

Cramer, Elena; Krumreihn, Aline; Schaffert, Corinna; Stiene, Meike

Research Report

Metadaten, Metadatenregister und Datenverknüpfungen im Gesundheitsbereich: Eine Einordnung

Forschung Aktuell, No. 03/2024

Provided in Cooperation with:

Institute for Work and Technology (IAT), Westfälische Hochschule, University of Applied Sciences

Suggested Citation: Cramer, Elena; Krumreihn, Aline; Schaffert, Corinna; Stiene, Meike (2024) :
Metadaten, Metadatenregister und Datenverknüpfungen im Gesundheitsbereich: Eine Einordnung,
Forschung Aktuell, No. 03/2024, Institut Arbeit und Technik (IAT), Gelsenkirchen,
<https://doi.org/10.53190/fa/202403>

This Version is available at:

<https://hdl.handle.net/10419/284725>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Forschung aktuell

März 2024

DOI: <https://doi.org/10.53190/fa/202403>

Ausgabe 03/2024

Metadaten, Metadatenregister und Datenverknüpfungen im Gesundheitsbereich – eine Einordnung

Elena Cramer, Aline Krumreihn, Corinna Schaffert, Meike Stiene



Auf den Punkt

- Im Gesundheitswesen werden zahlreiche und unterschiedlichste Daten erhoben und für die Gesundheitsversorgung von Patientinnen und Patienten, aber auch für die Forschung und Entwicklung und zur Berichterstattung genutzt.
- Insbesondere im Ruhrgebiet leben mehr als fünf Millionen Menschen und dort sind viele Primärdatenhalter ansässig.
- Für die Beantwortung von sozialmedizinischen Fragestellungen benötigt es die Zusammenführung von verschiedenen Datenkategorien, welche von unterschiedlichsten Datenhaltern stammen.
- Metadatenregister beschreiben detailliert vorhandene Datensätze und können so die Zusammenführung von Datensätzen unterstützen.

Inhalt

1	Einleitung und Problemstellung	1
2	Zielsetzung	3
3	Methodik	4
4	Überblick und Einordnung	5
5.	Diskussion	12
6.	Ausblick & Forschungsagenda	14

Abbildungsverzeichnis

Abbildung 1: Übersicht über Datenkategorien, Daten und mögliche Datenhalter im Kontext von sozialmedizinischen Fragestellungen	9
--	---

Tabellenverzeichnis

Tabelle 1: Suchstrings für die Literaturrecherche	4
Tabelle 2: Metamodellkonstrukt für das angestrebte Metadatenregister mit Beispielen	11

Schlüsselwörter:

Metadaten, Gesundheitsversorgung, FAIR Prinzipien, Datennutzung, Ruhrgebiet

1 Einleitung und Problemstellung

Im Gesundheitssektor werden zahlreiche Daten für unterschiedlichste Zwecke erhoben, verarbeitet und genutzt; von der Patientenversorgung bis hin zur Verwaltung und Forschung. So existieren beispielsweise Daten in Form von Gesundheits- und Patientenakten, aber auch auf anderen Ebenen, wie beispielsweise bei den Gesundheitsämtern, Krankenkassen, Kassenärztlichen Vereinigungen, medizinischen Registern, Forschungseinrichtungen, Apotheken und vielen weiteren Einrichtungen liegen Daten vor. Dabei können insbesondere für die Beantwortung von sozialmedizinischen Fragestellungen verschiedene Datensätze relevant sein. Sozialmedizinische Fragestellungen beschäftigen sich dabei mit dem Zusammenspiel von sozialen Faktoren und Gesundheit, wie beispielsweise sozialen Determinanten von Gesundheit, gesundheitlichen Ungleichheiten, der Gesundheit von Minderheiten, der Gesundheitsförderung von Gemeinschaften oder sozialen Gruppen. Dabei stehen beispielsweise Faktoren wie Arbeit und Arbeitslosigkeit, Migration oder Umweltbedingungen und das Wohn- und Lebensumfeld in Zusammenhang zu Krankheitsaufkommen eine bedeutende Rolle, da in der Sozialmedizin die medizinische Sichtweise mit der sozialwissenschaftlichen Sicht auf Krankheit und Behinderung verbunden wird (Waller, 2007).

Neben Daten aus dem Gesundheits- und Sozialwesen sind ebenfalls raumbezogene und umweltbezogene Daten sowie Daten für die Stadtentwicklung für sozialmedizinische Fragestellungen von Relevanz. Sowohl für die medizinische und gesundheitswissenschaftliche Forschung und die Patientenversorgung, als auch für die Gesundheitspolitik und Verwaltung von Gesundheitseinrichtungen werden diese Daten genutzt. Die Anwendungs- und Erhebungsmöglichkeiten sind dabei insgesamt betrachtet vielfältig. Klinische Daten und Patientendaten umfassen Informationen über Diagnosen, Behandlungspläne, die medizinische Versorgung oder Bildungsergebnisse. Im Rahmen von Forschungsprojekten liegen Daten über Erkenntnisse zu Therapiemöglichkeiten und Krankheiten vor, aber auch zur Untersuchung von neuen Versorgungsformen und -angeboten. Auch für politische Dimensionen und die Gesunderhaltung der Bevölkerung werden Daten erhoben und gesammelt, wie beispielsweise über das Gesundheitsmonitoring vom Robert-Koch-Institut (Kurth, 2012). Darüber hinaus verpflichten einige gesetzliche Vorgaben Primärdatenhalter zur Datenerhebung, -speicherung und -übermittlung. So müssen beispielsweise Krankenhäuser gemäß Paragraph 21 des Krankenhausentgeltgesetzes maschinenlesbare Datenträger mit umfassenden Informationen über das vorangegangene Kalenderjahr (z.B. über Anzahl des Personals, abgerechnete DRG-Fälle, Aufnahme- und Entlassdatum von Patient:innen, gestellte Haupt- und Nebendiagnosen, etc.) an das Institut für das Entgeltsystem im Krankenhaus übermitteln. In diesem Kontext sind auch Abrechnungsdaten zu nennen, welche Informationen über die Abrechnung von medizinischen, pflegerischen und therapeutischen Leis-

tungen beinhalten, aber auch Informationen über Gesundheitsausgaben und Versicherungsansprüche beinhalten können. Darüber hinaus werden auch umweltbezogene Daten, wie beispielsweise Luftqualität oder Lärmbelastungen, regelmäßig erhoben.

Durch aktuelle Digitalisierungsentwicklungen im Gesundheitswesen entstehen durch die elektronische Patientenakte oder telemedizinische Anwendungen zunehmend digitale Datenaufzeichnungen und digitale Datenverfügbarkeiten. Auch über Gesundheits-Apps und Wearables können personenindividuelle Gesundheitsdaten (z.B. Aktivitätsniveau, Herzfrequenz, Schlaf, etc.) erhoben und gespeichert werden.

