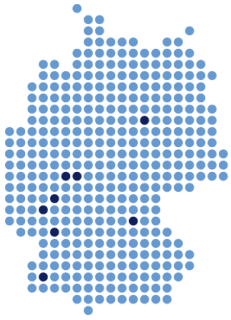

Datenschutzkonzept MIRACUM

MIRACUM - Medical Informatics in Research and Care in University Medicine

Version 0.8 vom 17.05.2018

Versionshistorie:

Version	Datum	Confluence	Einfluss/Autor	Änderung
0.1	16.03.2018	v.41	Marvin Kampf	Erster Aufschlag auf Basis GBA-DS-Konzept
0.2	20.03.2018	v.49	Andreas Borg	Kommentare/Anregungen
0.3	28.03.2018	v.69	Ulli Prokosch	Überarbeitung
0.4	29.03.2018	v.75	Manfred Brunner Ulli Prokosch Thomas Ganslandt Marvin Kampf	Erste Absprache Überarbeitung
0.5	05.04.2018	v.100	Manfred Brunner Marvin Kampf	Vorläufige Finalisierung
0.6	17.04.2018	v.132	Marvin Kampf Thomas Ganslandt Martin Boeker Dennis Kadioglu Tim Herrmann	Markieren der zukünftigen Verfahren als Ausblick Konkretisieren tatsächlicher Konzepte und Verantwortlichkeiten Kleine Korrekturen
0.7	19.04.2018	v.1	Ulli Prokosch Marvin Kampf	Prefinale Überarbeitung vor Abstimmung mit dem DS-Beauftragten des Landes Bayern Kleine Korrekturen
0.8	17.05.2018	v.24	Marvin Kampf Ulli Prokosch	Finale Korrekturen nach Abstimmung mit bayerischen Datenschützern

Herausgeber:

Friedrich-Alexander-Universität Erlangen Nürnberg
Medizinische Fakultät, Lehrstuhl Medizinische Informatik
Wetterkreuz 13, 91058 Erlangen

Prof. Dr. Hans-Ulrich Prokosch
Tel.: 09131 85 26721, 09131 85 26758 / Fax: 09131 85 26754
E-Mail: ulli.prokosch@uk-erlangen.de

Inhalt

1 Art und Ziele des Projekts

MIRACUM[1] (Medical Informatics in Research and Care in University Medicine) ist eines der vier Konsortien, die im Rahmen der BMBF Medizininformatik-Initiative (BMBF MI-I) zur Förderung von IT-Innovationen in Krankenversorgung und medizinischer Forschung finanziert werden. Ziel der Initiative ist es, die Digitalisierung in der Medizin voranzutreiben und zu nutzen, um klinische Daten standortübergreifend gemeinsam zu nutzen und damit medizinisches Wissen generieren zu können. Hierfür haben sich acht Universitäten und ihre jeweiligen Universitätskliniken, zwei Hochschulen für Angewandte Wissenschaften und ein Industriepartner verteilt auf fünf Bundesländer zu dem Konsortium MIRACUM zusammengeschlossen. Vereinbart wurde die Errichtung von acht modularen und skalierbaren Datenintegrationszentren (DIZ) innerhalb einer Infrastruktur, die innerhalb der vierjährigen Förderphase eine gemeinsame Datennutzung von Daten aus den Bereichen Routineversorgung, bildgebende Verfahren und OMICS-Daten ermöglicht. Darauf aufbauend werden insbesondere die Verbesserung von patientenzentrierter, kollaborativer Forschung, die Optimierung der klinischen Versorgungsprozesse und die Stärkung Biomedizinischer Informatik in Forschung und Lehre angestrebt.

Die Entwicklung der Datenintegrationszentren basiert auf MIRACOLIX (**M**edical **I**nformatics **R**eus**A**ble e**C**o-system of **O**pen source **L**inkable and **I**nteroperable software tools), einer Sammlung an größtenteils Open-Source-Werkzeugen, die den Fluss der Daten aus den jeweiligen primären klinischen Systemen (in denen sie in der Regel für Versorgungszwecke in Verbindung mit identifizierenden Daten vorliegen) über ein innerklinisches Datenrepository (Data Warehouse) hin zu pseudonymisierten/anonymisierten Forschungsdatenbanken, unter Berücksichtigung der Datenschutz-rechtlichen Vorgaben und des Patientenwillens (Informierte Einwilligung) ermöglichen sollen. Das MIRACOLIX Ökosystem gliedert sich in die Ebenen der Infrastruktur, der Basisdienste und der Anwendungen für den Nutzer auf. Es folgt dabei dem Grundsatz, pragmatisch, modular, wiederverwendbar, interoperabel und föderiert einsetzbar zu sein. Während der Konzeptphase (Ausarbeitung des Forschungsantrags) konnte an allen MIRACUM-Standorten eine erste Version eines DIZ (aufsetzend auf einer Version MIRACOLIX 0.9) etabliert werden.

Darüber hinaus sollen in den vier Jahren der aktuellen MIRACUM-Förderung (2018 - 2021) verschiedene Anwendungsszenarien (Use Cases) zum Nachweis des Nutzens der standortübergreifenden DIZ-Infrastruktur umgesetzt werden. Diese MIRACUM Use Cases sind:

- Use Case 1: Alerting in Care – IT Support for Patient Recruitment
- Use Case 2: From Data to Knowledge – Clinico-molecular Predictive Knowledge Tool
- Use Case 3: From Knowledge to Action – Support for Molecular Tumor Boards

Eine detaillierte Beschreibung der einzelnen Use Cases befindet sich im Anhang 4: Beschreibung der MI-I-Use-Cases.

Dieses Datenschutzkonzept beschreibt in einer generischen Form die an allen MIRACUM DIZ-Standorten zu etablierenden Strukturen/Komponenten. Es ist denkbar und Teil des MIRACUM-Konzepts, dass einzelne MIRACUM-Standorte bestimmte Funktionen ihres

DIZ durch andere Produkte/Systeme umsetzen, vorausgesetzt, dass diese im Hinblick auf die Gesamtfunktionalität die an die jeweilige Funktion gestellten Anforderungen erfüllen und über entsprechende interoperable Schnittstellen kompatibel zu der Gesamtstruktur des DIZ sind. Entsprechende lokale Anpassungen sind jeweils getrennt vom hier vorgelegten **generischen MIRACUM-Datenschutzkonzept** in standortspezifischen Ausprägungen dieses Datenschutzkonzeptes zu beschreiben.

Hinweis:

Der hiermit vorgelegte erste Entwurf eines Datenschutzkonzepts erhebt **nicht** den Anspruch, **alle** in den Jahren 2018 bis 2021 zu etablierenden DIZ-Strukturen und die darauf umzusetzenden Use Cases bereits vollständig abzudecken. Mit dieser Version ist ebenfalls **noch nicht** das langfristige Ziel des BMBF und des NSG erreichbar, dass Daten verschiedener Datenintegrationszentren eventuell für neue, zukünftige Projekte in einer gemeinsamen Datenhaltung zusammengeführt werden könnten.

Es ist bewusst pragmatisch darauf ausgerichtet diejenigen Abläufe und Strukturen zu beschreiben, die innerhalb von MIRACUM im ersten Projektjahr (2018) gemäß Projektplan umzusetzen sind bzw. benötigt werden:

1. Die Etablierung von Forschungsdaten-Repositories **lokal an jedem MIRACUM-Standort**, in denen jeweils auf die entsprechende spätere Verwendung zugeschnitten klinische Daten, Bilddaten und OMICS-Daten zusammengeführt werden (es sind dies die Repositories i2b2, OMOP, XNAT, tranSMART, cBioportal).
2. Die Ermöglichung **verteilter Machbarkeitsstudien** (Fallzahlabfragen) über die **lokal an den MIRACUM-Standorten** vorgehaltenen Forschungsdaten-Repositories hinweg.
3. Die Ermöglichung **verteilter Auswertungen** über die **lokal an den MIRACUM-Standorten** vorgehaltenen Forschungsdaten-Repositories hinweg.

Das wesentliche in MIRACUM verfolgte und in dieser Version des Datenschutzkonzeptes zugrunde gelegte Konzept ist das einer lokalen Datenhaltung, jeweils an dem MIRACUM Standort, an dem diese Daten auch im Versorgungskontext erhoben und dokumentiert wurden. **Die Patienten bezogenen Daten eines Standorts verlassen also nie das Hoheitsgebiet dieses Standorts.**

MIRACUM bringt die Analysen zu den Daten und nicht die Daten zu den Analysen.

