

# Medizininformatik-Initiative

Begleitstruktur – Geschäftsstelle des Nationalen Steuerungsgremiums



## MI-I-Metadaten zur Verfügbarkeit von Daten, Auswertungsmöglichkeiten und Kooperationsoptionen

Stand: 23.03.2017

# Medizininformatik-Initiative

Begleitstruktur – Geschäftsstelle des Nationalen Steuerungsgremiums



## Inhaltsverzeichnis

1. Einleitung und Zielsetzung	3
2. Methodik	4
3. Relevante Informationen	4
Mitglieder der Redaktionsgruppe und Mitwirkende	8

# Medizininformatik-Initiative

Begleitstruktur – Geschäftsstelle des Nationalen Steuerungsgremiums

## 1. Einleitung und Zielsetzung

Im Rahmen der Medizininformatik-Initiative (MI-I) sollen Daten zwischen Datenintegrationszentren (DIZ) und Datenempfängern ausgetauscht werden. Um bestimmen zu können, ob die Daten eines Standorts für ein Vorhaben relevant sind und in einer dem Vorhaben angemessenen Form verfügbar gemacht werden können, sind Informationen über die Daten und die Rahmenbedingungen des Datenaustauschs notwendig.

Es besteht eine Vielzahl von Querbeziehungen zu den Arbeiten anderer AGs und Redaktionsgruppen der Begleitstruktur der MI-I. Bzgl. „typischer“ Metadaten, also solcher die syntaktische, strukturelle und semantische Eigenschaften einzelner Schnittstellen, Datenelemente oder Datensammlungen beschreiben, wird eine Abstimmung mit den Ergebnissen der Redaktionsgruppe „Kerndatensatz“ angestrebt. Aspekte der Data Provenance lassen (mit gewissen Einschränkungen) Rückschlüsse auf die Qualität der Daten zu und enthalten Informationen über den Datenverantwortlichen, Rechte an den Daten und deren Urheberschaft. Hier besteht die Notwendigkeit zur Abstimmung mit den Arbeiten der AG „Data Sharing“, insbesondere mit der dort entwickelten einheitlichen Nutzungsordnung. Wesentliche Rahmenbedingungen des Datenaustauschs werden auch durch die konkrete Ausgestaltung der Einwilligungserklärungen festgelegt. Eine Abstimmung mit der in der AG „Consent“ entwickelten einheitlichen Einwilligungserklärung ist deshalb zwingend notwendig.

Die in den AGs entwickelten Rahmenbedingungen können und müssen zunächst einfache Minimalbedingungen festlegen, die nicht alle Möglichkeiten der Datennutzung abbilden werden. So könnten z.B. höher aufgelöste Daten oder besondere Datenarten (Bilddaten, genetische Daten) nur unter Beachtung bestimmter Sicherheitsvorkehrungen (Safe Setting) zur Verfügung gestellt werden. Solche spezifische Formen des Datenaustauschs können aber erst dann in der Breite eingesetzt werden, wenn sie systematisiert und, inklusive aller relevanten Rahmenbedingungen, strukturiert durch Metadaten transparent gemacht werden. Ein weiterer zentraler Aspekt ist, dass die Rahmenbedingungen des Datenaustauschs auch direkte Auswirkungen auf die Inhalte der ausgetauschten Daten haben können. So kann es beispielsweise vorkommen, dass durch unterschiedlich weitreichende Einwilligungserklärungen, Daten über manche Patienten nicht ausgetauscht werden können. In diesen Fällen wäre es wünschenswert, wenn entsprechende Informationen, bspw. wie viele Datensätze auf eine Anfrage aus welchen Gründen (fehlende Einwilligung vs. fehlende Datenerhebung) nicht zurückgegeben werden können, als Metadaten verfügbar wären.

Aus diesen Gründen erscheint es notwendig, im Rahmen der MI-I strukturierte, standardisierte und automatisch auswertbare Metadaten bzgl. der Verfügbarkeit von Daten zu entwickeln. Die

# Medizininformatik-Initiative

Begleitstruktur – Geschäftsstelle des Nationalen Steuerungsgremiums

wesentliche Motivation für die Entwicklung einer maschinenlesbaren Repräsentation besteht darin, diese für große Datenmengen sowie für einzelne Teildatensätze feingranular angeben und prüfen zu können und so aufwändige manuelle Prozesse zu vermeiden oder zu vereinfachen und organisatorische Komplexität zu reduzieren. Langfristig könnte evtl. sogar eine weitgehend automatisierte Form des Data Sharings auch unter komplexen Rahmenbedingungen erreicht werden (Stichwort: „Smart Contracts“).