Diese Auflistung skizziert nur exemplarisch die Vielfältigkeit gesundheitsbezogener Daten und deren Nutzung.

Von besonderem Interesse für die Beantwortung von sozialmedizinischen Fragestellungen sind vielmals Metropolregionen mit einer hohen Bevölkerungsdichte. Das Ruhrgebiet ist die drittgrößte Metropolregion in Europa und beheimatet mehr als fünf Millionen Menschen (Regionalverband Ruhr, 2021). Es ist der größte Ballungsraum Europas mit stabilen Bevölkerungszahlen und einer leichten Verjüngung der Gesellschaft sowie rund 150 verschiedenen Nationalitäten in dieser Region. Insgesamt ist das Ruhrgebiet geprägt von sozialen und kulturellen Unterschieden, sowie von großen Wohlstands- und Bildungsunterschieden der Bewohner:innen (ebd.). Insbesondere viele Primärdatenhalter, wie beispielsweise Krankenhäuser und niedergelassene Ärzt:innen, aber auch Krankenkassen sind im Ruhrgebiet ansässig. Das Ruhrgebiet steht mit seinen zahlreichen Städten vor der Herausforderung, eine Vielzahl kommunaler und regionaler Datenhalter ohne einheitliche Vorgaben zur Erhebung und Dokumentation der Daten zu haben.

Es muss also hervorgehoben werden, dass Daten datenhalterspezifisch erhoben, gespeichert und verarbeitet werden und dass eine Zusammenführung der Daten für sozialmedizinische Forschungsfragen unter Umständen nicht bis nur schwer möglich ist. Damit Daten regional, aber auch regionsübergreifend verwertbar und nutzbar sind und damit sie zwischen Datenempfänger:innen und Datenzentren ausgetauscht werden können und für die Bestimmung, ob die jeweiligen Daten für ein geplantes Vorhaben relevant sind und verfügbar gemacht werden können, benötigt es qualitativ hochwertige Metadaten. Metadaten sind Daten, welche die Merkmale von Datenelementen spezifizieren (Medizininformatik-Initiative, 2017; TMF e.V., 2023). Über Metadaten wird festgelegt, welche Daten in welcher Form zu erheben sind (Erfassungsmerkmale). Über ein Metadatenregister oder Metadaten Repository existiert eine zentrale Datenbank, über welche die Daten verwaltet werden können, und die Vergleichbarkeit der Daten verbessert wird. Da es sich, wie eingangs beschrieben, bei gesundheitsbezogenen Daten oftmals um sensible Daten handelt, ist der Datenschutz von entscheidender Bedeutung. Die Privatsphäre von bspw. Patient:innen und Sicherheit der Daten wird in Deutschland über die europäische Datenschutzgrundverordnung (DSGVO) geregelt. Aus diesem Grund dürfen Gesundheitsdaten von nationalen Behörden nicht auf zu kleinräumiger Ebene

ausgegeben werden. Daten müssen besonders geschützt werden und dürfen auch für Forschungszwecke nicht beliebig unter Primärdatenerheber:innen und Institutionen ausgetauscht werden. Dennoch kann es insbesondere für regions- und zielgruppenspezifische Versorgungs- und Interventionsmaßnahmen hilfreich sein, wenn eben diese kleinräumigen Daten nutzbar gemacht werden können. In Regionen wie dem Ruhrgebiet mit starkem Sozialgradienten und Wohlhabens- und auch Gesundheitsunterschieden ist es sinnvoll, wenn Unterstützungsmaßnahmen zielgerichtet entwickelt und verbreitet werden können.

Im Rahmen der „Digital Health Factory Ruhr“¹, gefördert durch das Bundesministerium für Bildung und Forschung (BMBF), ist daher die Entwicklung eines sozialmedizinischen Metadatenregisters für das Ruhrgebiet geplant. Das Institut Arbeit und Technik sowie die Hochschule für Gesundheit und MedEcon Ruhr haben das Projekt „Data4Health.Ruhr – Sozialmedizinisches Metadatenregister“² durchgeführt, um in einem ersten Vorbereitungsschritt Datenbestände für den sozialmedizinischen Datenverbund in der Gesundheitsmetropole Ruhr zu erfassen und zu bewerten. Für sozialmedizinische Fragestellungen können verschiedenste Datensätze relevant sein, welche zum einen oftmals dezentral gehalten werden und zum anderen in ihren Zielsetzungen differieren und daher nicht einfach zusammenführbar sind. Als Vorbereitung für ein sozialmedizinisches Metadatenregister wurde im Rahmen des Projektes ein systematischer Überblick erarbeitet, welcher der Grundlagenschaffung und Erfassung von Metadatenprojekten, Datenarten und -potenzialen im Ruhrgebiet dient.

2 Zielsetzung

Im Rahmen dieser Übersichtsarbeit wird eine mehrfache Zielsetzung verfolgt: Zum einen werden über eine Literaturrecherche Studien über die Entwicklung von Metadatenregistern identifiziert, zum anderen werden relevante Daten und Datenkategorien gesammelt und auch rechtliche Vorgaben recherchiert. Demzufolge sollen vier Forschungsfragen beantwortet werden:

1. Werden in der aktuellen Literatur Metadatenregister im Gesundheitsbereich entwickelt und gibt es ein Metadatenregister für das Ruhrgebiet?
2. Welche Daten und Datenkategorien sind für die Forschung relevant?
3. Welche gesetzlichen Vorgaben gibt es für Primärdatenhalter bezüglich der Datenerhebung?

¹ <https://digital-health-factory.ruhr/>

² <https://digital-health-factory.ruhr/projekte/sozialmedizinisches-metadatenregister/>

4. Wie kann eine Matrix-Struktur für ein Metadatenregister im Ruhrgebiet aussehen?

3 Methodik

Für die Beantwortung der ersten Forschungsfrage wurde eine semistrukturierte Literaturrecherche durchgeführt. Die Literaturrecherche wurde in insgesamt sechs Datenbanken (Web of Science, GoogleScholar, Livivo, PubPsych, PubMed, BASE) durchgeführt. Die relevanten Suchstrings wurden in deutscher und englischer Sprache definiert (Tabelle 1). Die Treffer wurden im Titel und Abstract auf inhaltlich relevante Ergebnisse gescreent. Relevante Studien wurden auf Volltext geprüft und die Inhalte zusammengefasst.