Durch die zukünftig wachsenden Strukturen und aufkommenden Forschungsszenarien entstehen schrittweise auch neue Anforderungen an dieses Datenschutzkonzept. Diesen soll das Datenschutzkonzept durch iterative Anpassungen in Abstimmung mit den entsprechenden Datenschutzbeauftragten gerecht werden, so dass dann die zum jeweiligen Zeitpunkt etablierten Strukturen und Prozesse davon abgedeckt sind. Ein grober Ausblick auf geplante Ausbaustufen ist innerhalb des vorliegenden Datenschutzkonzeptes in grüner, kursiver Schrift skizziert. Diese Ausblick-Abschnitte sind jeweils am Ende eines Abschnitts des entsprechenden Themengebietes zu finden.

2 Organisatorische Struktur, Lenkungsform, Verantwortlichkeiten

Die organisatorische Struktur und Verwaltung innerhalb von MIRACUM setzt sich aus (1) dem Lenkungsausschuss (Steering Board), (2) sechs Arbeitsgruppen (Working Groups), (3) einer zentralen Geschäftsstelle (Central Office, ansässig am Standort des MIRACUM Koordinators an der Friedrich-Alexander-Universität Erlangen-Nürnberg), die von den örtlichen, administrativen Koordinationsbüros jedes der acht DIZ-Standorte unterstützt wird, (4) der MIRACUM Hauptversammlung (General Assembly) und (5) einem internationalen, wissenschaftlichen Beratungsausschuss (International Scientific Advisory Board) zusammen (vgl. Abb. 1). Eine solche Managementstruktur zielt auf effiziente Entscheidungsfindung, produktive und zufriedenstellende interne Kommunikation und technische sowie administrative Projektkontrolle ab. Das Projekt wird vom Lenkungsausschuss geleitet und gesteuert. Jeder DIZ-Standort wird darin durch zwei ernannte Personen repräsentiert, dem Principal Partner Coordinator (PI) und dessen Stellvertreter (Co-PI). Gemeinsam sollen diese nicht nur exzellente Kompetenzen aus der medizinischen Informatik und der Gesundheitsversorgung einbringen, sondern ebenso eine hohe Entscheidungskompetenz in ihrer jeweiligen Organisation innehaben (beispielsweise Dekan der Fakultät, Medizinischer Direktor des Klinikums, Lehrstuhlinhaber für Medizinische Informatik, CIO des Klinikums).

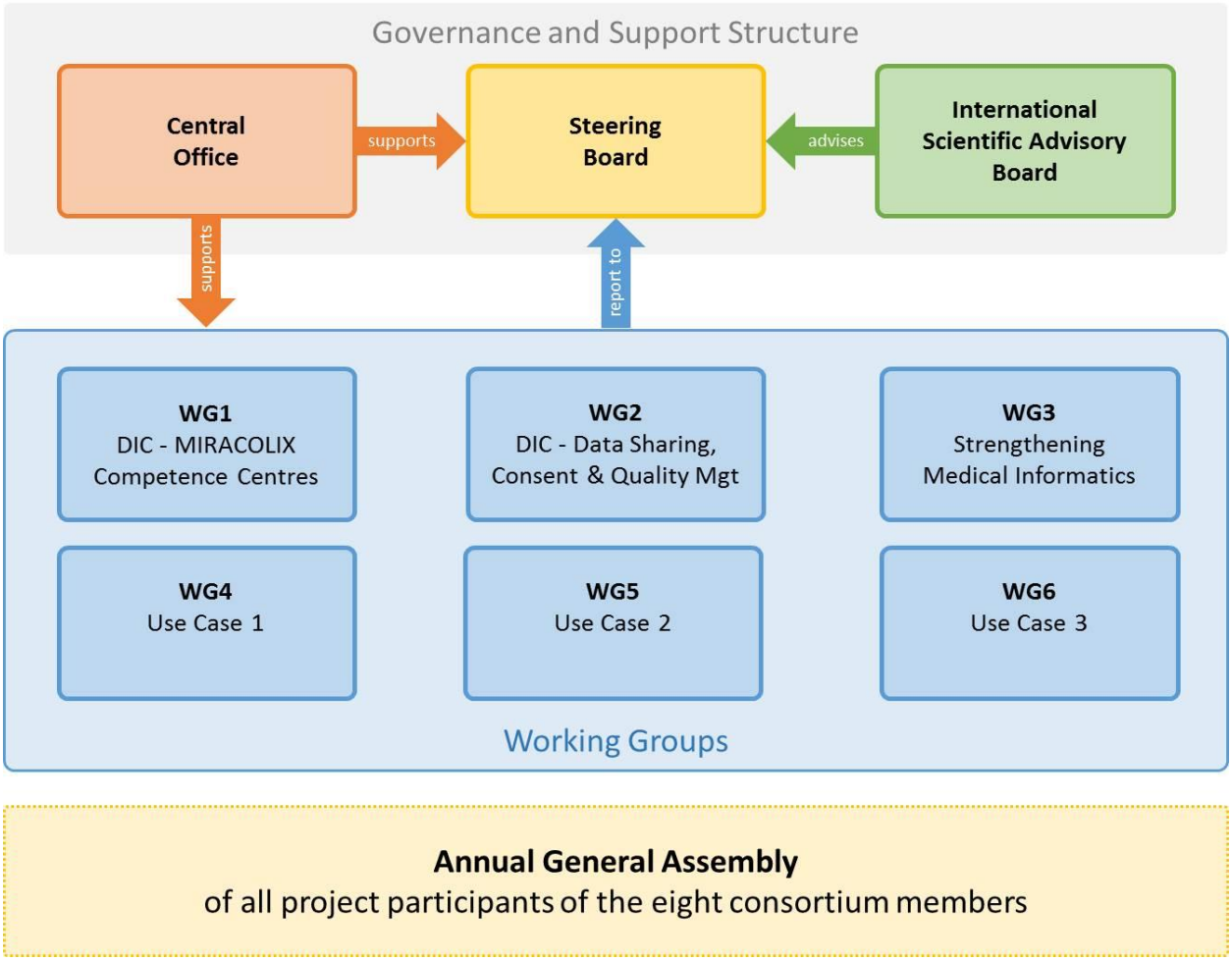

Abbildung 1 MIRACUM Managementstruktur

Der Betrieb der Datenintegrationszentren erfolgt durch die in MIRACUM vertretenen DIZ-Standorte in deren eigener Verantwortung:

- Erlangen: Medizinisches Zentrum für Informations- und Kommunikationstechnik (MIK) des Universitätsklinikums Erlangen
- Frankfurt: Dezernat 7 Informations- und Kommunikationstechnologie (DICT) des Universitätsklinikum Frankfurt
- Freiburg: Medizinische Fakultät der Albert-Ludwigs-Universität Freiburg, Universitätsklinikum Freiburg
- Gießen: Justus-Liebig-Universität Gießen, Universitätsklinikum Gießen/Marburg
- Magdeburg: Institut für Biometrie und Medizinische Informatik der medizinischen Fakultät der Otto-von-Guericke-Universität Magdeburg, Medizinisches Rechenzentrum (MRZ) des Universitätsklinikums Magdeburg
- Mainz: Johannes-Gutenberg-Universität Mainz, Universitätsmedizin der Johannes-Gutenberg-Universität Mainz
- Mannheim: Medizinische Fakultät Mannheim der Universität Heidelberg, Universitätsklinikum Mannheim
- Marburg: Fachbereich Medizin der Philipps-Universität Marburg

Weitere Mitglieder im Konsortium sind das Unternehmen Averbis GmbH in Freiburg, die Technische Hochschule Mittelhessen in Gießen und die Hochschule Mannheim (welche aber jeweils kein eigenes DIZ etablieren).

An jedem MIRACUM DIZ Standort wurde ein Use & Access Komitee etabliert welches auf Basis einer lokalen Datenzugriffs- und Nutzungsordnung (Use & Access Policy; UAP) über die Teilnahme eines Standorts an einer vorgeschlagenen föderierten Konsortialauswertung entscheidet (vgl. 3.3; Abschnitt Föderierte Auswertungen).

