



Kollaborativer Datenraum Gesundheitswesen

Vorstudie über die konzeptionellen Grundlagen



BDI

Bundesverband der
Deutschen Industrie e.V.



Institute for
Digital Transformation
in Healthcare

Inhaltsverzeichnis

1. Einleitung	2
2. Grundlagen des Data Sharings	3
2.1 Ausgangslage & Problemstellung.....	3
2.1.1 Potenziale des Data Sharings.....	3
2.1.2 Hürden des Data Sharings.....	4
2.1.3 Rechtliche Rahmenbedingungen.....	5
2.1.4 Besonderheiten des Gesundheitssektors.....	9
2.2 Marktanalyse.....	12
2.2.1 Framework Initiativen.....	13
2.2.2 Data Sharing Ökosysteme.....	16
2.2.3 Typisierung von Data Sharing Ökosystemen.....	27
2.2.4 Fazit.....	31
2.3 Aufbau & Funktionen eines Datenraums.....	33
2.3.1 Notwendige Rollen und Komponenten.....	34
2.3.2 Grundlegender Datenfluss.....	38
2.3.3 Optionale Rollen.....	39
3. Konzept für einen kollaborativen Datenraum Gesundheitswesen	44
3.1 Vision & Strategie.....	44
3.2 Anwendungsfälle.....	46
3.2.1 Typisierung.....	46
3.2.2 Merkmale relevanter Anwendungsfälle.....	50
3.3 Kollaborations- und Organisationsstruktur.....	51
4. Fazit und Ausblick	54
5. Literaturverzeichnis	55
Anhang	58

1. Einleitung

Mit dem sektorübergreifenden Austausch von Daten im Gesundheitswesen sind große Erwartungen verbunden, denn er ermöglicht die Schaffung gemeinsamer Anwendungen und Lösungen für die gesundheitsbezogenen Herausforderungen von heute und morgen. Die Vision reicht hier von der Prävention von Krankheiten über die Optimierung bestehender Abläufe und Prozesse bis hin zu einer personalisierten Gesundheitsversorgung. Dazu sollen die Daten aller Stakeholder des ersten und zweiten Gesundheitsmarktes, also die von Ärzt:innen, Patient:innen, Forscher:innen oder Industrieunternehmen, miteinander vernetzt und so die Optimierung der Gesundheitsversorgung des Einzelnen und der Gesellschaft ermöglicht werden. Patient:innen profitieren hier u. a. von einer verbesserten Anamnese und Diagnostik sowie individuellen Therapie- und Betreuungsmöglichkeiten. Darüber hinaus können sich gänzlich neue Geschäftsfelder und Wachstumschancen ergeben.

Ein Blick auf die heute verfügbaren Optionen und Rahmenbedingungen für das Teilen von Gesundheitsdaten offenbart jedoch noch großes Entwicklungspotential. Mangelnde Infrastrukturen und Interoperabilität sowie regulatorische Schranken gehören zu den Herausforderungen, die es zu adressieren gilt.

Um die eingangs beschriebene Vision zu verwirklichen, braucht es eine digitale Infrastruktur, die auf europäischen Werten wie Freiheit, Demokratie und Gleichberechtigung aufgebaut ist und Innovation durch Kollaboration ermöglicht. So soll langfristig die Wirtschaftlichkeit der Gesundheitsversorgung und der Gesundheitsindustrie in Deutschland und Europa gesichert werden. Für den sensiblen Bereich der Gesundheitsdaten ist zudem von Anfang an Vertrauen durch Datensouveränität zu gewährleisten, um die Interessen von Unternehmen und Patient:innen gleichermaßen zu schützen. Im Kern bedeutet Datensouveränität, dass die Kontrolle über Daten bei dem Stakeholder verbleibt, welcher die Daten erhebt bzw. den sie betreffen. Dieser geschützte Raum, über den der Datenaustausch abgewickelt werden kann, soll als kollaborativer Datenraum Gesundheitswesen bezeichnet werden.

