem REST-Paradigma folgend findet die Kommunikation über HTTP statt und verarbeitet verschiedene Ressourcen. In einem ersten Schritt fordert ein Client bei der Mainzelliste eine Session an, die eine Nutzersitzung repräsentiert und Tokens gruppiert. Ein Token wird im zweiten Schritt mit einer Session angefordert und repräsentiert eine Berechtigung zum Zugriff auf verschiedene Funktionen der Schnittstelle. Dritter Schritt ist das Einlösen eines Tokens; Create, Read, Update und Delete (CRUD)

[14] *Patient* ist ein definierter Datentyp der Mainzelliste.

sind die Operationen, mit denen der Datensatz eines Patienten oder einer Patientin angelegt, gelesen, aktualisiert oder gelöscht werden kann. Die vollständige Beschreibung der Ressourcen sowie die durch die Schnittstelle angebotenen Methoden sind der Schnittstellenbeschreibung[15] der Mainzelliste zu entnehmen.

Für die Implementierung der Schnittstelle der Mainzelliste zum Probandenmanagementsystem wurden verschiedene Klassen implementiert, die Datenobjekte und Ressourcen der Mainzelliste repräsentieren und bei der Interaktion mit der Schnittstelle unterstützten (UML Klassendiagramme: Abbildung 4.19, Abbildung 4.20 und Abbildung 4.21).

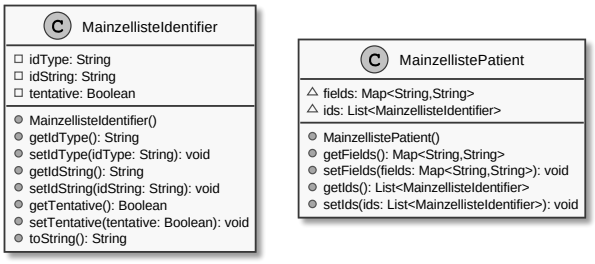

Abbildung 4.19: UML Klassendiagramm der Klassen *MainzellisteIdentifier* und *MainzellistePatient* die Teil der Schnittstellenimplementierung der Mainzelliste zum Probandenmanagementsystem sind.

Die Klasse *MainzellisteIdentifier* repräsentiert ein Pseudonym der Mainzelliste mit den Attributen: *idType* dem Typ des Pseudonyms, *idString* dem Wert des Pseudonyms und *tentative* der Angabe, ob das Pseudonym bei Feststellung eines unsicheren Matches als vorläufig zu betrachten ist. Die Klasse *MainzellistePatient* repräsentiert einen Patienten oder eine Patientin der Mainzelliste mit den Attributen: *fields* einer Liste von identifizierenden Daten eines Patienten oder eine Patientin als Schlüssel-Wert-Paare und *ids* einer Liste von *MainzellisteIdentifier*, den einem Patienten oder einer Patientin zugewiesenen Pseudonymen.

[15] Schnittstellendokumentation Mainzelliste (Version 3.1.0): `https://bitbucket.org/medicalinformatics/mainzelliste/downloads/Mainzelliste_Schnittstelle_v3.1.0.pdf` [Stand: 01.03.2022]

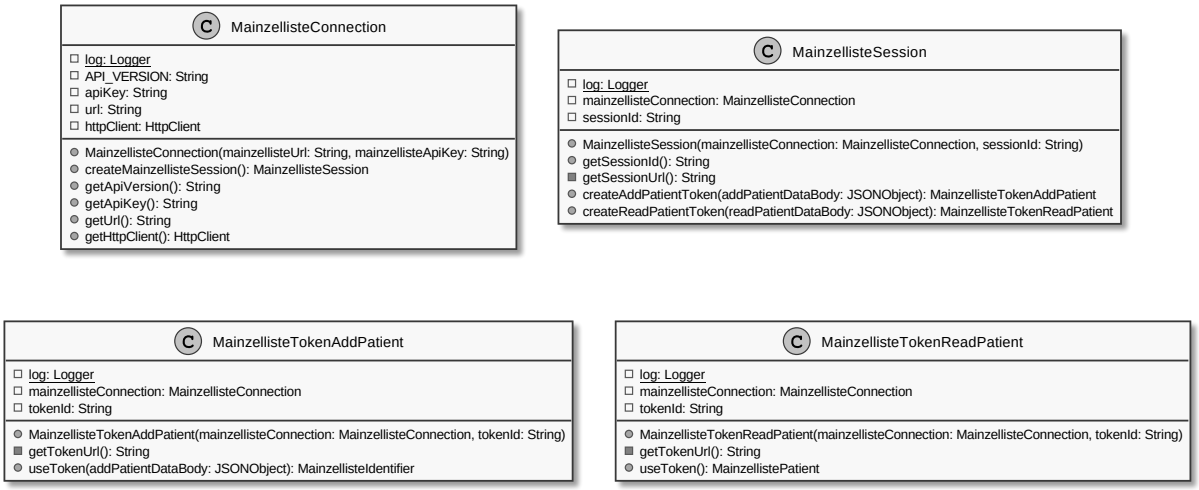

Abbildung 4.20: UML Klassendiagramm der Klassen der Schnittstellenimplementierung des Probandenmanagementsystems die den Kommunikationsablauf mit der Mainzelliste steuern.

Die Klassen *MainzellisteConnection*, *MainzellisteSession*, *MainzellisteTokenAddPatient* und *MainzellisteTokenReadPatient* dienen der Kommunikation mit der Mainzelliste. *MainzellisteConnection* etabliert eine Verbindung zur Mainzelliste und erstellt eine Session für den weiteren Kommunikationsablauf. Die Klasse *MainzellisteSession* repräsentiert eine Session und ist in der Lage zwei Arten von Tokens abzurufen, die für die Kommunikation mit dem Probandenmanagementsystem relevant sind. Ein Token vom Typ *AddPatient* wird durch die Klasse *MainzellisteTokenAddPatient* repräsentiert und ein Token vom Typ *ReadPatient* durch die Klasse *MainzellisteTokenReadPatient*. Das Einlösen der Token über die Methode *useToken* liefert entweder ein Pseudonym vom Typ der Klasse *MainzellisteIdentifier* oder angefragte Daten der Patientin oder des Patienten vom Typ der Klasse *MainzellistePatient* zurück. Für den Anwendungsfall der Pseudonymabfrage sind die angefragten Daten der Klasse *MainzellistePatient* immer die Liste von weiteren Pseudonymen einer Patientin oder eines Patienten (Variable *ids*) aus einer Liste von *MainzellisteIdentifier* (vgl. Klassendiagramm Abbildung 4.19).

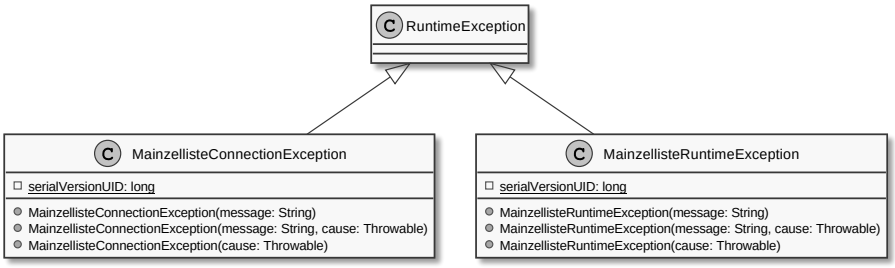

Abbildung 4.21: UML Klassendiagramm der Klassen *MainzellisteConnectionException* und *MainzellisteRuntimeException* die von *RuntimeException* erben, zur Fehlerbehandlung bei der Kommunikation zwischen Probandenmanagementsystem und Mainzelliste.

Die Klassen *MainzellisteConnectionException* und *MainzellisteRuntimeException* erben von der Oberklasse *RuntimeException* und dienen der Ausnahmebehandlung in Bezug auf die Kommunikation mit der Schnittstelle der Mainzelliste im Probandenmanagementsystem.
Alle zuvor genannten Klassen unterstützen den *MainzellisteAdapter* in seiner Funktion als Schnittstellenimplementierung der Mainzelliste im Probandenmanagementsystem. Die Klasse *MainzellisteAdapter* implementiert das Interface *ExternalSystemProvider* und setzt die beiden vom Interface geforderten Methoden *assignNewIdentifier(...)* und *retrieveIdentifier(...)* für die Schnittstelle der Mainzelliste um (siehe Abbildung 4.22).

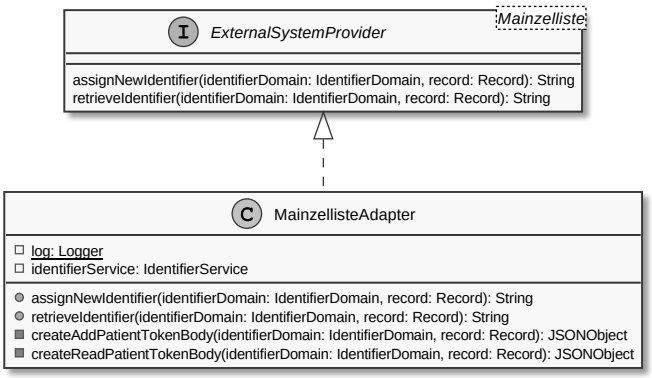

Abbildung 4.22: UML Klassendiagramm der Klasse *MainzellisteAdapter* und dem Interface *ExternalSystemProvider*. *MainzellisteAdapter* implementiert über die Methoden *assignNewIdentifier(...)* und *retrieveIdentifier(...)* Zugang zur Schnittstelle der Mainzelliste, wie sie für die im Probandenmanagementsystem für die Anwendungsfälle der Pseudonymzuweisung und –abfrage benötigt werden.

Die Mainzelliste unterstützt alle für das Konzept dieser Arbeit benötigten Funktionen: Eine Pseudonymisierung erster Stufe, die Verwaltung multipler Pseudonyme, sowie das

Auflösen von Pseudonymen. Voraussetzung für die Nutzung der Mainzelliste ist eine Konfiguration der Mainzelliste (siehe auch Konfigurationshandbuch[16] der Mainzelliste) die alle Systeme eines Forschungsprojektes einbezieht. Mit der vorgestellten prototypischen Implementierung der Schnittstelle zur Mainzelliste konnte gezeigt werden, dass die Mainzelliste ein geeigneter Kandidat für standortübergreifende Forschungsprojekte mit Verwendung eines Probandenmanagementsystems ist.

4.3 Zugangs- und Zugriffskontrolle

Der dritte Teil dieses Kapitels beschreibt die prototypische Umsetzung der Zugangs- und Zugriffskontrolle für das Probandenmanagementsystem, dessen Konzept in Iteration 3 (Unterkapitel 3.5, S. 63) erarbeitet wurde. Mit dem Ziel einen vertraulichen Zugang und Zugriff auf die Probandendaten sicherzustellen wurden fünf Anforderungen benannt und verschiedene Maßnahmen zur Umsetzung konzeptioniert. Diese dienen dem folgenden Teil des Implementierungskapitels zur Gliederung der Unterkapitel. Erste Anforderung und damit das erste Unterkapitel behandelt die Umsetzung einer sicheren Authentifizierung am Probandenmanagementsystem. Es folgt im zweiten Unterkapitel die Implementierung der benötigten verschlüsselten Verbindung zum Zugriff auf das Probandenmanagementsystem. Das dritte Unterkapitel dient zur Vorstellung der prototypischen Implementierung der Zugriffskontrolle, die gleichzeitig das Berechtigungskonzept umsetzt. Abschließend wird der Anforderung eines protokollierten Zugriffs nachgekommen und die Implementierung der konzeptionierten Maßnahmen beschrieben.

4.3.1 Sichere Authentifizierung

Für eine verbesserte Sicherheit des Anmeldeverfahrens des Probandenmanagementsystems wurden verschiedene Maßnahmen zur Umsetzung konzeptioniert. Die folgenden Abschnitte beschreiben deren Implementierung, beginnend mit der Zugriffsbeschränkung auf bestimmte IP-Adressen, Maßnahmen

[16] Konfigurationshandbuch Mainzelliste (Version 1.8): `https://bitbucket.org/medicalinformatics/mainzelliste/downloads/Mainzelliste__Konfigurationshandbuch_v1.8.pdf` [Stand: 01.03.2022]

zur Verbesserung der Benutzerauthentifizierung und abschließend der Authentifizierung mit Clientzertifikaten.

IP-Adressen-Zugriffsbeschränkung

Grundsätzlich ist nur autorisierten IP-Adressen ein Verbindungsaufbau zum Probandenmanagementsystem erlaubt. IP-Adressen beziehungsweise IP-Adressräume der Nutzerinnen und Nutzer des Probandenmanagementsystems sind daher von der Administratorin oder dem Administrator des Probandenmanagementsystems in einer Konfigurationsdatei des Tomcat Webservers über einen *Valve Component*[17] anzugeben.

```
<Valve className="org.apache.catalina.valves.RemoteAddrValve"
allow="X.X.X.X,Y.Y.Y.Y,Z.Z.Z.Z" />
```

Listing 4.3: Parameter der Tomcat-Konfigurationsdatei /conf/server.xml um bestimmte IP-Adressen zulassen. X, Y, und Z, sind als Beispiele für beliebige Zahlen in einer IP-Adresse zu sehen.

Listing 4.4 zeigt die Einfachheit, mit der eine IP-Adressen-Zugriffsbeschränkung bei einem Apache Tomcat Webserver umzusetzen ist.

Benutzerauthentifizierung

Die Benutzerauthentifizierung des Probandenmanagementsystems besteht aus einem Login des Systembenutzenden über die Eingabe eines Benutzernamens und eines Passwortes. Sobald die Authentifizierung erfolgt ist, ist die Identität des Systembenutzenden bekannt und die Autorisierung auf weitere Funktionen des Probandenmanagementsystems kann erfolgen. Damit ist die Benutzerauthentifizierung wichtige Voraussetzung für die Implementierung der Zugriffskontrolle (Unterkapitel 4.3.3, S. 132) und des protokollierten Zugriffs (Unterkapitel 4.3.4, S. 140).
Der Login des Probandenmanagementsystems wurde mit Spring Security[18] umgesetzt. Spring Security ist ein Java-Framework, das Unterstützung bei Authentifizierung und Autorisierung sowie weiteren IT-Security-Maßnahmen bietet und den De-facto-Standard

[17] The Valve Component, Apache Tomcat Configuration Reference: `https://tomcat.apache.org/tomcat-9.0-doc/config/valve.html` [Stand: 01.03.2022]

[18] Spring Security: `https://docs.spring.io/spring-security/reference/` [Stand: 01.03.2022]

für die Absicherung Spring-basierter Anwendungen bildet. Gerade in Bezug auf Kryptografie ist es äußerst wichtig, auf ausgereifte und getestete Implementierungen und Algorithmen zu setzen [vgl. Rohr, 2018, S. 149].
Den Maßnahmen zur sicheren Authentifizierung des Konzeptes in Iteration 3 (Unterkapitel 3.5.3, Maßnahmen zur sicheren Authentifizierung, S. 72) folgend, muss das Probandenmanagementsystem die Passwortstärke eines durch den Systembenutzenden gewählten Passwortes sicherstellen. Das Passwort muss mindestens 8 Zeichen lang, Groß- und Kleinbuchstaben sowie Ziffern enthalten. Um die Passwortregeln nicht manuell implementieren zu müssen, wurde auf die Open Source-Bibliothek Passay[19] zurückgegriffen. Listing 4.4 zeigt einen Auszug des Programmcodes zur Überprüfung der Passwortstärke bei Anlage oder Änderung eines Passwortes im Probandenmanagementsystem.

```
1 ...
2     PasswordValidator validator = new PasswordValidator(Arrays.asList(
3     new LengthRule(8, 255),
4     new UppercaseCharacterRule(1),
5     new LowercaseCharacterRule(1),
6     new DigitCharacterRule(1)));
7 ...
```

Listing 4.4: Auszug des Programmcodes zur Überprüfung der Passwortstärke durch einen PasswordValidator mittels verschiedener Regeln der Passay-Bibliothek.

LengthRule, *UppercaseCharacterRule*, *LowercaseCharacterRule* und *DigitCharacterRule* bilden die Zusammenstellung der Passay-Regeln, die mit den gesetzten Paramenten die oben genannten Passwortregeln des Probandenmanagementsystems validieren. Bei Anlage oder Änderung eines Passwortes im Probandenmanagementsystem über die Benutzeroberfläche des Probandenmanagementsystems werden die Systembenutzenden über entsprechende Hinweise zur Vergabe eines besseren Passwortes gezwungen.
Da das Probandenmanagementsystem eine passwortbasierte Authentifizierung verwendet, müssen die Passwörter der Benutzenden im Probandenmanagementsystem verwaltet werden. Daher ist eine weitere umzusetzende Maßnahme des Konzeptes, die Passwörter sicher in der Datenbank des Probandenmanagementsystems zu verwahren (vgl.

[19] Passay - Password policy enforcement for Java: `http://www.passay.org/` [Stand: 01.03.2022]

Unterkapitel 3.5.3, Maßnahmen zur sicheren Authentifizierung, S. 72). Die Passwörter werden nicht im Klartext gespeichert, sondern beim Anlegen eines Benutzerkontos gehasht und der Hash in der Datenbank mit den Benutzerdaten gespeichert. Bei Anmeldung eines Benutzenden wird sein dabei eingegebenes Passwort gehasht und mit dem gespeicherten Hash verglichen, um den Benutzenden zu authentifizieren. Zusätzlich empfiehlt sich die Verwendung eines sogenannten Salts. Ein Passwort wird nicht mehr direkt gehasht, sondern zusammen mit dem Salt (eine willkürlich festgelegte Zahlen- und Ziffernkombination) in die Hashfunktion eingegeben. Identische Passwörter, die entsprechend denselben Hash hätten, werden durch den Salt individualisiert. Die Verwendung eines Salts verhindert so für bekannte Hashfunktionen vorbereitete Hashlisten, die die häufigsten Passwörter und die dazugehörigen Hashes abbilden, zur Entschlüsselung zu verwenden.
Das Spring Security Framework, das für das Probandenmanagementsystem verwendet wird, beinhaltet verschieden Implementierungen von Passwort-Hashfunktionen, mit denen Passwörter sicher gespeichert werden können. Ausgewählt zum Hashing der Passwörter wurde für das Probandenmanagementsystem der weitverbreitete bcrypt-Algorithmus, implementiert in der Klasse *BCryptPasswordEncoder*[20] des Spring Security Crypto Packages. Listing 4.5 zeigt einen Auszug des Programmcodes des Probandenmanagementsystems zur Umwandlung des Klartext-Passwortes in ein gehashtes Passwort, das anschließend in der Datenbank des Probandenmanagementsystems gespeichert wird.

```
1 ...
2 int strength = 10; // the log rounds in bcrypt
3 BCryptPasswordEncoder bCryptPasswordEncoder = new BCryptPasswordEncoder(strength,
4 new SecureRandom());
5 String encodedPassword = bCryptPasswordEncoder.encode(plainPassword);
6 ...
```

Listing 4.5: Auszug des Programmcodes zur Erzeugung von gehashten Passwörtern zur Ablage in der Datenbank des Probandenmanagementsystems.

[20] Spring Security - BCryptPasswordEncoder: `https://docs.spring.io/spring-security/site/docs/current/api/org/springframework/security/crypto/bcrypt/BCryptPasswordEncoder.html` [Stand: 01.03.2022]

Der Konstruktor der Klasse *BCryptPasswordEncoder* bekommt als ersten Parameter *strength* die Anzahl an Iterationen zur Berechnung des Hashes. Der Standardwert in Spring Security ist 10. Zweiter Parameter ist der Salt-Generator *SecureRandom* der eine kryptografisch starke Zufallszahl generiert.

Authentifizierung mit Clientzertifikat

Zur Umsetzung der Zwei-Faktor-Authentifizierung wurde als zweiter Faktor neben der Benutzerkennung und Passwort, das Softwarezertifikat zur Clientauthentifizierung gewählt (vgl. Unterkapitel 3.5.3, Anforderung A1: Sichere Authentifizierung, S. 72). Zum Erstellen der Clientzertifikate wird ein Wurzelzertifikat benötigt. Während beim SSL-Server-Zertifikat (vgl. Unterkapitel 4.3.2, Verschlüsselte Verbindung, S. 130) das Zertifikat durch die Zertifizierungsstelle des DFN-Vereins unterschrieben werden musste, kann das hier benötigte Wurzelzertifikat selbst unterschrieben werden („selfsigned“). In diesem Fall wird das Zertifikat nicht zum Authentifizieren des Webservers und Verschlüsseln der Verbindung verwendet, sondern zum Ausstellen weiterer Clientzertifikate zur Authentifizierung der Benutzer/Clients. Durch die eigene Verwaltung des Wurzelzertifikats (= eigene *Certificate Authority*) kann der Administrator oder die Administratorin des Probandenmanagementsystems eigenständig beliebig viele Clientzertifikate erstellen und verwalten. Sowohl das Wurzelzertifikat als auch Clientzertifikate zu Testzwecken wurden mit OpenSSL[21] auf dem Probandenmanagementsystemserver erstellt.
Die Konfiguration für die Clientauthentifizierung ist nicht komplexer als die Konfiguration des SSL-Server-Zertifikates. Auch hier wurde auf dem Probandenmanagementsystemserver ein Keystore erstellt, der mit dem Wurzelzertifikat sowie den ausgestellten Clientzertifikaten befüllt wurde. Der Keystore wurde anschließend zur Verwendung durch den Tomcat Webserver konfiguriert (siehe Listing 4.6).

[21] OpenSSL: `https://www.openssl.org` [Stand: 01.03.2022]

```
[...]
clientAuth="true"
truststoreType="JKS"
truststoreFile="/opt/tomcat/conf/keystore/server-truststore.jks"
truststorePass="*"
[...]
```

Listing 4.6: Parameter der Tomcat-Konfigurationsdatei /conf/server.xml um Clientauthentifizierung zu aktivieren.