Alle Aktivitäten der BMBF MI-I werden von einem Nationalen Steuerungsgremium (NSG) übergreifend koordiniert, um die Interoperabilität von Datenintegrationszentren und IT-Lösungen zwischen den Konsortien sicherzustellen. Die Zusammenarbeit und Abstimmung der Konsortien wird im Rahmen eines Begleitprojekts koordiniert, das gemeinsam von der TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung (TMF), vom Medizinischen Fakultätentag (MFT) und vom Verband der Universitätsklinika Deutschlands (VUD) durchgeführt wird.

3 IT- und Netzarchitektur, Daten, Prozesse und Kommunikationswege

3.1 IT- und Netzarchitektur

An jedem DIZ-Standort wird mit Hilfe der MIRACOLIX Werkzeugsammlung ein Datenintegrationszentrum etabliert. Es besteht aus Hard- und Software-Komponenten, die unter der Hoheit jedes Standorts installiert und betrieben werden, und dient dazu, die Daten des jeweiligen Standorts in ein standortübergreifend kompatibles Format zu überführen und für verteilte Abfragen sowie verteilte Analysen nutzbar zu machen.

Ein DIZ besteht in seiner aktuellen Form (MIRACOLIX 0.9) aus den folgenden, lokal an einem DIZ-Standort installierten Softwarekomponenten (vgl. hierzu auch Abb. 2):

- **Lokales Klinisches Data Warehouse**: Stellt die in den lokalen Primärsystemen prinzipiell vorliegenden Datenbestände in einer integrierten Form zur Verfügung.
- **Lokales ID-Management**: Stellt für die Pseudonymisierung von Patientendaten die in Abschnitt 3.3 beschriebene Funktionalität bereit.
- **Lokale Forschungsdaten Repositories**: Kern-Komponenten zur Haltung, Abfrage, Auswertung und Visualisierung der Daten
    - **i2b2** (für klinische Daten)
    - **OMOP DB** (für klinische Daten)
    - **XNAT** (für Bilddaten)
    - **tranSMART** (für klinische Daten gemeinsam mit OMICS Daten)
    - **cBioPortal** (für Genomdaten zur Krebsforschung)
- **li2b2**: lokale Schnittstelle (Broker) zur Abholung der Abfragelogik föderiert zu verarbeitender Machbarkeitsabfragen, die mit dem zentralen i2b2 Web-Client (vgl. Abschnitt 3.3) erstellt werden

Die oben aufgeführten Software-Komponenten stehen innerhalb eines DIZ jeweils unter Kontrolle des entsprechenden Standortes. Somit stehen auch die in diesen Komponenten gespeicherten Daten immer unter der Hoheit der Institution, in der sie erhoben werden. **Es verlassen keine Patienten bezogenen oder beziehbare Daten den lokalen Standort.**

Ausblick

In zukünftigen Ausbaustufen wird ein DIZ unter anderem um die folgenden Komponenten erweitert:

1. ***Datenharmonisierung und -integration**: Alle Daten werden in ein für die jeweils eingesetzten Forschungsdaten Repositories kompatibles, standortübergreifend einheitliches Format überführt und anschließend in dem jeweiligen Forschungsdaten Repository gespeichert.*
2. ***Consent-Management**: Die im Rahmen der Patientenaufnahme an einem Klinikum erhobenen Einwilligungen (sowie späterer Widersprüche) zur Datenverarbeitung werden zukünftig (nach Abstimmung eines deutschlandweiten Mustertextes für einen Forschungsconsent über die NSG AG Consent mit dem Arbeitskreis der deutschen*

Ethikkommissionen und den Landesdatenschützern) über dieses Modul automatisiert abfragbar und für die Filterung des Patientendatenflusses nutzbar gemacht. Für die mit dem hier vorgelegten Datenschutzkonzept zu unterstützenden Szenarien, die lediglich auf einer lokalen Datenhaltung beruhen, ist eine Patienteneinwilligung noch nicht erforderlich. Die Datennutzung ist aktuell noch durch die jeweilige Krankenhausgesetzgebung bzw. das entsprechende Landesdatenschutzgesetz abgedeckt.

3. ***Semantische Anreicherung**: Mittelfristig ist es geplant auch unstrukturierte klinische Texte mittels Verfahren des "Natural Language Processing - NLP" für eine strukturierte Weiterverwendung von einzelnen Datenelementen zu analysieren und aufzubereiten, so dass sie in einer entsprechend angereicherten Form in ein Forschungsdaten Repository übernommen werden können.*

Diese Komponenten sind im ersten Projektjahr für den Routineeinsatz eines DIZ noch nicht erforderlich, weshalb nachfolgend nicht weiter auf sie eingegangen wird.

3.2 Daten

Im Rahmen der BMBF MI-I wurde durch die Arbeitsgruppe Interoperabilität des Nationalen Steuerungsgremiums der sogenannte MI-I-Kerndatensatz definiert. Dieser Kerndatensatz beinhaltet all jene Datenelemente, die von den teilnehmenden DIZ-Standorten zu bestimmten Zeitpunkten im Projektverlauf vorgehalten werden müssen. Der Kerndatensatz ist in verschiedene Module aufgeteilt: sieben Basismodule und verschiedene Erweiterungsmodule (vgl. Anhang 1, MI-I-Kerndatensatz). Die Basismodule sind

- Person, Demographie, Falldaten, Diagnosen, Prozeduren, Laborbefunde und Medikation

Neben den durch den MI-I Kerndatensatz vorgegebenen Erweiterungsmodulen wird es in MIRACUM zu einem späteren Zeitpunkt noch Erweiterungsmodule geben, die Datenelemente beinhalten, die für die MIRACUM Use Cases benötigt werden.

In den ersten 9 Monaten des Jahres 2018 beschränkt sich der Datenumfang in den MIRACUM DIZ auf die Basismodule Person, Demographie, Falldaten, Diagnosen und Prozeduren. Zur Erfüllung der BMBF MI-I Vorgaben wird dies im vierten Quartal 2018 um die Module Laborbefunde und Medikation ergänzt.

Grundsätzlich teilen sich die datenschutzrelevanten Daten in medizinische und identifizierende Daten auf, die im Folgenden in Anlehnung an den "TMF-Leitfaden zum Datenschutz in medizinischen Forschungsprojekten"[3] als MDAT und IDAT bezeichnet werden.

Die Basismodule Person und Demographie enthalten die besonders schützenswerten IDAT. Diese werden im Rahmen des Transformationsprozesses der Daten vom lokalen klinischen Data Warehouse (Datenvorhaltung noch im Versorgungskontext) mittels Pseudonymisierung abgeschnitten, so dass eine Re-Identifizierung faktisch ausgeschlossen werden kann. Genauere Informationen zu diesem Verarbeitungsschritt der Daten folgen in Abschnitt 3.3 ID-Management.

Generell gilt immer das Prinzip der Datenminimierung, so dass im Projektverlauf Daten lediglich in dem Umfang und zu dem Zeitpunkt bereitgestellt werden, zu dem Sie für die jeweiligen Teilprojekt-Anforderungen benötigt werden.

3.3 Prozesse und Kommunikationswege

ID-Management

In der aktuellen Projektphase werden die patientenidentifizierenden Daten durch nichtsprechende Pseudonyme ersetzt. Die MDAT der Patienten werden verknüpft mit dem jeweiligen Pseudonym in die Forschungsdaten Repositories importiert. Eine Anonymisierung dieser zu übertragenden Daten ist hierbei nicht zielführend, da - gemessen an der hohen Anzahl der Merkmale - eine grundlegende Generalisierung die Daten in den Forschungsdaten Repositories unbrauchbar machen würde. Da in 2018 keinerlei Standort übergreifende Datenzusammenführung vorgesehen ist, ist derzeit auch noch kein übergreifendes Record Linkage erforderlich.