Mit der vorliegenden, gemeinsamen Studie der BDI Initiative Gesundheit digital und des Institute for Digital Transformation in Healthcare sollen die Chancen und Umsetzungsmöglichkeiten für einen solchen kollaborativen Datenraum Gesundheitswesen aufgezeigt werden. Dazu werden in Kapitel 2 zunächst die Grundlagen des Data Sharings erarbeitet. So beschreibt Kapitel 2.1 sowohl Potenziale als auch Hürden im Hinblick auf den Austausch von Daten im Allgemeinen und von Gesundheitsdaten im Speziellen. Auch die rechtlichen Rahmenbedingungen erfahren eine Erläuterung. Um das aktuelle Angebot für das Teilen von Daten zu überblicken, erfolgt in Kapitel 2.2 eine umfangreiche Marktanalyse bestehender Initiativen, Ökosysteme, Projekte und kommerzieller Angebote. Dabei lässt sich eine Differenzierung in Framework Initiativen und Data Sharing Ökosysteme beobachten. Letztere werden hinsichtlich ihrer gemeinsamen Merkmale typisiert, um daraus das Grundgerüst für einen kollaborativen Datenraum Gesundheitswesen abzuleiten. Kapitel 2.3 widmet sich dann dem Aufbau und den Funktionen eines Datenraums. Dabei werden sowohl die notwendigen Rollen und Komponenten beschrieben als auch optionale Rollen vorgestellt. Im Anschluss wird in Kapitel 3 ein Konzept für einen kollaborativen Datenraum Gesundheitswesen erörtert. Hier liegt besonderes Augenmerk auf der Vision & Strategie, den möglichen Anwendungsfällen und den Grundzügen einer Kollaborations- und Organisationsstruktur.

2. Grundlagen des Data Sharings

2.1 Ausgangslage & Problemstellung

Der Wert und die Relevanz von Daten werden im Narrativ der digitalen Transformation gerne mit Metaphern umschrieben. Häufig werden sie als „das neue Öl“ oder „Gold des digitalen Zeitalters“ bezeichnet. Unabhängig davon, dass diese Vergleiche in sich unstimmt und daher missverständlich sind, zeigt die wirtschaftliche Relevanz großer Technologieunternehmen nur allzu deutlich, dass heutige Wertschöpfungsketten nicht mehr ohne die Ressource „Daten“ auskommen. Mittlerweile setzt die Hälfte der 10 wertvollsten Unternehmen der Welt (nach Marktkapitalisierung) auf datenbasierte Geschäftsmodelle¹. Ob als Geschäftsgrundlage oder für die Entwicklung neuer Produkte und Dienstleistungen: Unternehmen jeglicher Branchen werden auf Daten und ihre wertschöpfende Verwendung angewiesen sein, um wettbewerbsfähig zu bleiben. Dabei sind verschiedene Arten von Daten zu unterscheiden. Neben personenbezogenen Daten wie Kund:innen- und Patient:innendaten fallen Daten u. a. in Produktionsprozessen, bei Transaktionen oder im Kontext von Lagerbeständen an. Auch sogenannte Metadaten, also Daten, die andere Daten beschreiben, sind zu nennen. Je nach Sektor steigert die Analyse dieser Daten die Effizienz in der Produktion, verbessert die Kundenansprache oder ermöglicht völlig neue Geschäftsmodelle. Der Bedarf nach nutzbaren Daten ist dementsprechend groß. Die EU-Kommission geht im Kontext des im Februar 2022 vorgelegten Vorschlags für den „Data Act“ davon aus, dass das weltweite Datenvolumen von 33 Zettabyte im Jahre 2018 auf voraussichtlich 175 Zettabyte im Jahre 2025 ansteigen wird.²