Nach der Konfiguration wurde die Funktion über den in Listing 4.7 angegebenen Kommandozeilenbefehl überprüft. Der Befehlt baut eine verschlüsselte Verbindung auf und listet die Zertifikatseinstellungen auf. Listing 4.8 zeigt nach Ausführen des Befehls auf dem Probandenmanagementsystemserver, die Liste der akzeptierten Zertifikate mit ihren Parametern an.

```
openssl s_client -connect localhost:443 -showcerts
```

Listing 4.7: Kommandozeilenbefehl zum Kontrollieren der Zertifikatskonfiguration des Webservers.

```
[...]
Acceptable client certificate CA names
/C=DE/ST=NI/L=Goettingen/O=Universitaetsmedizin Goettingen
/OU=Institut fuer Medizinische Informatik/CN=Probandenmanagementsystem
/CN=Johannes Pung/OU=MI/O=UMG/L=Goettingen/ST=NI/C=DE
/emailAddress=johannes.pung@med.uni-goettingen.de
Client Certificate Types: RSA sign, DSA sign, ECDSA sign
Requested Signature Algorithms: ECDSA+SHA512:RSA+SHA512:ECDSA+SHA384:RSA+SHA384:ECDSA
+SHA256:RSA+SHA256:DSA+SHA256:ECDSA+SHA224:RSA+SHA224:DSA+SHA224:ECDSA+SHA1:RSA+SHA1:DSA
+SHA1
Shared Requested Signature Algorithms: ECDSA+SHA512:RSA+SHA512:ECDSA+SHA384:RSA
+SHA384:ECDSA+SHA256:RSA+SHA256:DSA+SHA256:ECDSA+SHA224:RSA+SHA224:DSA
+SHA224:ECDSA+SHA1:RSA+SHA1:DSA+SHA1
Peer signing digest: SHA512
Server Temp Key: ECDH, P-256, 256 bits
[...]
```

Listing 4.8: Ausgabe des Befehls aus Listing 4.7 zur Kontrolle der Zertifikatskonfiguration

Wird das Probandenmanagementsystem im Webbrowser aufgerufen, so schlägt die Verbindung vorerst fehl. Der Webserver antwortet auf die Anfrage mit einem *TLS alert, bad certificate* Fehler. Erst nach dem Import des für die Benutzerin oder den Benutzer ausgestellten Clientzertifikates im Webbrowser kann eine Verbindung hergestellt werden. Bei der Übergabe der Clientzertifikate an die Benutzerin oder den Benutzer ist darauf zu achten, einen sicheren Weg zur Übermittlung zu wählen und das Zertifikat mit einem Passwort zu schützen.

4.3.2 Verschlüsselte Verbindung

Die umzusetzenden Maßnahmen des erarbeiteten Konzeptes zur *Anforderung A2: Verschlüsselte Verbindung* lassen sich wie folgt zusammenfassen (vgl. Unterkapitel 3.5.3, S. 76):

1. Über HSTS erzwungene HTTPS Verbindung
2. SSL-Server-Zertifikat über den DFN-PKI mit RSA-Schlüssel (4096 Bit)
3. Ausschließlich Verschlüsselungsprotokolle TLS 1.2 oder TLS 1.3
4. Überprüfung der Einstellungen mit SSL Report von Qualys SSL Labs

In einem ersten Schritt wurde das SSL-Server-Zertifikat erstellt. Dabei wurde entsprechend der Anleitung für Zertifikatsanträge[22] der DFN-PKI vorgegangen. Nach der Installation der Software OpenSSL[23] auf dem Probandenmanagementsystemserver, wurde der private RSA-Schlüssel sowie der *Certificate Signing Request* (CSR) zur Beantragung des Zertifikates beim DFN-PKI generiert (siehe Listing 4.9).

```
1 openssl req -newkey rsa:4096 -sha256 -keyout priv-key.pem -out certreq.pem
```

Listing 4.9: Kommandozeilenbefehl zum Erzeugen des CSR und dem privaten RSA-Schlüssel mit OpenSSL.

[22] DFN-PKI: Anleitung Webschnittstelle Nutzer und Administratoren: `https://www.pki.dfn.de/fileadmin/PKI/Anleitung_DFN-Test-PKI.pdf` [Stand: 01.03.2022]

[23] OpenSSL: `https://www.openssl.org` [Stand: 01.03.2022]

Die Schlüssellänge beträgt wie konzeptioniert 4096 Bit. Entsprechend der Zertifizierungsrichtlinie[24] der DFN-PKI wurden alle Zertifikatsanforderungsangaben gemacht und an den DFN-PKI übermittelt. Nach Prüfung durch den sogenannten Teilnehmerservice konnte das Zertifikat ausgestellt werden. Mit dem Zertifikat sowie den abhängigen Wurzelzertifikaten wurde auf dem Probandenmanagementsystemserver ein Keystore erstellt. Der Keystore wurde anschließend zur Verwendung durch den Tomcat Webserver konfiguriert.
Die Verschlüsselungsprotokolle wurden ausschließlich auf TLS 1.2 oder TLS 1.3 eingeschränkt (siehe Listing 4.10).

```
sslProtocol="TLS" sslEnabledProtocols="TLSv1.2+TLSv1.3"
```

Listing 4.10: Parameter der Tomcat-Konfigurationsdatei /conf/server.xml um auf die TLS Versionen 1.2 und 1.3 einzuschränken.

Um die Sicherheit weiter zu erhöhen, wurde die Verwendung des HSTS-Headers aktiviert. HSTS ermöglicht, dem Webbrowser mitzuteilen, dass immer über HTTPS zugegriffen werden muss (siehe Listing 4.11 und Listing 4.12).

```
<security-constraint>
    <web-resource-collection>
        <web-resource-name>pms</web-resource-name>
        <url-pattern>/*</url-pattern>
    </web-resource-collection>
    <user-data-constraint>
        <transport-guarantee>CONFIDENTIAL</transport-guarantee>
    </user-data-constraint>
</security-constraint>
```

Listing 4.11: Parameter der Tomcat-Konfigurationsdatei WEB-INF/web.xml um den Datenverkehr auf eine sicheren Verbindung zu erzwingen.

[24] DFN-PKI: Zertifizierungsrichtlinie: `https://www.pki.dfn.de/fileadmin/PKI/DFN-PKI_CP.pdf` [Stand: 01.03.2022]

```
<filter>
    <filter-name>httpHeaderSecurity</filter-name>
    <filter-class>org.apache.catalina.filters.HttpHeaderSecurityFilter</filter-class>
    <async-supported>true</async-supported>
</filter>
[...]
<filter-mapping>
    <filter-name>httpHeaderSecurity</filter-name>
    <url-pattern>/*</url-pattern>
    <dispatcher>REQUEST</dispatcher>
</filter-mapping>
```

Listing 4.12: Parameter der Tomcat-Konfigurationsdatei /conf/web.xml um HSTS zu aktivieren.

Die Prüfung der Konfiguration wurde, wie konzeptioniert, mit dem frei verfügbaren Online-Scanner SSL Report von Qualys SSL Labs[25] durchgeführt. Die zu diesem Zeitpunkt von außen überprüfbaren Einstellungen erhielten durch den Report[26] die bestmögliche Bewertung von A+.

4.3.3 Zugriffskontrolle

Die Verwendung ausgereifter und gepflegter Bibliotheken von vertrauenswürdigen Verfassern ist von großer Wichtigkeit, gerade wenn es um die Sicherheitsfunktionen einer Anwendung geht. Für Anwendungen, die sich auf das Spring Framework stützen, ist für die Implementierung der Sicherheit auf Anwendungsebene Spring Security[27] eine gute Wahl. Das Security-Framework wurde daher für die Implementierung der Zugriffskontrolle im Probandenmanagementsystem gewählt. Die Verwendung des Frameworks wurde bereits im vorangehenden Unterkapitel erwähnt (vgl. Unterkapitel 4.3.1, Benutzerauthentifizierung, S. 125), denn die Authentifizierung ist ein wichtiger Teil der Zugriffskontrolle. Das Probandenmanagementsystem prüft die Identität registrierter Benutzerinnen und Benutzer, bevor der Zugriff auf geschützte Ressourcen erlaubt wird. In welcher Form Zugriff gewährt wird, wurde durch das Zugriffskontrollkonzept und die damit verbundene Auswahl des Zugriffskontrollmodells im

[25] Qualys, Inc. - SSL Server Test: `https://www.ssllabs.com/ssltest/` [Stand: 01.03.2022]

[26] SSL Server Rating Guide: `https://www.ssllabs.com/projects/rating-guide/index.html` [Stand: 01.03.2022]

[27] Spring Security: `https://docs.spring.io/spring-security/reference/` [Stand: 01.03.2022]

Probandenmanagementsystem definiert (vgl. Unterkapitel 3.5.3, Iteration 3, Anforderung A4: Zugriffskontrollmodell, S. 84).

Die Zugriffssteuerung im Probandenmanagementsystem erfolgt zum einen über Rollen, an die verschiedene Zugriffsrechte geknüpft sind, zum anderen über die Attribute *Forschungsprojekt* und *Studienstandort*. Diese Attribute, die zur Zugriffssteuerung jedem Systembenutzenden zugewiesen werden, sind auch Attribute einer jeden Forschungsakte im Probandenmanagementsystem. Jede Forschungsakte eines Probanden gehört zu einem Forschungsprojekt (vgl. auch Unterkapitel 4.2.1, Datenmodell für Probandenidentifikatoren, S. 112) und gleichzeitig zu einem Studienstandort. In Abbildung 4.23 wird der Zusammenhang zwischen Benutzern, Rollen und Rechten sowie Forschungsprojekten, Studienstandorten und Forschungsakten verdeutlicht.

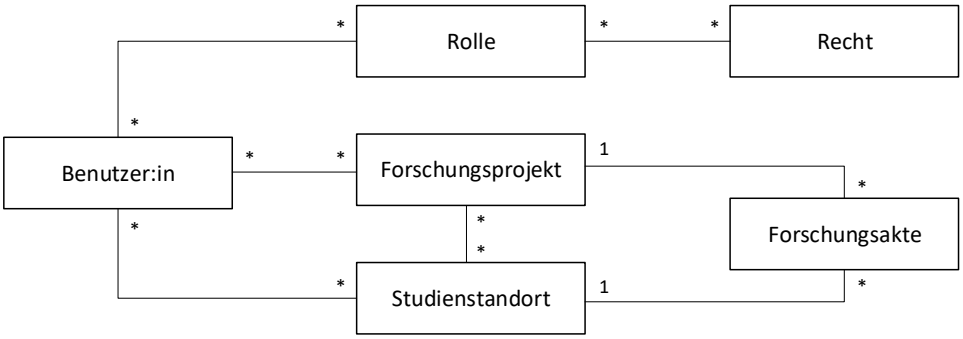

Abbildung 4.23: Konzeptuelles Datenmodell zur Darstellung der Beziehungen zwischen Benutzern, Rollen und Rechten sowie Forschungsprojekten, Studienstandorten und Forschungsakten im Probandenmanagementsystem.

Dieses konzeptuelle Datenmodell überführt in das objektrelationale Datenmodell des Probandenmanagementsystems ist in Abbildung 4.24 dargestellt.

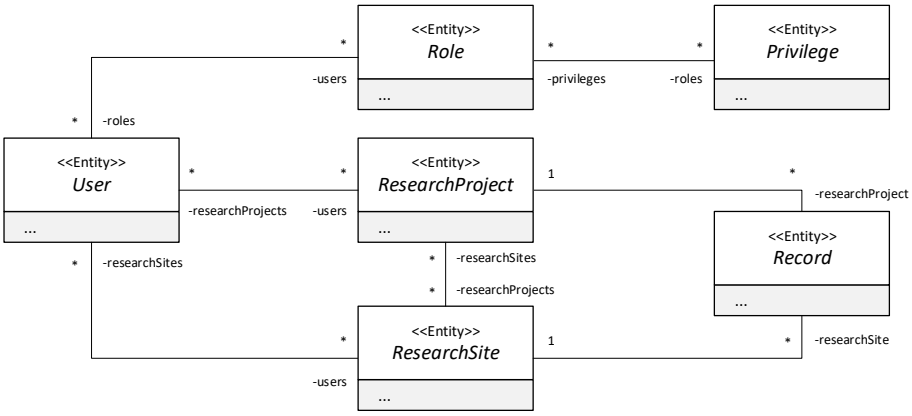

Abbildung 4.24: UML Klassendiagramm der Entitäten User, Role, Privilege, ResearchProject, ResearchSite und Record und ihre Beziehungen zueinander. Sie bilden die Entitäten, die zur Abbildung der Daten zur Zugriffskontrolle im Probandenmanagementsystem benötigt werden.

Zentrale Figur der Zugriffssteuerung ist die oder der Systembenutzende. Die Entität *User* beschreibt die objektrelationale Abbildung Systembenutzender in die relationale Datenbank. Im Kontext der Implementierung der Zugriffssteuerung mit Spring Security wurde zur Trennung der Logik von der Datenhaltung die Klasse *SecurityUser* erstellt. Die Klasse implementiert das Spring Security Interface *UserDetails*[28], um Systembenutzende in die Spring Security-Architektur zu integrieren. *SecurityUser* implementiert somit alle erforderlichen Methoden des Security-Frameworks zur Beschreibung eines Systembenutzenden (vgl. Abbildung 4.25) und nutzt die bestehende Entität *User*, ohne Programmcode der JPA-Entität mit securitybezogenem Code zu mischen.

[28] Spring Security - UserDetails: `https://docs.spring.io/spring-security/site/docs/current/api/org/springframework/security/core/userdetails/UserDetails.html` [Stand: 01.03.2022]

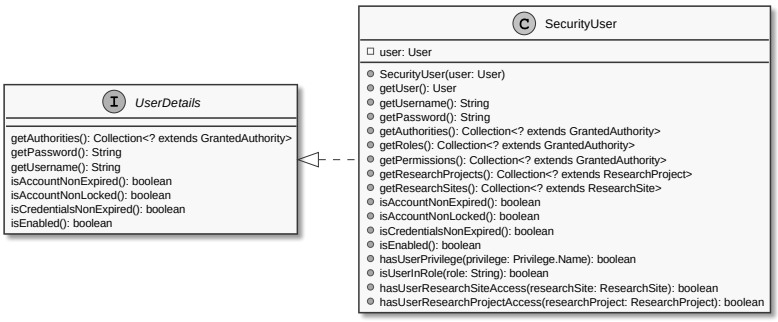

Abbildung 4.25: UML-Klassendiagramm des Interface *UserDetails* und der Klasse *SecurityUser*. *SecurityUser* implementiert verschiedene Methoden, um Eigenschaften eines Systembenutzenden des Probandenmanagementsystems abzufragen.

Eine weitere essenzielle Klasse, die zur Nutzung von Spring Security implementiert wurde, ist das Interface *UserDetailsService*[29]. Ein Objekt der Klasse, welche das Interface implementiert, verwaltet die Details von Systembenutzenden, die über die Klasse *UserRepository* geladen werden. Während des Authentifizierungsprozesses erhält der *UserDetailsService* alle Details über den Systembenutzenden, einschließlich der Berechtigungen (Interface *GrantedAuthority*). Abbildung 4.26 zeigt die Beziehung zwischen *UserDetailsService*, *UserDetails* und *GrantedAuthority*.

Abbildung 4.26: Beziehung zwischen dem Interface *UserDetailsService*, *UserDetails* und *GrantedAuthority* als UML-Klassendiagramm. Der *UserDetailsService* bezieht die Details der Systembenutzenden des Probandenmanagementsystems während des Authentifizierungsprozesses. Systembenutzenden haben eine oder mehrere Berechtigungen.

Zusammen mit einem *PasswordEncoder* (vgl. Unterkapitel 4.3.1, Benutzerauthentifizierung, Listing 4.5, S. 127) ist der *AuthenticationProviderService* in der Lage, bei einem Login von Systembenutzenden ins Probandenmanagementsystem eine Authentifizierung durchzuführen und ein Objekt der Klasse *SecurityContext* zu befüllen. Diese vorgestellten Interfaces und Klassen bilden die Hauptkomponenten eines Authentifizierungsprozesses mit Spring Security (siehe Abbildung 4.27, ↦ Login).

[29] Spring Security - UserDetailsService: `https://docs.spring.io/spring-security/site/docs/current/api/org/springframework/security/core/userdetails/UserDetailsService.html` [Stand: 01.03.2022]

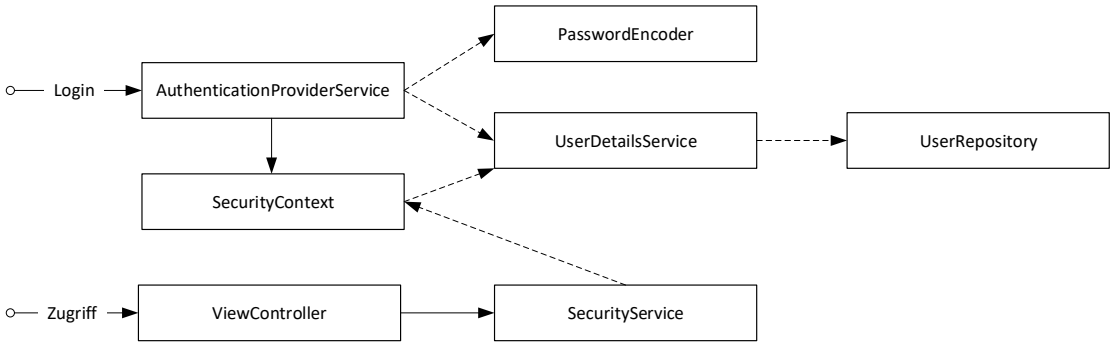

Abbildung 4.27: Hauptkomponenten des Authentifizierungsprozesses mit Spring Security und ihre Beziehungen zueinander. Diese Architektur stellt das Rückgrat der Implementierung der Authentifizierung mit Spring Security dar. Eigene erweiterte Abbildung angelehnt an [Spilca, 2020, S. 39].

Ein Objekt der Klasse *SecurityContext* behält die Authentifizierungsdaten auch nach dem Authentifizierungsprozess. Auf diesen greift der *SecurityService* zu, ein Service, der Informationen über eingeloggte Systembenutzende und ihre Rechte und Rollen für weitere Komponenten zur Verfügung stellt. Zu diesen Komponenten gehören die verschiedenen Ansichten der Benutzeroberfläche (*Views*) des Probandenmanagements (siehe Abbildung 4.27, $\mapsto$ Zugriff).
Spring Security bietet mehrere Möglichkeiten zur Prüfung von Zugriffsrechten und Benutzerrollen im Programmcode. Die verschiedenen Varianten wurden je nach Anforderung der Implementierung im Probandenmanagementsystem angewandt. Listing 4.13, Listing 4.14 und Listing 4.15 zeigen Beispiele für die Verwendung von Spring Security Annotationen im Programmcode des Probandenmanagementsystems.

```
1 @RolesAllowed("ROLE_DATAMANAGER")
2 public Record getRecord(Integer recordId) {
3     // Abruf einer Forschungsakte anhand der übergebenen ID
4 }
```

Listing 4.13: Beispiel für die Nutzung der Spring Security Annotation @*RolesAllowed*. Zugriff auf die Methode wird nur gewährt, wenn die oder der aktuell eingeloggte Systembenutzende die Rolle *DATAMANAGER* besitzt.

Für den Fall, dass komplexere Berechtigungen geprüft werden müssen, eignen sich Spring Security Expression[30]-basierte Annotationen. Sie akzeptieren einen Ausdruck der Spring Expression Language[31] (SpEL). Listing 4.14 zeigt ein Beispiel zur Überprüfung eines einzelnen Zugriffsrechtes über die Spring Security Built-In Expression[32] *hasPermission(...)*.

```
@PreAuthorize("hasPermission(#recordId, 'OPEN_RECORD')")
public Record getRecord(Integer recordId) {
    // Abruf einer Forschungsakte anhand der übergebenen ID
}
```

Listing 4.14: Beispiel für die Nutzung der *@PreAuthorize* Spring Security Annotation mit der *hasPermission(...)* Expression. Der Ausdruck überprüft die Berechtigung *OPEN_RECORD* von Systembenutzenden zum Abrufen von Forschungsakten im Probandenmanagementsystem.

Die Annotation *@PreAuthorize* entscheidet entsprechend der Prüfung des Ausdrucks, ob die Methode aufgerufen werden darf oder nicht. Im Hintergrund prüft ein eigens implementierter *PermissionEvaluator*[33] die Berechtigung. Bei der Autorisierungsentscheidung werden auch Übergabeparameter der Methodenaufrufe berücksichtigt.

Neben den durch Spring Security vorgegebenen Ausdrücken lassen sich auch eigene Ausdrücke definieren. Durch den selbst festgelegten Namen des Ausdrucks lässt sich die Bedeutung der zugrunde liegenden Zugriffsprüfung individueller festlegen. Listing 4.15 zeigt ein Beispiel für den eigenen Ausdruck *isProjectMember(...)*, der die Zugehörigkeit von Systembenutzenden zu einem Forschungsprojekt im Probandenmanagementsystem prüft.

---

[30] Spring Security - Expression-Based Access Control: `https://docs.spring.io/spring-security/reference/servlet/authorization/expression-based.html` [Stand: 01.03.2022]

[31] Spring Expression Language (SpEL): `https://docs.spring.io/spring-framework/docs/current/reference/html/core.html#expressions` [Stand: 01.03.2022]

[32] Spring Security - Common Built-In Expressions: `https://docs.spring.io/spring-security/reference/servlet/authorization/expression-based.html#el-common-built-in` [Stand: 01.03.2022]

[33] Spring Security - PermissionEvaluator: `https://docs.spring.io/spring-security/reference/servlet/authorization/expression-based.html#el-permission-evaluator` [Stand: 01.03.2022]

```
@PreAuthorize("isProjectMember(#researchProject"))
public void editProjectName (ResearchProject researchProject) {
    // Ändert den Namen eines Projektes
}
```

Listing 4.15: Beispiel für die Nutzung eines selbst definierten SpEL-Ausdrucks mit der der @*PreAuthorize* Spring Security Annotation. *isProjectMember(#researchProject)* prüft die Zugehörigkeit von Systembenutzenden zu einem Forschungsprojekt bevor die Änderungen am Projektnamen akzeptiert werden.

Allen Autorisierungsmechanismen in Spring Security ist gemein, dass sie sich auf das zentrale *SecurityContext*-Objekt stützen, das die sicherheitsspezifischen Informationen über die Systembenutzenden enthält, die vom Probandenmanagementsystem authentifiziert wurden. Um diesen Kontext auch ohne die zuvor beschriebenen Berechtigungsprüfungen mittels Annotationen zu nutzen, denn dieses Vorgehen ist nicht für alle Zugriffsabfragen im Probandenmanagementsystem geeignet, wurde ein zentraler Service für den Zugriff auf sicherheitsrelevante Funktionen im Probandenmanagementsystem implementiert. Die Klasse *SecurityService* kann in andere Klassen injiziert werden und liefert über den *SecurityContext* die oder den eingeloggten Systembenutzenden mittels eines *SecurityUser*-Objekts. Die verschiedenen Services des Probandenmanagementsystems, die die Anwendungslogik hinter der grafischen Oberfläche des Systems steuern, erhalten so die Möglichkeit, die Zugriffssteuerung in ihre Logik zu implementieren. Beispielsweise der *ResearchRecordService* implementiert wichtige Methoden, um nach Prüfung der Zugriffsberechtigung Forschungsakten aus dem Repository zu laden und bereitzustellen. Listing 4.16 zeigt den Programmcode der Methode *isRecordOpenAllowed()* des *ResearchRecordService* zur Entscheidung, ob eine Forschungsakte durch eine oder einen Systembenutzenden geöffnet werden darf oder nicht.

```
public boolean isRecordOpenAllowed() {
    SecurityUser authenticatedUser = securityService.getAuthenticatedUser();
    if (authenticatedUser.hasUserPrivilege(Privilege.Name.OPEN_RECORD)) {
        return true;
    }
    return false;
}
```

Listing 4.16: Programmcode der Methode *isRecordOpenAllowed()* der Service-Klasse *ResearchRecordService* zur Überprüfung des Zugriffsrechts *OPEN_RECORD* im Probandenmanagementsystem.

