



INTEROP COUNCIL
for digital health in Germany

Einreichung eines
Interoperabilitätsproblems

Version

Dokumentversion	1.0
Datum	15.06.2022

Zweck

Sie haben einen Vorschlag für ein Thema, mit dem sich das Interop Council auseinandersetzen sollte? Wir freuen uns über Ihre Anregungen und diskutieren diese in unseren Besprechungen. Zur effektiven Bearbeitung bitten wir Sie, Ihren Vorschlag strukturiert einzubringen und das dafür vorgesehene Formblatt zu nutzen.

Die Koordinierungsstelle der gematik GmbH prüft Ihren Vorschlag auf Vollständigkeit sowie Plausibilität und setzt sich bei Rückfragen mit Ihnen in Verbindung. Im Anschluss wird der Vorschlag in den regulären Diskussions- und ggf. Abstimmungsprozess des Interop Councils eingebracht.

Über den Bearbeitungsstand der eingebrachten Vorschläge wird im Rahmen der regulären Sitzungen des Interop Councils informiert.

Ziel des Dokuments soll es sein, eine einheitliche Form zu finden, um Interoperabilitätsprobleme an das Interop Council zu melden. Es dient einer übersichtlichen Form auf die wesentlichen Punkte und ermöglicht eine strukturierte Einordnung.

Allgemeine Informationen

Kurzzusammenfassung	Die korrekte Interpretation und damit auch Nutzung von medizinischen Daten erfordert das Vorliegen eines Data Dictionaries (Variablenbeschreibungen), für die es Stand heute keinen einfach verständlichen und in der Breite etablierten Standard gibt.
Autor	AG DDMIM
Datum der Einreichung	29.11.2022
Datum der letzten Bearbeitung	29.11.2022
Unterstützung aus dem Expertenkreis	Löbe, Matthias Majeed, Raphael Kadioglu, Dennis Vengadeswaran, Abishaa Stöhr, Mark (wenn Bewerbung erfolgreich)

Problembeschreibung & Kontext

Beschreibung	<p>Die korrekte Interpretation und damit auch die Nachnutzung von medizinischen Daten erfordert das Vorliegen eines Data Dictionaries, d.h. eines Katalogs der enthaltenen Datenelemente und ihrer Ausprägungen. Es resultiert die Notwendigkeit nach einer Beschreibung eines solchen Datensatzes, mit Hilfe dessen konkrete Fragen der Implementierung im biomedizinischen Umfeld beantwortet werden können.</p> <p>Auf den Datensatz selbst bezogen bspw.:</p> <ul style="list-style-type: none"> - Wie lautet die menschenlesbare Bezeichnung eines Datenfeldes? - Woher kommen die Daten? (Messmethode, Gerätetyp, Datenaufbereitung, ...) - Welche Ausprägungen, Datentypen und Formate werden erwartet? - Ist ein bestimmtes Datenfeld in dem gegebenen Datensatz enthalten? <p>In Zusammenhang mit weiteren Partnern:</p> <ul style="list-style-type: none"> - Welcher Datensatz enthält ein konkretes Datenfeld? - Welche Datenfelder enthält ein Datensatz mit ähnlichem Fokus gewöhnlicherweise und welche Annotationen kann ich sinnvoll übernehmen? <p>Diese Data Dictionaries sollten entweder zusätzlich zu den Daten bereitgestellt oder an anderer bekannter Stelle veröffentlicht werden. Für die Erstellung eines Data Dictionaries</p>
--------------	--