Data Sharing unter Wahrung der Datensouveränität

Unter dem Begriff „Data Sharing“ lässt sich allgemein der Prozesse des Austausches von Daten zwischen zwei Parteien fassen. „Austausch“ kann sich dabei sowohl auf das bidirektionale Teilen von Daten als auch die Gewährung eines Datenzugangs oder -zugriffs im Rahmen des einseitigen Teilens von Daten beziehen. Entscheidend ist, dass Daten, die in der Regel siloartig erhoben und genutzt werden, von unterschiedlichen Parteien gemeinsam genutzt werden können. Übliche Konstellationen sind Business-to-Business (B2B), Business to Consumer (B2C) oder Business-To-Government (B2G). Im Kontext des Data Sharings über Datenräume und Dateninfrastrukturen gewinnt zudem der Begriff der Datensouveränität an Relevanz. „Datensouveränität [bezeichnet] die Fähigkeit einer juristischen oder natürlichen Person zur Selbstbestimmung über ihre Datengüter. [...] Datenräume [wie der kollaborative Datenraum Gesundheitswesen] und Dateninfrastrukturen zielen darauf, die Fähigkeit der Datensouveränität softwaretechnisch zu unterstützen“ und so ein Gleichgewicht zwischen dem Schutz von Daten und der kollektiven Nutzung in geschäftlichen Kontexten zu schaffen.³⁴

2.1.1 Potenziale des Data Sharings

Trotz aller Relevanz von Daten für das eigene Geschäft besteht derzeit jedoch ein Problem: Nicht alle Daten, aus denen sich neue Erkenntnisse gewinnen ließen, können sinnvoll zusammengeführt werden – weder von Unternehmen noch von öffentlichen Stellen oder der Forschung. Allein die Diversität der möglichen internen und externen Datenquellen, etwa Geräte des Internet of Things (IoT), Open Data oder smarte Endgeräte von Kund:innen sowie deren Nutzung verschiedenster Anwendungen, erhöht die Komplexität der qualitativen Datenerhebung. Zudem stellen Kontrolle und Verfügung über Daten eine komplizierte Frage dar, die die Wertschöpfung aus diesen Daten behindert. Da relevante Erkenntnisse für die Geschäftsentwicklung in der Regel vor allem aus der Aggregation und Analyse größerer Datenmengen aus unterschiedlichen Quellen zu erwarten sind, deren Verfügbarkeit jedoch wie geschildert in den meisten Fällen eingeschränkt ist, rückt das Teilen von Daten zwischen

¹ Companies Market Cap (2022)

² Europäische Kommission (2022a)

³ Otto & Burmann (2021), S.284

⁴ Otto (2016)

Unternehmen immer mehr in den Fokus. Daten mit anderen Unternehmen zu teilen ist nicht nur als Empfänger von Daten attraktiv, sondern bietet auch die Möglichkeit, mit der Monetarisierung der Weitergabe von Daten ein weiteres Geschäftsfeld zu erschließen. Schließlich verfügen Unternehmen nicht selten auch über Daten, aus denen sie selbst keinen direkten Mehrwert gewinnen, die aber wiederum für andere Unternehmen von Interesse sein könnten. Neben der Erschließung potenzieller neuer Geschäftsfelder liegt der Nutzen des Datenaustausches in der Effizienzsteigerung bestehender Prozesse. Dies gilt insbesondere in Netzwerkstrukturen, wie es in der Forschung der Fall ist. Der konstante und zeitnahe Datenaustausch zwischen den einzelnen Knotenpunkten kann über Erfolg und Misserfolg entscheiden – z. B. bei der Generierung und Testung von Hypothesen zur schnelleren Entwicklung neuer Arzneimittel, der Entwicklung personalisierter Produkte, bei Gentherapeutika oder individualisierten Diagnostika zur Verlaufskontrolle.