Ausblick

Mittelfristig wird eine im Projektverlauf zu entwickelnde, dedizierte ID-Management-Komponente die folgenden Funktionalitäten in der DIZ-Infrastruktur übernehmen:

- ***Pseudonymisierung**: An die Stelle von identifizierenden Daten bzw. IDAT treten nichtsprechende Pseudonyme, um ein hohes Datenschutzniveau aufrecht zu erhalten. In bestimmten Fällen soll eine Depseudonymisierung durch autorisiertes Personal möglich sein, z.B. im Zuge der Patienten-Rekrutierung, im Falle einer nötigen Intervention, oder ähnlichen Szenarien, in denen Rückschluss auf die Person hinter den Daten gezogen werden muss.*
- ***Record Linkage**: Dies ist ein automatisches Verfahren zur Identifikation von Dubletten. Gewisse Ähnlichkeitsvergleiche erlauben dabei Übereinstimmungen zwischen zwei Einträgen mit einer bestimmten Wahrscheinlichkeit zu benennen. Ab einem zu definierenden Wahrscheinlichkeits-Grenzwert gelten die beiden Einträge automatisch als Dubletten, beziehen sich also auf denselben Patienten und werden in diesem Zuge zusammengeführt.*

Die MIRACOLIX Toolbox wird zukünftig eine solche lokale ID-Management-Komponente für die beschriebenen Aufgaben bereitstellen. Alternativ können an den Standorten eigene Pseudonymisierungs-/Record-Linkage-Tools mit gleicher Funktionalität zum Einsatz kommen. Die Pseudonymisierungs-/Record-Linkage-Verfahren unterliegen der Kontrolle des jeweiligen Standorts und damit den lokal geltenden Datenschutzrichtlinien.

Import in Forschungsdaten Repositories und Pseudonymisierung mittels lokalem ID-Management

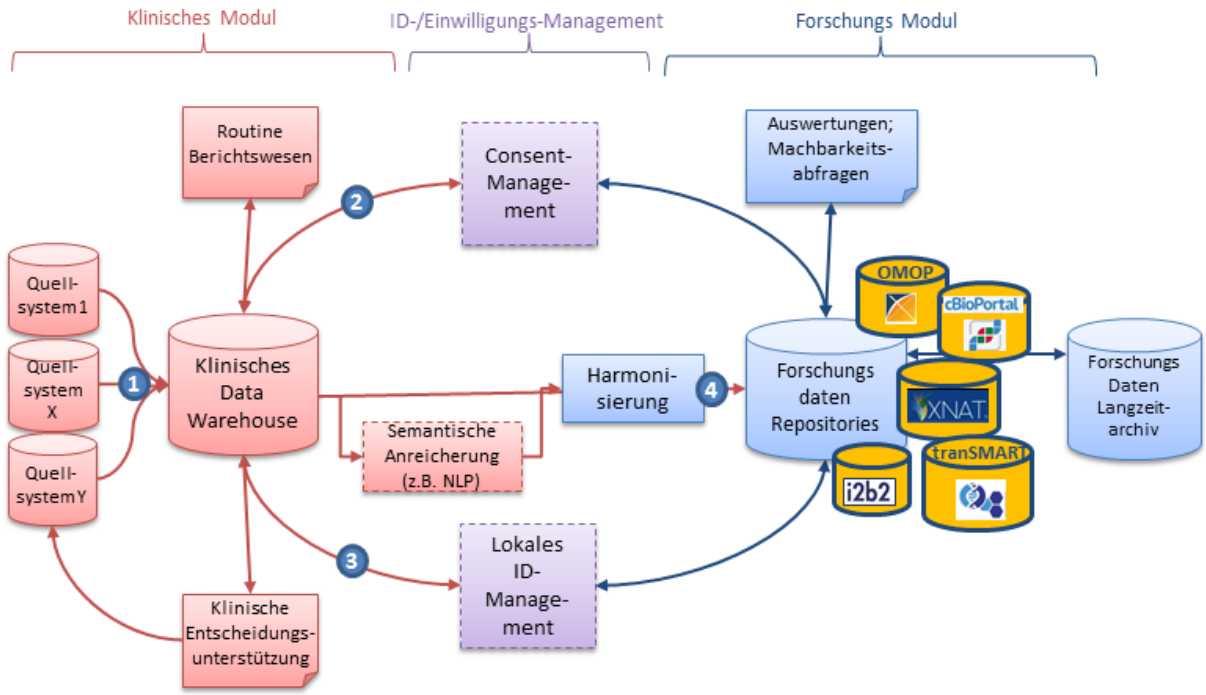

Abbildung 2 Datenfluss innerhalb eines MIRACUM DIZ (generische Illustration) - jegliche Patienten bezogene/beziehbare Datenhaltung erfolgt nur an dem Standort, an dem die Daten auch dokumentiert wurden.

Abbildung 2 zeigt die drei Domänen eines DIZ entsprechend dem TMF-Leitfaden zum Datenschutz in medizinischen Forschungsprojekten: das klinische Modul, das Forschungsmodul und dazwischen die Komponenten des ID-/Einwilligungs-Managements. IDAT und MDAT werden aus verschiedenen Quellsystemen eines DIZ-Standorts durch ETL-Prozesse in das jeweils lokale klinische Data Warehouse extrahiert (1). Die ID-Management-Komponente sorgt für das Ersetzen der IDAT auf dem Weg der Daten in die Forschungsdaten Repositories mit generierten eindeutigen Pseudonymen erster Stufe (3). Letztlich lassen sich so alle Daten eines Patienten aus verschiedenen Quellsystemen eindeutig und datenschutzgerecht in einem Forschungsdaten Repository des Forschungsmoduls zusammenführen (4). Diese Forschungsdaten Repositories verbleiben immer im Hoheitsgebiet des jeweiligen Standorts.

Ausblick

In weiteren Ausbaustufen wird ein lokales Consent-Management-System die Einwilligungen zur Datenverarbeitung von entsprechenden Patienten überprüfen und den jeweiligen Stand der Einwilligung berücksichtigen (2). Das lokale ID-Management wird um die Funktionalitäten der Depseudonymisierung und des Record Linkages erweitert werden (3). Dabei soll ein Matching-Algorithmus mögliche Dubletten desselben Patienten erkennen und zusammenführen (lokales Record Linkage). In späteren Projektstufen werden in den Schritt (4) noch Funktionen integriert werden, die eine Extraktion strukturierter Datenelemente aus Freitexten (NLP) sowie eine Harmonisierung unterschiedlicher Terminologien auf eine in

MIRACUM gemeinsam vereinbarte Terminologie ermöglichen. Weiterhin ist mittelfristig auch die Etablierung eines standortübergreifenden "Privacy Preserving Record Linkage (PPRL)" vorgesehen. Da diese Funktionalität langfristig auch konsortienübergreifend funktionieren soll, ist für die Definition dieser Funktionalität auf NSG-Ebene die Einrichtung einer entsprechenden Taskforce vorgesehen.

Föderierte Machbarkeitsstudien (Fallzahlabfragen)

Für die Durchführung föderierter Machbarkeitsstudien bzw. föderierter Auswertungen (nachfolgender Abschnitt) wird die Softwarearchitektur eines jeden lokalen DIZ um einen sogenannten lokalen li2b2-Such-Client erweitert. Weiterhin wird eine zentrale Software-Komponente (in der aber **keinerlei Patienten bezogene/beziehbare Daten** vorgehalten werden), das i2b2 Query & Analysis Tool bereit gestellt (vgl. Abbildung 3)