Ein weiteres Beispiel zur Zugriffssteuerung im Probandenmanagementsystem wird in Listing 4.17 vorgestellt. Der Programmcode der Methode dient zur Überprüfung eines Systembenutzenden, ob identifizierende Daten in einer Listenansicht von Studienteilnehmenden angezeigt werden dürfen. Die Methode prüft sowohl, ob das Recht zur Einsicht von identifizierenden Daten vorliegt, als auch, ob die Berechtigung für den übergebenen Standort der Forschungsakte für die oder den Systembenutzenden vorliegt.

```
public boolean isIDATAccessAllowed(ResearchSite researchSite) {
    SecurityUser authenticatedUser = securityService.getAuthenticatedUser();
    boolean idatAccess = authenticatedUser.hasUserPrivilege(Privilege.Name.SITE_IDAT);
    boolean siteAccess = authenticatedUser.hasUserResearchSiteAccess(researchSite);
    if (idatAccess && siteAccess) {
        return true;
    }
    return false;
}
```

Listing 4.17: Programmcode der Methode *isIDATAccessAllowed(...)* der Service-Klasse *ResearchRecordService*. Es wird das Zugriffsrecht *SITE_IDAT* als auch die Berechtigung zum Zugriff auf den Forschungsstandort der Forschungsakte überprüft.

Auf Basis der in Listing 4.16 und Listing 4.17 vorgestellten Überprüfungen steuert die in Abbildung 4.28 gezeigte Oberfläche des Probandenmanagementsystems benutzerabhängig, ob Forschungsakten geöffnet werden dürfen und ob identifizierende Daten der Studienteilnehmenden angezeigt werden.

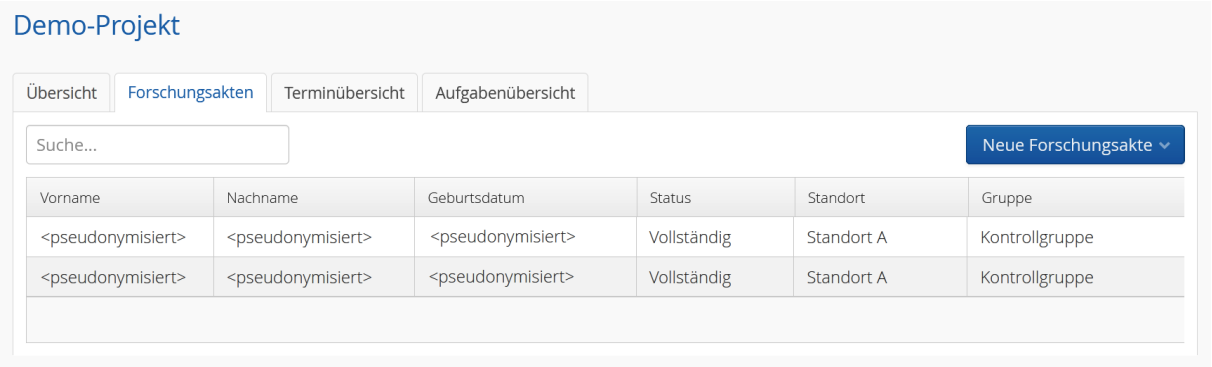

Abbildung 4.28: Screenshot der Benutzeroberfläche des Probandenmanagementsystems zur Übersicht der Forschungsakten eines Forschungsprojektes.

Zusammenfassend bietet Spring Security alle benötigten Funktionen, um die in dieser Arbeit konzeptionierte Zugriffssteuerung des Probandenmanagementsystems umzusetzen. Rollenbasierte und attributbasierte Zugriffskontrolle ließen sich in der prototypischen Implementierung kombiniert umsetzen.

4.3.4 Protokollierter Zugriff

Das Speichern von Datenänderungen ist eine häufige Anforderung an Anwendungssysteme, die auch Schwanke [2015] als Randbedingung in seinen Anforderungen an das Probandenmanagementsystem aufführte [vgl. Schwanke, 2015, Randbedingung A-R-4, S. 42]. Für diese häufig benötigte Funktion der Änderungsnachverfolgung existieren bereits erprobte Implementierungen, wie beispielsweise das Modul Envers[34] als Teil von Hibernate. Hibernate findet bereits für die objektrelationale Abbildung der Daten im Probandenmanagementsystem Anwendung (vgl. Unterkapitel 4.1.2, Softwarestack des Probandenmanagementsystems, S. 105). Hibernate Envers ermöglicht auf einfache Weise die Erfassung aller Datenänderungen in Revisionstabellen. Einträge in eine Tabelle bekommen eine eindeutige Revisionsnummer und einen Zeitstempel, der zu jeder Änderung gespeichert wird [vgl. Bauer u. a.,

[34] Hibernate Envers: `https://hibernate.org/orm/envers/` [Stand: 01.03.2022]

2015, S. 330 ff.], [vgl. auch Schwanke, 2015, S. 55]. Auf einen Datensatz bezogen lässt sich so jeder Schritt der Datennachbearbeitung nachvollziehen. Die in dieser Arbeit konzeptionierte Zugriffsprotokollierung (vgl. Unterkapitel 3.5.3, Iteration 3, Anforderung A5: Protokollierter Zugriff, S. 86) geht über die bisherige Umsetzung einer einfachen Änderungsnachverfolgung hinaus. Das Protokoll der Änderungshistorie ist Teil einer Protokollierung im Sinne eines Audit Trails[35] zum Nachweis der ordnungsgemäßen Verarbeitung der personenbezogenen Daten im Probandenmanagementsystem. Umgesetzt mit Hibernate Envers, wurde für alle Datensätze mit personenbezogenen Daten der Studienteilnehmenden das Protokoll der Änderungshistorie um die Angabe des im Probandenmanagementsystem eingeloggten Systembenutzenden erweitert. So lässt sich nachvollziehen, wer eine Änderung zu welchem Zeitpunkt an einem Datensatz durchgeführt hat. Listing 4.18 zeigt einen Auszug des Programmcodes der dazu notwendigen Implementierung der Klasse *Revision*, die mit dem Attribut *UserAccount* einen Verweis auf die oder den Systembenutzenden in eine Revisionstabelle ergänzt. *Revision* erbt von der Klasse *DefaultRevisionEntity* der Hibernate Envers Implementierung und bringt damit die Grundimplementierung mit *RevisionNumber* und *RevisionTimestamp* mit.

```
[...]
@Entity
@Table(name = "revision")
@RevisionEntity(RevisionListener.class)
public class Revision extends DefaultRevisionEntity {

    @JoinColumn(name = "user_account_id", updatable = false, referencedColumnName = "id")
    private UserAccount userAccount;
[...]
```

Listing 4.18: Programmcode-Auszug der Klasse *Revision*, die eine für das Probandenmanagementsystem angepasste *RevisionEntity* für die Datenänderungshistorie mit Hibernate Envers implementiert. Zu den erbenden Attributen der Klasse *DefaultRevisionEntity* wurde das Attribut *UserAccount* hinzugefügt.

[35] Fachbegriff *Audit Trail*: Werkzeug der Qualitätssicherung zur Kontrolle und Aufzeichnung von Änderungen, die in Prozessen vorgenommen werden (vgl. Definition Audit Trail, In: Gabler Wirtschaftslexikon Online: `https://wirtschaftslexikon.gabler.de/definition/audit-trail-53401/version-276494/` [Stand: 01.03.2022])

Listing 4.19 zeigt exemplarisch die Klasse *Participant*, die über die *@Audited*-Annotation angibt, dass alle Datensätze dieser Klasse von Hibernate Envers in die Auditierung einbezogen werden.

```
[...]
@Entity
@Table(name = "participant")
@Audited
public class Participant extends AbstractPersistable {

    @NotNull
    @Size(min = 1, max = 255)
    @Column(name = "first_name", nullable = false, length = 255)
    private String firstName;
[...]
```

Listing 4.19: Programmcode-Auszug der Klasse *Participant*, die eine Participant-Entität beschreibt. Exemplarisch mitdargestellt das Pflichtfeld eines Probandendatensatzes *firstName* mit seinen Metadaten, das durch die *@Audited*-Annotation der gesamten Klasse in die Änderungshistorie einbezogen wird.

Neben der Änderungshistorie mit Zeitpunkt und Bezug zu Systembenutzenden wurde eine Protokollierung von Zugriffen auf die personenbezogenen Daten der Studienteilnehmenden implementiert. Ziel dieser Implementierung ist, Zugriffe auf die durch die Zugriffskontrolle geschützten Funktionen des Probandenmanagementsystems für eine Datenschutzkontrolle zu protokollieren (vgl. Unterkapitel 3.5.3, Iteration 3, Anforderung A5: Protokollierter Zugriff, S. 86). Für eine programmiertechnisch elegante Lösung, die für die gewünschte Protokollierung wenig in die Kerngeschäftslogik des Probandenmanagementsystems eingreift und den Programmcode übersichtlich hält, wurde das Prinzip benutzerdefinierter AspectJ[36]-Annotation genutzt. Das Programmierframework Spring bietet eine gute Unterstützung[37] für AspectJ. Eine eigene Annotation *@LogAuditTrail* wurde erstellt, um Informationen zum Zugriff zu protokollieren (siehe Listing 4.20). Zu den protokollierten Informationen gehört der Name der Aktion, die ID der Forschungsakte innerhalb der die Aktion ausgeführt wurde, der Zeitpunkt der Durchführung und die oder der durchführende Systembenutzende.

[36] AspectJ: `https://www.eclipse.org/aspectj/` [Stand: 01.03.2022]

[37] Aspect Oriented Programming with Spring: `https://docs.spring.io/spring-framework/docs/5.2.9.RELEASE/spring-framework-reference/core.html#aop/` [Stand: 01.03.2022]

```
@Retention(RetentionPolicy.RUNTIME)
@Target(ElementType.METHOD)
public @interface LogAuditTrail {
}
```

Listing 4.20: Programmcode der Definition der LogAuditTrail-Annotation mit weiteren AspectJ-Annotationen. Die *@Target-Annotation* mit *ElementType.Method* gibt an, dass die Annotation nur auf Methoden angewendet werden darf. *@Retention* gibt an, ob die Annotation der Java Virtual Machine zur Laufzeit zur Verfügung steht.

Im weiteren Programmcode des Probandenmanagementsystems werden relevante Methoden mit der Annotation markiert und kennzeichnen so datenschutzrelevante Aktionen, die protokolliert werden sollen. Bei Ausführung der markierten Methode erfolgt die Protokollierung. Listing 4.21 zeigt ein Beispiel für die Nutzung der Annotation im Probandenmanagementsystem.

```
@LogAuditTrail
public String retrieveIdentifier(Parameter parameter) {
    // Programmcode zur Abfrage von Pseudonymen
}
```

Listing 4.21: Programmcode zur Verdeutlichung der Nutzung der LogAuditTrail-Annotation, hier für die Methode *retrieveIdentifier(...)* zum Abruf von Pseudonymen.

Das Protokoll der Änderungshistorie und das Protokoll von sensiblen Zugriffen im Probandenmanagementsystem helfen, eine ordnungsgemäße Verarbeitung der personenbezogenen Daten der Studienteilnehmenden überprüfen zu können. Doch auch die Protokolldaten unterliegen einer strikten Zweckbindung und müssen entsprechend des Protokollierungskonzeptes verarbeitet werden (vgl. Unterkapitel 3.5.3, Iteration 3, Anforderung A5: Protokollierter Zugriff, S. 86). So ist es in der Regel angemessen, die Speicherdauer der Protokolldaten auf einen bestimmten Zeitraum zu beschränken. Mittels eines Datenbanktriggers (MariaDB Event Scheduler[38]) wurde die Speicherdauer des Zugriffsprotokolls exemplarisch auf einen Monat begrenzt. Listing 4.22 zeigt den zugehörigen Programmcode, der die Löschung für alle Logeinträge die älter als einen Monat sind, in der Tabelle *audittrail* vornimmt.

[38] MariaDB Event Scheduler: `https://mariadb.com/kb/en/event-scheduler/` [Stand: 01.03.2022]

```
CREATE EVENT pm.eventclearexpiredlog
    ON SCHEDULE
        EVERY 1 DAY
        STARTS CURRENT_TIMESTAMP
DO
    DELETE FROM pm.audittrail WHERE creationdate >= (NOW() - INTERVAL 1 MONTH);
```

Listing 4.22: Programmcode des Datenbank Events zur Löschung von Logeinträgen, die älter als einen Monat sind. Das Event wird täglich durch den Event Scheduler ausgeführt.

Kapitel 5

Diskussion und Ausblick

5.1 Diskussion

Ein Probandenmanagement- oder Probandenverwaltungssystem ist eine Softwarekategorie, die in der klinischen Forschung Anwendung findet, um Prozesse bei der Verwaltung von Studienteilnehmenden in klinischen Studien zu unterstützen. Diese eigene Definition schließt die Begriffslücke, der eine Literaturrecherche zur Herkunft und Bestimmung des Begriffs *Probandenmanagementsystem* vorausging. In Form einer zweiten systematischen Literaturrecherche wurden zahlreiche Literaturquellen untersucht, um die Hauptprozesse der Probandenverwaltung und ihre IT-Unterstützung zu identifizieren: Rekrutierung (1), Erfassung und Verwaltung der Einwilligungen (2), Pseudonymisierung (3) und Verwaltung administrativer Daten (4) der Studienteilnehmenden. Die gewonnenen Erkenntnisse mündeten in der Einordnung der einzelnen IT-Komponenten mit Probandenverwaltungsanteil in einer medizinischen Forschungs-IT-Infrastruktur. Das Göttinger Probandenmanagementsystem kann als eine Anwendung betrachtet werden, die Teile der vorgestellten Komponenten vereint.
Die ausführliche Aufarbeitung des Themas Probandenverwaltung in klinischen Forschungsprojekten im ersten Teil dieser Arbeit trägt zu einem besseren Verständnis der Möglichkeiten der IT-gestützten Verwaltung von Studienteilnehmenden bei. Das erklärte Ziel, durch die Begriffsuntersuchung und die Einordnung in die medizinische Forschungs-IT-Infrastruktur wichtige Grundlagen für den weiteren Verlauf dieser Arbeit zu legen, wurde erreicht. Die durchgeführte Literaturrecherche lieferte jedoch

nicht nur die benötigten Informationen, sondern ließ auch eine Einschätzung des Umfangs der verfügbaren Forschungsliteratur zu. So zeigte die Literaturrecherche, dass IT-unterstützte Probandenmanagementprozesse vermutlich ein Schattendasein fristen. Nur wenige Publikationen beschreiben explizit Probandenmanagementprozesse und ihre IT-Unterstützung. Es ist zwar davon auszugehen, dass mehr Lösungen zur Verwaltung von Studienteilnehmenden in Forschungsprojekten existieren als über eine Literatur- und Internetrecherche zu ermitteln war, doch wäre wünschenswert, wenn in der medizinischen Forschung mehr über genutzte IT-Infrastrukturen publiziert würde. Auch würde Literatur zu den Prozessen innerhalb von klinischen Forschungsprojekten helfen, über deren mögliche Verbesserung durch Digitalisierung zu diskutieren. Um die gewonnenen Erkenntnisse zum Stand der Forschungsliteratur in den wissenschaftlichen Diskurs einzubringen, wurden diese in Form eines Scoping Reviews im Journal JAMIA Open veröffentlicht. Auch wenn die Relevanz der Verwaltung von Studienteilnehmenden sich bislang nicht in der Literatur widerspiegelt, so ist sie elementarer Teil jeder medizinischen Forschung am Menschen. Das Institut für Medizinische Informatik der Universitätsmedizin Göttingen beschäftigte sich daher seit 2012 mit der Entwicklung eines Probandenmanagementsystems. Dieses System für die Nutzung in klinischen Verbundforschungsprojekten auszubauen stand im Mittelpunkt dieser Dissertation.
Ein fachliches Konzept, das insbesondere die Datenschutz- und IT-Sicherheitsanforderungen für eine standortübergreifende Nutzung berücksichtigt, wurde als Lösungsansatz für den Ausbau des Probandenmanagementsystems gewählt. Unter Berücksichtigung des Umfeldes, in dem das Probandenmanagementsystem zum Einsatz kommt, wurden vier zentrale Anforderungen an das System formuliert, um in standortübergreifenden Kooperationen eingesetzt zu werden:

1. Ein vertraulicher Zugang und Zugriff auf die Probandendaten muss sichergestellt sein.
2. Eine Datentrennung (informationelle Gewaltenteilung) bei der Probandenverwaltung muss gegeben sein.
3. Das Probandenmanagementsystem muss sich in eine IT-Infrastruktur einbinden lassen und Schnittstellen zu Pseudonymisierungsdienst und Identitätsmanagement bereitstellen.

4. Für den Betrieb des Probandenmanagementsystems muss ein Datenschutzkonzept vorhanden sein.

Zur Erarbeitung des Konzeptes wurde das Vorgehensmodell *Design and Creation* eingesetzt. Das Gesamtkonzept gliedert sich angelehnt an die identifizierten Anforderungen in die Teile: (1) Datentrennung bei der Probandenverwaltung, (2) Vernetze Probandenverwaltung, (3) Zugangs- und Zugriffskontrolle und (4) Datenschutzkonzept. Um eine strikte Datentrennung bei der Probandenverwaltung mit einem Probandenmanagementsystem zu gewährleisten dürfen keine externen Pseudonyme der Clinical Data Management Systeme in der Datenbank des Probandenmanagementsystems gespeichert werden. Stattdessen muss eine temporäre Pseudonymabfrage über eine Schnittstelle zum Pseudonymisierungsdienst eingerichtet sein. Pseudonymzuweisung und Pseudonymabfrage sind die zentralen Anwendungsfälle die zur Schnittstellenimplementierung spezifiziert wurden. Zur bestmöglichen Wahrung der informationellen Gewaltenteilung wird das Probandenmanagementsystem und der Pseudonymisierungsdienst unter getrennten Verantwortlichkeiten in separaten Netzwerken betrieben. Der sichere Zugriff von verschiedenen Studienstandorten auf das Probandenmanagementsystem wird durch technische Datenschutzmaßnahmen zur Zugangs- und Zugriffskontrolle möglich. Diese wurden in fünf Bereichen zur Implementierung vorbereitet: (1) Sichere Authentifizierung, (2) Verschlüsselte Verbindung, (3) Berechtigungskonzept, (4) Zugriffskontrollmodell und (5) Protokollierter Zugriff. Vorbereitend auf ein konkretes Forschungsprojekt wurden alle für ein Datenschutzkonzept relevanten Punkte zusammengetragen.

Der Erhebung von Anforderungen kommt im Softwareentwicklungsprozess eine wichtige Bedeutung zu, da die erhobenen Anforderungen die Basis für die weitere Entwicklung bilden. So auch der Anforderungserhebung im Rahmen des Konzeptes für den Ausbau des Probandenmanagementsystems in dieser Dissertation. Eine Besonderheit ist die Art der Anforderungen, die sich aus dem Ziel einer standortübergreifenden Nutzung des Probandenmanagementsystems ergeben haben. Diese sind Datenschutz- und IT-Sicherheitsanforderungen, die einen erheblichen Einfluss auf die Entwicklung des Systems hatten. Zwar wird die Unterteilung in funktionale und nichtfunktionale Anforderungen aufgrund der Schwierigkeit ihrer Unterscheidung in der Praxis oft nicht vorgenommen, an dieser Stelle führt sie aber zu einer wichtigen Erkenntnis. Schwanke

[2015] erfasste zur prototypischen Entwicklung des Probandenmanagementsystems zahlreiche funktionale Anforderungen, aber auch nichtfunktionale Anforderungen als Randbedingungen. Konformität mit den niedersächsischen Datenschutzbestimmungen wurde als Randbedingung aufgenommen, ebenso wie die Empfehlungen des generischen Datenschutzkonzeptes der TMF zu berücksichtigen [vgl. Schwanke, 2015, S. 42]. Beide Randbedingungen wurden allerdings in der gesamten Arbeit nicht weiter erwähnt, die Entwicklung der funktionalen Anforderungen stand im Vordergrund. Für die standortübergreifende Nutzung des Probandenmanagementsystems sind Randbedingungen jedoch außerordentlich wichtig. Bei der Konzeption des standortübergreifenden Probandenmanagementsystems zeigte sich, dass funktionalen und nichtfunktionalen Anforderungen in Form von Randbedingungen gleichwertige Aufmerksamkeit beigemessen werden muss.

Das zur Erarbeitung des Konzeptes genutzte Vorgehensmodell *Design and Creation* eignete sich gut, um die aufeinander aufbauenden Anforderungen schrittweise in mehreren Iterationen zu bearbeiten. Jede Iteration diskutierte bereits ihre gewonnenen Ergebnisse und verfeinerte so das Konzept. Das gesamte Konzept erfüllt das gesteckte Ziel, Studienteilnehmende von unterschiedlichen Studienstandorten in Verbundforschungsprojekten sicher und datenschutzkonform in einem System zu verwalten. Pseudonyme schützen die Identität der Studienteilnehmenden außerhalb einer vertrauenswürdigen Umgebung, in der Forschungsdaten gespeichert und ausgetauscht werden müssen, während durch eine Kontrolle des Datenzugriffs innerhalb einer vertrauenswürdigen Umgebung persönliche identifizierende Daten verwendet werden können. Die Verarbeitung der persönlichen Daten der Studienteilnehmenden stellt einen Kompromiss dar, der in den vorgestellten Grenzen gerechtfertigt und durch besonders sorgfältige Schutzmaßnahmen kompensiert wird.

Prototypisch implementiert wurden die konzeptionierten Maßnahmen zur Datentrennung und zur Zugangs- und Zugriffskontrolle des Probandenmanagementsystems. Die Implementierung für eine datenschutzkonforme Datentrennung setzt sich aus einem neuen Datenmodell für das Probandenmanagementsystem und einer Schnittstellenimplementierung zusammen. Das Datenmodell erlaubt die Verwaltung von Identitätsmanagementmetadaten sowie Probandenidentitäten und Identifikatoren im Probandenmanagementsystem. Die Implementierung der

Schnittstellen zur Mainzelliste setzt die Anwendungsfälle zur Pseudonymzuweisung und Pseudonymabfrage um. Zusammen wird die datenschutzkonforme Datentrennung in einer IT-Infrastruktur mit Probandenmanagementsystem möglich. Zur sicheren Zugangs- und Zugriffskontrolle des Probandenmanagementsystems wurden die erarbeiteten Anforderungen des Konzeptes umgesetzt: Eine sichere Authentifizierung wird über eine Zugriffsbeschränkung auf bestimmte IP-Adressen, eine Benutzerauthentifizierung über Benutzername und sicher verwaltete Passwörter mit der zusätzlichen Authentifizierung über Clientzertifikate gewährleistet. Eine verschlüsselte Verbindung nach den Empfehlungen des BSI gewährleistet eine sichere Datenübertragung. Eine rollen- und attributbasierte Zugriffskontrolle ließ sich in der prototypischen Implementierung kombiniert verwirklichen. Umgesetzt wurde auch eine Zugriffsprotokollierung, mit der sich eine ordnungsgemäße Verarbeitung der personenbezogenen Daten im Probandenmanagementsystem nachvollziehen lässt.