2.1.2 Hürden des Data Sharings

Es bleibt zu konstatieren, dass der Zugang zu qualitativ hochwertigen Daten zu einem essenziellen Bestandteil für das nachhaltige Wachstum in allen Branchen wird. Trotzdem gibt es aktuell noch einige Hindernisse, die den interorganisationellen Datenaustausch behindern. So haben längst nicht alle Unternehmen erkannt, welch enormes Potenzial in der Ressource „Daten“ steckt. Eine Befragung von 500 Unternehmen mit Sitz in Deutschland im Auftrag des BDI aus dem Jahr 2021⁵ ergab, dass lediglich „28 Prozent der Unternehmen [...] einen hohen Digitalisierungsstand hinsichtlich des eigenen Datenmanagements“⁶ aufweisen. Die große Mehrheit der Unternehmen ist entsprechend weniger digital einzustufen. Nur ein Drittel dieser sehen Vorteile darin, externe Daten zu verwenden. Auch die Bereitschaft, eigene Daten mit anderen Unternehmen zu teilen, lag insgesamt bei lediglich knapp 12 Prozent. Als größte Hürde für das Teilen von Daten gaben sowohl digitale als auch weniger digitale Unternehmen die „Sorge vor unautorisiertem Zugriff Dritter auf ihre Daten“ an. Hinzu kommen „datenschutzrechtliche Grauzonen“ und „Unklarheiten bezüglich der Nutzungsrechte an den Daten“.⁷ Lösungen für das Teilen von Daten sollten daher besonderen Fokus auf Sicherheit und Vertrauenswürdigkeit legen.

Die Skepsis, Daten mit Dritten auszutauschen, ist zudem auch auf Bedenken hinsichtlich einer Übervorteilung durch Wettbewerber oder der Schaffung neuer Konkurrenten zurückzuführen.⁸ Die Bereitschaft, gerade aus Unternehmenssicht besonders schützenswerte Daten wie „Einkaufs- und Verkaufspreise, Stammdaten wie Losgrößen, Wiederbeschaffungszeiten und Sicherheitsbestände, Bewegungsdaten, die Prognosen über zukünftige Verkäufe, Verkaufshistorien und Bestellungen beinhalten sowie Key-Account Informationen“⁹ zu teilen, ist daher gering und bedarf einer Compliance-rechtlichen Prüfung. Technische Maßnahmen zur Kontrolle der Nutzungsbedingungen von Daten sind entsprechend wichtige Bestandteile von Angeboten zum Teilen von Daten.

Im Gesundheitswesen ist die Weitergabe von sensiblen, personenbezogenen Gesundheitsdaten zudem an noch strikere Bedingungen geknüpft, da es sich um besonders schützenswerte Daten handelt, wenn diese nicht ausreichend anonymisiert werden. Hinzu kommt, dass viele Daten für die (forschende) Industrie derzeit nicht zugänglich sind, da sie auf dem geschlossenen ersten Gesundheitsmarkt entstehen. Weitere Hürden sind in diesem Zusammenhang in Deutschland auf die Trägheit des Gesundheitssystems, die Verteilung der Gesetzgebungskompetenzen auf Bund und Länder, die komplexe Stakeholder-Situation oder die Skepsis von Patient:innen zurückzuführen. Diese Besonderheiten werden in Kapitel 2.1.4.2 im Detail betrachtet.

Neben dem noch geringen Bewusstsein für den Mehrwert der Datennutzung im Allgemeinen und des Datenteilens im Speziellen, mangelt es an einer entsprechenden Infra- und Kollaborationsstruktur zum

⁵ Röhl, Bolwin, Hüttl (2021)

⁶ ebd., S.4

⁷ ebd., S.40

⁸ Lindner, Straub, Kühne (2021)

⁹ INFORM (2017)

Teilen von Daten. So gibt es noch keine Option, Daten zwischen Unternehmen sicher und souverän auszutauschen, die sich branchenübergreifend durchgesetzt hätte. Zudem fehlen eine gemeinsame Semantik und Syntax sowie international akzeptierte Standards, um eine allgemeine Interoperabilität zu ermöglichen. Die Landschaft der technischen Lösungen, die einen sicheren Datenaustausch versprechen, ist aktuell noch im Aufbau und entsprechend fragmentiert. Bestrebungen, entsprechende Standards zu etablieren, sind aber bereits zu beobachten (siehe Kapitel 2.2 für bestehende Ökosysteme für das Teilen von Daten).