Tabelle 1 Suchstrings für die Literaturrecherche

Deutsch:

(Datenkategorien ODER Metadatenregister ODER Datenregister ODER Datensammlung ODER Datenpooling ODER Datenarten ODER Datenvorhaltung) UND (Daten ODER Gesundheitsdaten ODER Metadaten ODER Geodaten ODER Umweltdaten ODER Sozialdaten ODER Diagnosestatistik ODER Abrechnungsdaten) UND (Gesundheit ODER Sozialraum) UND (NRW ODER Ruhrgebiet ODER Deutschland)

Englisch:

(Data categories OR Metadata registry OR Data registry OR Data collection OR Data pooling OR Data types OR Data storage) AND (Data OR Health data OR Metadata OR Geodata OR Environmental data OR Social data OR Diagnosis statistics OR Accounting data) AND (Health OR Social space) AND (North Rhine Westphalia OR Ruhr area OR Germany)

Für die Beantwortung der zweiten Forschungsfrage wurden die Studien aus der Literaturrecherche auf genutzte Daten, Datenkategorien und Datenhalter gescreent. Zusätzlich wurde für die Erstellung einer orientierenden Übersicht über Daten und Datenhaltung auf Erfahrungswissen aus den wissenschaftlichen Arbeiten der Projektpartner zurückgegriffen.

Im Rahmen der dritten Forschungsfrage wurde für eine orientierende Übersicht über gesetzliche Vorgaben eine Online-Recherche durchgeführt. Dabei wurde gezielt nach gesetzlichen Vorgaben in Deutschland und Nordrhein-Westfalen gesucht, welche Primärdatenhalter gesetzlich verpflichten, Daten zu erheben, zu speichern und zu übermitteln. Die Recherche wurde für die unterschiedlichsten Sektoren durchgeführt.

Für die vierte Forschungsfrage wurden die identifizierten Studien auf Hinweise zum Aufbau der Metadatenregister untersucht und nach standardisierten Vorgaben für die Erstellung einer Matrixstruktur für das Ruhrgebiet gesucht.

4 Überblick und Einordnung

Für die Literaturrecherche konnten insgesamt sechs Studien und Berichte identifiziert werden, welche über Metadaten und Metadatenregister im Gesundheitskontext in Deutschland berichten. An dieser Stelle kann jedoch bereits hervorgehoben und festgehalten werden, dass keine Studie eine ähnliche oder synonyme Zielsetzung wie die Digital Health Factory Ruhr verfolgt hat und zurzeit kein Metadatenregister für das Ruhrgebiet vorliegt.

Damit Daten sinnvoll genutzt werden können, wurden die FAIR Datenprinzipien entwickelt, welche dazu beitragen sollen, ein nachhaltiges Forschungsdatenmanagement zu gewährleisten. Mit FAIR sind die Eigenschaften „auffindbar“ („findable“), „zugänglich“ („accessible“), „interoperabel“ („interoperable“) und „wiederverwendbar“ („re-usable“) gemeint (Wilkinson et al., 2016). Diese Prinzipien wurden entwickelt, um sicherzustellen, dass Daten effizient genutzt werden können und wissenschaftliche Forschung transparent und nachvollziehbar ist. Diese Prinzipien finden wiederkehrende Beachtung.

Das NFDI4Health Projekt³ (the National Research Data Infrastructure for Personal Health Data) befasst sich mit Daten aus klinischen, epidemiologischen und Public-Health Studien. Das Projekt beruft sich auf die Tatsache, dass viele Daten für die Entwicklung neuer Therapien, Versorgungskonzepte und Präventionsmaßnahmen dringend notwendig sind, diese jedoch vielmals nicht den FAIR-Prinzipien genügen. Dies bedeutet, dass die Daten aufgrund fehlender internationaler Standards oftmals nicht auffindbar sind, aber auch, dass die Regelung für die Datennutzung durch Dritte oft unklar ist, oder dass Datenbanken (wegen methodischer Heterogenität) nicht interoperabel sind. Das Ziel des genannten Projektes ist es, neue Möglichkeiten für die Nutzung personenbezogener Gesundheitsdaten unter Wahrung des Datenschutzes zu schaffen. Auf der Grundlage der FAIR-Grundsätze (Wilkinson et al., 2016) sollen Forscher:innen Dienste zur Auffindbarkeit von und Zugang zu strukturierten Gesundheitsdaten ermöglicht werden, ein föderaler Rahmen für datenhaltende Organisationen geschaffen werden, der Austausch und die Verknüpfung von Datensätzen (mit personenbezogenen Gesundheitsdaten unter Einhaltung der Datenschutzbestimmungen) erleichtert werden, automatisierte Dienste (z.B. über Such- und Analysetools) errichtet werden, die dauerhafte Speicherung von Daten ermöglichen, die Interoperabilität gefördert werden, die Zusammenführung von Daten aus verschiedenen Quellen ermöglichen, sowie die Zusammenarbeit

³ <https://www.nfdi4health.de/en/>

zwischen Forschungseinrichtungen verbessert werden. Die detaillierten Leistungen, welche in dem Projekt angestrebt werden, werden nach den Zielgruppen – Datenanalysten und Datenhalter – unterschieden. Für Datenanalysten sollen dabei beispielsweise die Funktionen einer zentralen Suchmaschine für die Auffindbarkeit von Studien und Datensätzen sowie FAIR-Kennzahlen, eine zentrale Datenzugangsstelle zur Einreichung von Anträgen auf Datenzugang, und viele weitere Funktionen ermöglicht werden. Für die Datenhalter werden weitere Funktionen erarbeitet, wie beispielsweise ein kurzes Benutzerhandbuch, wie Daten nach den FAIR-Prinzipien aufbereitet werden können, eine Vorlage für einen allgemeinen Plan zur Verwaltung von Forschungsdaten, Leitfäden und weitere Funktionen. Aktuelle Ergebnisse und Zugänge zu Portalen o.ä. waren online noch nicht verfügbar (Fluck, 2019).

Die Studie von Stausberg und Harkener (2019) befasst sich mit der Datenbündelung von Datenspezifikationen aus 15 Projekten. Dabei wurden die Daten in den Metadaten Standard ISO/IEC 11179-3 übertragen. Ein Datenelement wird durch eine konzeptuelle Domäne (z.B. Geschlecht) und in Kombination mit einem Wert (Value-Domain; z.B. männlich, weiblich, divers) definiert. Darüber hinaus wird beispielsweise nach eindeutigen Werten (enumerated value domain) und textuellen Definitionen (described value domain) unterschieden. Von insgesamt 15 Projekten wurden unterschiedlichste Daten (Excel, Access, Word Dateien) in deutscher oder englischer Sprache bereitgestellt, sodass insgesamt 3.935 Datenelemente untersucht werden konnten. Die Ergebnisse zeigen, dass sich die Projekte uneinig über die Bezeichnungen von Dokumentationsobjekten, Datenelementen und Werten waren. Sie waren sich auch uneinig über die Wertelisten, die durch die Aggregation von Einzelwerten entstehen. Eine evidenzbasierte Empfehlung über einen geeigneten Dokumentationsstandard war nicht ausreichend, um homogene Metadaten zu etablieren. Demnach schlussfolgern Stausberg und Harkener (2019), dass die Ergebnisse deutlich den Bedarf an Metadaten-Repositories unterstreichen. Die Studie hebt hervor, dass Metadaten-Repositories genutzt werden können, um die Formulare und Modelle für die Datenerhebung zu vereinfachen und die Qualität zu verbessern, um Variablen und Wertelisten zu harmonisieren und das Mapping von Daten und Metadaten zu unterstützen. Dennoch zeigt sich auch, dass es für komplexe Gesundheitsdaten eine große Herausforderung darstellt, die Modelle zu definieren. Dies zeigt sich an folgendem Beispiel: einerseits kann die Hauptdomäne “Krankheit” sein und ein “Herzinfarkt” könnte ein dazugehöriges Attribut sein, andererseits könnte der “Herzinfarkt” auch die Hauptdomäne sein und das Attribut könnte “Vorhandensein” sein. Es zeigt sich durch dieses Beispiel, dass die Definition von Gesundheitsdaten komplex und herausfordernd ist (Stausberg & Harkener, 2019).