Die prototypische Implementierung des standortübergreifenden Probandenmanagementsystems bestätigte die Machbarkeit des Konzeptes. Die bestehende Software- und Systemarchitektur des Probandenmanagementsystems erlaubte eine erfolgreiche Umsetzung der neuen Anforderungen für ein standortübergreifend nutzbares Probandenmanagementsystem. Schwanke [2015] wählte eine Architektur, die Anpassungen und Erweiterungen ermöglichte. Anpassungsfähigkeit ist eine wesentliche Voraussetzung zeitgemäßer Softwarearchitektur, denn es wird immer wichtiger, dass Software sich schnell verändern kann, um auf neue Rahmenbedingungen zu reagieren. Im Vordergrund dieser Dissertation stand die konzeptionelle Arbeit, eine vollständige Funktionsprüfung und Erprobung in der Praxis muss in zukünftigen Forschungsprojekten erfolgen. Der entwickelte Demonstrator belegt jedoch die präsentierte prototypische Implementierung und bestätigt die Machbarkeit des Konzeptes.

5.2 Ausblick

Die Durchführung klinischer Studien mit IT-Systemen zu optimieren, wird zunehmend eine wichtige Rolle spielen, geht es doch am Ende darum, die Erkenntnisse der medizinischen Forschung möglichst schnell und effektiv zu Patientinnen und Patienten zu bringen. Ein leistungsfähiges Management von Forschungsdaten und eine effiziente Prozessunterstützung sind hierfür grundlegend. Die Ausgangslage im universitären klinischen Forschungsumfeld ist jedoch häufig die, dass eine Versorgung der Forschenden mit niedrigschwelligen IT-Services für die Durchführung klinischer Studien fehlt. Die Komplexität der Systemlandschaft, die hohen Anforderungen an Datenschutz und Datensicherheit, der Einsatz qualifizierten Personals und die befristete Finanzierung durch Drittmittel sind Teile der Herausforderung, IT-Services für klinische Studien zu etablieren. Pionierarbeiten wie die prototypische Entwicklung des Probandenmanagementsystems sind häufig von projektförmigen Initiativen geprägt, bei denen nach Projektabschluss viel innovatives Potenzial verloren gehen kann. Wichtig ist daher, dass Ansätze aus der Forschung den Übergang in den Regelbetrieb schaffen. Im Sinne einer Basisinfrastruktur für klinische Studien sollten an universitätsmedizinischen Standorten unabhängige Treuhandstellen mit der Aufgabe des Identitäts- und Pseudonymmanagements betraut und ein Probandenmanagementsystem zur Unterstützung der probandenverwaltenden Studienprozesse unter dauerhafter Finanzierung eingerichtete werden.

Denkbar sind auch Entwicklungspartnerschaften mit der Softwareindustrie. Studienmanagementsysteme, die in kommerziellen klinischen Studien der Biotechnologie- und Pharmaindustrie Anwendung finden, könnten das Konzept des standortübergreifen Probandenmanagementsystems aufnehmen. Auch könnte durch eine Partnerschaft die professionelle Entwicklung eines Probandenmanagementsystems für den wissenschaftlichen Einsatz vorangetrieben werden.

Generell sollte der Bereich der Studienverwaltung aus wissenschaftlicher Sicht stärker beleuchtet werden. Während die Ergebnisse aus klinischen Studien zahlreich publiziert werden, zeigte sich im ersten Teil dieser Arbeit, dass der Bereich der Studienverwaltung und der Einsatz von IT-Systemen in klinischen Studien in Publikationen unterrepräsentiert sind. Aufschluss über den aktuellen Stand der Nutzung von IT-Systemen in klinischen Studien würde eine Bestandsaufnahme unter den

deutschen Universitätskliniken geben. Wünschenswert wäre, dass in der medizinischen Forschung mehr über genutzte IT-Infrastrukturen publiziert und so die Innovation von Systemen wie dem Probandenmanagementsystem gefördert würde. Des Weiteren wäre ein Vergleich digitaler und nicht-digitaler Verwaltung von Studienteilnehmenden und deren Auswirkungen auf den Studienerfolg für die Weiterentwicklung prozessunterstützender Software für klinische Studien hoch interessant.

Moderne Forschungsansätze in der Medizin zielen immer häufiger darauf ab, verschiedene Datenquellen für eine gemeinsame Analyse zusammenzuführen. Bindeglied zwischen den Datensätzen ist meist die Identitäten der Patientinnen und Patienten, der Studienteilnehmenden oder der Datenspendenden. Wo immer eine Verarbeitung personenbeziehbarer Forschungsdaten anfällt, werden administrative identitätsverwaltende Prozesse nötig. Identitätsmanagement und Pseudonymisierungsdienste sind dafür essenziell und ein prozessunterstützendes System wie das Probandenmanagementsystem ist in diesem Kontext hoch relevant. Welche Prozesse in Forschungsprojekten zukünftig mit dem Probandenmanagementsystem abzubilden sind, sollte weiter untersucht werden.

Die Relevanz der Probandenverwaltung für medizinische Verbundforschungsprojekte wurde bereits in der Einleitung betont. Nicht zu vernachlässigen sind aber auch kleinere Forschungsprojekte, die ohne große Infrastrukturförderung Studienteilnehmende verwalten müssen. Auch bei dieser Herausforderung kann das Probandenmanagementsystem als zentrale Software einfache Unterstützung bieten. Den Nutzen des Probandenmanagementsystems in konkreten kleinen wie großen Forschungsprojekten zu bewerten ist eine wichtige Aufgabe für zukünftige Untersuchungen.

Ebenso wichtig ist es, die hohe Dynamik der Entwicklung in Wissenschaft, Technik und Datenschutz zu beachten. Die für das Probandenmanagementsystem umgesetzten Maßnahmen zum Datenschutz als auch zur Datensicherheit berücksichtigen die aktuellen rechtlichen Vorgaben und den Stand der Technik. Da jedoch beides ein „Moving Target“ ist, kann lediglich eine Beurteilung der momentanen Situation gegeben werden, die für die Zukunft nur unter regelmäßiger Überprüfung eine Lösung darstellt. Die Verwaltung sensibler identifizierender Gesundheitsdaten in einem IT-System ist daher ein Forschungsthema, das stets aktuelle Entwicklung verfolgen muss.

Mit dem in dieser Arbeit vorgestellten Ausbau des Probandenmanagementsystems für standortübergreifende Kooperationen steht klinischen Forschungsprojekten ein leistungsfähiges Werkzeug zur unkomplizierten wie sicheren digitalen Probandenverwaltung zur Verfügung. Dies kommt der universitären klinischen Forschung zugute, da Kooperationen zwischen Studienzentren vereinfacht und groß angelegte, multizentrische klinische Studien unterstützt werden.

Literaturverzeichnis

[Aamot u. a. 2013] Aamot, Harald ; Kohl, Christian D. ; Richter, Daniela ; Knaup-Gregori, Petra: Pseudonymization of patient identifiers for translational research. In: *BMC Medical Informatics and Decision Making* 13 (2013), Nr. 75. – DOI 10.1186/1472-6947-13-75

[Almeida u. a. 2019] Almeida, João R. ; Gini, Rosa ; Roberto, Giuseppe ; Rijnbeek, Peter ; Oliveira, José L.: TASKA: A modular task management system to support health research studies. In: *BMC Medical Informatics and Decision Making* 19 (2019), Nr. 121. – DOI 10.1186/s12911-019-0844-6

[Ammon u. a. 2020] Ammon, Danny ; Backer-Heuveldop, Andrea ; Isele, Christoph ; Kadi, Hasan ; Letter, Michael ; Rüdlin, Mark ; Schlütter, Johannes ; Schütze, Bernd ; Wichterich, Eric: Protokollierung und Protokollierungskonzept - Eine Einführung in die Thematik. (2020). – URL `https://gesundheitsdatenschutz.org/download/protokollierungskonzept_2020.pdf`, Zugriffsdatum: 01.04.2021

[Antony u. a. 2017] Antony, Gisela ; Bialke, Martin ; Pommerening, Klaus ; Repp, Ronny: Checkliste zur Erstellung eines Datenschutzkonzeptes / TMF - Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. (2017). – URL `https://www.toolpool-gesundheitsforschung.de/sites/default/files/2018-11/Checkliste_Datenschutzberatung-deutsch-v1.0.pdf`, Zugriffsdatum: 14.07.2021

[Bauer u. a. 2015] Bauer, Christian ; King, Gavin ; Gregory, Gary: *Java Persistence with Hibernate*. Manning Publications, Shelter Island, NY, 2015

[Bauer u. a. 2018] Bauer, Christoph ; Eickmeier, Frank ; Eckard, Michael: *E-Health: Datenschutz und Datensicherheit*. Springer Gabler, Wiesbaden, 2018

[Bialke u. a. 2018] Bialke, Martin ; Bahls, Thomas ; Geidel, Lars ; Rau, Henriette ; Blumentritt, Arne ; Pasewald, Sandra ; Wolff, Robert ; Steinmann, Jonas ; Bronsch, Tobias ; Bergh, Björn ; Tremper, Galina ; Lablans, Martin ; Ückert, Frank ; Lang, Stefan ; Idris, Tarik ; Hoffmann, Wolfgang: MAGIC: once upon a time in consent management – a FHIR tale. In: *Journal of Translational Medicine* 16 (2018), Nr. 256. – DOI 10.1186/s12967-018-1631-3

[Bialke u. a. 2015a] Bialke, Martin ; Bahls, Thomas ; Havemann, Christoph ; Piegsa, Jens ; Weitmann, Kerstin ; Wegner, Tim ; Hoffmann, Wolfgang: MOSAIC - A Modular Approach to Data Management in Epidemiological Studies. In: *Methods of Information in Medicine* 54 (2015), Nr. 04, S. 364–371. – DOI 10.3414/me14-01-0133

[Bialke u. a. 2015b] Bialke, Martin ; Penndorf, Peter ; Wegner, Tim ; Bahls, Thomas ; Havemann, Christoph ; Piegsa, Jens ; Hoffmann, Wolfgang: A workflow-driven approach to integrate generic software modules in a Trusted Third Party. In: *Journal of Translational Medicine* 13 (2015), Nr. 176. – DOI 10.1186/s12967-015-0545-6

[Boutin u. a. 2016] Boutin, Natalie T. ; Mathieu, Kathleen ; Hoffnagle, Alison G. ; Allen, Nicole L. ; Castro, Victor M. ; Morash, Megan ; O'Rourke, P. P. ; Hohmann, Elizabeth L. ; Herring, Neil ; Bry, Lynn ; Slaugenhaupt, Susan A. ; Karlson, Elizabeth W. ; Weiss, Scott T. ; Smoller, Jordan W.: Implementation of Electronic Consent at a Biobank: An Opportunity for Precision Medicine Research. In: *Journal of Personalized Medicine* 6 (2016), Nr. 2. – DOI 10.3390/jpm6020017

[BSI 2020a] BSI: IT-Grundschutz: OPS.1.1.5 Protokollierung / Bundesamt für Sicherheit in der Informationstechnik. (2020). – URL `https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/Kompendium_Einzel_PDFs/04_OPS_Betrieb/OPS_1_1_5_Protokollierung_Edition_2020.pdf`, Zugriffsdatum: 21.09.2020

[BSI 2020b] BSI: Technische Richtlinie BSI TR-02102-1 Kryptographische Verfahren: Empfehlungen und Schlüssellängen / Bundesamt für Sicherheit in der Informationstechnik. (2020). – URL `https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/TechnischeRichtlinien/TR02102/BSI-TR-02102.html`, Zugriffsdatum: 21.09.2020

[BSI 2020c] BSI: Technische Richtlinie BSI TR-02102-2 Kryptographische Verfahren: Verwendung von Transport Layer Security (TLS) / Bundesamt für Sicherheit in der Informationstechnik. (2020). – URL `https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/TechnischeRichtlinien/TR02102/BSI-TR-02102-2.html`, Zugriffsdatum: 21.09.2020

[BSI 2021] BSI: *IT-Grundschutz-Kompendium.* Bundesamt für Sicherheit in der Informationstechnik, Bonn, 2021

[Buchner u. a. 2021] Buchner, Benedikt ; Haber, Anna C. ; Hahn, Horst K. ; Prasser, Fabian ; Kusch, Harald ; Sax, Ulrich ; Schmidt, Carsten O.: Das Modell der Datentreuhand in der medizinischen Forschung. In: *Datenschutz und Datensicherheit - DuD* 45 (2021), Nr. 12, S. 806–810. – DOI 10.1007/s11623-021-1534-y

[Chen 1976] Chen, Peter Pin-Shan: The Entity-Relationship Model – toward a Unified View of Data. In: *ACM Transactions on Database Systems* 1 (1976), Nr. 1, S. 9–36. – DOI 10.1145/320434.320440

[Cockburn 2005] Cockburn, Alistair: The Pattern: Ports and Adapters (Object Structural). (2005). – URL `https://alistair.cockburn.us/hexagonal-architecture/`, Zugriffsdatum: 06.07.2020

[Dangl u. a. 2010] Dangl, Andrea ; Demiroglu, Sara Y. ; Gaedcke, Jochen ; Helbing, Krister ; Jo, Peter ; Rakebrandt, Fabian ; Rienhoff, Otto ; Sax, Ulrich: The IT-Infrastructure of a Biobank for an Academic Medical Center. In: *Studies in Health Technology and Informatics* Bd. 160 MEDINFO 2010, 2010, S. 1334–1338. – DOI 10.3233/978-1-60750-588-4-1334

[Demiroglu u. a. 2013] Demiroglu, Sara Y. ; Quade, Matthias ; Schwanke, Jens ; Gusky, Linda ; Skrowny, Daniela ; Schulze, Thomas G.: Datenschutzkonzept für die Klinische Forschergruppe 241. (2013). – URL http://www.kfo241.de/materials/20120718_datenschutzkonzept.pdf, Zugriffsdatum: 15.01.2020

[Demiroglu u. a. 2012] Demiroglu, Sara Y. ; Skrowny, Daniela ; Quade, Matthias ; Schwanke, Jens ; Budde, Monika ; Gullatz, Verena ; Reich-Erkelenz, Daniela ; Jakob, J. J. ; Falkai, Peter ; Rienhoff, Otto ; Helbing, Krister ; Heilbronner, Urs ; Schulze, Thomas G.: Managing sensitive phenotypic data and biomaterial in large-scale collaborative psychiatric genetic research projects: practical considerations. In: *Molecular Psychiatry* 17 (2012), S. 1180–1185. – DOI 10.1038/mp.2012.11

[Eckert 2008] Eckert, Claudia: *IT-Sicherheit : Konzepte, Verfahren, Protokolle.* 5., überarbeitete Auflage. Oldenbourg, München, Wien, 2008

[Eigner u. a. 2012] Eigner, Martin ; Gerhardt, Florian ; Gilz, Torsten ; Nem, Fabrice M.: *Informationstechnologie für Ingenieure.* Springer, Berlin, Heidelberg, 2012

[Faldum und Pommerening 2005] Faldum, Andreas ; Pommerening, Klaus: An optimal code for patient identifiers. In: *Computer Methods and Programs in Biomedicine* 79 (2005), Nr. 1, S. 81–88. – DOI 10.1016/j.cmpb.2005.03.004

[Ferraiolo und Kuhn 1992] Ferraiolo, David F. ; Kuhn, D. R.: Role-Based Access Controls. In: *Proceedings of the 15th National Computer Security Conference.* Baltimore, Maryland, United States, 1992, S. 554–563. – URL https://csrc.nist.gov/CSRC/media/Publications/conference-paper/1992/10/13/role-based-access-controls/documents/ferraiolo-kuhn-92.pdf

[Ferraiolo u. a. 2001] Ferraiolo, David F. ; Sandhu, Ravi ; Gavrila, Serban ; Kuhn, D. R. ; Chandramouli, Ramaswamy: Proposed NIST Standard for Role-Based Access Control. In: *ACM Transactions on Information and System Security* 4 (2001), August, Nr. 3, S. 224–274. – DOI 10.1145/501978.501980

[Geißler und Ostler 2019] Geissler, Otto ; Ostler, Ulrike: Was ist ein Software-Stack? / DataCenter-Insider. (2019). – URL https://www.datacenter-insider.de/was-ist-ein-software-stack-a-841053/, Zugriffsdatum: 10.12.2021

[Good Clinical Practice 2016] Good Clinical Practice: Guideline for Good Clinical Practice ICH E6 (R2). (2016). – URL https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R2__Step_4_2016_1109.pdf, Zugriffsdatum: 25.01.2019

[Gordon u. a. 2002] Gordon, Neil F. ; English, Carla D. ; Contractor, Aashish S. ; Salmon, Richard D. ; Leighton, Richard F. ; Franklin, Barry A. ; Haskell, William L.: Effectiveness of three models for comprehensive cardiovascular disease risk reduction. In: *The American Journal of Cardiology* 89 (2002), Nr. 11, S. 1263–1268. – DOI 10.1016/S0002-9149(02)02323-8

[Grady u. a. 2017] Grady, Christine ; Cummings, Steven R. ; Rowbotham, Michael C. ; McConnell, Michael V. ; Ashley, Euan A. ; Kang, Gagandeep: Informed Consent. In: *New England Journal of Medicine* 376 (2017), Nr. 9, S. 856–867. – DOI 10.1056/NEJMra1603773

[Gregorczyk 2014] Gregorczyk, David: *Technologien für eine interoperable und automatisierte Vernetzung von medizinischen IT-Systemen*, Universität zu Lübeck, Dissertation, 2014

[Haak u. a. 2017] Haak, Daniel ; Deserno, Verena ; Deserno, Thomas: *Informationsmanagement und Kommunikation in der Medizin.* Kap. Datenmanagement für Medizinproduktestudien, S. 145–164, Springer, Berlin, Heidelberg, 2017

[Hagenhoff 2008] Hagenhoff, Svenja: *Innovationsmanagement für Kooperationen.* Universitätsverlag Göttingen, Göttingen, 2008

[Hasselbring 2006] Hasselbring, Wilhelm: Software-Architektur. In: *Informatik-Spektrum* 29 (2006), S. 48–52. – DOI 10.1007/s00287-005-0049-5

[Haussen u. a. 2017] Haussen, Diogo C. ; Doppelheuer, Shannon ; Schindler, Kiva ; Grossberg, Jonathan A. ; Bouslama, Mehdi ; Schultz, Meagan ; Perez, Hilarie ; Hall, Alex ; Frankel, Michael ; Nogueira, Raul G.: Utilization of a Smartphone Platform for Electronic Informed Consent in Acute Stroke Trials. In: *Stroke* 48 (2017), Nr. 11, S. 3156–3160. – DOI 10.1161/STROKEAHA.117.018380

[Heinze 2015] Heinze, Oliver: *Persönliche, einrichtungsübergreifende, elektronische Patientenakte (PEPA) - Konzept, Systemarchitektur und Umsetzungserfahrungen*, Ruprecht-Karls-Universität Heidelberg, Dissertation, 2015

[Hellige 2004] Hellige, Hans D. (Hrsg.): *Geschichten der Informatik.* Springer, Berlin, Heidelberg, 2004

[Hu u. a. 2014] Hu, Vincent C. ; Ferraiolo, David ; Kuhn, Rick ; Schnitzer, Adam ; Sandlin, Kenneth ; Miller, Robert ; Scarfone, Karen: Guide to Attribute Based Access Control (ABAC) Definition and Considerations / National Institute of Standards and Technology Special Publication 800-162. (2014). – URL `https://doi.org/10.6028/NIST.SP.800-162`, Zugriffsdatum: 01.04.2021

[ISDSG 2021] ISDSG: Defintion Berechtigungskonzept / Institut für Sicherheit und Datenschutz im Gesundheitswesen. (2021). – URL `https://www.isdsg.de/beratung/konzepterstellung/berechtigungskonzept`, Zugriffsdatum: 21.11.2020

[Kashmar u. a. 2019] Kashmar, Nadine ; Adda, Mehdi ; Atieh, Mirna: From Access Control Models to Access Control Metamodels: A Survey. In: Arai, Kohei (Hrsg.) ; Bhatia, Rahul (Hrsg.): *Lecture Notes in Networks and Systems* Bd. 70. Cham : Springer, 2019, S. 892–911. – DOI 10.1007/978-3-030-12385-7_61

[Köpcke und Prokosch 2014] Köpcke, Felix ; Prokosch, Hans-Ulrich: Employing Computers for the Recruitment into Clinical Trials: A Comprehensive Systematic Review. In: *Journal of Medical Internet Research* 16 (2014), Nr. 7, S. e161. – DOI 10.2196/jmir.3446

[Lablans u. a. 2015] Lablans, Martin ; Borg, Andreas ; Ückert, Frank: A RESTful interface to pseudonymization services in modern web applications. In: *BMC Medical Informatics and Decision Making* 15 (2015), Nr. 2. – DOI 10.1186/s12911-014-0123-5

[LfDNI 2020] LfDNI: Handlungsempfehlung sichere Authentifizierung / Die Landesbeauftragte für den Datenschutz Niedersachsen. (2020). – URL https://lfd.niedersachsen.de/download/61775/Handlungsempfehlung_sichere_Authentifizierung_Stand_Februar_2020_.pdf, Zugriffsdatum: 27.08.2020

[Matoni 2014] Matoni, Markus: *Abbildung und Verwaltung von Einwilligungserklärungen in einem Probandenmanagementsystem*, Georg-August-Universität Göttingen, Masterarbeit, 2014

[McDonald u. a. 2006] McDonald, Alison M. ; Knight, Rosemary C. ; Campbell, Marion K. ; Entwistle, Vikki A. ; Grant, Adrian M. ; Cook, Jonathan A. ; Elbourne, Diana R. ; Francis, David ; Garcia, Jo ; Roberts, Ian ; Snowdon, Claire: What influences recruitment to randomised controlled trials? A review of trials funded by two UK funding agencies. In: *Trials* 7 (2006), Nr. 9. – DOI 10.1186/1745-6215-7-9

[MITRE Corporation 2014] MITRE Corporation: Electronic Consent Management: Landscape Assessment, Challenges, and Technology. (2014). – URL https://www.healthit.gov/sites/default/files/privacy-security/ecm_finalreport_forrelease62415.pdf, Zugriffsdatum: 16.02.2019