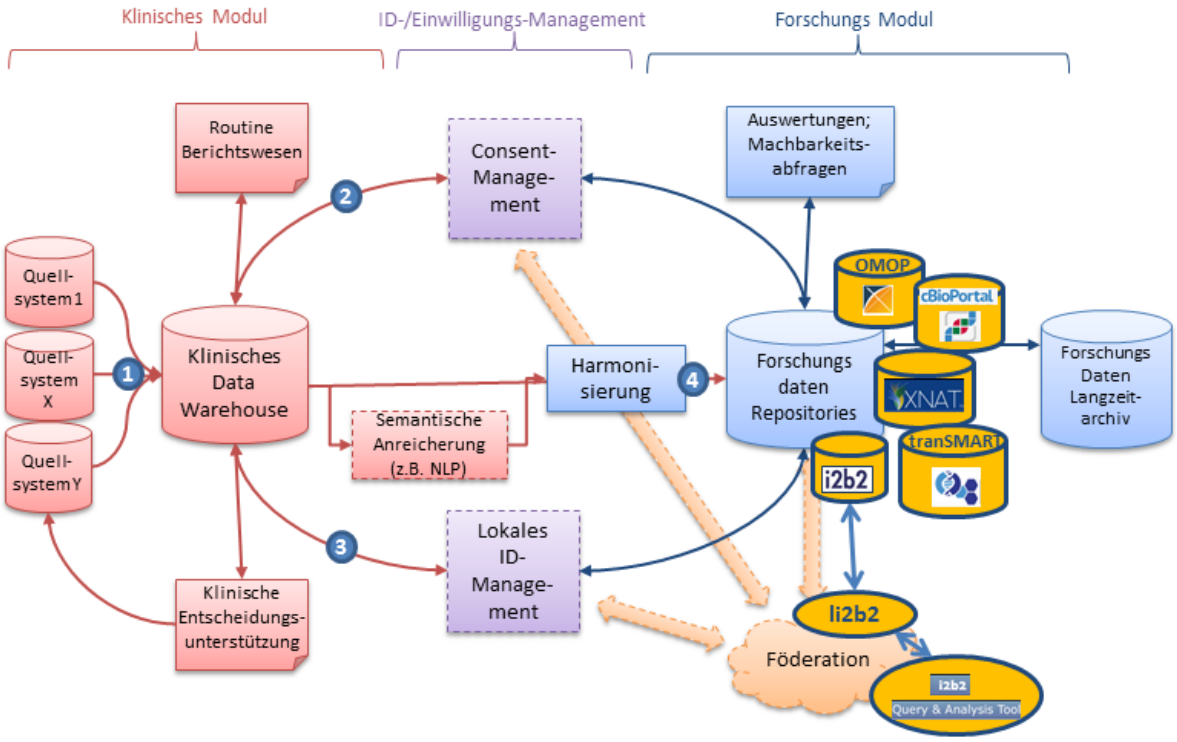

Abbildung 3 Erweiterung der lokalen DIZ Architektur um eine lokale (li2b2) und eine zentrale Softwarekomponente (i2b2 Query & Analysis Tool) zur Unterstützung föderierter Abfragen/Auswertungen

Föderierte Fallzahlabfragen dienen der Identifikation bzw. dem Auffinden von geeigneten Patientenkohorten für in Planung befindliche Forschungsvorhaben und liefern dem Anfragenden lediglich Ergebnisse in aggregierter Form. Das zentrale i2b2 Query & Analysis Tool (Web Client) stellt eine graphische Benutzeroberfläche zur Formulierung von Anfragen (Boole'sche Ausdrücke zur Eingrenzung von Patientenkohorten anhand klinischer Parameter) zur Verfügung und verwaltet diese Anfragen. In diesem Tool werden keine personenbezogenen Daten von Patienten verarbeitet. Die personenbezogenen Daten von Benutzern hingegen, die Abfragen formulieren bzw. einreichen, können im Rahmen der

Protokollierung (vgl. Abschnitt 5) gespeichert werden. Darauf werden Benutzer beim Login in den i2b2 Web Client hingewiesen.

Die Verbindung zwischen dem zentralen i2b2 Web Client und den an den jeweiligen MIRACUM-Standorten etablierten i2b2 Forschungsdaten Repositories wird über sogenannte "light" i2b2 (li2b2) Komponenten hergestellt. Dies sind der zentrale li2b2-Suchbroker und die jeweils lokalen li2b2-Such-Clients.

Nachdem Wissenschaftler mittels des i2b2 Web Clients anhand klinischer Merkmale Ein- und Ausschlusskriterien eine Kohorte definiert haben, wird diese Abfrage zunächst zentral im li2b2-Suchbroker gespeichert. Die lokalen li2b2-Such-Clients der Standorte rufen in regelmäßigen Abständen neu hinzugekommene Abfragen vom li2b2-Suchbroker ab (vgl. Abbildung 4) und ermitteln, welche Datensätze in ihrem jeweiligen lokalen Forschungsdaten Repository den Suchkriterien entsprechen. Der Inhalt der Abfrage sowie die lokal gefundenen Datensätze können an jedem Standort von einer dazu berechtigten Person eingesehen werden.

Gemäß der vorliegenden Version dieses Datenschutz-Konzepts werden **NIE** patientenbezogene Daten das Hoheitsgebiet des jeweiligen Standorts verlassen. Lediglich aggregierte Fallzahl-Ergebnisse werden zentral zusammengeführt.

Ein illustriertes Beispiel hierzu befindet sich im Anhang 5: Illustriertes Beispiel einer föderierten Fallzahlabfrage.

Die lokalen li2b2-Such-Clients liefern über den li2b2-Suchbroker lediglich die aggregierte Anzahl gefundener Datensätze, die den Suchkriterien entsprechen, an den Web Client des i2b2 Query & Analysis Tools zurück, so dass die Suchergebnisse aller antwortenden MIRACUM Standorte dort für den abfragenden Benutzer visualisiert werden können. In der aktuellen Projektphase (2018) und somit auch in der hier vorliegenden Version des Datenschutzkonzepts ist noch keinerlei Herausgabe von patientenbezogenen Daten aus einem lokalen DIZ vorgesehen.

Ausblick

Ziel der Durchführung solcher föderierter Machbarkeitsabfragen ist es, heraus zu finden, ob für eine nachfolgend geplante Auswertung zur Beantwortung einer bestimmten Forschungsanfrage in der Gesamtheit der im Konsortium zusammen geschlossenen Standorte ausreichend Patienten für eine statistisch signifikante Auswertung verfügbar sind. Sollte für einen Forscher die zurückgemeldete Gesamtanzahl ausreichend sein, so sind - je nach geplantem Forschungsprojekt und zugrunde liegender Forschungsfrage - zwei mögliche Vorgehensweisen denkbar.

1. *Die Forschungsfrage lässt sich mit einer föderierten Auswertung beantworten. In diesem Fall kann das im nachfolgenden Abschnitt beschriebene Vorgehen zum Tragen kommen.*
2. *Für die komplexere Fragestellung ist evtl. eine zentrale Zusammenführung von Daten und evtl. sogar eine Integration dieser Daten mit noch anderen Datenquellen (falls hierfür die Patienteneinwilligung vorliegt) erforderlich:* ***In einem solchen Fall muss nun ein völlig abgetrenntes neues Verfahren initiiert werden****, in dem die Freigabe*

der Daten durch die jeweils lokalen Use & Access Komitees genehmigt werden muss und im Falle einer Genehmigung auch die Genehmigung der jeweiligen Ethikkommissionen erforderlich ist, sowie letztendlich ein Vertrag zur gemeinsamen Datennutzung zwischen allen Parteien zu schließen ist. ***Für das geplante gemeinsame Forschungsprojekt ist dann ein neues - vom hier vorgelegten generischen MIRACUM Datenschutzkonzept völlig unabhängiges - Datenschutzkonzept zu erstellen*** *und von den jeweils zuständigen Datenschutzbehörden genehmigen zu lassen.*

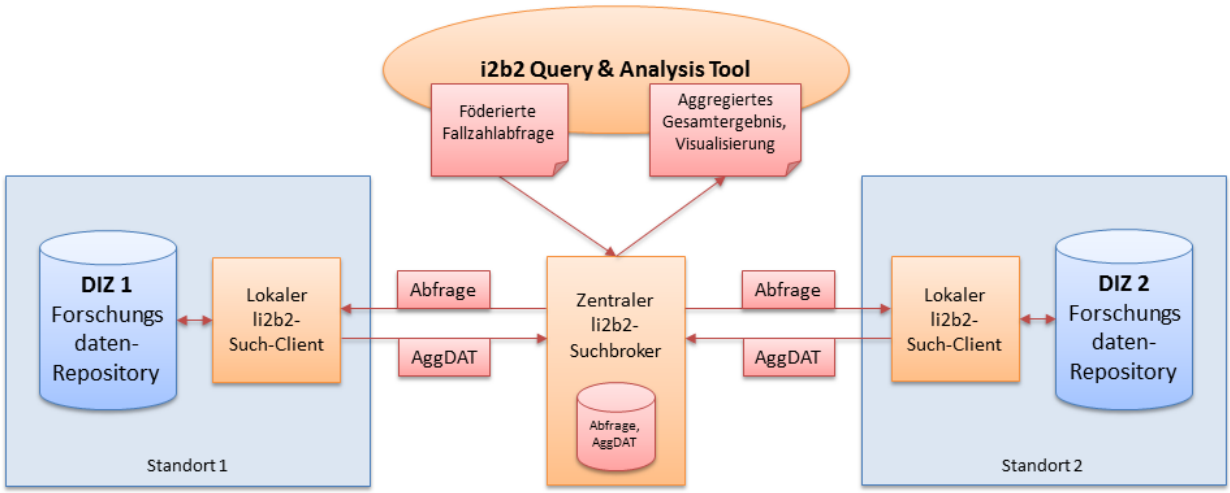

Abbildung 4 Datenfluss für föderierte Machbarkeitsstudien (beispielhaft nur für 2 Datenintegrationszentren dargestellt)

Der zentrale Web Client des i2b2 Query & Analysis Tools sowie der zentrale li2b2-Suchbroker werden vom DIZ des MIRACUM-Partners Gießen betrieben.