2.1.3 Rechtliche Rahmenbedingungen

Auch wenn Daten mittlerweile als wertvolle Assets angesehen werden, sind sie rechtlich nicht mit klassischen Ressourcen vergleichbar. Daten stellen keine körperlichen Gegenstände im Sinne des § 90 BGB dar und können somit nicht im Eigentum stehen. Dementsprechend kann das Bürgerliche Gesetzbuch nicht direkt auf Daten angewandt werden, sondern lediglich auf die Verträge zwischen Unternehmen, die entsprechende Nutzungsbedingungen beinhalten. Personen und Unternehmen, die de facto Eigentümer von Datensätzen sind, haben daher kein absolutes, d. h. für jedermann geltendes Recht, Dritte vom Zugang zu den Daten oder deren ausschließlicher Nutzung auszuschließen.¹⁰ Vielmehr gilt: „Derjenige der die Verfügungsgewalt über Daten hat, darf diese auch nutzen.“¹¹ Werden Daten gehandelt, räumt der Datengeber durch den Vertrag lediglich eine de facto Rechtsposition ein. Der Datenempfänger wird aufgrund der fehlenden Möglichkeit, eine absolute Rechtsposition zu schaffen, nicht Eigentümer der Daten¹², sondern erhält lediglich einen faktischen Zugang zu den Daten. Vertragliche Regelungen zur Festlegung von Nutzungsbedingungen stellen daher ein zentrales Element für das Teilen von Daten zwischen Unternehmen dar. Auch wenn der Spielraum der Vertragsgestaltung groß ist, gilt es, einige rechtliche Rahmenbedingungen zu berücksichtigen. Neben dem Datenschutz haben auch das Wettbewerbs- und Urheberrecht sowie das Gesetz zum Schutz von Geschäftsgeheimnissen Einfluss auf die Gestaltung von Verträgen über das Teilen von Daten.

2.1.3.1 Datenschutz

Ein großer Faktor beim Teilen von Daten ist die automatisierte Verarbeitung von personenbezogenen Daten. Diese fällt unter die seit 2018 wirksame Datenschutzgrundverordnung (DSGVO). Daten dürfen somit nur zweckgebunden und mit Einverständnis der betreffenden Person verarbeitet werden, es sei denn, es liegen andere juristische Legitimationsgründe wie etwa der des öffentlichen Interesses im Bereich der öffentlichen Gesundheit (z. B. Krebsregister) vor. Dem Teilen mit Dritten muss aus diesem Grund in den meisten Fällen explizit zugestimmt werden. Dazu müssen der genaue Zweck aufgezeigt sowie ausreichende Informationen vorgelegt werden, damit die betreffende Person eine möglichst freie Entscheidung treffen kann. Im Rahmen von datenzentrierten Wertschöpfungsketten kommt der Festlegung der Zweckbindung zum Zeitpunkt der Datenerhebung eine entscheidende Rolle für die weitere Verarbeitung zu. Mögliche zukünftige Verwendungszwecke müssen entsprechend mitgedacht und plausibel beschrieben werden, um Daten rechtssicher teilen zu können. Zudem besteht ein Kopplungsverbot, sodass die Erlaubnis zur Weitergabe von personenbezogenen Daten an Dritte keine Bedingung für die Erfüllung von Verträgen darstellen darf. Diese und weitere Einschränkungen gelten beim Teilen von anonymisierten Daten nicht, allerdings lassen sich in der Praxis personenbezogene und anonymisierte Daten nicht so einfach voneinander trennen, wie dies in der Theorie der Fall ist. Je nach Verfahren, welches zur Anonymisierung genutzt wird, ist es möglich, Personen zu re-identifizieren. „Um festzustellen, ob eine natürliche Person identifizierbar ist, sollten alle Mittel berücksichtigt werden, die von dem Verantwortlichen oder einer anderen Person nach allgemeinem Ermessen wahrscheinlich genutzt werden, um die natürliche Person direkt oder indirekt zu identifizieren, wie beispielsweise das Aussondern.“¹³ Ob ein anonymisierter Datensatz de-anonymisierbar ist, muss durch das datenverarbeitende Unternehmen beurteilt werden. Die Urteilsfähigkeiten hängt jedoch vom Know-how