In der Studie von Meissner et al. (2019) wurde der Fokus auf die Datenverfügbarkeit von Daten in Schwangerschaftsregistern gelegt, um durch die Bündelung von Daten, das Wissen über Schwangerschafts-Outcomes für Frauen mit rheumatischen Erkrankungen zu untersuchen. Vier Schwangerschaftsregister aus Frankreich, der Schweiz, Norwegen und Deutschland wurden einbezogen, sodass insgesamt Daten

zu 2.200 Schwangerschaften vorlagen. Bezüglich der rheumatischen Erkrankungen zeigte sich, dass alle Register übereinstimmend nur drei Krankheitsbilder (rheumatoide Arthritis, psoriatische Arthritis und Spondyloarthritis) eingeschlossen hatten. Eine breite Konsistenz wurde bei den nicht krankheitsspezifischen Datenelementen festgestellt, wobei auch Unterschiede in den Instrumenten und Kategorien sowie der Häufigkeit der Datenerhebung festgestellt wurden. Krankheitsspezifische Datenelemente werden weniger homogen erhoben. Die Autoren schlussfolgern, dass, obwohl die Register ähnlich aufgebaut sind, zahlreiche Unterschiede in den erhobenen Daten existieren. Insbesondere der Detailgrad und die Informationen über krankheitsspezifische Merkmale sind heterogen. Eine Harmonisierung der Daten würde zukünftige Analysen ermöglichen und erleichtern. Dafür benötigt es Schlüsseldaten, welche einheitlich definiert werden (Meissner et al., 2019).

Die Arbeit von Löbe et al. (2019) beschreibt das Konzept des SMITH-Konsortiums, welches sich unter anderem mit Konventionen, Prozessen und Tools zur Beschreibung und Verwaltung von Metadaten unter Verwendung gemeinsamer Standards für semantische Interoperabilität befasst. Insbesondere auf die Kette der Verarbeitungsschritte von Daten aus bestehenden Informationssystemen und den Standard für die klinische Dokumentenarchitektur wurde ein Fokus gelegt. Viele bestehende Systeme in Krankenhäusern können diese Standards nicht unterstützen und ihre Daten müssen umgewandelt werden. Es wurde darauf hingearbeitet, dass in Aachen, Jena und Leipzig jeweils ein Metadaten-Repository erstellt und genutzt wurde, um die Interoperabilität der Daten zu gewährleisten. Das technische Konzept und die Architektur sind der entsprechenden Studie zu entnehmen (Löbe et al., 2019).

Zens et al. (2020) thematisieren darüber hinaus den Bereich des “Capacity Buildings” als zentralen Schlüssel für die Entwicklung und möglichst nachhaltige Implementierung von strukturellen Kapazitäten wie beispielsweise der koordinierten Datenerhebung, Abstimmungen über Sektorengrenzen hinweg und der dauerhaften Bereitstellung von Ressourcen. Der Begriff des “Capacity Buildings” umfasst verschiedene Dimensionen, dennoch richten sich die meisten Ansätze auf eine Infrastrukturkomponente aus. Es wird konstatiert, dass für eine landesweite Kapazitätsentwicklung kommunaler Gesundheitsförderung eine valide Datenlage unumgänglich ist. An dieser Stelle soll Digital Public Health als Hebel für die Datenerhebung, Datenintegration und Datenverfügbarkeit genutzt werden können. Die Datenerhebung soll durch digitale Technologien einfacher, kostengünstiger und schneller möglich sein. Bezüglich der Datenintegration wird die Verknüpfung von Datenbeständen über Metadaten als Lösung vorgestellt, sofern die Attribute präzise und detailliert genug vorliegen. Um regional Daten zu analysieren wird der Einbezug von geografischen Informationssystemen (GIS) empfohlen. Auch die Datenverfügbarkeit soll durch digitale Technologien beschleunigt und erleichtert werden, um gewonnene Informationen zeitnah verfügbar zu machen. Die Arbeit untermau-

ert den Bedarf nach gebündelten und zugänglichen Daten und appelliert an die verstärkte Nutzung von Digital Public Health und “Capacity Building” als entsprechende Maßnahmen (Zens et al., 2020).

Ein gutes Beispiel für den Aufbau einer einheitlichen Dateninfrastruktur ist die INSPIRE-Richtlinie der Europäischen Union (EU) (INSPIRE: INfrastructure for SPatial InfoRmation in Europe), welche darauf abzielt, eine Geodateninfrastruktur als Grundlage einer gemeinsamen Umweltpolitik für die EU zu schaffen (European Commission, 2023). Praktisch fordert INSPIRE die einheitliche Beschreibung und ein einheitliches Format von Geodaten, die Bereitstellung der Daten online mit Diensten für Suche und Visualisierung und der Möglichkeit zum Download. Die Geodateninfrastrukturen der Mitgliedstaaten sollen in einem gemeinschaftlichen und grenzüberschreitenden Kontext kompatibel und nutzbar sein. Im Rahmen der Richtlinie ist auch eine Metadatenstruktur festgelegt. Die Mitgliedstaaten verpflichten sich dazu, Metadaten für die Geodatenätze und -dienste zu erstellen und auf dem neuesten Stand zu halten (Bundesministerium des Innern & Bundesministerium für Umwelt, Naturschutz, Bau und Reaktorsicherheit, 2016).

An dieser Stelle lässt sich zusammenfassen, dass Metadaten Repositories im Gesundheitswesen von großem aktuellem Interesse sind und durchweg vorteilhaft und sinnstiftend beurteilt werden. In der praktischen Umsetzung zeigt sich jedoch, dass insbesondere Gesundheitsdaten sehr stark in ihrem Detailgrad und der Zuordnung differieren und eine einheitliche Beschreibung herausfordernd ist.

Im zweiten Arbeitsschritt wurde untersucht, welche möglichen Datenkategorien und Datenarten für sozialmedizinische Untersuchungen relevant sein können. Dabei ergibt sich eine Vielzahl an möglichen Datenkategorien und Datenhaltern (Abbildung 1). Es ist anzumerken, dass die dargestellte Übersicht nur der Orientierung dient und keinen Anspruch auf Vollständigkeit hat.