Föderierte Auswertungen

Analog zum Konzept der föderierten Fallzahlabfragen können in MIRACUM auch föderierte Auswertungen durchgeführt werden. Dazu werden Auswertungsalgorithmen zunächst von einem die Auswertung initiierenden Standort auf der Basis seines jeweils eigenen lokalen Forschungsdaten Repositories getestet und validiert. Ein solcher validierter Auswertungsalgorithmus kann dann von diesem MIRACUM Standort an einer dafür vorgesehenen Stelle in der MIRACUM Kollaborationsplattform (Confluence, Atlassian) abgelegt werden (an gleicher Stelle ist auch eine entsprechende Textbeschreibung dieser Auswertung und die Zielsetzung/medizinische Fragestellung die der Auswertung zugrunde liegt einzutragen). Daraufhin erhalten die entsprechenden DIZ-Verantwortlichen der anderen MIRACUM Standorte eine Mail mit dem Hinweis, dass eine föderierte Auswertungsanfrage vorliegt. Diese Auswertungsanfrage wird dann dem jeweiligen Use & Access Komitee (UAC) des Standorts vorgelegt, so dass dieses entscheiden kann, ob sich der jeweilige Standort an der vorgeschlagenen Konsortial-Auswertung beteiligt. Falls diese Entscheidung des lokalen UAC positiv ausfällt, lädt der verantwortliche Standort-Administrator des entsprechenden Forschungsdaten Repositories die Auswertungsalgorithmen von der MIRACUM Kollaborationsplattform in seine lokale Umgebung und startet den Auswertungsprozess.

Grundkonzept: **MIRACUM bringt die Analysen zu den Daten und nicht die Daten zu den Analysen.**

Die Ergebnisse eines Auswertungsprozesses sind jeweils nur aggregierte Ergebniswerte, aber niemals patientenbezogene Einzeldaten. Da allerdings durch bilden von Subgruppen in den Auswertungen sehr geringe aggregierte Ergebnisse (z.B. x=1) in einzelnen Subgruppen der durchgeführten Auswertung entstehen könnten, besteht ein potentielles Risiko, auch durch aggregierte Ergebnisse Informationen über die zugrundeliegenden Datensätze zu erhalten. Aus diesem Grund kommt das sogenannte **Cell-Obfuscation-Verfahren** zum Einsatz. Hierbei werden Ergebnisse im Bereich 1 <= x < OG, mit OG als parametrierbare Obergrenze des potentiell riskanten Wertebereichs, auf x = OG gesetzt und somit eine gezielte Abfrage einzelner Datensätze verhindert.

Der Standort-Administrator des entsprechenden Forschungsdaten-Repositories lädt dann die Datei mit diesen aggregierten Auswertungsergebnissen an die entsprechende Stelle auf der zentralen MIRACUM Kollaborationsplattform zurück. Der Forscher des MIRACUM Standorts, an dem die Auswertung initiiert wurde, kann dann - nach Bereitstellung aller Einzelergebnisse der an der Auswertung teilnehmenden MIRACUM Standorte - die Einzelergebnisse zu einem Gesamt-Auswertungsergebnis zusammenführen.

Jeder MIRACUM Standort ist frei in seiner Entscheidung ob er sich an einer solchen Konsortialauswertung beteiligt oder nicht. Entscheidet das UAC eines Standorts, dass sich der Standort an einer angefragten Auswertung nicht beteiligen möchte, so teilt dies der entsprechende Standort-Administrator über die MIRACUM Kollaborationsplattform mit und lädt sich die bereit gestellten Auswertungsalgorithmen *nicht* herunter. Dieser Standort bleibt dann von allen Folgeaktivitäten einer Konsortialauswertung ausgeschlossen.

4 Organisatorische Maßnahmen

Zugriff durch Systemadministratoren

Die in den lokalen Data Warehouses der DIZ-Standorte gespeicherten Daten können prinzipiell von den Administratoren der verwendeten IT-Infrastruktur eingesehen werden. Zugriffe auf die Daten durch Administratoren dürfen nur erfolgen, wenn dies zur Erfüllung ihrer Aufgaben zwingend erforderlich ist. Alle Administratoren sind schriftlich zu diesem Grundsatz und zur Verschwiegenheit zu verpflichten. Eine solche schriftliche Verschwiegenheitserklärung ist durch den jeweiligen DIZ-Standort zu regeln und kann ggf. durch die bereits bei der Einstellung des Mitarbeiters unterzeichneten Erklärungen abgedeckt sein.

Authentifizierung von Benutzern

Die Authentifizierung von Benutzern an zentralen Komponenten von MIRACUM erfolgt mithilfe eines zentralen Authentifizierungsdienstes, der eine einheitliche, einmalige Anmeldung unter Nutzung entsprechender Zugangsdaten bereitstellt. Die Prüfung von Identität und Berechtigung von Benutzern erfolgt durch die MIRACUM Geschäftsstelle im Rahmen der Freischaltung einer Benutzerkennung.

Die Authentifizierung von Benutzern für die Nutzung von lokalen Komponenten obliegt den Standorten selbst.

Ausblick

Perspektivisch wird die Authentifizierung an den MIRACUM Komponenten mit Hilfe eines föderierten Authentifizierungsdienstes wie DFN-AAI[4] realisiert werden.

Authentifizierung von Komponenten

Zugriffe von DIZ-Komponenten über das Internet auf zentrale Infrastrukturkomponenten von MIRACUM erfolgen ausschließlich nach erfolgreicher Authentifizierung (Identität der Komponente) sowie Autorisierung (Berechtigung der Komponente). Dafür werden etablierte Verfahren wie OAuth2 und OpenID Connect angewendet.

5 Technische Maßnahmen

Sicherheit der gespeicherten Daten

Die Sicherheit der gespeicherten Daten obliegt dem jeweiligen Standort und wird im standortinternen Datenschutzkonzept des jeweiligen DIZ-Standorts dargelegt. Ebenso unterliegen die ETL-Prozesse zur Befüllung der lokalen klinischen Data Warehouse Instanzen aus den Quellsystemen, sowie die Funktionsweise des ID-Managements den lokalen Datenschutzrichtlinien und sind nicht Bestandteil dieses generischen Datenschutzkonzepts. Details dieser Schritte werden innerhalb von MIRACUM durch entsprechende Verfahrensanweisungen (SOPs) geregelt.

Sicherheit der Kommunikation

Die DIZ werden prinzipbedingt verteilt betrieben und kommunizieren über das öffentliche Internet. Die Vertraulichkeit ihrer Kommunikation wird durch folgende Maßnahmen sichergestellt:

- Die Kommunikation zwischen den einzelnen Komponenten erfolgt grundsätzlich über verschlüsselte Verbindungen (HTTPS). Die dafür eingesetzten Schlüssel und Zertifikate sind so zu erstellen, dass sie den aktuell anerkannten Anforderungen entsprechen (z.B. Schlüssellänge, Algorithmus, siehe BSI TR-02102[5]).
- Durch Firewalls ist sichergestellt, dass die Server, auf denen die zentralen Komponenten laufen, nur über diejenigen Protokolle und Ports erreichbar sind, die für die Kommunikation mit Benutzern oder anderen Komponenten unbedingt erforderlich sind (in der Regel HTTPS-Verbindungen). Der administrative Zugang ist auf das Intranet des jeweiligen Betreibers beschränkt.
- Alle Kommunikationsvorgänge zwischen den DIZ und den zentralen Komponenten werden von den DIZ initiiert. Diese können dadurch hinter einer Firewall oder einem Proxyserver betrieben werden, ohne über eine öffentliche Adresse aus dem Internet erreichbar zu sein.

Protokollierung

Es erfolgt eine Protokollierung der Zugriffe von Forschern auf die Komponenten sowie zwischen den Komponenten untereinander. Der Nutzer wird bei Erstzugriff durch den Authentifizierungsdienst in Form eines "Terms of use" darüber informiert und seine Zustimmung eingeholt. Das Protokoll enthält mindestens:

- Die Identität der zugreifenden Person oder Komponente.
- Datum und Uhrzeit des Zugriffs.
- Den Inhalt des Zugriffs (die übermittelten aggregierten Daten) oder Informationen, aus denen dieser rekonstruiert werden kann (z.B. Verweis auf einen Datenbankeintrag o.ä.).