## 2. Methodik

Für die weiteren Arbeiten wird ein iteratives, mehrstufiges Vorgehen vereinbart. Die folgenden Arbeitsschritte können in dem Zuge auch mehrfach durchlaufen werden:

- 1.) Zunächst ist abzustimmen, welche Informationsarten und Informationen im Rahmen des Strukturierungs- und Standardisierungsvorhabens berücksichtigt werden sollen. Hier wird auch eine zeitliche sowie inhaltliche Priorisierung erfolgen. Eine wesentliche Voraussetzung für diese Abstimmung ist eine Analyse der in den Konsortien und auf der nationalen Ebene vorgesehenen Anwendungsfälle und Anfragen. Die Zeitpunkte der Durchführung entsprechender Analysen orientieren sich daher an den Meilensteinen der Roadmap (siehe auch Punkt 3).
- 2.) Für ausgewählte Informationsarten und Informationen ist zu recherchieren, inwiefern bereits internationale Standards zur Repräsentation, Kommunikation und Verarbeitung existieren. Nicht an internationalen Standards und Entwicklungen ausgerichtete Eigenarbeiten sollen weitgehend ausgeschlossen werden. Wo notwendig sollen existierende Standards auf der Basis vorgegebener Weiterentwicklungspfade angepasst oder nationale Extensions etabliert werden. In manchen Bereichen wird man sich allenfalls an nicht formal standardisierten Best Practices orientieren können. Nicht zuletzt wird auch von der sorgfältigen Umsetzung dieses Schrittes die spätere internationale Anschlussfähigkeit der Ergebnisse der MI-I abhängen.
- 3.) Abschließend erfolgt die Umsetzung auf Basis der gewählten, angepassten und entwickelten Lösungen. Für alle Ergebnisse soll eine Versionierung implementiert werden. Dazu gehört auch eine technische voll rückwärtskompatible Umsetzung aller verabschiedeten Versionsstände, um eine flexible Kommunikation standort- oder konsortienübergreifend zu gewährleisten.

## 3. Relevante Informationen

Die folgende Auflistung stellt eine erste und nicht abschließende oder gar vollständige Sammlung relevanter Informationen zur Verfügbarkeit dar, die perspektivisch in Form strukturierter, standardisierter und automatisch auswertbarer Metadaten bereitgestellt werden sollen. Diese Sammlung ist jedoch im weiteren Verlauf der Medizininformatik-Initiative im Sinne eines iterativen Vorgehens weiterzuentwickeln und anzupassen. Erste Hinweise zu relevanten Vorarbeiten oder

# Medizininformatik-Initiative

Begleitstruktur – Geschäftsstelle des Nationalen Steuerungsgremiums

beispielhaften Standards wurden aufgenommen, allerdings auch ohne Anspruch auf Vollständigkeit.

- **Consent-Informationen**

- Aspekte (**Zeithorizont: Roadmap M 1.2 (Q4/18, Q4/20)**)
  - Proben vs. Daten bzw. Datenarten
  - Zweck- bzw. Verwendungseinschränkungen
  - Rekontaktierbarkeit
  - Transparenzgebote
- Standards
  - IHE BPPC, APPC, HL7 FHIR (z.B. Ressource ConsentDirective)
- Vorarbeiten
  - “Automatable Discovery and Access Matrix” der Global Alliance for Genomics and Health<sup>1</sup>

- **Organisatorische Rahmenbedingungen**

- Provenance-Aspekte (**Zeithorizont: Roadmap M 5.3 (Q3/20)**)
  - Aus welchen Abteilungen kommen die Daten
  - Rechte an den Daten / Urheberschaft
  - Beteiligte/Relevante IT-Systeme
- Aspekte der Nutzung (**Zeithorizont: Roadmap M 3.2 (Q3/19)**)
  - Abbildung zentraler Aspekte der Nutzungsordnung
  - Abbildung der relevanten Gremien/Entscheidungsprozesse
- Vorarbeiten
  - Organisationsstruktur im Qualitätsbericht der Krankenhäuser (Qb-R)<sup>2,3</sup>
  - “Automatable Discovery and Access Matrix” der Global Alliance for Genomics and Health<sup>1</sup>

- **Kontext der Datennutzung (Zeithorizont: Roadmap M 3.2 (Q3/19), M 4.2. (Q3/19))**

- für eine konkrete Anfrage geltende Rahmenbedingungen
  - z.B. Eigenschaften der Anfrage, der anfragenden Einrichtung, Authentifizierung des Anfragers, Vorvereinbarungen, vorliegende Genehmigungen und Gremien-Entscheidungen

- **Datenarten (z.B. klinische Phänotypdaten vs. genom. Daten) (Zeithorizont: Roadmap M 5.1 (Q4/17, Q4/18 und Q4/19))**

<sup>1</sup> <https://genomicsandhealth.org/work-products-demonstration-projects/automatable-discovery-and-access-matrix>

<sup>2</sup> [https://www.g-ba.de/institution/themenswerpunkte/qualitaetssicherung/qualitaetsdaten/qualitaetsbericht/servicedateien/\\_ \(XML-Schema\)](https://www.g-ba.de/institution/themenswerpunkte/qualitaetssicherung/qualitaetsdaten/qualitaetsbericht/servicedateien/_ (XML-Schema))