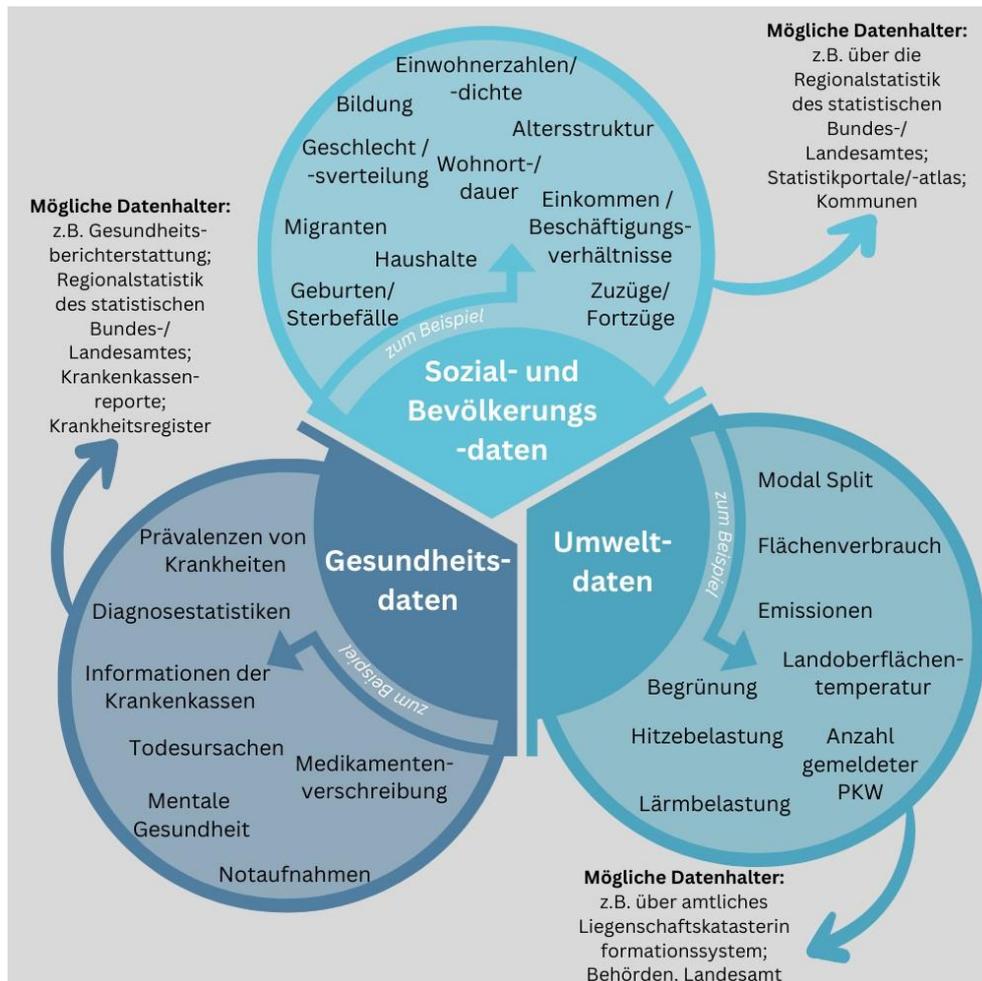


Abbildung 1: Übersicht über Datenkategorien, Daten und mögliche Datenhalter im Kontext von sozialmedizinischen Fragestellungen

Die Recherche über gesetzliche Vorgaben ergänzt die Abbildung 1 und zeigt, dass Leistungserbringer, die maßgeblich an der vertragsärztlichen Versorgung beteiligt sind, nach § 294 SGB V in der Pflicht sind, Daten an die Krankenkassen sowie die Kassenärztlichen Vereinigungen weiterzugeben (Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - § 294 Pflichten der Leistungserbringer/1988). Dies gilt unabhängig davon, ob Patienten die Einwilligung zur Datenübermittlung gegeben haben. Darunter fallen Angaben, die aus der Leistungserbringung, der Verordnung ebendieser sowie der Abgabe von Versicherungsleistungen entstehen und für die weitere Verarbeitung bei den Kassen benötigt werden.

Betrachtet man die Datenweitergabe der Krankenhäuser oder Krankenhausträger ebenso wie Rehabilitationseinrichtungen, wird diese nach § 301 SGB V näher definiert. Wichtige Anhaltspunkte für die Krankenkassen und Kassenärztlichen Vereinigungen sind demnach Angaben zur Aufnahme des/der Patient:in, Einweise- und Aufnahmediagnosen, Angaben zur voraussichtlichen Behandlungsdauer, beteiligte

Fachabteilungen und Angaben des ärztlichen Personals, ebenso wie Angaben zur Entlassung, etwaige Weiterbehandlungen und Überweisungen und die schlussendliche Darstellung der Diagnose sowie Empfehlungen zu weiteren Behandlungsmaßnahmen (Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - § 301 Krankenhäuser und Rehabilitationseinrichtungen/1988). Die Vollständigkeit der Datenübertragung wird durch die Angaben nach § 291a Abs. 2 SGB V, die alle Informationen zur Identität der Patienten im Rahmen der Elektronischen Gesundheitskarte liefern, d.h. Angaben zum Namen, Geburtsdatum, Wohnort, Krankenversicherungsnummer, Versichertenstatus u.a., abgebildet (Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - § 291a - Elektronische Gesundheitskarte als Versicherungsnachweis und Mittel zur Abrechnung/1988).

Leistungserbringer für die Versicherten der Krankenkassen, wie beispielsweise Vertragsärzt:innen oder Einrichtungen, die pflegerische Leistungen oder häusliche Krankenpflege erbringen, sind nach § 104 SGB XI in der Pflicht, Daten der Patient:innen zu sammeln und an die Pflegekassen zu übermitteln, insofern sie die Notwendigkeit von Pflegehilfsmitteln, die Qualität der Leistungserbringung oder die Abrechnung pflegerischer Leistungen umfassen. Biographische Datenspeicherung darf auch hier nur mit der Einwilligung der Patienten erfolgen (Sozialgesetzbuch (SGB) Elftes Buch (XI) - Soziale Pflegeversicherung - § 104 - Pflichten der Leistungserbringer/1988).