Das Protokoll wird zusammen mit den Nutzdaten des entsprechenden Servers auf diesem Server gespeichert. Nach zwölf Monaten werden entsprechende Protokolldateien gelöscht. Die aufgezeichneten Daten werden nur für folgende Zwecke verarbeitet und eingesehen:

- Im Rahmen der technischen Administration (insbesondere zur Fehlersuche).
- Zur Aufdeckung möglicher Missbrauchsfälle (durch Stichproben und Suchen nach auffälligen Zugriffsmustern)

- Zur Erstellung anonymisierter Nutzungsstatistiken.

6 Wahrung von Betroffenenrechten

Rechtsgrundlage

Daten von Patienten verbleiben unter der Hoheit des jeweiligen DIZ-Standorts. In Hinblick auf die Rechtsgrundlage sowie die institutionsinternen Prozesse der Datenverarbeitung wird auf die Datenschutzkonzepte der teilnehmenden Standorte verwiesen. Diese sehen vor, dass Daten entweder auf Basis der entsprechend gültigen Rechtslage (LKHG, LDSG, BDSG, sowie EU-DSGVO) ausgewertet werden, oder die Patienten eine Einwilligung nach erfolgter Aufklärung gegeben haben, die eine Verwendung ihrer Daten im Rahmen der hier beschriebenen standortübergreifenden Prozesse erlaubt. In der Aufklärung werden Patienten nachvollziehbar über die Verarbeitung Ihrer Daten informiert.

Im DIZ lokal gespeicherte Datensätze können über die föderierten Fallzahlabfragen oder aber mittels lokal ausgeführter Auswertungsalgorithmen abgefragt werden. Dabei verlassen immer nur aggregierte Fallzahlen bzw. Auswertungsergebnisse den Standort (vgl. Abschnitt 3.3). Grundlage für die lokale Abfrage/Auswertung und die Rückmeldung der aggregierten Ergebnisse sind immer die am Standort anwendbaren landes- und bundesrechtlichen Datenschutzbestimmungen, sowie die EU-DSGVO. Die beschriebenen Daten werden durch nachvollziehbare Zugriffsrechte geschützt. Die Beschreibung der dafür notwendigen Komponenten und Prozesse ist in Abschnitt 4 gegeben.

Datenschutzrechtliche Abgrenzung zentraler Komponenten zu den lokalen DIZ-Strukturen

Die Erhebung und Speicherung der Daten von Patienten erfolgt nur in der jeweiligen Institution. Dort ist zunächst zu prüfen, ob eine lokale Einwilligung in die Verwendung und Weitergabe von klinischen Daten und/oder Biomaterialen als Rechtsgrundlage vorliegt. Falls dies nicht zutrifft, sind die jeweiligen landes - und bundesrechtlichen Regelungen mit den entsprechenden Ausnahmetatbeständen zu prüfen (siehe Anhang 2: Tatbestände für lokale Forschungsdaten Repositories). Sollten diese im Einzelfall eine Verwendung von Bestandsdaten aus dem Behandlungskontext für die medizinische Forschung zulassen, können diese auch ohne Einwilligung verwendet werden. Ebenso ist die Speicherung von Daten im DIZ unbedenklich, sofern sichergestellt ist, dass diese Daten auch nach der Speicherung im DIZ nur von der jeweiligen Institution eingesehen werden können (vgl. Abschnitt 7).

Die MDAT der Patienten werden in lokalen Komponenten, den Forschungsdaten Repositories, in pseudonymisierter Form gespeichert. Das DIZ steht unter lokaler Kontrolle des jeweiligen Standorts, und auch nur dort kann (mit erheblichem technischen Aufwand) mithilfe des Pseudonyms auf die Identität des Patienten geschlossen werden.

Falls es für einen der beteiligten DIZ-Standorte keine spezialgesetzlichen Ermächtigungsregelungen für die Verwendbarkeit der Daten aus dem Behandlungskontext zu Forschungszwecken (z.B. LKHG) gibt, so kann ggf. eine Erhebung der Daten auf Basis der Forschungsklauseln des jeweiligen LDSG (für öffentliche Stellen der Länder) oder des BDSG (für Stellen in privater Trägerschaft) möglich sein.

Aufklärung und Einwilligung (Ausblick)

Eine informierte Einwilligung, in der der Patient auch über sein Recht auf Auskunft und Widerruf aufgeklärt wird, wird zukünftig eine Rechtsgrundlage der Datenverarbeitung werden. Ein durch die AG Consent der MI-I erstellter und vom Arbeitskreis Medizinischer Ethik-Kommissionen in der Bundesrepublik Deutschland e.V. geprüfter Mustertext befindet sich in Anhang 3. Es handelt sich hierbei um eine noch nicht freigegebene Version, die noch mit dem Arbeitskreis der Datenschützer abgestimmt wird. Der Einwilligungsstatus eines Patienten wird im ETL-Prozess elektronisch an das Consent-Management-System übergeben werden.

Zum aktuellen Zeitpunkt der Projektphase ist ein solches Consent-Management noch nicht Bestandteil der DIZ-Infrastruktur, so dass die Datenverarbeitung aktuell ausschließlich auf den o.g. Ausnahmetatbeständen (siehe Anhang 2: Tatbestände für lokale Forschungsdaten Repositories) basiert.

Auskunft über gespeicherte Daten

Alle Patienten, deren Daten in den technischen Komponenten verwendet werden, haben das Recht, gemäß Art. 15 EU-DSGVO Auskunft über die über sie gespeicherten Daten zu erhalten. Der Antrag auf Auskunft ist schriftlich an die behandelnde Klinik/verantwortliche Institution zu stellen.

Berichtigung unrichtiger Daten

Sämtliche Patienten haben das Recht, gemäß Art. 16 EU-DSGVO unverzüglich eine Berichtigung bzw. Vervollständigung unrichtiger bzw. unvollständiger Daten zu verlangen. Der Antrag auf Berichtigung ist schriftlich an die behandelnde Klinik/verantwortliche Institution zu stellen und ggf. durch die Abgabe einer Erklärung zu ergänzen.

Widerruf, Löschung (Ausblick)

Sämtliche Patienten, deren Daten in den technischen Komponenten eines DIZ auf Basis der Einwilligung gespeichert und verwendet werden, haben jederzeit das Recht, ihre Einwilligung in die Verarbeitung der Daten zu widerrufen. Infolgedessen wird eine Löschung der Daten dieses Patienten in den Forschungsdaten Repositories eines DIZ-Standorts vorgenommen, soweit dies nicht durch vorherige Anonymisierung der Daten unmöglich gemacht wurde. Der Widerruf ist schriftlich an den Standort zu richten, der die Daten hält.

Diese Löschung der Daten im DIZ ist von den zuständigen Betreibern am jeweiligen Standort in einem angemessenen Zeitraum vorzunehmen und der Patient vom Vollzug zu unterrichten. Die meist impraktikable Löschung in Datensicherungen ist verzichtbar, sofern die Sicherungen nur durch den zuständigen Systemadministrator eingesehen werden können und alte Sicherungen regelmäßig gelöscht werden. Für den Sonderfall einer eventuell notwendig werdenden Datenwiederherstellung aus einer Datensicherung wird an jedem Standort Sorge getragen, dass Daten von Patienten, die ihre Einwilligung widerrufen haben, nicht wieder in das DIZ zurückgespielt werden.

In der aktuellen Projektphase werden Daten noch nicht auf Basis einer Einwilligung, sondern auf Basis der jeweiligen lokal geltenden Gesetzesgrundlage (siehe Anhang 2: Tatbestände für lokale Forschungsdaten Repositories) in den Forschungsdaten Repositories gespeichert. Aus diesem Grund kommt auch das Widerrufkonzept an dieser Stelle noch nicht zum Tragen.

Dauer der Speicherung

Die erhobenen Daten bleiben in den lokalen Komponenten der DIZ gespeichert, wie es im Rahmen der Patienteneinwilligung bzw. der angewendeten Gesetzesgrundlage (siehe Anhang 2: Tatbestände für lokale Forschungsdaten Repositories) vorgesehen ist. Falls die Daten nicht mehr in der vorgesehenen Form genutzt werden können (z.B. falls der Betrieb des DIZ nicht fortgeführt wird), ist von jedem DIZ-Standort individuell eine Entscheidung über den weiteren Umgang mit den Daten zu treffen, da diese möglicherweise weiter für eigene Forschungsvorhaben genutzt werden können.