¹⁰ Lauf et al. (2021)

¹¹ Lindner, Straub, Kühne (2021), S.22

¹² Paal & Hennemann (2017)

¹³ DSGVO Erwägungsgrund 26

und den Mitteln dieses Unternehmens ab. Da sich beides von Unternehmen zu Unternehmen unterscheidet, muss die Frage nach dem Personenbezug von Fall zu Fall neu beantwortet werden.¹⁴ Bei unzureichender Anonymisierung muss die DSGVO Anwendung finden – mit allen zuvor beschriebenen Konsequenzen. Für den Umgang mit Gesundheitsdaten gelten zudem weitere Regelungen, auf die in Kapitel 2.1.4.2 noch einmal explizit eingegangen wird.

2.1.3.2 Wettbewerbsrecht

Durch das Teilen von Daten mit Unternehmen, die als Wettbewerber anzusehen sind, kann es zu neuen Kooperationsformen kommen, die dem Wettbewerbsrecht unterliegen. Nach dem Gesetz gegen Wettbewerbsbeschränkungen (GWB) § 1 sind „Vereinbarungen zwischen Unternehmen, Beschlüsse von Unternehmensvereinigungen und aufeinander abgestimmte Verhaltensweisen, die eine Verhinderung, Einschränkung oder Verfälschung des Wettbewerbs bezwecken oder bewirken, [...] verboten.“ Werden also gemeinsame Ökosysteme, Projekte und kommerzielle Angebote für das Teilen von Daten geschaffen, ist darauf zu achten, dass entsprechende Verträge keine wettbewerbsbeschränkenden Regelungen enthalten.

2.1.3.3 Urheberrecht

Daten können unter bestimmten Umständen unter das Urheberrecht fallen und es so Datengebern ermöglichen, Dritten über ein ausschließliches Verwertungsrecht Nutzungsrechte einzuräumen oder zu verwehren. Da es im Urheberrecht jedoch vor allem um den Schutz schöpferischer Leistung geht, gilt dies nicht für einzelne Datensätze, wenn diese automatisiert erstellt wurden. Werden aber Daten in Datenbanken gesammelt, kann es sich um ein „Sammelwerk [handeln], dessen Elemente systematisch oder methodisch angeordnet und einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sind.“¹⁵ Solche Datenbanken fallen unter das Urheberrecht, sofern ein Mensch die Struktur der Datenbank festgelegt und somit ein schützenswerter schöpferischer Akt stattgefunden hat. Alternativ ist es auch möglich, Rechtsansprüche als Datenbankhersteller ohne diesen schöpferischen Akt abzuleiten, wenn zur Erstellung der Datenbank eine wesentliche Investition getätigt wurde. Wichtig ist, dass hier immer nur die Datenbank als Ganzes oder ein substanzieller Teil betroffen ist. Rechtsansprüche auf einzelne Daten aus einer solchen Datenbank lassen sich – wie eingangs deutlich gemacht – nicht ableiten.