Seit 2016 ist das Landeskrebsregistergesetz Nordrhein-Westfalen (LKRGR NRW) in Kraft getreten und verankert die rechtlichen Grundlagen des Landeskrebsregisters NRW und der Meldepflicht für neu erkannte Krebserkrankungen in Folge des Krebsfrüherkennungs- und registergesetzes nach § 65 c SGB V (Krebsfrüherkennungs- und -registergesetz (KFRG), 2016). Nach dem LKRGR NRW ist das Landeskrebsregister NRW verantwortlich für die epidemiologische und klinische Krebsregistrierung des Landes Nordrhein-Westfalens. Ärzt:innen und Zahnärzt:innen unterliegen einer elektronischen Meldepflicht für festgestellte Krebserkrankungen und Informationen über die angestrebte Behandlung. Jede/e Patient:in mit einer Krebs-erkrankung erhält eine „Krebsakte“ innerhalb des Registers, die nach höchsten Datenschutzstandards bereits bei dem/der meldenden Ärzt:in verschlüsselt werden und letztendlich auch nur vollumfänglich von den meldenden Leistungserbringer:innen einsehbar sind. Nach § 23 LKRGR NRW kann beim Landeskrebsregister NRW ein Antrag auf Verfügbarmachung der Daten für weitere Forschungsvorhaben und für die Gesundheitsberichterstattung gestellt, die nach Darstellung des Nutzungsvorhabens pseudonymisiert übermittelt werden (Krebsfrüherkennungs- und -registergesetz (KFRG), 2016).

Im Blick gehalten werden muss auch das auf den Weg gebrachte Gesundheitsdatennutzungsgesetz (GDNG), welches die erleichterte Nutzung von Primärdaten für Sekundärforschungszwecke zu gemeinwohlorientierten Zwecken reglementiert und eine dezentrale Gesundheitsdateninfrastruktur in den Fokus rückt. Durch zentrale

Datenzugangs- und Koordinierungsstellen wird nicht nur der Datenzugang unter einer datenschutzrechtlichen Aufsicht erleichtert, es werden auch bürokratische Hürden für die Sekundärdatennutzung abgebaut. Es ist ebenfalls angestrebt, dass die Daten aus der Elektronischen Patientenakte (ePA) nach Freigabe durch den/die Patient:in für Forschungszwecke freigegeben und integriert werden (Gesundheitsdatennutzungsgesetz (GDNG)/2023).

Somit zeigt sich insgesamt, dass es viele datenhaltende Institutionen gibt, welche bereits durch gesetzliche Vorgaben zur Datenerhebung und -speicherung verpflichtet sind. Dadurch existiert bereits ein großer Datenpool, welcher für sozialmedizinische Forschungszwecke zwar sehr wertvoll, aber nicht unbedingt leicht zugänglich und übersichtlich ist.

Um auf dieser Grundlage erste Hinweise für den Aufbau eines Metadatenregisters im Ruhrgebiet sammeln zu können, wurden weitere Standards untersucht. Metadaten beschreiben detailliert Daten hinsichtlich ihrer Struktur, den Komponenten und weiteren Informationen, damit Daten erfolgreich gemeinsam genutzt werden können. Damit auch Metadatenregister einheitlich aufgebaut sind, gibt es die bereits erwähnte ISO/IEC 11179-3:2023 (Information technology: Metadata registries (MD)) (ISO 2023). Diese ISO soll in diesem Projekt ebenfalls als Grundlage genutzt werden, um erste Hinweise für den Aufbau der Matrixstruktur für das “Sozialmedizinische Metadatenregister” im Ruhrgebiet zu erhalten. Das Metamodell des Registers ist als konzeptionelles Datenmodell spezifiziert, d. h. es beschreibt, wie relevante Informationen in der natürlichen Welt strukturiert sind. In diesem Fall wurden 19 Metamodellkonstrukte als relevant identifiziert. Eine Übersicht über die Meta-modellkonstrukte mit entsprechenden Beispielen ist in Tabelle 2 ersichtlich:

Tabelle 2 Metamodellkonstrukt für das angestrebte Metadatenregister mit Beispielen

Metamodellkonstrukt	Beispiel
Organisation	<i>Institut, Krankenkasse, Register, etc.</i>
Kontakt	<i>Kontaktdaten falls verfügbar</i>
Tätigkeitsregion	<i>Bundesweit, Ruhrgebiet, Kommune, etc.</i>
Klassen	<i>Sozio-Demografische Daten / Umweltdaten / Gesundheitsdaten</i>
Assoziationsklasse	<i>Weitere Spezifikation der Klasse, z.B. Daten zu Bevölkerung, Wohnort, Luft, Diagnosedaten, etc.</i>
Attribute	<i>Weitere Spezifikation der Assoziationsklasse über die Ausprägung des Datenobjektes, z.B. Anzahl der Einwohner, Feinstaubbelastung, Hauptdiagnosen, etc.</i>
Weitere Attribute	<i>Weitere Spezifikation der Attribute (insbesondere für Gesundheitsdaten), z.B. zum Setting oder der Art der erbrachten Leistung</i>
Räumliche Ebene	<i>Verfügbarkeit der Daten auf Bundes-, Landes-, Kreis-, Bezirksebene, etc.</i>

Zeitraum	<i>Zeitraum der vorliegenden Daten, z.B. für ein Jahr, Quartal, Monate, Wochen, etc.</i>
Aktualisierungsrhythmus	<i>Aktualisierungsmöglichkeit des jeweiligen Datensatzes, z.B. jährlich, monatlich, tagesaktuell, etc.</i>
Zeitliche Verfügbarkeit	<i>Zeitpunkt, ab wann die Daten verfügbar/abrufbar sind, z.B. nach einem Jahr, nach einigen Stunden, etc.</i>
Art der Daten	<i>Beschreibung, ob es sich um modellierte, erhobene oder prognostizierte Daten handelt</i>
Dateiformat	<i>Datenformat, z.B. Datenbankformat, Excel, Statistikformate, etc.</i>
Speicherort	<i>Ort der Speicherung der Daten z.B. lokales Rechenzentrum, Cloud, etc.</i>
Sprache	<i>Verwendete Sprache im Datensatz</i>
Verfügbarkeit für Externe	<i>Kennzeichnung, ob der Datensatz frei verfügbar ist (open data) oder nur für Forschung oder hoheitliche Aufgaben bestimmt ist oder nicht frei verfügbar ist</i>
Kosten	<i>Information über Preise, falls Datensätze nicht frei verfügbar sind</i>
Speicherdauer	<i>Angaben, wie lange die jeweiligen Daten gespeichert werden</i>
Weiterführende Definitionen/Informationen	<i>Weitere beschreibende Angaben zum Datensatz</i>

5. Diskussion

Die primäre Zielsetzung dieser Übersichtsarbeit war es, einen Einblick in die wissenschaftliche Perspektive über Metadatenregister und insbesondere die Identifikation von Metadatenregistern im Ruhrgebiet zu geben. Zudem wurde die Sammlung von relevanten Daten und Datenkategorien und deren Eignung für Metadatenregister, der Recherche nach gesetzlichen Vorgaben zur Datenhaltung und Datenübertragung sowie die Erstellung eines ersten Entwurfes für eine Matrix Struktur im Forschungsverbund “Digital Health Factory Ruhr” angestrebt.