	<p>muss jedoch explizit Aufwand investiert werden, was häufig nicht oder nicht ausreichend erfolgt. Dies liegt nach Ansicht der Autoren hauptsächlich daran, dass es keinen hinreichend einfach verständlichen Standard gibt, der sowohl die inhaltlichen als auch technischen Anforderungen abdeckt. Somit entstehen divergente tabellenbasierte Modelle, denen wichtige Informationsartefakte fehlen.</p>
Wer oder was ist betroffen	<ul style="list-style-type: none"> - Datennutzer im Sinne einer dritten Partei, d.h. Personen/Institutionen, die nicht der Datenerzeuger sind - Datenerzeuger im Sinne einer Partei, die eigenverantwortlich, mit einer gewissen Zielsetzung bestimmte Daten sammelt
Warum ist das Problem vorhanden	<p>Die Beschreibung von Datensätzen geschieht aktuell meist in tabellarischer Form, wobei die Variablen durch erklärende Spalten näher beschrieben werden. In den Spalten werden Attribute wie textuelle Beschreibungen, Maßeinheiten, zulässige Ausprägungen oder Formatierungsinformationen hinterlegt. Auswahl, Umfang und Bedeutung der einzelnen Spalten unterscheiden sich jedoch häufig von Anwendungsfall zu Anwendungsfall, oder abhängig von der eingesetzten Software zur Datensammlung. Daher ist die Interpretation mit Aufwand/Missverständnissen verbunden.</p> <p>Verbreitete EDC-Systeme wie REDCap, OpenClinica, etc. stellen diese Informationen in tabellarischer Form dar (Excel/CSV) und sind nicht unmittelbar kompatibel. Die Erstellung und Pflege eines Data Dictionaries ist aufwändig und zunächst an keine formalen Kriterien gebunden. Orientierung können einerseits ausformulierte Metadaten-Konventionen wie beispielsweise der Dublin Core oder andererseits die ISO-Norm 11179-3 für den Aufbau von Metadaten-Registern im Allgemeinen bieten. Auch gibt es verfügbare Standards zur Formalisierung und Kommunikation von Metadaten im Studienkontext, z.B. CDISC ODM/SDM oder HL7 FHIR. Allen gemein ist eine hohe Hürde der Einarbeitung und Implementierung. Für die genannten Ziele sind die Standards entweder zu allgemein (z.B. Dublin Core, SKOS) oder zu speziell (z.B. CDISC ODM).</p>
Betroffene Use Cases	<ul style="list-style-type: none"> - Datenerfassungen in prospektiven Forschungsvorhaben - Secondary Use von medizinischen Routinedaten - Datenintegrationsprojekte, auch im Hinblick auf EHDS und EOSC-A - Nachnutzung von medizinischen Forschungsdaten - Allgemeine Spezifikationsdokumente für Metadatenschemas, Kerndatensätze, Vokabulare
Ende-zu-Ende Betrachtung	<p>Medizinische Daten, die ursprünglich für einen bestimmten Zweck erhoben worden sind, werden zunehmend darüber hinaus verwendet. Dementsprechend wird die technische Herausforderung, Datenquellen miteinander interoperabel zu machen, als eine der vordringlichsten betrachtet. Eine entscheidende Voraussetzung ist jedoch das Vorliegen von ausreichend Informationen über die zu integrierenden Daten.</p>

Ebene der Interoperabilität	Sowohl die syntaktische als auch die semantische Ebene der Interoperabilität sind betroffen. Die Autoren dieser Problembeschreibung fokussieren sich zunächst jedoch auf die semantische Ebene.
Betroffene Bereiche	Sowohl die medizinische Forschung als auch der Versorgungsbereich sind gleichermaßen von dem Problem betroffen.
Auswirkungen	<ul style="list-style-type: none"> - Zur Interpretation/Nutzung von Datensätzen ist die Einarbeitung in anwendungsspezifische Datenbeschreibungen nötig. - Dateninterpretationen und -zusammenführungen sind möglicherweise unmöglich. - Dateninterpretationen und -zusammenführungen sind möglicherweise fehlerhaft. - Kollaborationsvorhaben kommen unter Umständen nicht zustande. - Einzelne Attribute aus Data Dictionaries sind nicht vergleichbar. Hierzu zählen z.B. Festlegungen zu Pflichtfeldern, Formaten, Genauigkeiten und Kodierungen.
Fristen / gesetzliche Abhängigkeiten (optional)	<i>keine</i>
Relevante Akteure / Stakeholder	<p>Die relevanten Akteure, Stakeholder oder auch Fachgruppen, die bei der Bearbeitung des Problems beachtet werden müssen, sind folgende:</p> <ul style="list-style-type: none"> - Interop-Council - GMDS AG Terminologien, FAIR Data, Secondary Use - DGEpi - TMF AG IT-QM

Lösungsansatz

Mögliche Lösung	<p>Aufbauend auf einer umfassenden Sichtung existierender Data Dictionaries werden verbreitet genutzte Attribute herausgearbeitet. In einem zweiten Schritt wird ein "Minimales Informationsmodell" entwickelt, in welchem typische Beschreibungsfelder aufgelistet, terminologisch gesäubert und in einem modularen System geordnet werden sollen. Es werden Crosswalks zu bestehenden Vokabularen und Konventionen bzw. zu internationalen Standards bereitgestellt. Des Weiteren soll eine Vorlage entwickelt werden, die für zukünftige Data Dictionaries verwendet werden kann und auf dem Modell basiert. Darauf aufbauend sollen prototypisch technische</p>
-----------------	---

	<p>Lösungen für bestehende Informationssysteme aufgesetzt werden. Der Arbeitskreis soll sich dabei an dem üblichen Aufbau von Minimal Information Models [https://fairsharing.org/search?q=minimal%20information] orientieren.</p>
Aufteilung in Teilaspekte	<p>Zunächst gilt der Fokus der inhaltlichen/semantischen Ebene, um daraus dann konkrete technische Spezifikationen und Implementierungen abzuleiten.</p>
Vorarbeiten (optional)	<p>Vorarbeiten der AG DDMIM (Data Dictionary Minimal Information Model)</p> <ul style="list-style-type: none"> - Module des Minimalen Informationsmodells für Data Dictionaries - Genereller Ansatz und Struktur des DDMIM - Erfahrungen bei der Spezifikation von Datensätzen in Forschungsvorhaben