2.1.3.4 Gesetz zum Schutz von Geschäftsgeheimnissen

Je nach Sensibilität der Daten, die mit anderen Unternehmen geteilt werden sollen, können diese unter das Gesetz zum Schutz von Geschäftsgeheimnissen (GeschGehG) fallen. Dies gilt bspw. für Daten von Produktionsanlagen, die etwa im Rahmen von Predictive Maintenance Services geteilt werden und aus denen sich die Auslastung und damit die Auftragslage ableiten ließen. Unter Umständen ergeben sich aufgrund solcher ableitbaren Informationen Einschränkungen bei der Nutzung der Daten durch den Datenempfänger. Ein Geschäftsgeheimnis ist eine Information, die von wirtschaftlichem Wert ist, durch angemessene Maßnahmen geheim gehalten wird und „bei der ein berechtigtes Interesse an der Geheimhaltung besteht“.¹⁶ Erlangt und nutzt eine unberechtigte Partei diese Information oder legt diese offen, ergibt sich ein Anspruch auf Beseitigung der Beeinträchtigung gegenüber dem Rechtsverletzer.¹⁷ „Diese Ausgangssituation führt dazu, dass datengebende Unternehmen sicherstellen müssen, dass die bereitgestellten Daten, [sic!] auf Seiten des Datenempfängers durch angemessene Geheimhaltungsmaßnahmen geschützt werden.“¹⁸ Entsprechende Non-Disclosure Agreements sollten daher im Falle von Daten, die Informationen mit Geschäftsgeheimnischarakter enthalten, im Rahmen des Teilens von Daten geschlossen werden.

¹⁴ BMWK (2020)

¹⁵ § 4 Abs. 2 UrhG

¹⁶ § 2 Nr. 1 GeschGehG

¹⁷ § 6 GeschGehG

¹⁸ Lindner, Straub, Kühne (2021), S.26

2.1.3.5 Ausblick: EU-Regulierungen

Neben den rechtlichen Fragen, die sich rund um die vertragliche Ausgestaltung und die Eigentumsverhältnisse von Daten drehen, wurden auf europäischer Ebene verschiedene Regulierungen angestoßen, um einen Binnenmarkt für Daten zu schaffen. Im Rahmen der europäischen Datenstrategie von 2020, die das Ziel hat, „die EU an die Spitze einer datengesteuerten Gesellschaft zu bringen“¹⁹, einigten sich das europäische Parlament und der europäische Rat Ende 2021 bereits vorläufig auf den „Data Governance Act“ und bereiten aktuell den sogenannte „Data Act“ vor. Während Ersterer darauf abzielt, die strukturelle Grundlage für eine „gemeinsame Datennutzung von Unternehmen, Einzelpersonen und der öffentlichen Hand [zu] vereinfachen“²⁰, wird Letzterer Bedingungen für die Wertschöpfung aus Daten festlegen. So sieht die EU-Datenstrategie unter anderem die Schaffung von zehn gemeinsamen europäischen Datenräumen in maßgeblichen Sektoren wie Gesundheit, Energie und Mobilität vor, um die Weitergabe von Daten zwischen Sektoren und Mitgliedstaaten zu erleichtern.

¹⁹ ebd.

²⁰ Europäische Kommission (2022b)

European Health Data Space

Die EU-Kommission treibt die Schaffung von europäischen Datenräumen, wie sie in der „Datenstrategie“ vorgesehen sind, voran und nimmt u. a. das Gesundheitswesen in den Blick. Einem vorab veröffentlichten Verordnungsvorschlag für den European Health Data Space zufolge soll so der Zugang zu primären und sekundären Gesundheitsdaten innerhalb der EU geregelt werden.^{21,22} Bürger:innen sollen Zugang zu ihren primären personenbezogenen Gesundheitsdaten bekommen, um etwa Impfnachweise, elektronische Verschreibungen, Bilder, Laborergebnisse, Entlassungsberichte etc. einsehen zu können. Zudem soll dieser Zugang individuell eingeschränkt oder unentgeltlich für Dritte freigegeben werden können. Auch die Sekundärnutzung von Gesundheitsdaten, Verwaltungsdaten, genetischen und genomische Daten, Sozialdaten, öffentlichen Registern, klinischen Studien, biomedizinischen Daten wie Biobanken und Fragebögen zu Forschungszwecken soll ermöglicht werden. Für welche Zwecke die Sekundärdaten verwendet werden dürfen und für welche nicht, wird ebenfalls genau definiert. So sind Zugriffe u. a. im Rahmen der öffentlichen Gesundheitsforschung, der Unterstützung von Behörden bei der Erfüllung ihrer Aufgaben, der wissenschaftlichen Forschung, der Entwicklung und des Trainings von algorithmischen Systemen für medizinische Anwendungen erlaubt. Ein Zugriff auf Sekundärdaten etwa für die Berechnung von Versicherungsprämien, Werbung oder den Verkauf an Dritte wird hingegen ausgeschlossen.