Zusammenfassend kann festgehalten werden, dass für das Ruhrgebiet kein Metadatenregister im Gesundheitsbereich existiert. Die Literaturrecherche hebt die aktuelle Relevanz des Themas eindeutig hervor und untermauert sowohl den steigenden Bedarf nach gebündelten und verfügbaren Daten als auch die Nützlichkeit von Metadatenregistern für die Wissenschaft, Forschung, Medizin, Prävention, Rehabilitation und Gesundheitsförderung. Insbesondere die FAIR Prinzipien (Wilkinson et al., 2016) werden in diesem Kontext wiederkehrend genannt. Für den Bereich der Auffindbarkeit ist es grundlegend, dass Daten sowohl für Menschen als auch Maschinen leicht zu finden sind. Maschinenlesbare Metadaten sind für das automatische Auffinden von Datensätzen und Diensten zwingend notwendig und daher ein

wesentlicher Bestandteil des FAIRen-Datenprozesses. Dafür benötigt es einen weltweit eindeutigen und dauerhaften Bezeichner, eine umfangreiche Beschreibung aller Datensätze mit Metadaten, und dass (Meta-)Daten in einer durchsuchbaren Ressource registriert oder indiziert werden. Für den Zugang zu Daten müssen (Meta-)Daten über ihren Identifikator mit Hilfe eines standardisierten Kommunikationsprotokolls abrufbar sein, Protokolle frei und universell implementierbar sein und Metadaten zugänglich sein, auch wenn die Daten nicht mehr verfügbar sind. Um die Interoperabilität, also die Integration mit anderen Daten zu ermöglichen, müssen Daten mit Anwendungen oder Arbeitsabläufen für die Analyse, Speicherung und Verarbeitung kompatibel sein und zusammenarbeiten. Dafür ist es relevant, dass (Meta-)Daten eine formale, zugängliche, gemeinsame und breit anwendbare Sprache zur Wissensdarstellung verwenden und die Vokabulare den FAIR-Grundsätzen entsprechen. Ein wesentliches Ziel der FAIR-Prinzipien ist die Wiederverwendung von Daten. Damit Daten in verschiedenen Kontexten genutzt oder kombiniert werden können, müssen Daten und Metadaten detailliert beschrieben sein. Dafür müssen (Meta-)Daten reichhaltig beschrieben sein mit einer Vielzahl von genauen und relevanten Attributen oder sie mit einer klaren und zugänglichen Datennutzungslizenz freigegeben werden.

Während der Literaturrecherche konnten nur wenige thematisch passende Paper für Deutschland (n=6) identifiziert werden. Die eingeschlossenen Studien schlussfolgern allesamt, dass Metadatenregister in Deutschland dringend benötigt werden. Zudem untermauert die Analyse der vorhandenen Daten und zu gesetzlichen Rahmenbedingungen, dass zahlreiche Datensätze und Datenhalter existieren. Dennoch klingen insbesondere Herausforderungen und Hemmnisse für Gesundheitsdaten, wie die einheitliche Definition von Domänen und Attributen und dem Detailgrad der Informationen, an. Sowohl der Bedarf als auch der Mehrwert von Metadatenregistern sind eindeutig gegeben. Darüber hinaus gibt es aktuell bereits weitere technische Paper, welche sich mit der Umsetzung von Metadatenregistern beschäftigen. Die übersichtsartige Darstellung in dieser Arbeit lässt erkennen, wie groß und vielfältig das Feld der Daten und Datenkategorien sowie Datenhalter ist. Wie durch die FAIR-Prinzipien gefordert, bedarf es dabei jedoch einheitlicher Metadaten-Standards, welche aktuell noch selten genutzt werden. Wie bereits von Stausberg and Harkener (2019) konstatiert, ist eben diese Definition von einheitlichen Hauptdomänen und Attributen eine Herausforderung an sich, da dabei je nach Perspektive und Zielsetzung unterschiedliche Vorgehensweisen sinnvoll erscheinen. Auf Ebene der europäischen Union ist es gelungen, durch die INSPIRE-Richtlinie einen Schritt in Richtung einer EU-weiten einheitlichen Geodateninfrastruktur zu schaffen.

Das in Tabelle 4 aufgelistete Matrixmodell gibt erste Anhaltspunkte für die benötigten Metamodellkonstrukte, wobei dem Problem der fehlenden Homogenität ebenfalls nicht begegnet werden kann. Dennoch zeigt diese Projektarbeit, dass vorhandene Datensätze im Ruhrgebiet durch das entworfene Metadatenregister sichtbar gemacht werden können und übersichtlich beschrieben werden können.

6. Ausblick & Forschungsagenda

Die vorliegende Arbeit zeigt die dringende Notwendigkeit von Metadatenregistern in Deutschland und auch auf internationaler Ebene auf. Zahlreiche wissenschaftliche Arbeiten, sowie Standards und Prinzipien (vgl. FAIR oder ISO 11179-3:2023) verweisen auf den dringenden Handlungsbedarf, Daten transparent und verfügbar zu machen und verweisen auf die daraus resultierenden vielfältigen Handlungsoptionen für Medizin und Wissenschaft. Darüber hinaus kann die systematische Analyse dazu beitragen, das Gesundheitssystem insgesamt effizienter zu gestalten, indem sie die Ressourcenallokation verbessert und den Zugang zu qualitativ hochwertiger Versorgung erleichtert. Durch die Analyse großer Datenmengen können Gesundheitseinrichtungen nicht nur prädiktive Modelle erstellen, um Krankheitsausbrüche vorherzusagen und frühzeitig intervenieren zu können, vielmehr können die Erkenntnisse genutzt werden, um Strategien der Verhaltens- und Verhältnisprävention zu etablieren.

Es wird dabei jedoch auch deutlich, dass insbesondere im Gesundheitswesen eine Einigung über die Beschreibung von Krankheitsdaten eine große Herausforderung darstellt. Der Missbrauch oder die unsachgemäße Weitergabe dieser Daten könnte schwerwiegende Folgen für die individuelle Privatsphäre und die Autonomie haben. Um die Vorteile der Sammlung und Zusammenführung von Gesundheitsdaten zu maximieren und gleichzeitig ethische und rechtliche Bedenken zu adressieren, ist es entscheidend, robuste Datenschutz- und Sicherheitsstandards zu etablieren. Es bedarf klarer rechtlicher Rahmenbedingungen, die den Schutz der Privatsphäre und die Einwilligung der Patienten sicherstellen. Zudem ist eine kontinuierliche Überwachung und Regulierung der Datennutzung erforderlich, um sicherzustellen, dass sie ausschließlich zum Wohl der Individuen und der Gesellschaft eingesetzt werden. Letztendlich muss bei der Nutzung von Gesundheitsdaten ein ausgewogenes Verhältnis zwischen dem Streben nach Fortschritt in der medizinischen Forschung und dem Schutz der individuellen Rechte und Freiheiten gefunden werden.