[Moher u. a. 2009] Moher, David ; Liberati, Alessandro ; Tetzlaff, Jennifer ; Altman, Douglas G.: Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. In: *BMJ* 339 (2009). – DOI 10.1136/bmj.b2535

[Oates 2006] Oates, Briony J.: *Researching Information Systems and Computing.* SAGE Publications, London, 2006

[Ohmann und Speer 2016] Ohmann, Christian ; Speer, Ronald: *IT-Infrastrukturen in der patientenorientierten Forschung: Aktueller Stand und Handlungsbedarf - 2016.* Kap. Klinische Studien, S. 25–45, IT-Reviewing Board der TMF - Technologie und Methodenplattform für die vernetzte medizinische Forschung e.V., Akademische Verlagsgesellschaft AKA, Berlin, 2016

[Pommerening u. a. 2017] Pommerening, Klaus ; Altmann, Udo ; Jöckel, Karl-Heinz ; Kieschke, Joachim ; Müller, Thomas H. ; Pigeot-Kübler, Iris ; Schütze, Bernd: Memorandum zum Datenschutz in der medizinischen Forschung. (2017). – URL https://gmds.de/fileadmin/user_upload/170511_Memorandum_zum_Datenschutz.pdf, Zugriffsdatum: 19.02.2021

[Pommerening u. a. 2014] Pommerening, Klaus ; Drepper, Johannes ; Helbing, Krister ; Ganslandt, Thomas: *Leitfaden zum Datenschutz in medizinischen Forschungsprojekten - Generische Lösungen der TMF 2.0.* Schriftenreihe der TMF Band 11. MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin, 2014

[Posch u. a. 2011] Posch, Thorsten ; Birken, Klaus ; Gerdom, Michael: *Basiswissen Softwarearchitektur.* 3., aktualisierte und erweiterte Auflage. dpunkt.verlag, Heidelberg, 2011

[Priebe u. a. 2005] Priebe, Torsten ; Dobmeier, Wolfgang ; Muschall, Björn ; Pernul, Günther: ABAC – Ein Referenzmodell für attributbasierte Zugriffskontrolle. In: Federrath, Hannes (Hrsg.): *Sicherheit 2005, Sicherheit – Schutz und Zuverlässigkeit.* Bonn : Gesellschaft für Informatik e.V., 2005, S. 285–296. – URL https://dl.gi.de/handle/20.500.12116/28384

[Pung u. a. 2021] Pung, Johannes ; Krefting, Dagmar ; Rienhoff, Otto: Access control policy for a study participant management system for multi-site clinical research. In: *GMDS 2020. 65th Annual Meeting of the German Association for Medical Informatics, Biometry and Epidemiology (GMDS). Berlin, 06.-09.09.2020* DocAbstr. 42, Düsseldorf: German Medical Science GMS Publishing House, 2021. – DOI 10.3205/20gmds154

[Pung und Rienhoff 2020] Pung, Johannes ; Rienhoff, Otto: Key components and IT assistance of participant management in clinical research: a scoping review. In: *JAMIA Open* 3 (2020), Nr. 3, S. 449–458. – DOI 10.1093/jamiaopen/ooaa041

[Reich-Erkelenz u. a. 2012] Reich-Erkelenz, Daniela ; Schmitt, Angelika ; Gruber, Oliver ; Dechent, Peter ; Brockmöller, Jürgen ; Tzvetkov, Mladen ; Fischer, André ; Bickeböller, Heike ; Rossner, Moritz ; Nave, Klaus A. ; Falkai, Peter ; Schulze, Thomas G.: Neurobiology of severe psychiatric disorders - new approaches in the molecular research landscape. In: *Die Psychiatrie* 9 (2012), Nr. 01, S. 14–24. – DOI 10.1055/s-0038-1671755

[RfII 2016a] RfII: Begriffsklärungen No. 1. (2016). – URL `http://nbn-resolving.de/urn:nbn:de:101:1-201607146410`, Zugriffsdatum: 20.02.2019

[RfII 2016b] RfII: Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland. (2016). – URL `https://d-nb.info/1104292440/34`, Zugriffsdatum: 20.02.2019

[Rohr 2018] Rohr, Matthias: *Sicherheit von Webanwendungen in der Praxis*. Springer Vieweg, Wiesbaden, 2018

[Rothwell u. a. 2014] Rothwell, Erin ; Wong, Bob ; Rose, Nancy C. ; Anderson, Rebecca ; Fedor, Beth ; Stark, Louisa A. ; Botkin, Jeffrey R.: A Randomized Controlled Trial of an Electronic Informed Consent Process. In: *Journal of Empirical Research on Human Research Ethics* 9 (2014), Nr. 5, S. 1–7. – DOI 10.1177/1556264614552627

[Röhrig u. a. 2009] Röhrig, Bernd ; Prel, Jean-Baptist d. ; Blettner, Maria: Study Design in Medical Research. In: *Deutsches Ärzteblatt International* 106 (2009), Nr. 11, S. 184–189. – DOI 10.3238/arztebl.2009.0184

[Samarati und de Vimercati 2001] Samarati, Pierangela ; Vimercati, Sabrina C. de: Access Control: Policies, Models, and Mechanisms. In: Focardi, Riccardo (Hrsg.) ; Gorrieri, Roberto (Hrsg.): *Lecture Notes in Computer Science* Bd. 2171. Berlin, Heidelberg : Springer, 2001, S. 137–196. – DOI 10.1007/3-540-45608-2_3

[Sandhu u. a. 1996] Sandhu, Ravi S. ; Coyne, Edward J. ; Feinstein, Hal L. ; Youman, Charles E.: Role-based access control models. In: *Computer* 29 (1996), Nr. 2, S. 38–47. – DOI 10.1109/2.485845

[Schonschek 2019] Schonschek, Oliver: Schritt für Schritt zum Berechtigungskonzept / Datenschutz Praxis. (2019). – URL `https://www.datenschutz-praxis.de/fachartikel/berechtigungskonzept-die-wasserdichte-pruefung`, Zugriffsdatum: 21.11.2020

[Schreiweis 2016] Schreiweis, Björn: *Einrichtungsübergreifendes Patientenrekrutierungssystem - Konzeption und Systemarchitektur*, Ruprecht-Karls-Universität Heidelberg, Dissertation, 2016

[Schwanke 2015] Schwanke, Jens: *Konzeption und Entwicklung eines Probandenmanagementsystems am Beispiel der Universitätsmedizin Göttingen*, Georg-August-Universität Göttingen, Dissertation, 2015

[Schwanke u. a. 2012] Schwanke, Jens ; Buckow, Karoline ; Demiroglu, Sara ; Gusky, Linda ; Helbing, Krister ; Quade, Matthias ; Rienhoff, Otto: Eignung eines Customer-Relationship-Management-Systems als Probandenmanagementsystem für klinische Studien. In: *GMDS 2012. 57. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. (GMDS). Braunschweig, 16.-20.09.2012* DocAbstr. 211, Düsseldorf: German Medical Science GMS Publishing House, 2012. – DOI 10.3205/12gmds211

[Schwanke u. a. 2013a] Schwanke, Jens ; Nussbeck, Sara Y. ; Helbing, Krister ; Rienhoff, Otto: Entwicklung eines Probandenmanagementsystems für ein multizentrisches Forschungsprojekt. In: *GMDS 2013. 58. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. (GMDS). Lübeck, 01.-05.09.2013* DocAbstr. 151, Düsseldorf: German Medical Science GMS Publishing House, 2013. – DOI 10.3205/13gmds046

[Schwanke u. a. 2013b] Schwanke, Jens ; Rienhoff, Otto ; Schulze, Thomas G. ; Nussbeck, Sara Y.: Suitability of Customer Relationship Management Systems for the Management of Study Participants in Biomedical Research. In: *Methods of Information in Medicine* 52 (2013), Nr. 4, S. 340–350. – DOI 10.3414/ME12-02-0012

[Schwarzmaier 2014] Schwarzmaier, Martin: *Erweiterung eines Probandenmanagementsystems zur Generierung von übergreifenden Forschungspseudonymen*, Georg-August-Universität Göttingen, Bachelorarbeit, 2014

[SEI 2010] SEI: What Is Your Definition of Software Architecture? / Software Engineering Institute (SEI), Carnegie Mellon University. (2010). – URL `http://www.sei.cmu.edu/architecture/start/glossary/community.cfm`, Zugriffsdatum: 12.02.2020

[Seufert 2002] Seufert, Steffen E.: Der Entwurf strukturierter rollenbasierter Zugriffskontrollmodelle. In: *Informatik Forschung und Entwicklung* 17 (2002), S. 1–11. – DOI 10.1007/s004500100083

[Sommerville 2018] Sommerville, Ian: *Software Engineering.* 10., aktualisierte Auflage. Pearson Studium, München, 2018

[Soni u. a. 2017] Soni, Hiral ; Grando, Adela ; Murcko, Anita ; Bayuk, Mike ; Chandrashekar, Pramod ; Mukundan, Madhumita ; Abrams, Meredith ; Aliste, Marcela P. ; Hiestand, Megan ; Varkey, Julia ; Zhou, Wentao ; Horrow, Caroline ; Saks, Michael ; Sharp, Richard ; Whitfield, Mary J. ; Callesen, Mark ; Dye, Christy ; Chern, Darwyn: Current State of Electronic Consent Processes in Behavioral Health: Outcomes from an Observational Study. In: *AMIA Annual Symposium Proceedings 2017*, S. 1607–1616. – URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5977724/`

[Spilca 2020] Spilca, Laurentiu: *Spring Security in Action.* Manning Publications, Shelter Island, NY, 2020

[Stahmann u. a. 2014a] Stahmann, Alexander ; Bauer, Christian R. ; Schulze, Thomas G. ; Schwanke, Jens: Standortbezogene Datenzusammenführung in Verbundforschungsvorhaben durch ein Probandenmanagementsystem. In: *GMDS 2014. 59. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. (GMDS). Göttingen, 07.-10.09.2014* DocAbstr. 200, Düsseldorf: German Medical Science GMS Publishing House, 2014. – DOI 10.3205/14gmds118

[Stahmann u. a. 2014b] Stahmann, Alexander ; Bauer, Christian R. ; Schwanke, Jens: Concept for Linking De-identified Biomedical Research Data Using a Study Participant Management System. In: *Studies in Health Technology and Informatics* Bd. 205 e-Health - For Continuity of Care, 2014, S. 848–852. – DOI 10.3233/978-1-61499-432-9-848

[Starke 2015] Starke, Gernot: *Effektive Softwarearchitekturen.* 9., überarbeitete Auflage. Carl Hanser Verlag, München, 2015

[Sully u. a. 2013] Sully, Ben G. O. ; Julious, Steven A. ; Nicholl, Jon: A reinvestigation of recruitment to randomised, controlled, multicenter trials: a review of trials funded by two UK funding agencies. In: *Trials* 14 (2013), Nr. 166. – DOI 10.1186/1745-6215-14-166

[Thun und Dewenter 2017] Thun, Sylvia ; Dewenter, Heike: *E-Health-Ökonomie.* Kap. Syntaktische und semantische Interoperabilität, S. 669–682, Springer Gabler, Wiesbaden, 2017

[Tsolkas und Schmidt 2017] Tsolkas, Alexander ; Schmidt, Klaus: *Rollen und Berechtigungskonzepte.* Springer Vieweg, Wiesbaden, 2017

[Vanaken 2016] Vanaken, Hilde: eConsent Study Provides Insights to Shape Industry Adoption. In: *Applied Clinical Trials* 25 (2016), Nr. 8/9. – URL `http://www.appliedclinicaltrialsonline.com/econsent-study-provides-insights-shape-industry-adoption`

[Vernon 2013] Vernon, Vaughn: *Implementing Domain-Driven Design.* Addison-Wesley Professional, Upper Saddle River, NJ, 2013

[Vogel u. a. 2009] Vogel, Oliver ; Arnold, Ingo ; Chughtai, Arif ; Ihler, Edmund ; Kehrer, Timo ; Mehlig, Uwe ; Zdun, Uwe: *Software-Architektur.* Spektrum Akademischer Verlag, Heidelberg, 2009

[Walters u. a. 2017] Walters, Stephen J. ; Bonacho dos Anjos Henriques-Cadby, Inês ; Bortolami, Oscar ; Flight, Laura ; Hind, Daniel ; Jacques, Richard M. ; Knox, Christopher ; Nadin, Ben ; Rothwell, Joanne ; Surtees, Michael ; Julious, Steven A.: Recruitment and retention of participants in randomised controlled trials: a review of trials funded and published by the United Kingdom Health Technology Assessment Programme. In: *BMJ Open* 7 (2017), Nr. 3, S. e015276. – DOI 10.1136/bmjopen-2016-015276

[Weng u. a. 2013] Weng, Chunhua ; Li, Yu ; Berhe, Solomon ; Boland, Mary R. ; Gao, Junfeng ; Hruby, Gregory W. ; Steinman, Richard C. ; Lopez-Jimenez, Carlos ; Busacca, Linda ; Hripcsak, George ; Bakken, Suzanne ; Bigger, J. T.: An Integrated Model for Patient Care and Clinical Trials (IMPACT) to support clinical research visit scheduling workflow for future learning health systems. In: *Journal of Biomedical Informatics* 46 (2013), Nr. 4, S. 642–652. – DOI 10.1016/j.jbi.2013.05.001

[Wunder und Grosche 2009] WUNDER, Michael (Hrsg.) ; GROSCHE, Jürgen (Hrsg.): *Verteilte Führungsinformationssysteme*. Springer, Berlin, Heidelberg, 2009

Abkürzungsverzeichnis

ABAC	Attribute Based Access Control
API	Application Programming Interface
BDSG	Bundesdatenschutzgesetz
BMBF	Bundesministerium für Bildung und Forschung
BSI	Bundesamt für Sicherheit in der Informationstechnik
CA	Certifcate Authority
CDMS	Clinical Data Management System
CRF	Case Report Form
CRM	Customer-Relationship-Management
CRMS	Clinical Research Management System
CRO	Contract Research Organisation
CT	Computertomografie
CTMS	Clinical Trial Management System
CTRSS	Clinical Trial Recruitment Support System
DAC	Discretionary Access Control
DFG	Deutsche Forschungsgemeinschaft
DFN-Verein	Verein zur Förderung eines Deutschen Forschungsnetzes (DFN-Verein)
DMZ	Demilitarisierte Zone
DSA	Digital Signature Algorithm
DSGVO	Datenschutz-Grundverordnung
ECDSA	Elliptic Curve Digital Signature Algorithm
eCRF	Electronic Case Report Forms
EDC	Electronic Data Capture
GCP	Good Clinical Practice
GMDS	Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie
HSTS	HTTP Strict Transport Security

HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
ICF	Informed Consent Form
ICH	International Council for Harmonisation
ICL	Identifcation Code List
IDAT	Personenidentifzierenden Daten
IHE	Integrating the Healthcare Enterprise
IP	Internet Protocol
ISF	Investigator Site File
IT	Informationstechnik
JPA	Jakarta Persistence API (früher Java Persistence API)
KFO	Klinische Forschungsgruppe (Förderform der DFG)
LAN	Local Area Network
LIMS	Laboratory Information Management System
MAC	Mandatory Access Control
MDAT	Medizinischen Daten
PACS	Picture Archiving and Communication System
PID	Patientenidentifkator / Personenidentifkator
PKI	Public Key Infrastructure
PRS	Patient Recruitment System
PSN	Pseudonym
RBAC	Role Based Access Control
REST	Representational State Transfer
RFC	Request for Comments
RfII	Rat für Informationsinfrastrukturen
RSA	Rivest-Shamir-Adleman
SHA	Secure Hash Algorithm
SOAP	Simple Object Access Protocol
SOP	Standard Operating Procedure
SpEL	Spring Expression Language
SSL	Secure Sockets Layer
TLS	Transport Layer Security
TMF	Technologie- und Methodenplattform für die vernetzte medizinische Forschung

TOM	Technisch-organisatorischen Maßnahmen
TTP	Trusted Third Party
UML	Unified Modeling Language
URL	Uniform Resource Locator
WMA	World Medical Association

Abbildungsverzeichnis

2.1 Funktionsbereiche klinischer Studien 9
2.2 Verlauf einer klinischen Studie mit Probandenverwaltungsprozessen 16
2.3 Bestandteile einer Forschungs-IT-Infrastruktur mit Probandenverwaltung . 21

3.1 Prozessmodell Design and Creation 29
3.2 Netzwerksegmente und IT-Systeme für medizinische Forschungsprojekte am Institut für Medizinische Informatik 31
3.3 Module des TMF-Datenschutzleitfaden 39
3.4 Module des TMF-Datenschutzleitfadens mit Probandenmanagementsystem 41
3.5 Funktionsweise der Patientenliste 44
3.6 Prinzip mehrstufiger Pseudonymisierung 45
3.7 Daten und Datenflüsse im Studienkontext 47
3.8 Daten und Datenflüsse des Probandenmanagementsystems 49
3.9 Netzwerksegmente und IT-Systeme für medizinische Forschungsprojekte am Institut für Medizinische Informatik (erweitert) 50
3.10 Anwendungsfalldiagramm Schnittstelle zum Identitätsmanagement 55
3.11 Aktivitätsdiagramm Anlage eines Studienteilnehmenden im Probandenmanagementsystem 56
3.12 Sequenzdiagramm Pseudonymzuweisung 58
3.13 Sequenzdiagramm Pseudonymzuweisung und Pseudonymabfrage 60
3.14 Grundprinzip der rollenbasierten Zugriffskontrolle 70
3.15 Grundprinzip der attributbasierten Zugriffskontrolle 71
3.16 Mockup Übersicht für Studienkoordinierende im Probandenmanagementsystem 80
3.17 Mockup Übersicht für Studienmitarbeitende im Probandenmanagementsystem 81
3.18 Rollenmanagement für einen Standort 82
3.19 Rechtemanagement für einen Standort 82
3.20 Rechte- und Rollenmanagement für mehrere Standorte 83
3.21 Vergleich Vergabekonzepte für Rechte im Probandenmanagementsystem . . 85

4.1 Architekturmuster Ports and Adapters 103
4.2 Übersicht Systemarchitektur Probandenmanagementsystem 104
4.3 Übersicht Softwarestack Probandenmanagementsystem 105
4.4 Domänenmodell Bereich Pseudonymisierung 108

4.5 Entity-Relationship-Modell Identitätsmanagementmetadaten 109
4.6 Erweitertes Domänenmodell Bereich Pseudonymisierung 110
4.7 Tabellendiagramm project, research_system, external_system, identifier_domain und identifier 110
4.8 Screenshot Projektverwaltung Externe Systeme 111
4.9 Screenshot Projektverwaltung Forschungssysteme 111
4.10 Entity-Relationship-Modell Proband:in, Forschungsakte und -projekt . . . 112
4.11 Tabellendiagramm participant, record, participant 112
4.12 UML-Klassendiagramm Persistable, AbstractPersistable und Participant . 113
4.13 Entity-Relationship-Modell Proband:in, Pseudonym, Forschungsakte, -system und -projekt 114
4.14 Tabellendiagramm identifier, identifier_domain, research_system, external_system, participant, record und project 115
4.15 Screenshot Projektverwaltung Pseudonymdomänen 116
4.16 Tabellendiagramm identifier_domain, domain_interface_mapping und interface_identifier 117
4.17 UML-Klassendiagramm ExternalSystemProvider und Adapter 119
4.18 Sequenzdiagramm Kommunikationsablauf Mainzelliste 120
4.19 UML Klassendiagramm Mainzelliste-Identifier und -Patient 121
4.20 UML Klassendiagramm Mainzelliste Kommunikation 122
4.21 UML Klassendiagramm Mainzelliste Exceptions 123
4.22 UML Klassendiagramm MainzellisteAdapter 123
4.23 Konzeptuelles Datenmodell Zugriffskontrolle 133
4.24 UML Klassendiagramm der Entitäten zur Zugriffskontrolle im Probandenmanagementsystem 134
4.25 UML-Klassendiagramm UserDetails und SecurityUser 135
4.26 UML-Klassendiagramm UserDetailsService, UserDetails und GrantedAuthority 135
4.27 Hauptkomponenten des Authentifizierungsprozesses mit Spring Security . . 136
4.28 Screenshot Forschungsaktenübersicht 140

Listings

3.1 Liste der durch den Programmierer des Probandenmanagementsystems vorgegebenen Rechte zur Zugriffssteuerung der Systembenutzenden des Probandenmanagementsystems. 79
4.1 Programmcode-Auszug der Klasse *Participant*, die eine Participant-Entität beschreibt und im Datenmodell die Tabelle *participant* erzeugt. Die Klasse *Participant* erbt von der Klasse *AbstractPersistable*. Exemplarisch mitdargestellt das Pflichtfeld eines Probandendatensatzes *firstName* mit seinen Metadaten. 113
4.2 Auszug des Programmcodes der Klasse *AbstractPersistable*, die das Hibernate Interface *Persistable* implementiert und über die Variable *id* allen ableitenden Klassen einen Primärschlüssel erzeugt. 113
4.3 Parameter der Tomcat-Konfigurationsdatei /conf/server.xml um bestimmte IP-Adressen zulassen. X, Y, und Z, sind als Beispiele für beliebige Zahlen in einer IP-Adresse zu sehen. 125
4.4 Auszug des Programmcodes zur Überprüfung der Passwortstärke durch einen PasswordValidator mittels verschiedener Regeln der Passay-Bibliothek. 126
4.5 Auszug des Programmcodes zur Erzeugung von gehashten Passwörtern zur Ablage in der Datenbank des Probandenmanagementsystems. 127
4.6 Parameter der Tomcat-Konfigurationsdatei /conf/server.xml um Clientauthentifizierung zu aktivieren. 129
4.7 Kommandozeilenbefehl zum Kontrollieren der Zertifikatskonfiguration des Webservers. 129
4.8 Ausgabe des Befehls aus Listing 4.7 zur Kontrolle der Zertifikatskonfiguration 129
4.9 Kommandozeilenbefehl zum Erzeugen des CSR und dem privaten RSA-Schlüssel mit OpenSSL. 130
4.10 Parameter der Tomcat-Konfigurationsdatei /conf/server.xml um auf die TLS Versionen 1.2 und 1.3 einzuschränken. 131
4.11 Parameter der Tomcat-Konfigurationsdatei WEB-INF/web.xml um den Datenverkehr auf eine sicheren Verbindung zu erzwingen. 131
4.12 Parameter der Tomcat-Konfigurationsdatei /conf/web.xml um HSTS zu aktivieren. 132
4.13 Beispiel für die Nutzung der Spring Security Annotation @*RolesAllowed*. Zugriff auf die Methode wird nur gewährt, wenn die oder der aktuell eingeloggte Systembenutzende die Rolle *DATAMANAGER* besitzt. 136