7 Vergleich mit dem TMF-Datenschutzleitfaden

Das Konzept folgt dem Klinischen Modul und dem Forschungsmodul des TMF-Datenschutzleitfadens.

8 Abkürzungsverzeichnis

BDSG – Bundesdatenschutzgesetz

CIO – Chief Information Officer

DIZ – Datenintegrationszentrum

ETL – Extraction, Transformation, Loading (Data Warehouse Prozess zur Datenintegration)

EU-DSGVO – Europäische Datenschutz-Grundverordnung

IDAT – Identifzierende Daten (Personenmerkmale, Stammdaten)

LDSG – Landesdatenschutzgesetz

LKHG – Landeskrankenhausgesetz

MDAT – Medizinische Nutzdaten ohne direkten Personenbezug (aber ggf. Quasi-Identifikatoren enthalten)

MI-I – Medizininformatik-Initiative des BMBF

MIRACOLIX – Medical Informatics ReusAble eCo-system of Open source Linkable and Interoperable software tools

MIRACUM – Medical Informatics in Research and Care in University Medicine

PI – Principal Investigator/Principal Partner Coordinator

SOP – Standard Operating Procedure (Formale, standardisierte Verfahrensbeschreibung)

UAC – Use & Access Committee

9 Glossar

Zu Definitionen siehe v.a. DSGVO Art. 4 "Begriffsbestimmungen".

Anonymisierung – Veränderung personenbezogener Daten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem

unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können. (§3 BDSG)

Kommentare zur DSGVO: zur Pseudonymisierung der Daten (s.u. Art. 4 Abs 5): es fehlt eine Entsprechung in der Einwilligung in Art. 6. Jedoch wird es in der Praxis i.d.R. so gehandhabt, dass pseudonymisierte Daten für die verantwortliche Stelle, die keinen Zugang zum personalisierenden Schlüssel besitzt und auch sonst die Daten wenn überhaupt nur mit unverhältnismäßigem Aufwand personalisieren könnte, als *anonymisierte* Daten und nicht mehr als personenbezogene Daten gelten. Interessante Erwägungsgründe bez. der Pseudonymisierung: 28 und 29. Es fehlt eine Definition der *Anonymisierung* bzw. von anonymen Daten. Jedoch wird in Erwägungsgrund 26 auf diese eingegangen.

Einwilligung – Vereinbarung zwischen Patient und datenerhebender Stelle betreffs Erhebung und Verarbeitung personenbezogener Daten

DSGVO Art 4. Abs 11: [Eine] „*Einwilligung*“ der betroffenen Person [bezeichnet] jede freiwillig für den bestimmten Fall, in informierter Weise und unmissverständlich abgegebene Willensbekundung in Form einer Erklärung oder einer sonstigen eindeutigen bestätigenden Handlung, mit der die betroffene Person zu verstehen gibt, dass sie mit der Verarbeitung der sie betreffenden personenbezogenen Daten einverstanden ist;

Ermächtigung – Erlaubnisgewährung gegenüber Dritten, ein üblicherweise nicht zustehendes Recht im eigenen Namen auszuüben.

DSGVO Art 4. Abs. 8. „*Auftragsverarbeiter*“ eine natürliche oder juristische Person, Behörde, Einrichtung oder andere Stelle, die personenbezogene Daten im Auftrag des Verantwortlichen verarbeitet;

Pseudonymisierung – Ersetzung des Namens und anderer Identifikationsmerkmale durch ein Kennzeichen zu dem Zweck, die Bestimmung des Betroffenen auszuschließen oder wesentlich zu erschweren. (§3 BDSG).

DSGVO Art. 4 Abs 5. "*Pseudonymisierung*" die Verarbeitung personenbezogener Daten in einer Weise, dass die personenbezogenen Daten ohne Hinzuziehung zusätzlicher Informationen nicht mehr einer spezifischen betroffenen Person zugeordnet werden können, sofern diese zusätzlichen Informationen gesondert aufbewahrt werden und technischen und organisatorischen Maßnahmen unterliegen, die gewährleisten, dass die personenbezogenen Daten nicht einer identifizierten oder identifizierbaren natürlichen Person zugewiesen werden;

Widerruf – Zu jeder Zeit mögliche Zurücknahme einer Einwilligung

10 Literaturverzeichnis

[1]	K. Pommerening, J. Drepper, K. Helbing und T. Ganslandt, „Leitfaden zum Datenschutz in medizinischen Forschungsprojekten. Generische Lösungen der TMF 2.0,“ MWV, Berlin, 2014.

[2]	J. Drepper, „tmf-ev.de," 23 05 2014. [Online]. Available: http://www.tmf-ev.de/Desktopmodules/Bring2Mind/DMX/Download.aspx?EntryId=24476&PortalId=0. [Zugriff am 20 06 2014].
[3]	B. Kurth, H. Hense und W. Hoffmann, „gesundheitsforschung-bmbf.de," 2004. [Online]. Available: http://www.gesundheitsforschung-bmbf.de/_media/Empfehlungen_GEP.pdf. [Zugriff am 07 10 2013].
[4]	U. K. Schneider, Sekundärnutzung klinischer Daten - Rechtliche Rahmenbedingungen (TMF-Schriftenreihe Band 12), Berlin: MWV, 2015.
[5]	T. Hillegeist, Rechtliche Probleme der elektronischen Langzeitarchivierung wissenschaftlicher Primärdaten, Göttingen: Universitätsverlag Göttingen, 2012.
[6]	S. Wirth, „datenschutz-hamburg.de," 2012. [Online]. Available: http://www.datenschutz-hamburg.de/uploads/media/Hinweise_zur_Risikoanalyse_und_Vorabkontrolle.pdf. [Zugriff am 20 9 2013].
[7]	Bundesamt für Sicherheit in der Informationstechnik, „bsi.bund.de," 2013. [Online]. Available: https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzKataloge/itgrundschutzkataloge_node.html. [Zugriff am 20 9 2013].
[8]	K. Pommerening, „Wie soll eine Anleitung zur Erstellung eines Datenschutzkonzepts anhand des Datenschutzleitfadens der TMF aussehen?," in *Vortragsfolien zur TMF-Sitzungswoche vom 16. September 2014*, Berlin, 2014.

11 Anhang

Anhang 1: MI-I-Kerndatensatz

Der in MIRACUM zur Verwendung angestrebte, konsortiumübergreifend geltende und durch die NSG Arbeitsgruppe Interoperabilität definierte Datensatz.

Anhang 2: Tatbestände für lokale Forschungsdaten Repositories

Eine tabellarische Übersicht über alle Rechtsgrundlagen für die lokale Nutzung der Versorgungsdaten zu Forschungszwecken ohne Patienteneinwilligung.

Anhang 3: Derzeitiger Entwurf des Mustertextes "Broad Consent"

Der durch die AG Consent der MI-I erstellte und vom Arbeitskreis Medizinischer Ethik-Kommissionen in der Bundesrepublik Deutschland e.V. geprüfte Mustertext der informierten Einwilligung. Es handelt sich hierbei um eine noch nicht freigegebene Version, die noch mit dem Arbeitskreis der Datenschützer abgestimmt wird.

Anhang 4: Beschreibung der MI-I-Use-Cases

Eine Beschreibung der drei konsortiumübergreifenden und durch die MI-I vorgegebenen Use Cases.

Anhang 5: Illustriertes Beispiel einer föderierten Fallzahlabfrage

Eine illustrierte Beschreibung des Vorgehens für eine föderierte Fallzahlabfrage.

Danksagung

Dieses Datenschutzkonzept beruht auf dem generischen Datenschutzleitfaden der TMF und dem Datenschutzkonzept der German Biobank Alliance (GBA), welches unter maßgeblicher Mitwirkung von Dr. Martin Lablans et al. entstand.

---

[1] http://www.miracum.org/

[2] Grundwerte nach §21 LDSG-MVs

[3] Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V., "Leitfaden zum Datenschutz in medizinischen Forschungsprojekten" (http://www.tmf-ev.de/Publikationen/www.tmf-ev.de/datenschutz-leitfaden)

[4] Authentifikations- und Autorisierungs-Infrastruktur des Deutschen Forschungsnetzes (https://www.aai.dfn.de/)

[5] Kryptographische Verfahren: Empfehlungen und Schlüssellängen (https://www.bsi.bund.de/DE/Publikationen/TechnischeRichtlinien/tr02102/index_htm.html)