<sup>3</sup> [https://www.g-ba.de/downloads/62-492-1235/Qb-R\\_2016-07-21\\_iK\\_2016-08-06.pdf](https://www.g-ba.de/downloads/62-492-1235/Qb-R_2016-07-21_iK_2016-08-06.pdf) (Anlage 1)

# Medizininformatik-Initiative

Begleitstruktur – Geschäftsstelle des Nationalen Steuerungsgremiums

- Verweis auf Ergebnisse der Redaktionsgruppe „Kerndatensatz“
- Perspektivisch auch Metadaten zur Verfügbarkeit von Bioproben (in Abstimmung mit der German Biobank Alliance, GBA, und den Informationen aus dem Kerndatensatz)
- **Art des Zugangs zu den Daten / Anfragen**
  - Aspekte (**Zeithorizont: Roadmap M 7.1 (Q4/18), M 4.2. (Q3/19)**)
    - Safe Setting (z.B. Query-Schnittstelle)
    - Safe Data (z.B. Herausgabe von Mikrodaten)
  - Vorarbeiten
    - Regelungen der statistischen Ämter<sup>4</sup>
    - Regelungen des DIMDI zum Datentransparenzverfahren
- **Datenqualität**
  - Provenance-Aspekte (**Zeithorizont: Roadmap M 5.3 (Q3/20)**)
    - Zweck der Datenerhebung (Studien, Behandlung usw.)
    - Vorverarbeitungen (z.B. zu Datenschutzzwecken)
  - Datenqualität ist letztlich immer abhängig von der Fragestellung (**Zeithorizont: Roadmap M 7.1 (Q4/18)**)
    - Ggf. Typisierung von Fragestellungen notwendig
  - Vorarbeiten
    - Leitlinie zur Datenqualität von Registern und Kohorten (Band 4 der TMF-Schriftenreihe)
    - DFG-Projekt zur praktischen Erfassung der Datenqualität in Kohortenstudien (Ltg.Herr C. Schmidt aus Greifswald)
    - Begleitstruktur hat in der Konzeptphase die Aufgabe, einen Anforderungskatalog für eine Expertise zur Datenqualität und Bewertung der Nutzbarkeit von Datenbeständen zu erstellen (Teil des TMF-Antrags zur Begleitstruktur). Die eigentliche Expertise soll in der Aufbauphase erstellt werden.
    - Erfahrungen mit der Datenqualität im Rahmen der externen Qualitätssicherung nach § 137 SGB V
- **Möglichkeiten der Rückmeldung von Informationen (Zeithorizont: Roadmap M 3.2 (Q3/19), M 1.2 (Q4/18 und Q4/20))**
  - Die folgenden Anlässe erfordern eine geeignete Rückmeldung zum Ort bzw. zum Anwendungssystem der primären Erfassung/Anreicherung/Kuratierung von Daten:
    - Es werden Fehler in den bereitgestellten Daten festgestellt.
    - Bei der Verwendung bereitgestellter Daten werden Hinweise auf schwerwiegende Erkrankungen festgestellt, und aus dem vorliegenden

<sup>4</sup> <http://www.forschungsdatenzentrum.de/datenzugang.asp>

# Medizininformatik-Initiative

Begleitstruktur – Geschäftsstelle des Nationalen Steuerungsgremiums



Auswertungsszenario / UseCase lässt sich nicht ableiten, dass eine entsprechende Diagnose bereits gestellt wurde.

- o Der Empfänger von Daten / die auswertende Stelle ist dabei nicht verpflichtet, Analysen/Bewertungen der bereitgestellten Daten vorzunehmen, die über den unmittelbaren Verwendungszweck hinausgehen.
- **Repräsentativität (Inklusions- / Exklusionskriterien)**
  - o DIZ-bezogene Aspekte (**Zeithorizont: Roadmap M 7.1 (Q4/18)**)
    - Vorhaltezeit von Daten
    - Anteil von Nicht-Einwilligungen
    - usw.
  - o Anfragebezogene Aspekte (**Zeithorizont: Roadmap M 7.1 (Q4/18)**)
    - Welche Daten sind vorhanden, können aber für die Anfrage nicht zur Verfügung gestellt werden, z.B. aufgrund von nicht ausreichenden Einwilligungen

# Medizininformatik-Initiative

Begleitstruktur – Geschäftsstelle des Nationalen Steuerungsgremiums



## Mitglieder der Redaktionsgruppe und Mitwirkende

Johannes Drepper

Thomas Ganslandt

Josef Ingenerf

Fabian Prasser (Leiter der Redaktionsgruppe Metadaten zur Verfügbarkeit von Daten,  
Auswertungsmöglichkeiten und Kooperationsoptionen)

Ulrich Sax

Thomas Wendt

Sven Zenker