4.14 Beispiel für die Nutzung der @*PreAuthorize* Spring Security Annotation mit der *hasPermission(...)* Expression. Der Ausdruck überprüft die Berechtigung *OPEN_RECORD* von Systembenutzenden zum Abrufen von Forschungsakten im Probandenmanagementsystem. 137
4.15 Beispiel für die Nutzung eines selbst definierten SpEL-Ausdrucks mit der der @*PreAuthorize* Spring Security Annotation. *isProjectMember(#researchProject)* prüft die Zugehörigkeit von Systembenutzenden zu einem Forschungsprojekt bevor die Änderungen am Projektnamen akzeptiert werden. 138
4.16 Programmcode der Methode *isRecordOpenAllowed()* der Service-Klasse *ResearchRecordService* zur Überprüfung des Zugriffsrechts *OPEN_RECORD* im Probandenmanagementsystem. 139
4.17 Programmcode der Methode *isIDATAccessAllowed(...)* der Service-Klasse *ResearchRecordService*. Es wird das Zugriffsrecht *SITE_IDAT* als auch die Berechtigung zum Zugriff auf den Forschungsstandort der Forschungsakte überprüft. 139
4.18 Programmcode-Auszug der Klasse *Revision*, die eine für das Probandenmanagementsystem angepasste *RevisionEntity* für die Datenänderungshistorie mit Hibernate Envers implementiert. Zu den erbenden Attributen der Klasse *DefaultRevisionEntity* wurde das Attribut *UserAccount* hinzugefügt. 141
4.19 Programmcode-Auszug der Klasse *Participant*, die eine Participant-Entität beschreibt. Exemplarisch mitdargestellt das Pflichtfeld eines Probandendatensatzes *firstName* mit seinen Metadaten, das durch die *@Audited*-Annotation der gesamten Klasse in die Änderungshistorie einbezogen wird. 142
4.20 Programmcode der Definition der LogAuditTrail-Annotation mit weiteren AspectJ-Annotationen. Die *@Target-Annotation* mit *ElementType.Method* gibt an, dass die Annotation nur auf Methoden angewendet werden darf. *@Retention* gibt an, ob die Annotation der Java Virtual Machine zur Laufzeit zur Verfügung steht. 143
4.21 Programmcode zur Verdeutlichung der Nutzung der LogAuditTrail-Annotation, hier für die Methode *retrieveIdentifier(...)* zum Abruf von Pseudonymen. 143
4.22 Programmcode des Datenbank Events zur Löschung von Logeinträgen, die älter als einen Monat sind. Das Event wird täglich durch den Event Scheduler ausgeführt. 144

Anhang A

JAMIA Open, 3(3), 2020, 449–458
doi: 10.1093/jamiaopen/ooaa041
Advance Access Publication Date: 14 October 2020
Review

Review

Key components and IT assistance of participant management in clinical research: a scoping review

Johannes Pung and Otto Rienhoff

Department of Medical Informatics, University Medical Center Göttingen, Göttingen, Germany

Corresponding Author: Johannes Pung, MSc, Department of Medical Informatics, University Medical Center Göttingen, Robert-Koch-Straße 40, 37075 Göttingen, Germany; johannes.pung@med.uni-goettingen.de

Received 12 December 2019; Revised 16 July 2020; Editorial Decision 16 August 2020; Accepted 24 August 2020

ABSTRACT

Objectives: Managing participants and their data are fundamental for the success of a clinical trial. Our review identifies and describes processes that deal with management of trial participants and highlights information technology (IT) assistance for clinical research in the context of participant management.
Methods: A scoping literature review design, based on the Preferred Reporting Items for Systematic Reviews and Meta-analyses statement, was used to identify literature on trial participant-related proceedings, work procedures, or workflows, and assisting electronic systems.
Results: The literature search identified 1329 articles of which 111 were included for analysis. Participant-related procedures were categorized into 4 major trial processes: recruitment, obtaining informed consent, managing identities, and managing administrative data. Our results demonstrated that management of trial participants is considered in nearly every step of clinical trials, and that IT was successfully introduced to all participant-related areas of a clinical trial to facilitate processes.
Discussion: There is no precise definition of participant management, so a broad search strategy was necessary, resulting in a high number of articles that had to be excluded. Nevertheless, this review provides a comprehensive overview of participant management-related components, which was lacking so far. The review contributes to a better understanding of how computer-assisted management of participants in clinical trials is possible.

Key words: clinical trials as topic, information systems, data management, research subjects, medical informatics

LAY SUMMARY

Clinical trials involving participants play a critical role in the development of new treatments. Management of participants and their data is fundamental for the success of a clinical trial. Information technology (IT) offers opportunities to approach clinical trial methodology in new ways, but so far there have been no comprehensive reviews of how computer-supported management of participants' data is disseminated in trial management processes. We identified 4 main process categories of participant-related management and evaluated their supporting IT systems: recruitment, administration of informed consent, management of participants' identities, and management of administrative participant data. Our results demonstrated that management of trial participants is considered in every step of clinical trials. Furthermore, we found that IT was successfully introduced to all participant-related areas of a clinical trial to facilitate processes. Our review contributes to a better understanding of how computer-assisted management of participants is possible. We anticipate our review to be a starting point for further research, providing a comprehensive overview of participant management-related components by categorizing existing literature in the field, and thereby giving an assessment of potential size and scope of available research literature.

© The Author(s) 2020. Published by Oxford University Press on behalf of the American Medical Informatics Association.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

INTRODUCTION

Clinical research is a key component of medical progress. Clinical trials are used to test efficacy and safety of drugs, new forms of treatment, medical interventions, or medical devices. After the trial planning and design phase, it is necessary to recruit volunteers as participants and inform them about the planned research with its aims, potential risks, and advantages. Tasks of the staff at a study site include the administration of contact data and obtaining informed consent of participants. Furthermore, scheduling and organization of examinations in the study context have to be managed. At many stages of the starting clinical trial, collaboration of the participants is necessary. The management of participants is, therefore, a core component of every successful person-related research project.

The organization and execution of modern clinical trials is not possible without the use of IT-assisted procedures.[1–3] Clinical Data Management Systems have become an accepted part of clinical trial data management and are considered efficient in terms of data management requirements.[4–8] For other trial management processes, such as management of participants, it is likely that new technologies to facilitate and accelerate workflows are used as well. However, to date, there have been no comprehensive reviews of how computer-supported management of participants is disseminated in trial management processes, and it is currently unknown what evidence exists to substantiate the presumed benefits of a computer-supported participant management.

This scoping review aimed to (1) identify and describe processes that deal with management of trial participants and (2) highlight IT assistance for clinical research in the context of participant management along the process chain. It provides a broad overview of published research regarding the management of participants for clinical trials, providing a basis for an in-depth analysis.

METHODS

We followed the *Preferred Reporting Items for Systematic Reviews and Meta-analyses* (PRISMA) guidelines[9] for searching and selecting relevant publications. In 4 phases, articles were first identified by a literature database search, then pre-selected based on title and abstract, checked for suitability by means of the entire text, and finally the validated publications were included in the literature analysis.

Search strategy

A systematic search of literature regarding *participant management in clinical research* is challenging, due to a lack of consistently used standard terms. Medical Subject Headings, which is the controlled vocabulary thesaurus of the United States National Library of Medicine (NLM), defines "Persons who are enrolled in research studies or who are otherwise the subjects of research" as *human subjects* or as its preferred term *research subjects*.[10] On the other hand, the website ClinicalTrials.gov, a database for clinical studies conducted around the world, and also provided by NLM, defines in their protocol registration data element definitions, that the term *participant* is used to refer to *human subjects*.[11] When we started with a broad internet search to identify further keywords and indexing terms, we have seen the use of the terms *subject*, *study-*, *trial-*, *research-participant* or *-patient*, and *-volunteer* to describe someone who enrolls in a clinical research study. Which term is most appropriate will not be discussed here, but it shows that many keywords have to be included in the search strategy.

From the keywords we identified, we developed a database search strategy (Table 1). Databases were searched using a combination of (1) keywords for process flows in the management of participants, (2) *participant* and similar search terms, and lastly (3) *clinical trial* and synonyms. Additionally, we included keywords for electronic workflows that we already assigned to the topic of participant management, like *electronic informed consent*, *recruitment system*, or *pseudonymization services*. The challenge was to include all relevant keywords but avoid generic search terms such as *informatics* or *data management*, which yielded a too massive result as also off-topic literature was included. We searched a medical and a multidisciplinary science database to find relevant articles for this study: PubMed and Web of Science. The search strategy was mapped to the respective search query syntax and rules of the databases (Supplementary Appendix SA). As search categories, [Title/Abstract] in PubMed and [TS=Topic] in Web of Science were chosen to keep the search broad and include all relevant literature. The search was conducted at the end of July 2019.

In addition, we looked for practice guidelines and available software products, and searched sources of gray literature. The gray literature search was based on the methodology outlined by Godin et al[12] for which the same keywords were used as for the systematic literature search, but the search was customized to the syntax of the Google search engine.

Selection of articles and synthesis of results

The literature selection process according to PRISMA is shown in Figure 1. Web of Science and PubMed searches identified the potentially relevant articles. A reference management software was used to manage the results of the literature search and the subsequent selection process of relevant articles. After removing duplicates from the results, the articles were screened by titles and abstracts. Articles were kept for further analysis, if they met one or more of the following 3 inclusion criteria:

- Processes: We included articles that described processes, work procedures, or workflows from clinical trials that are trial participant-related. We excluded studies that only mentioned the use of participant management but did not describe how it was conducted. Processes in clinical trials vary depending on the design and aim of the study. Therefore, only procedures common across trials are taken into account.
- IT assistance: We were interested in articles that described or evaluated electronic systems used in an administrative context with trial participants. This included systems that were either standalone systems or software as part of a bundled system.
- Publication type: We were looking for systematic reviews and original articles in English.

Articles that did not meet the inclusion criteria were excluded. Articles meeting the inclusion criteria and, thereby, considered relevant for this review, were read and tagged with key terms of content to classify articles regarding the clinical trial processes that were covered. This classification was used to identify the main trial processes. In addition, gray literature allowed us to identify topics that are not covered by scientific publications.

RESULTS

The searches from Web of Science and PubMed yielded 1321 potentially relevant articles. Of those, 175 were selected for full-text screening, resulting in a total of 103 articles to be finally included in

Table 1. Database search strategy

1 ("participant management" OR "trial patient management" OR "study patient management" OR "volunteer management" OR "subject management" OR "trial management" OR "research management" OR "study management" OR "workflow management" OR "task management" OR "project management" OR "identity management" OR "pseudonymization service*" OR "pseudonymisation service*" OR "trusted third party" OR "consent management" OR "electronic consent" OR "electronic informed consent" OR "digital consent" OR "digital informed consent" OR "econsent" OR "e-consent" OR "participant management system*" OR "patient management system*" OR "recruitment system*" OR "enrollment system*" OR "eClinical")
AND (
2 (participant* AND ("clinical study" OR "clinical studies" OR "clinical trial*" OR "clinical research") OR patient* AND (study OR studies OR trial* OR research)) OR volunteer* AND ("clinical study" OR "clinical studies" OR "clinical trial*" OR "clinical research") OR "human subject*" OR "research subject*")
3 OR ("clinical trial*" OR "clinical research" OR "clinical study" OR "clinical studies")
)

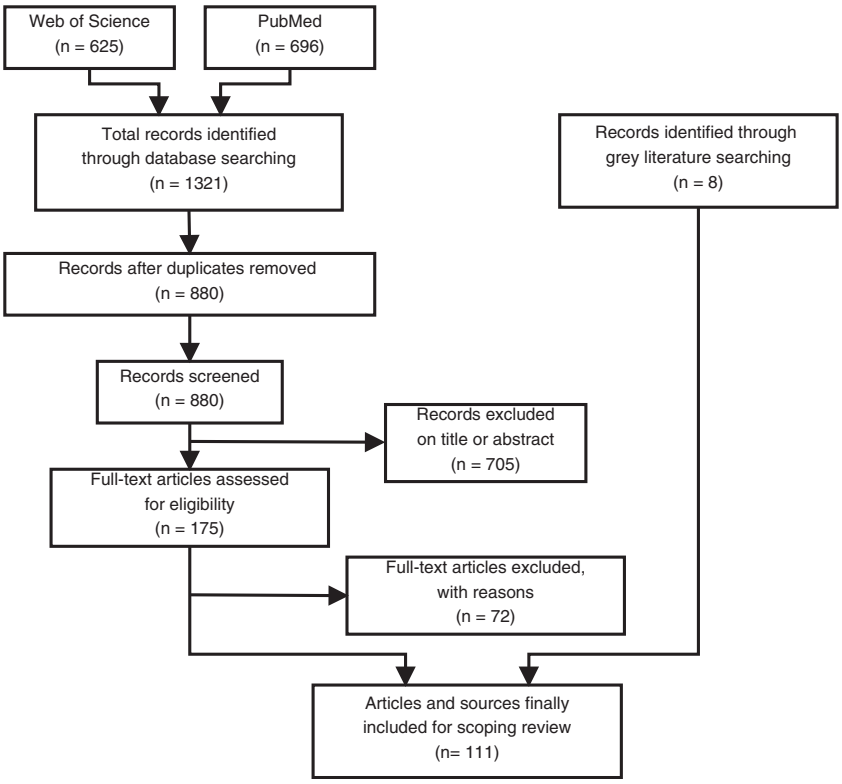

Figure 1. Flow diagram of the literature selection process.

the review. Most of the excluded articles were not within our scope, did not describe the process of a clinical study but routine care, or described managing of patients or participants in an unrelated context for this review. In addition, 8 results of the gray literature search were included in the review,[13–20] which provided information that were not accessible via scientific publications.

The screened articles were grouped into categories that represent the processes related to management of trial participants throughout a clinical trial execution, from concept development to trial result reporting. In Figure 2, we summarized these processes and highlight possible IT assistance for clinical research in the context of participant management along the process chain. Management of trial participants needs to be considered in nearly every step of the clinical trial—from recruitment of eligible participants and obtaining informed consent, to planning their participation by scheduling tasks and exams, to managing their contact data to inform them about study results and possible follow-up periods.

The characteristics of the reviewed articles are presented in Table 2. Out of the 103 selected articles, 36 concerned recruitment workflows, 36 dealt with informed consent processes during enrollment, 12 described pseudonymization workflows during enrollment, and 19 articles dealt with processes during trial participation. As shown in Table 2, articles considered different types of IT assistance: for example, recruitment support system, digital consent management, and identity management.

The selected literature addresses the various aspects of participant involvement in clinical trial processes. In the following chapters, we will reflect each process within the clinical trial workflow taken from the grouped articles with background information on the execution and how IT assistance can contribute.

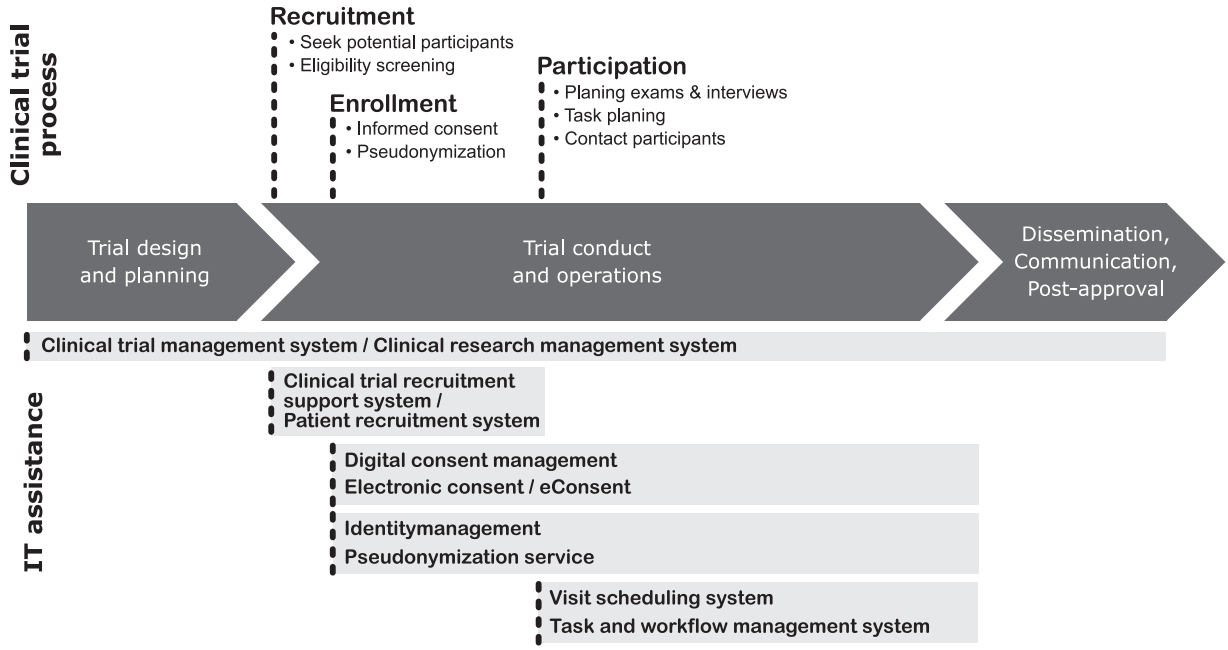

Figure 2. Clinical trial management process from start to finish, highlighting processes with trial participant involvement over time. IT systems provide support for the corresponding processes.

Table 2. Characteristics of reviewed articles

Article	Trial process	IT assistance
Afrin et al,[21] Ahmad et al,[22] Breitfeld et al,[23] Butte et al,[24] Cuggia et al,[25] Dowling et al,[26] Dugas et al,[27,28] Ferranti et al,[29] Fink et al,[30] Grundmeier et al,[31] Harris et al,[32] Heinemann et al,[33] Khosropour et al,[34] Kost et al,[35] Kotoulas et al,[36] Köpcke et al,[37] Lagor et al,[38] McDonald et al,[39,40] Mattingly et al,[41] Nielsen et al,[42] Rollman et al,[43] Schmickl et al,[44] Schreiweis et al,[45–47] Straube et al,[48] Sully et al,[49] Thadani et al,[50] Thompson et al,[51] Trinczek et al,[52,53] Treweek et al,[54] Walters et al,[55] Zimmerman et al[56]	Recruitment	Recruitment support system
Bergmann et al,[57] Bethune et al,[58] Bialke et al,[59] Boutin et al,[60] Buckley et al,[61] Chen et al,[62] Chhin et al,[63] Doerr et al,[64] Fink et al,[65] Grady et al,[66] Greenhalgh et al,[67] Harle et al,[68,69] Haussen et al,[70] Iafrate et al,[71] Kim et al,[72] Kondylakis et al,[73] Nijhawan et al,[74] Phillippi et al,[75] Ramos,[76] Rothwell et al,[77] Schreiweis et al,[78] Shelton,[79] Simon et al,[80–82] Sommer et al,[83] Soni et al,[84] St John et al,[85] Stevens et al,[86] Suarez et al,[87] Vanaken,[88] Vanaken and Masand,[89] Warriner et al,[90] Welch et al,[91] Wilbanks[92]	Enrollment	Electronic informed consent, digital consent management
Aamot et al,[93] Bialke et al,[94] Bruland et al,[95] Chevrier et al,[96] Deserno et al,[97] Ebner et al,[98] Jonas et al,[99] Lablans et al,[100] Lautenschläger et al,[101] Nurmi et al,[102] Sahi et al,[103] Schwaneberg et al[104]	Enrollment	Identity management, pseudonymization service
Abshire et al,[105] Almeida et al,[106] Berard et al,[107] Bose and Das,[108] Campion et al,[109] Cramon et al,[110] Durkalski et al,[111] Geyer et al,[112] Gupta et al,[113] Leroux et al,[114] Müller et al,[115] Nadkarni et al,[116] Park et al,[117] Raptis et al,[118] Schobel et al,[119] Schwanke et al,[120] Solomon et al,[121] Weng et al,[122] Zhao and Pauls[123]	Participation—managing administrative data	Task- and workflow management system, visit scheduling system, clinical trial management system, clinical research management system

Recruitment

Patient recruitment describes the process of finding suitable participants for clinical trials. Potential trial participants are usually identified by the treating physician, the study nurse, or the principal investigator. Conventional clinical trials often occur within the confines of health care settings, through the identification of participants during their clinical encounters, or prior to visits via electronic and paper healthcare records. Other channels to recruit participants are used as well such as advertisements, phone calls, use of third-party data, processing information from clinical trial registries or patient registries. Prior to formal enrollment in clinical trials, patients, interested in participating, will go through a screening process. The eligibility criteria, defined in the study protocol, describe characteristics that must be met by all participants.

A common problem with clinical trials is the achievement of the recruitment target. Trials often fail to achieve the required recruitment numbers or to recruit the necessary number of patients within the planned recruitment time.[39] Between 50% and 60% of random-

Table 3. e-Consent management divided into 3 levels, describing the options of collecting and using structured data in e-consent

Level	Type	Structure	Explanation
Level I	Paper consent form	No structured data	Consent on a paper form, which is scanned and kept electronically
Level II	Paper and e-consent form	Some structured data	Consent on a paper form, which is transferred into an electronic format (manually or by automatic parsing of paper-based consent scans)
Level III	e-consent form	Structured data	Consent in an electronic form. IT systems can interpret and process participants consent decisions from structured data

Note: In accordance with ref.[20]

ized clinical trials do not meet their original recruitment targets or face significant delays.[49,55] Poor recruitment can lead to the prolongation of trials, increasing the costs, and can lead to an underpowered study from which wrong conclusions may be drawn. Several potential limiting factors have been identified in the literature, including organizational, financial, or trial management-related difficulties.[40,48] Barriers to patients' involvement include lack of knowledge, general lack of trust in trials, and disagreement with the assignment to a certain treatment group.[40] There are several proposals how to increase recruitment rates in trials. A recent systematic review of methods to improve recruitment numbers in randomized controlled trials compared 68 recruitment strategies.[54]

IT assistance offers the potential to improve the recruiting processes and meet the recruitment target. Clinical Trial Recruitment Support Systems (CTRSS), or sometimes called Patient Recruitment Systems, have been designed to support the patient recruitment process by suggesting potentially suitable study participants. These CTRSS use electronic patient data, typically from electronic health records, to assess patient eligibility for a clinical trial. The system alerts of a patients' study eligibility or provides a list of potential trial participants to a study investigator. Our search provided a total of 30 relevant articles on CTRSS,[21–38,41–47,50–53,56] which indicates that CTRSS are potentially beneficial in managing trial participants.