Für das Ruhrgebiet konnte bisher kein "Sozialmedizinisches Metadatenregister" identifiziert werden, sodass die vorliegenden Bestrebungen im Rahmen der Digital Health Factory alleinstehend sind. Es zeigt sich, dass insbesondere für eine Metropolregion, wie dem Ruhrgebiet, ein Metadatenregister mannigfaltige Potenziale bieten würde. Ferner wird dies gestützt durch die hohe Akzeptanz von medizinischen Registern in der Bevölkerung (Haug et al., 2023). Es zeigt sich, dass beispielsweise über 80% ihre Daten in der elektronischen Patientenakte für die Forschung freigeben würden (ebd.).

Literatur

- Sozialgesetzbuch (SGB) Elftes Buch (XI) - Soziale Pflegeversicherung - § 104 - Pflichten der Leistungserbringer, Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung 1988).
- Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - § 291a - Elektronische Gesundheitskarte als Versicherungsnachweis und Mittel zur Abrechnung, Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung 1988).
- Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - § 294 Pflichten der Leistungserbringer, Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung 1988).
- Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - § 301 Krankenhäuser und Rehabilitationseinrichtungen, Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung 1988).
- Bundesministerium des Innern, & Bundesministerium für Umwelt, Naturschutz, Bau und Reaktorsicherheit (Eds.). (2016). INSPIRE - Grenzen aufheben mit INSPIRE - Geodaten aus ganz Europa nutzen.
https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/moderne-verwaltung/geoinformationen/inspire-flyer.pdf?__blob=publicationFile&v=3
- Gesundheitsdatennutzungsgesetz (GDNG), Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung 2023).
- European Commission (Ed.). (2023). INSPIRE Knowledge Base. https://knowledge-base.inspire.ec.europa.eu/index_en
- Fluck, J. (2019). National Research Data Infrastructure for Personal Health Data (NFDI4Health): Proposal. Bonn: ZB MED–Information Centre for Life Sciences.
- Haug, S., Schnell, R., Raptis, G., Dotter, C., & Weber, K. (2023). Wissen und Einstellung zur Speicherung und Nutzung von Gesundheitsdaten: Ergebnisse einer Bevölkerungsbefragung. Zeitschrift Für Evidenz, Fortbildung Und Qualität Im Gesundheitswesen. Advance online publication.
<https://doi.org/10.1016/j.zefq.2023.11.001>
- ISO 2023. ISO/IEC 11179-3:2023: Information technology - Metadata registries (MDR). <https://www.iso.org/standard/78915.html>
- Kurth, B.-M. (2012). Das RKI-Gesundheitsmonitoring – was es enthält und wie es genutzt werden kann. Public Health Forum, 20(3), 4.
<https://doi.org/10.1016/j.phf.2012.06.001>
- Krebsfrüherkennungs- und -registergesetz (KFRG), 2016.
<https://www.landeskrebsregister.nrw/das-lkr-nrw/rechtsgrundlage/>
- Löbe, M., Beyan, O., Stäubert, S., Meineke, F., Ammon, D., Winter, A., Decker, S., Löffler, M., & Kirsten, T. (2019). Design of metadata services for clinical data interoperability in germany. In MEDINFO 2019: Health and Wellbeing e-Networks for All (pp. 1528–1529). IOS Press.

- Medizininformatik-Initiative. (2017). MI-I-Metadaten zur Verfügbarkeit von Daten, Auswertungsmöglichkeiten und Kooperationsoptionen.
https://www.medizininformatik-initiative.de/sites/default/files/inline-files/MII_04_Metadaten_Verf%C3%BCgbarkeit_1-0.pdf
- Meissner, Y., Strangfeld, A., Costedoat-Chalumeau, N., Förger, F., Goll, D., Molto, A., Özdemir, R., Wallenius, M., & Fischer-Betz, R. (2019). European Network of Pregnancy Registers in Rheumatology (EuNeP)—an overview of procedures and data collection. *Arthritis Research & Therapy*, 21, 1–9.
- Regionalverband Ruhr (2021). Metropole Ruhr: Die Region in Zahlen.
https://www.rvr.ruhr/fileadmin/user_upload/01_RVR_Home/03_Daten_Digitales/Regionalstatistik/03_Publikationen/2021-08_Metropole_Ruhr_Die_Region_in_Zahlen.pdf
- Stausberg, J., & Harkener, S. (2019). Bridging documentation and metadata standards: Experiences from a funding initiative for registries. In *MEDINFO 2019: Health and Wellbeing e-Networks for All* (pp. 1046–1050). IOS Press.
- TMF e.V. (2023). Metadaten und Metadata Repositories: Toolpool Gesundheitsforschung. Redaktion ToolPool. [https://www.toolpool-gesundheitsforschung.de/themen/metadaten-und-metadata-repositories#:~:text=Metadaten%20sind%20beschreibende%20Daten%20jener,%E2%80%9EItems%E2%80%9C\)%20spezifiziert%20werden.](https://www.toolpool-gesundheitsforschung.de/themen/metadaten-und-metadata-repositories#:~:text=Metadaten%20sind%20beschreibende%20Daten%20jener,%E2%80%9EItems%E2%80%9C)%20spezifiziert%20werden.)
- Waller, H. (2007). *Sozialmedizin: Grundlagen und Praxis*. W. Kohlhammer Verlag.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., van Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., . . . Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>
- Zens, M., Shajanian Zarneh, Y., Dolle, J., & Bock, F. de (2020). Digital Public Health—Hebel für Capacity Building in der kommunalen Gesundheitsförderung. *Bundesgesundheitsbl*, 63, 729–740.

/// Zitationshinweis

Cramer, E; Krumreihn, A; Schaffert, C & Stiene, M (2024). Metadaten, Metadatenregister und Datenverknüpfungen im Gesundheitsbereich – eine Einordnung Forschung aktuell, <https://doi.org/10.53190/fa/202403>

/// Autorinnen und Autoren

Elena Cramer ist Junior Researcherin im Forschungsschwerpunkt „Gesundheitswirtschaft & Lebensqualität“ des Instituts Arbeit und Technik der Westfälischen Hochschule Gelsenkirchen.

Aline Krumreihn ist wissenschaftliche Mitarbeiterin im Department of Community Health an der Hochschule für Gesundheit Bochum.

Corinna Schaffert ist Projektmanagerin der MedEcon Ruhr GmbH Bochum.

Meike Stiene ist Projektmanagerin der MedEcon Ruhr GmbH Bochum.

/// Impressum

Herausgeberin

Westfälische Hochschule Gelsenkirchen
Institut Arbeit und Technik
Munscheidstr. 14
45886 Gelsenkirchen

Redaktion

Claudia Braczko
Telefon: +49 (0)209.17 07-176
E-Mail: braczko@iat.eu

Bildnachweis

Titelbild: Pixabay.com

Mit dem Publikationsformat „Forschung aktuell“ sollen Ergebnisse der IAT-Forschung einer interessierten Öffentlichkeit zeitnah zugänglich gemacht werden, um Diskussionen und die praktische Anwendung anzuregen. Für den Inhalt sind allein die Autorinnen und Autoren verantwortlich, die nicht unbedingt die Meinung des Instituts wiedergeben.