A comprehensive review of employing IT for the recruitment for clinical trials was done by Köpcke and Prokosch.[37] Compared to a previous review paper by Cuggia et al,[25] which analyzed 28 CTRSS from articles published before October 2009, Köpcke and Prokosch reviewed 101 papers on 79 different systems and created an overview of all CTRSS reported until the end of 2013. They distinguished between 3 types of CTRSS: (1) systems for the retrospective identification of trial participants based on existing clinical data, (2) systems that monitored the appearance of a key event of an existing health IT component in which the occurrence of the event triggered an eligibility test of a patient, and (3) independent systems that required a user to enter patient data into an interface to trigger an eligibility assessment.

While in scientific publications the non-commercial part of clinical research is mainly considered, there is also an industrial view of patient recruitment. For profit-oriented companies—such as pharmaceutical and medical device enterprises—a short time to market is crucial. Accelerating patient recruitment for clinical trials with CTRSSs can be one effective component to reduce time until revenues are generated. The patient recruitment industry is valued at a total of $19 billion per year.[15] Some companies outsource their patient recruitment-related needs to specialty service providers.[14] IT has impacted this branch as well. Several companies have built proprietary software for a range of activities that rely on patient recruitment. The website capterra.com, an online peer-review site that aims to help businesses find software solutions, lists 36 products for *Recruiting Management* in the category *Clinical Trial Software*.[13]

Enrollment—informed consent

Before a participant can enroll in a clinical trial, they must be recruited, screened, and must give their informed consent.[16] Informed consent is one of the founding principles of modern research ethics. It is a process by which "a subject voluntarily confirms his or her willingness to participate in a particular trial, after having been informed of all aspects of the trial that are relevant to the subject's decision to participate," as stated in the International Council for Harmonisation's (ICH) Good Clinical Practice guidelines (GCP).[17] Informed consent is not only required for clinical trials, but is an essential prerequisite before enrolling a participant in any type of research involving human subjects.[74] The process of obtaining informed consent is tightly regulated. Requirements are defined in federal laws and regulations as well as in a worldwide harmonization approach, the ICH-GCP guidelines, and the guidelines for Data Protection of the Organisation for Economic Co-operation and Development.[17,18] Informed consent is documented by means of a written, signed, and dated informed consent form (ICF). In accordance with GCP, the signed ICF has to be stored in an Investigator Site File.[17]

In recent years, there has been a greater use of electronic methods to gain informed consent (called eIC or e-consent) in research studies.[57–65,67–92] The New England Journal of Medicine published in "The Changing Face of Informed Consent" innovative approaches to improve and expand the electronic informed consent process for researchers and participants.[66] e-Consents can include multimedia information (graphics or videos) and interactive components to increase the understanding of the study purpose and is less focused on signing legal documents. e-Consents can be divided into 3 levels of consent management maturity; from plain paper-based consent to combining paper and electronic means to solely relying on e-consent (Table 3).[20] e-Consent management solutions enable to manage a large number and variety of research projects in a hospital due to process automation.[59]

Positive results using e-consent are increasingly reported by clinical studies, in which electronic management is perceived by, both, participants and staff members as straightforward, efficient, and simplifying workflows.[77,88] As a reaction to the increasing use, the Food and Drug Administration and the Office of Human Research Protections published a joint guidance for the use of e-consent for institutional review boards, investigators, and sponsors in December 2016.[19] e-Consent has been accepted by central institutional review boards in the United States and some other countries. However, launching of e-consent is challenging in international trials as each country has different legal requirements for the use of e-consent.[89]

Enrollment—identity management and pseudonymization service

Preserving participants' privacy is crucial in clinical research.[102] *De-identification* and *anonymization* are the 2 most common terms

used to refer to the technical approaches that protect privacy.[96] However, anonymous data collection are only possible in very few clinical trials. Many clinical study designs require to capture and integrate data from heterogeneous data sources or from different locations, which requires the identification of participants. Here, one of the most important requirements is keeping acquired clinical data and personal information of participants apart to ensure confidentiality and data protection.[98,100,101] Therefore, an identification code (pseudonym) is assigned to each participant upon enrollment. The identification code is saved in a confidential list [subject identification code list (ICL)], also including the name of the participant, the date informed consent was obtained, and all additional information needed to relate a participant to their identification code (eg, gender, date of birth). Clinical data, documented in case report forms, are encoded with the identification code and cannot be directly related to the participant's identity. Only by means of the ICL can the acquired clinical data be connected to identifying information. Linking data and identity, however, is a necessity, as stated in the GCP guidelines,[17] to allow contacting participants in case medical irregularities during a trial emerge.

The described workflow is often conducted paper-based, especially in local (ie, single site) and small trials. However, the procedure quickly reaches its limits when dealing with larger projects or greater technical requirements. Technical requirements increase, for example, for collaborative acquisition of data (eg, in multi-site studies), which has become a core element of biomedical research over the last years. To advance data management and improve data quality in clinical trials, electronic data capture (EDC) systems are increasingly used.[4–8] Integrating pseudonymization services into EDC systems is, therefore, a chance to improve and facilitate the described procedure by making use of IT infrastructures.[95,97,104] In this context, digital solutions that are compliant with data protection legislation are needed to ensure (1) that participants who appear in multiple institutions obtain the same pseudonym across sites and are not handled as 2 different individuals (identity management) and that (2) unambiguous pseudonyms are assigned for each participant and each separate system (pseudonymization service). A comprehensive overview of current methods and applications for pseudonymization services can be found in the literature.[93,103] For identity management and pseudonymization services, a trusted third party (TTP) is often assigned to handle the ICLs.[94] Participants can then only be identified by the person in charge of the identifying data. A recent study has investigated a novel approach to pseudonymization without the necessary use of a TTP, but requires a higher accountability of the participant.[99]

Managing administrative participant data

Commitment of study participants is essential to ensure the power and internal validity of longitudinal research. One of the most commonly used strategies to avoid losing participants are reminder, contact, and regular scheduling methods.[105] The participant's contact information is securely stored at each clinical site for internal use during the study. The information is collected on a so-called *locator form*—a working document of all contact information that can possibly help the researcher team find the participant when needed.[121] At the end of the study, all records will be kept in a secure location for as long as dictated by local regulations. The clinical research staff is in charge of scheduling appointments for trial visits and may send reminders to participants. Depending on the trial protocol, the trial visits could include lab tests, X-rays, computed tomography scans, physical exams, etc.

Using a task and workflow management system for clinical studies, all the processes associated with a health research study can be simplified.[106] These systems support the setup of clinical trials, the management of participants, as well as the overall governance process. Electronic clinical research visit scheduling systems provide the potential of coordinating trial visits with patient care visits. Thereby, the efficiency at clinical sites as well as the likelihood of participants keeping their trial appointments can be increased.[122] Unnecessary or redundant visits or tests for patients, and a considerable administrative burden for involved institutions can, thus, be avoided. The ability to manage participant schedules displayed in calendar form can be a feature of a Clinical Trial Management System (CTMS).[117] Managing administrative participant data in software systems is often not a standalone application. CTMS[107,108,110–112,114,117–119,123] or Clinical Research Management Systems (CRMS)[109,115,116] support the whole clinical trial process. Although the functional scope of a CTMS or a CRMS is larger and varies from product to product, these systems also manage trial participants and support tasks in the administration of study participants. Systems that were intended for a different purpose, however, can be used for the management of study participants as well, such as custom-build case management systems for participant data[113] or customer relationship management systems for the management of study participants.[120]

DISCUSSION

Trial participant management in its generic concept stands for a number of processes that handle participants and their data in clinical research. However, there is no precise definition of participant management as a distinct subset of processes in clinical research. For the purposes of this review, we selected articles that either describe processes that deal with management of trial participants, or in which IT assistance for clinical research in the context of participants is used. This made a broad search strategy necessary, including many keywords (Table 1). In our opinion, using the determined terms was sufficient to provide a broad range of participant-related processes and workflows. Disadvantage of the broad search was the high number of articles ($n = 656$) that had to be excluded. Many of the excluded articles described clinical trials on specific diseases and focused on their results, but were only casually mentioning the participant managing processes. Based on the included articles, we identified the key components of participant management in clinical research, divided into 4 processes: recruitment, obtaining informed consent, managing identities, and managing administrative data. We found that IT was successfully introduced to all participant-related areas of clinical trials. However, our review is limited insofar that the processes identified are based on the literature and information publically accessible. In our experience, clinical trials use far more IT assistance than is mentioned in the available literature on clinical trials. Most medical publications, however, describe little about methods or systems that have supported the data collection in an operational context. Additionally, traditional methodologies that rely on paper-based processes are also rarely published. We intended to emphasize the benefits of using IT-based solutions and technologies, but similar to missing process descriptive literature, studies comparing paper-based and IT-assisted management are also lacking.

A good overview of research activities was found for the recruitment process of clinical trials. It shows the importance of the topic for medical research—patient recruitment is vital to the success of any trial. The use of CTRSSs dates back to the early 1990s. Since then, CTRSSs have become even more popular and numerous publi-

cations have appeared. Although CTRSSs have been established, some of the problems in recruitment still persist, which require further research.

The combination of organizational, legal, ethical, and technical approaches is necessary to protect health data. The first and most important basis is established by gaining informed consent. Working with pseudonymous data is another important element to protect privacy and confidentiality. Nevertheless, a central question that researchers have to explore is how to protect privacy and confidentiality in constantly changing clinical research practice, technical possibilities, and legislation. For instance, getting ethical approval for studies varies enormously across European countries.[124]

After conducting the literature review, we noted that few publications deal with the managing of administrative participant data in a clinical trial. This is surprising as these processes must be conducted in every clinical trial. It can be assumed that administrative patient management processes are often managed with the common office IT tools.

IT can assist clinical research in the context of participant management. However, the degree of digitalization depends on the respective clinical trial. While the biopharmaceutical industry often uses more complex software solutions to manage the whole process of a clinical trial, the non-commercial clinical study research is generally less well equipped.[8,125–127] This is also reflected in the fact that commercial software products were not mentioned at all in scientific publications, although there is a strong market for these products.

Attention has been paid to each participant-related topic individually in the literature, but a synthesis addressing all aspects of participant management in clinical research was lacking so far. This review provides a comprehensive overview of participant management-related components.

CONCLUSION

Conducting a clinical trial typically implies management processes handling participants and their data. Participant-related processes include recruitment, administration of informed consent, management participants' identities with pseudonyms, and the management of administrative participant data. Effective use of available technologies is a crucial advantage. This article categorizes existing literature in the field, and thereby gives an assessment of potential size and scope of available research literature. It contributes to a better understanding of how computer-supported management of participants is disseminated in trial management processes. Further research should focus on availability, interconnectivity, and costs of IT assistance for participant management presented here. Additionally, the comparison of various participant management methods and their impact on the trial success as well as the comparison of management in different trial types should find more attention in future studies.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. This publication was supported by the Open Access Publication Funds of the Göttingen University.

AUTHOR CONTRIBUTIONS

JP and OR developed the study design. JP was responsible for data acquisition, analysis, and interpretation, and wrote the first draft of the manuscript. OR participated in the revision and approved the final version of the manuscript.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Krishnankutty B, Bellary S, Kumar NBR, *et al*. Data management in clinical research: an overview. *Indian J Pharmacol* 2012; 44 (2): 168–72.
2. Bhatti P, Schemitsch EH, Bhandari M. Managing data in surgical trials: a guide to modern-day data management systems. *J Bone Joint Surg Am* 2012; 94 (Suppl 1): 45–8.
3. Payne PRO. The clinical research environment. In: Richesson RL, Andrews JE, eds. *Clinical Research Informatics*. Cham: Springer International Publishing; 2019: 27–47.
4. Sahoo U, Bhatt A. Electronic data capture (EDC)—a new mantra for clinical trials. *Quality Assurance* 2004; 10 (3–4): 117–21.
5. El Emam K, Jonker E, Sampson M, *et al*. The use of electronic data capture tools in clinical trials: web-survey of 259 Canadian trials. *J Med Internet Res* 2009; 11 (1): e8.
6. Schrimpf D, Haag M, Pilz LR. Possible combinations of electronic data capture and randomization systems. *Methods Inf Med* 2014; 53 (3): 202–7.
7. Ohmann C, Kuchinke W, Canham S, *et al*.; the ECRIN Working Group on Data Centres. Standard requirements for GCP-compliant data management in multinational clinical trials. *Trials* 2011; 12 (1): 85.
8. Ohmann C, Speer R. Klinische Studien. In: Drepper J, Semler SC, eds. *IT-Infrastrukturen in Der Patientenorientierten Forschung Aktueller Stand Und Handlungsbedarf - 2016*. Berlin: Akademische Verlagsgesellschaft AKA GmbH; 2016: 25–47.
9. Moher D, Liberati A, Tetzlaff J, *et al*.; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; 6 (7): e1000097.
10. Medical Subject Headings (MeSH): Heading: Research Subjects, Unique ID: D035842. https://meshb.nlm.nih.gov/record/ui?ui=D035842 Accessed July 26, 2019.
11. ClinicalTrials.gov: Protocol Registration Data Element Definitions for Interventional and Observational Studies. https://prsinfo.clinicaltrials.gov/definitions.html Accessed July 26, 2019.
12. Godin K, Stapleton J, Kirkpatrick SI, *et al*. Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada. *Syst Rev* 2015; 4 (1): 138.
13. capterra.com. https://www.capterra.com/clinical-trial-management-software/?utf8=%E2%9C%93&feature%5B7%5D=30992&commit=Apply+Filters Accessed September 05, 2019.
14. PR News Wire. Patient Recruitment and Retention Services Market, 2019-2030 2019. https://www.prnewswire.com/news-releases/patient-recruitment-and-retention-services-market-2019-2030-300834895.html Accessed September 05, 2019.
15. PR News Wire. Outcome Health Introduces Clinical Trial Solution to Boost Patient Recruitment 2017. https://www.prnewswire.com/news-releases/outcome-health-introduces-clinical-trial-solution-to-boost-patient-recruitment-300395346.html Accessed September 05, 2019.
16. EUPATI - European Patients' Academy on Therapeutic Innovation. Enrolling in clinical trials 2015. https://www.eupati.eu/clinical-development-and-trials/enrolling-clinical-trials/ Accessed September 05, 2019.

17. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). Guideline for Good Clinical Practice ICH E6 (R2), 2nd ed. 2016. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R2__Step_4_2016_1109.pdf Accessed January 25, 2019.
18. The Organisation for Economic Co-operation and Development (OECD). Guidelines on the Protection of Privacy and Transborder Flows of Personal Data: The OECD Privacy Framework 2013. http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf Accessed January 29, 2019.
19. U.S. Department of Health and Human Services Office for Human Research Protections (OHRP) Food and Drug Administration (FDA). Use of Electronic Informed Consent. Questions and Answers: Guidance for Institutional Review Boards, Investigators, and Sponsors 2016. https://www.fda.gov/downloads/drugs/guidances/ucm436811.pdf Accessed January 29, 2019.
20. MITRE Corporation. Electronic Consent Management: Landscape Assessment, Challenges, and Technology, 1st ed. 2014. https://www.healthit.gov/sites/default/files/privacy-security/ecm_finalreport_forrelease62415.pdf Accessed February 16, 2019.
21. Afrin LB, Oates JC, Boyd CK, *et al.* Leveraging of open EMR architecture for clinical trial accrual. *AMIA Annu Symp Proc* 2003; 2003: 16–20.
22. Ahmad F, Gupta R, Kurz M. Real time electronic patient study enrollment system in emergency room. *AMIA Annu Symp Proc* 2005; 2005: 881.
23. Breitfeld PP, Ullrich F, Anderson J, *et al.* Web-based decision support for clinical trial eligibility determination in an international clinical trials network. *Control Clin Trials* 2003; 24 (6): 702–10.
24. Butte AJ, Weinstein DA, Kohane IS. Enrolling patients into clinical trials faster using RealTime Recuiting. *Proc AMIA Symp* 2000; 111–5.
25. Cuggia M, Besana P, Glasspool D. Comparing semi-automatic systems for recruitment of patients to clinical trials. *Int J Med Inform* 2011; 80 (6): 371–88.
26. Dowling NM, Olson N, Mish T, *et al.* A model for the design and implementation of a participant recruitment registry for clinical studies of older adults. *Clinical Trials* 2012; 9 (2): 204–14.
27. Dugas M, Amler S, Lange M, *et al.* Estimation of patient accrual rates in clinical trials based on routine data from hospital information systems. *Methods Inf Med* 2009; 48 (03): 263–6.
28. Dugas M, Lange M, Berdel WE, *et al.* Workflow to improve patient recruitment for clinical trials within hospital information systems—a case-study. *Trials* 2008; 9 (1): 2.
29. Ferranti JM, Gilbert W, McCall J, *et al.* The design and implementation of an open-source, data-driven cohort recruitment system: the Duke Integrated Subject Cohort and Enrollment Research Network (DISCERN). *J Am Med Inform Assoc* 2012; 19 (e1): e68–75.
30. Fink E, Hall LO, Goldgof DB, eds., *et al.* Experiments on the automated selection of patients for clinical trials. In: SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme—System Security and Assurance (Cat. No.03CH37483) 2003; Washington, DC.
31. Grundmeier RW, Swietlik M, Bell LM. Research subject enrollment by primary care pediatricians using an electronic health record. *AMIA Annu Symp Proc* 2007; 2007: 289–93.
32. Harris PA, Scott KW, Lebo L, *et al.* ResearchMatch: a national registry to recruit volunteers for clinical research. *Acad Med* 2012; 87 (1): 66–73.
33. Heinemann S, Thuring S, Wedeken S, *et al.* A clinical trial alert tool to recruit large patient samples and assess selection bias in general practice research. *BMC Med Res Methodol* 2011; 11 (1): 16.
34. Khosropour CM, Dombrowksi JC, Hughes JP, *et al.* Evaluation of a computer-based recruitment system for enrolling men who have sex with men into an observational HIV behavioral risk study. *Am J Epidemiol* 2016; 184 (6): 477–83.
35. Kost RG, Corregano LM, Rainer T-L, *et al.* A data-rich recruitment core to support translational clinical research. *Clin Transl Sci* 2015; 8 (2): 91–9.
36. Kotoulas A, Lambrou G, Koutsouris D-D. Design and virtual implementation of a biomedical registry framework for the enhancement of clinical trials: colorectal cancer example. *BMJ Health Care Inform* 2019; 26 (1): e100008.
37. Köpcke F, Prokosch H-U. Employing computers for the recruitment into clinical trials: a comprehensive systematic review. *J Med Internet Res* 2014; 16 (7): e161.
38. Lagor C, Aronsky D, Fiszman M, *et al.* Automatic identification of patients eligible for a pneumonia guideline: comparing the diagnostic accuracy of two decision support models. *Stud Health Technol Inform* 2001; 84 (Pt 1): 493–7.
39. McDonald AM, Knight RC, Campbell MK, *et al.* What influences recruitment to randomised controlled trials? A review of trials funded by two UK funding agencies. *Trials* 2006; 7 (1): 9.
40. McDonald AM, Treweek S, Shakur H, *et al.* Using a business model approach and marketing techniques for recruitment to clinical trials. *Trials* 2011; 12 (1): 74.
41. Mattingly WA, Kelley RR, Wiemken TL, *et al.* Real-time enrollment dashboard for multisite clinical trials. *Contemp Clin Trials Commun* 2015; 1: 17–21.
42. Nielsen JS, Thomsen RW, Steffensen C, *et al.* The Danish Centre for Strategic Research in Type 2 Diabetes (DD2) study: implementation of a nationwide patient enrollment system. *Clin Epidemiol* 2012; 4 (Suppl 1): 27–36.
43. Rollman BL, Fischer GS, Zhu F, *et al.* Comparison of electronic physician prompts versus waitroom case-finding on clinical trial enrollment. *J Gen Intern Med* 2008; 23 (4): 447–50.
44. Schmickl CN, Li M, Li G, *et al.* The accuracy and efficiency of electronic screening for recruitment into a clinical trial on COPD. *Respir Med* 2011; 105 (10): 1501–6.
45. Schreiweis B, Bergh B. Requirements for a patient recruitment system. *Stud Health Technol Inform* 2015; 210: 521–5.
46. Schreiweis B, Gruber G, Bergh B. First experiences in implementing a software-based support for patient recruitment at Heidelberg university hospital. *Stud Health Technol Inform* 2012; 180: 1147–9.
47. Schreiweis B, Trinczek B, Kopcke F, *et al.* Comparison of electronic health record system functionalities to support the patient recruitment process in clinical trials. *Int J Med Inform* 2014; 83 (11): 860–8.
48. Straube C, Herschbach P, Combs SE. Which obstacles prevent us from recruiting into clinical trials: a survey about the environment for clinical studies at a German University Hospital in a Comprehensive Cancer Center. *Front Oncol* 2017; 7: 181.
49. Sully BGO, Julious SA, Nicholl J. A reinvestigation of recruitment to randomised, controlled, multicenter trials: a review of trials funded by two UK funding agencies. *Trials* 2013; 14 (1): 166.
50. Thadani SR, Weng C, Bigger JT, *et al.* Electronic screening improves efficiency in clinical trial recruitment. *J Am Med Inform Assoc* 2009; 16 (6): 869–73.
51. Thompson DS, Oberteuffer R, Dorman T. Sepsis alert and diagnostic system: integrating clinical systems to enhance study coordinator efficiency. *Comput Inform Nurs* 2003; 21 (1): 22–6.
52. Trinczek B, Köpcke F, Leusch T, *et al.* Design and multicentric implementation of a generic software architecture for patient recruitment systems re-using existing HIS tools and routine patient data. *Appl Clin Inform* 2014; 5 (1): 264–83.
53. Trinczek B, Schulte B, Breil B, *et al.* Patient recruitment workflow with and without a patient recruitment system. *Stud Health Technol Inform* 2013; 192: 1124.
54. Treweek S, Pitkethly M, Cook J, *et al.* Strategies to improve recruitment to randomised trials. *Cochrane Database Syst Rev* 2018; (2): 1465–858.
55. Walters SJ, Bonacho dos Anjos Henriques-Cadby I, Bortolami O, *et al.* Recruitment and retention of participants in randomised controlled trials: a review of trials funded and published by the United Kingdom Health Technology Assessment Programme. *BMJ Open* 2017; 7 (3): e015276.
56. Zimmerman LP, Goel S, Sathar S, *et al.* A novel patient recruitment strategy: patient selection directly from the community through linkage to clinical data. *Appl Clin Inform* 2018; 09 (1): 114–21.

57. Bergmann J, Bott OJ, Hoffmann I, *et al.* An eConsent-based system architecture supporting cooperation in integrated healthcare networks. *Stud Health Technol Inform* 2005; 116: 961–6.
58. Bethune A, Davila-Foyo M, Valli M, *et al.* e-Consent: approaching surgical consent with mobile technology. *Can J Surg* 2018; 61 (5): 339–44.
59. Bialke M, Bahls T, Geidel L, *et al.* MAGIC: once upon a time in consent management—a FHIR tale. *J Transl Med* 2018; 16 (1): 256.
60. Boutin TN, Mathieu K, Hoffnagle GA, *et al.* Implementation of electronic consent at a biobank: an opportunity for precision medicine research. *J Pers Med* 2016; 6 (2): 2075–4426.
61. Buckley MT, Lengfellner JM, Koch MJ, *et al.* The Memorial Sloan Kettering (MSK) electronic informed consent (eIC) platform for clinical trials: an operational model and suite of tools for obtaining informed consent, and managing consent documents. *J Clin Oncol* 2018; 36 (15_suppl): e18577.
62. Chen C, Turner SP, Sholle ET, *et al.* Evaluation of a REDCap-based workflow for supporting federal guidance for electronic informed consent. *AMIA Jt Summits Transl Sci Proc* 2019; 2019: 163–72.
63. Chhin V, Roussos J, Michaelson T, *et al.* Leveraging mobile technology to improve efficiency of the consent-to-treatment process. *JCO Clin Cancer Inform* 2017; 1 (1): 1–8.
64. Doerr M, Truong AM, Bot BM, *et al.* Formative evaluation of participant experience with mobile eConsent in the app-mediated Parkinson mPower study: a mixed methods study. *JMIR Mhealth Uhealth* 2017; 5 (2): e14.
65. Fink AS, Prochazka AV, Henderson WG, *et al.* Enhancement of surgical informed consent by addition of repeat back: a multicenter, randomized controlled clinical trial. *Ann Surg* 2010; 252 (1): 27–36.
66. Grady C, Cummings SR, Rowbotham MC, *et al.* Informed consent. *N Engl J Med* 2017; 376 (9): 856–67.
67. Greenhalgh T, Morris L, Wyatt JC, *et al.* Introducing a nationally shared electronic patient record: case study comparison of Scotland, England, Wales and Northern Ireland. *Int J Med Inform* 2013; 82 (5): e125.
68. Harle CA, Golembiewski EH, Rahmanian KP, *et al.* Patient preferences toward an interactive e-consent application for research using electronic health records. *J Am Med Inform Assoc* 2018; 25:360–8.
69. Harle CA, Golembiewski EH, Rahmanian KP, *et al.* Does an interactive trust-enhanced electronic consent improve patient experiences when asked to share their health records for research? A randomized trial. *J Am Med Inform Assoc* 2019; 26 (7): 620–9.
70. Haussen DC, Doppelheuer S, Schindler K, *et al.* Utilization of a smartphone platform for electronic informed consent in acute stroke trials. *Stroke* 2017; 48 (11): 3156–60.
71. Iafrate RP, Lipori GP, Harle CA, *et al.* Consent2Share: an integrated broad consenting process for re-contacting potential study subjects. *J Clin Transl Res* 2017; 2 (4): 113–22.
72. Kim H, Bell E, Kim J, *et al.* iCONCUR: informed consent for clinical data and bio-sample use for research. *J Am Med Inform Assoc* 2017; 24 (2): 380–7.
73. Kondylakis H, Koumakis L, Hanold S, *et al.* Donor's support tool: enabling informed secondary use of patient's biomaterial and personal data. *Int J Med Inform* 2017; 97: 282–92.
74. Nijhawan LP, Janodia MD, Muddukrishna BS, *et al.* Informed consent: Issues and challenges. *J Adv Pharm Technol Res* 2013; 4 (3): 134–40.
75. Phillippi JC, Doersam JK, Neal JL, *et al.* Electronic informed consent to facilitate recruitment of pregnant women into research. *J Obstet Gynecol Neonatal Nurs* 2018; 47 (4): 529–34.
76. Ramos SR. User-centered design, experience, and usability of an electronic consent user interface to facilitate informed decision-making in an HIV clinic. *Comput Inform Nurs* 2017; 35 (11): 556–64.
77. Rothwell E, Wong B, Rose NC, *et al.* A randomized controlled trial of an electronic informed consent process. *J Empir Res Hum Res Ethics* 2014; 9 (5): 1–7.
78. Schreiweis B, Bronsch T, Merzweiler A, *et al.* Implementing modular research consents using IHE advanced patient privacy consents. *Stud Health Technol Inform* 2018; 247: 840–4.
79. Shelton RH. Electronic consent channels: preserving patient privacy without handcuffing researchers. *Sci Transl Med* 2011; 3 (69): 69cm4.
80. Simon CM, Klein DW, Schartz HA. Traditional and electronic informed consent for biobanking: a survey of U.S. biobanks. *Biopreserv Biobank* 2014; 12 (6): 423–29.
81. Simon CM, Klein DW, Schartz HA. Interactive multimedia consent for biobanking: a randomized trial. *Genet Med* 2016; 18 (1): 57–64.
82. Simon CM, Schartz HA, Rosenthal GE, *et al.* Perspectives on electronic informed consent from patients underrepresented in research in the United States: a focus group study. *J Empir Res Hum Res Ethics* 2018; 13 (4): 338–48.
83. Sommer C, Zuccolin D, Arnera V, *et al.* Building clinical trials around patients: evaluation and comparison of decentralized and conventional site models in patients with low back pain. *Contemp Clin Trials Commun* 2018; 11: 120–26.
84. Soni H, Grando A, Murcko A, *et al.* Current state of electronic consent processes in behavioral health: outcomes from an observational study. *AMIA Annu Symp Proc* 2017; 2017: 1607–16.
85. St John ER, Scott AJ, Irvine TE, *et al.* Completion of hand-written surgical consent forms is frequently suboptimal and could be improved by using electronically generated, procedure-specific forms. *Surgeon* 2017; 15 (4): 190–95.
86. Stevens N, Edwards L, Balayah Z, *et al.* Risk based survey evidence supports electronic informed consent as a recruitment method for UK clinical trials. *J Clin Epidemiol* 2016; 77: 134–36.
87. Suarez A, Reilly C, Fajgenbaum DC. Quantitative analysis of a rare disease network's international contact database and E-repository provides insights into biobanking in the electronic consent era. *Orphanet J Rare Dis* 2019; 14 (1): 173.
88. Vanaken H. eConsent study provides insights to shape industry adoption. *Applied Clinical Trials* 2016; 25 (8). http://www.appliedclinicaltrialsonline.com/econsent-study-provides-insights-shape-industry-adoption Accessed February 18, 2019.
89. Vanaken H, Masand SN. Awareness and collaboration across stakeholder groups important for eConsent achieving value-driven adoption. *Ther Innov Regul Sci* 2019; 53 (6): 724–735.
90. Warriner AH, Foster PJ, Mudano A, *et al.* A pragmatic randomized trial comparing tablet computer informed consent to traditional paper-based methods for an osteoporosis study. *Contemp Clin Trials Commun* 2016; 3: 32–38.
91. Welch BM, Marshall E, Qanungo S, *et al.* Teleconsent: a novel approach to obtain informed consent for research. *Contemp Clin Trials Commun* 2016; 3: 74–79.
92. Wilbanks J. Design issues in E-consent. *J Law Med Ethics* 2018; 46 (1): 110–8.
93. Aamot H, Kohl CD, Richter D, *et al.* Pseudonymization of patient identifiers for translational research. *BMC Med Inform Decis Mak* 2013; 13 (1): 75.
94. Bialke M, Penndorf P, Wegner T, *et al.* A workflow-driven approach to integrate generic software modules in a trusted third party. *J Transl Med* 2015; 13 (1): 176.
95. Bruland P, Doods J, Brix T, *et al.* Connecting healthcare and clinical research: workflow optimizations through seamless integration of EHR, pseudonymization services and EDC systems. *Int J Med Inform* 2018; 119: 103–08.
96. Chevrier R, Foufi V, Gaudet-Blavignac C, *et al.* Use and understanding of anonymization and de-identification in the biomedical literature: scoping review. *J Med Internet Res* 2019; 21 (5): e13484.
97. Deserno TM, Haak D, Brandenburg V, *et al.* Integrated image data and medical record management for rare disease registries. A general framework and its instantiation to the German Calciphylaxis Registry. *J Digit Imaging* 2014; 27 (6): 702–13.
98. Ebner H, Hayn D, Falgenhauer M, *et al.* Piloting the European Unified Patient Identity Management (EUPID) concept to facilitate secondary use of neuroblastoma data from clinical trials and biobanking. *Stud Health Technol Inform* 2016; 223: 31–38.

99. Jonas S, Siewert S, Spreckelsen C. Privacy-preserving record grouping and consent management based on a public-private key signature scheme: theoretical analysis and feasibility study. *J Med Internet Res* 2019; 21 (4): e12300.
100. Lablans M, Borg A, Uckert F. A RESTful interface to pseudonymization services in modern web applications. *BMC Med Inform Decis Mak* 2015; 15 (1): 2.
101. Lautenschläger R, Kohlmayer F, Prasser F, *et al.* A generic solution for web-based management of pseudonymized data. *BMC Med Inform Decis Mak* 2015; 15 (1): 100.
102. Nurmi S-M, Kangasniemi M, Halkoaho A, *et al.* Privacy of clinical research subjects: an integrative literature review. *J Empir Res Hum Res Ethics* 2019; 14 (1): 33–48.
103. Sahi MA, Abbas H, Saleem K, *et al.* Privacy preservation in e-healthcare environments: state of the art and future directions. *IEEE Access* 2018; 6: 464–78.
104. Schwaneberg T, Weitmann K, Dosch A, *et al.* Data privacy management and data quality monitoring in the German Centre for Cardiovascular Research's multicentre TranslatiOnal Registry for CardiomyopatHies (DZHK-TORCH). *ESC Heart Fail* 2017; 4 (4): 440–47.
105. Abshire M, Dinglas VD, Cajita MIA, *et al.* Participant retention practices in longitudinal clinical research studies with high retention rates. *BMC Med Res Methodol* 2017; 17 (1): 30.
106. Almeida JR, Gini R, Roberto G, *et al.* TASKA: a modular task management system to support health research studies. *BMC Med Inform Decis Mak* 2019; 19 (1): 121.
107. Berard C, Cloutier LM, Cassivi L. Evaluating clinical trial management systems: a simulation approach. *Ind Manage Data Syst* 2012; 112 (1): 146–64.
108. Bose A, Das S. Trial analytics—a tool for clinical trial management. *Acta Pol Pharm* 2012; 69 (3): 523–33.
109. Campion TR, Blau, JR Brown, VLSW, *et al.* Implementing a clinical research management system: one institution's successful approach following previous failures. *AMIA Jt Summits Transl Sci Proc* 2014; 2014: 12–17.
110. Cramon P, Rasmussen AK, Bonnema SJ, *et al.* Development and implementation of PROgmatic: a clinical trial management system for pragmatic multi-centre trials, optimised for electronic data capture and patient-reported outcomes. *Clinical Trials* 2014; 11 (3): 344–54.
111. Durkalski V, Wenle Z, Dillon C, *et al.* A web-based clinical trial management system for a sham-controlled multicenter clinical trial in depression. *Clin Trials* 2010; 7 (2): 174–82.
112. Geyer J, Myers K, Vander Stoep A, *et al.* Implementing a low-cost web-based clinical trial management system for community studies: a case study. *Clinical Trials* 2011; 8 (5): 634–44.
113. Gupta A, Calfas KJ, Marshall SJ, *et al.* Clinical trial management of participant recruitment, enrollment, engagement, and retention in the SMART study using a Marketing and Information Technology (MARKIT) model. *Contemp Clin Trials* 2015; 42: 185–95.
114. Leroux H, McBride S, Gibson S. On selecting a clinical trial management system for large scale, multi-centre, multi-modal clinical research study. *Stud Health Technol Inform* 2011; 168: 89–95.
115. Müller J, Heiss KI, Oberhoffer R. Implementation of an open adoption research data management system for clinical studies. *BMC Res Notes* 2017; 10 (1): 252.
116. Nadkarni PM, Kemp R, Parikh CR. Leveraging a clinical research information system to assist biospecimen data and workflow management: a hybrid approach. *J Clin Bioinformatics* 2011; 1 (1): 22.
117. Park YR, Yoon YJ, Koo H, *et al.* Utilization of a clinical trial management system for the whole clinical trial process as an integrated database: system development. *J Med Internet Res* 2018; 20 (4): e103.
118. Raptis DA, Mettler T, Fischer MA, *et al.* Managing multicentre clinical trials with open source. *Inform Health Soc Care* 2014; 39 (2): 67–80.
119. Schobel M, Staubert S, Lobe M, *et al.* Requirements on clinical trial management systems for academic site management organizations. *Stud Health Technol Inform* 2016; 228: 292–96.
120. Schwanke J, Rienhoff O, Schulze TG, *et al.* Suitability of customer relationship management systems for the management of study participants in biomedical research. *Methods Inf Med* 2013; 52 (4): 340–50.
121. Solomon P, Cavanaugh MM, Draine J. *Randomized Controlled Trials*. New York: Oxford University Press; 2009.
122. Weng C, Li Y, Berhe S, *et al.* An Integrated Model for Patient Care and Clinical Trials (IMPACT) to support clinical research visit scheduling workflow for future learning health systems. *J Biomed Inform* 2013; 46 (4): 642–52.
123. Zhao W, Pauls K. Architecture design of a generic centralized adjudication module integrated in a web-based clinical trial management system. *Clin Trials* 2016; 13 (2): 223–33.
124. de Lange DW, Guidet B, Andersen FH, *et al.* Huge variation in obtaining ethical permission for a non-interventional observational study in Europe. *BMC Med Ethics* 2019; 20 (1): 39.
125. Lamberti MJ, Kush R, Kubick W, *et al.* An examination of eClinical technology usage and CDISC standards adoption. *Ther Innov Regul Sci* 2015; 49 (6): 869–76.
126. Skripcak T, Just U, Simon M, *et al.* Toward distributed conduction of large-scale studies in radiation therapy and oncology: open-source system integration approach. *IEEE J Biomed Health Inform* 2016; 20 (5): 1397–403.
127. Wilkinson M, Young R, Harper B, *et al.* Baseline assessment of the evolving 2017 eClinical landscape. *Ther Innov Regul Sci* 2019; 53 (1): 71–80.

Anhang B

Abstract zum Vortrag:

Konzept zur datenschutzkonformen Kontaktdatenverwaltung in klinischen Forschungsprojekten

Einleitung

Ein etabliertes Mittel zum Schutz von Daten in klinischen Forschungsprojekten ist die Pseudonymisierung und die damit verbundene Trennung identifizierender Daten (IDAT) und medizinischer Forschungsdaten (MDAT) der Studienteilnehmer:innen. Bei der so geschaffenen informationellen Gewaltenteilung wird die Verwaltung der IDAT sowie der Pseudonymlisten (Identitätsmanagement) häufig an einen Datentreuhänder übertragen. Ebenso wie die klassischen administrativen Daten einer Person wie Name, Geburtsdatum und Geschlecht sind die Kontaktdaten der Studienteilnehmer:innen sensible und schützenswerte IDAT. Kontaktdaten im Identitätsmanagement bei einem Datentreuhänder zu verwalten, ist eine mögliche, jedoch nicht immer praktikable Lösung. Forschungsprojekte, mit der Anforderung ein eigenes System zur Erfassung, Aktualisierung und Erweiterung einer großen Anzahl von Kontaktdaten im Verlauf eines Projektes einzusetzen, benötigen ein Konzept, welches Kontaktdaten in eine datenschutzkonforme Informationstrennung einbezieht.

Methodik

Seit vielen Jahren bilden die generischen Datenschutzkonzepte der TMF [1] die Grundlage für die datenschutzkonforme Planung und Umsetzung medizinischer Forschungsprojekte in Deutschland. Zur Umsetzung einer Kontaktdatenverwaltung bleiben die generischen Datenschutzkonzepte recht vage, jedoch lassen sich Prinzipien datenschutzgerechter Lösungen des Leitfadens auf das hier benötigte Konzept anwenden. Die Kontaktdatenverwaltung wurde als ein weiterer Baustein zur zentralen Infrastruktur-Komponente des Leitfadens, dem Identitätsmanagement, konzeptuell ergänzt. Zur sinnvollen Informationsaufteilung wurden Zweck und Anwendungsbereich sowie die zugehörigen Daten und Datenflüsse beider Komponenten untersucht.

Ergebnisse

Zentraler Punkt des erarbeiteten Konzepts ist die feinere Untergliederung der Datenkategorie IDAT bei der Verwaltung personenbezogener Daten unter dem Gesichtspunkt der informationellen Gewaltenteilung. Im Sinne der Datensparsamkeit werden dem Identitätsmanagement ausschließlich die IDAT übermittelt, die für die eindeutige Zuweisung von Pseudonymen erforderlich sind. IDAT die zur Kontaktierung der Studienteilnehmer:innen benötigt werden (Adresse, Telefonnummer etc.), werden im System zur Kontaktdatenverwaltung gespeichert. Beide Systeme werden im Idealfall unter einer getrennten Datenhoheit und in jedem Fall getrennt von der Verwaltung der MDAT geführt, ein externer Datentreuhänder übernimmt die Aufgabe des Identitätsmanagements. Zugriffsrechte werden nur an das Personal vergeben, dass für die Kontrolle des Identitätsmanagements oder die Kontaktdatenverwaltung verantwortlich ist.

Diskussion

Das vorgestellte Konzept weicht vom Paradigma ab, IDAT ausschließlich im Identitätsmanagement eines Treuhänders zu verwalten. Die vorgestellte Aufteilung der IDAT ermöglicht jedoch auch, eine weitere Ebene der Informationstrennung hinzuzufügen und so das Schutzprinzip der informationellen Gewaltenteilung zu stärken. Kontaktdaten werden gesondert von den Pseudonymlisten des Identitätsmanagements unter getrennter Datenhoheit - insbesondere auch in Bezug auf die Verwaltung der MDAT - sowie mit separaten Zugriffsrechten verwaltet. Grundsätzlich sind dies

gegenläufige Effekte: Mit der Verteiltheit der Daten steigt die Informationstrennung, jedoch auch die Zahl der Angriffspunkte. Diesem Nachteil muss mit den gleichen Maßnahmen begegnet werden, die auch zur Sicherung des Datentreuhänders unternommen werden.

Schlussfolgerung

Forschungsprojekte, die Systeme zur Kontaktdatenverwaltung einsetzen, haben für sich individuelle Vorteile durch deren Nutzung bestimmt. Der daraus entstehenden Herausforderung sensible und schützenswerte IDAT an mehreren Orten in einem Forschungsprojekt zu verwahren, kann mit dem vorgestellten Konzept begegnet werden. Das Konzept wendet Grundprinzipien datenschutzgerechter Lösungen wie Datensparsamkeit und informationelle Gewaltenteilung an und betrachtet Kontaktdaten als eine Kategorie von IDAT. Die Vereinbarkeit der gefundenen Lösung mit den Komponenten und Modulen der generischen Datenschutzkonzepte der TMF ermöglicht das Konzept in unterschiedlichen Forschungsprojekten anzuwenden.

Referenzen

[1] Pommerening K, Drepper J, Helbing K, Ganslandt T: Leitfaden zum Datenschutz in medizinischen Forschungsprojekten - Generische Lösungen der TMF 2.0. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft. 2014. DOI: 10.32745/9783954662951.

Anhang C

Access control policy for a study participant management system for multi-site clinical research

J Pung[1], D Krefting[1], O Rienhoff[1]

[1] Department of Medical Informatics, University Medical Center Göttingen, Göttingen

Background

The protection of sensitive data against unauthorized access is especially important in healthcare and medical research. Not only medical data but also identifying administrative data in medical research projects must be managed and protected. While Clinical Data Management Systems (CDMS) are well established in clinical trial data management, digitalization of study participant management is still underway. A Study Participant Management System (SPMS) was developed by the Department of Medical Informatics at the University Medical Center Göttingen as web application to manage administrative data in single-site clinical trials [1]. To expand the SPMS for the use in multi-site studies, however, stronger and specific regulations for participant data access are required to meet data protection requirements. Therefore, a comprehensive concept to control data access in multi-site clinical trials is currently developed. In this context, access control mechanisms were examined as a fundamental component for secure information sharing in the SPMS application.

Methods

Existing state-of-the-art access control models and architectures were investigated to meet access control requirements for a multi-site SPMS in general and for possible technical solutions. Only technical solutions were considered that fit into the software stack of the SPMS, which was implemented as a web application with the use of Java Enterprise Edition (Java EE) technology. Based on different usage scenarios specific for actors and locations, an authorization concept was developed. In particular, the functions and workflows required to process data were described. Furthermore, specific user roles were identified, and their access rights to different study participants and data items were determined.

Results

An access control architecture capable of providing data access control for a multi-center SPMS has been defined using a combination of Role-Based Access Control (RBAC) [2, 3] and Attribute-Based Access Control (ABAC) [4] models. To establish the architecture, tasks and necessary access permissions of each user role (e.g., principal investigator, clinical trial assistant) were described by collecting respective use cases for study participant data management. The authorization concept was implemented with the Spring Security framework [5], which provides comprehensive security services for Java EE-based software applications. Spring Expression Language (SpEL) [6] expressions allowed to implement fine-grained access controls. The implementation of the authorization concept was tested with JUnit [7] integration test cases to reduce the risk of software errors.

Conclusion

Strict regulations must be implemented to ensure that only authorized users can access sensitive data in multi-site studies. Data access requirements depend on software architecture, data model, and business requirements and cannot be easily simplified. In each case, the complexity should be kept as low as possible in order to minimize efforts. With a well-adjusted fine-grained access control, it is more likely to prevent breaches and improve data protection of participant management in multi-site studies.

[1] J. Schwanke, Konzeption und Entwicklung eines Probandenmanagementsystems am Beispiel der Universitätsmedizin

Göttingen, PhD dissertation, Georg-August-Universität Göttingen, 2015.

[2] D. F. Ferraiolo and D. R. Kuhn, Role-Based Access Control, Proceedings of the 15th National Computer Security Conference, (1992) 554-563.

[3] R. S. Sandhu, E. J. Coyne, H. L. Feinstein and C. E. Youman, Role-based access control models, Computer 29(2) (1996) 38–47.

[4] V. C. Hu, D. R. Kuhn and D. F. Ferraiolo, Attribute-based access control, Computer 48(2) (2015) 85–88.

[5] Spring Security. Available at: https://spring.io/projects/spring-security Accessed March 25, 2020.

[6] Spring Expression Language (SpEL). Spring Framework Documentation. Available at: https://docs.spring.io/spring/docs/4.3.10.RELEASE/spring-framework-reference/html/expressions.html Accessed March 25, 2020.

[7] JUnit - Unit testing framework for Java programming language. Available at: https://junit.org Accessed March 25, 2020.

Keywords: Clinical Trials, Information Systems, Access Control, Data Privacy, Research Subjects, Clinical Research Informatics

Information about submission 42:

- Last change: 13 Jul 2020 09:14
- Kind: 02 All other submissions
- Category: 05 Medical Informatics
- Sub-category: Other
- Submission type: Talk
- Submission declarations:

  Conflicts of interest:

  The corresponding author declares that there is no conflict of interest with the authors.

  Statement on ethics vote:

  No ethics vote is required.