Um den Austausch von primären und sekundären Gesundheitsdaten zu ermöglichen, soll ein dezentrales europäisches Ökosystem für Gesundheitsdaten geschaffen werden, „das aus Vorschriften, gemeinsamen Standards und Verfahren, Infrastrukturen und einem Governance-Rahmen besteht“: der „European Health Data Space“.²³ Dieser muss bestimmte Sicherheitsanforderungen erfüllen, zu denen etwa die Unterbindung des Zugangs von Unbefugten sowie des unerlaubten Lesens, Änderns, Kopierens oder Entfernens von elektronischen Gesundheitsdaten zählen. Zudem sind alle Zugriffe zu protokollieren. Des Weiteren wird großer Wert auf Interoperabilität bei der internationalen Datenübermittlung gelegt. Nationale Systeme für elektronische Gesundheitsakten müssen mit der europäischen Infrastruktur kompatibel sein, um einen Austausch von Gesundheitsdaten mit Fachleuten und Einrichtungen des Gesundheitswesens über Landesgrenzen hinaus zu gewährleisten. Dazu müssen sich entsprechende Systeme zertifizieren lassen.

Auf organisatorischer Ebene will die Kommission einen europäischen Ausschuss für Digital- und Gesundheitsdaten einrichten, der wiederum zwei Untergruppen zu Primär- beziehungsweise Sekundärdaten erhalten soll. Der Ausschuss soll sich mit der Ausarbeitung der Leitlinien und Anforderungen für die Kennzeichnung, Zertifizierung und Datenqualität, die für den Betrieb des europäischen Gesundheitsdatenraums erforderlich sind, beschäftigen und die Kommission zu diesem Thema beraten.

Auch finanziell wird die EU-Kommission den Aufbau des European Health Data Space unterstützen: Mit dem EU4Health-Förderprogramm soll das Pilotprojekt zur Entwicklung der neuen dezentralen EU-Infrastruktur für die Sekundärnutzung von Daten gefördert werden.

Mit der Einrichtung eines „European Data Innovation Board“, bestehend aus Vertretern aller EU-Staaten, sollen unter anderem Prioritäten für sektorübergreifende Standards, die für die Datennutzung und den sektorübergreifenden Datenaustausch verwendet und entwickelt werden sollen, festgelegt werden.²⁴ Im „Data Act“ werden unterschiedliche Maßnahmen vorgeschlagen, um verschiedenen Stakeholdern Zugriff auf Daten zu geben. So sollen u. a. Endkund:innen Zugang zu den von ihren vernetzten Geräten erzeugten Daten bekommen, kleine und mittlere Unternehmen vor Ungleichgewichten in Verträgen über die gemeinsame Datennutzung bewahrt werden und öffentlichen

²¹ Ankowitsch (2022)

²² Fortuna & Bertuzzi (2022)

²³ Europäische Kommission (2022b)

²⁴ Data Governance Act Proposal, Article 27

Stellen der Zugang zu und die Nutzung von Daten im Besitz des Privatsektors im Falle von besonderen Situationen ermöglicht werden. Insgesamt soll ein Wettbewerbsmarkt für Daten geschaffen werden, der das „Potenzial von Daten und Technologien im Einklang mit den Vorschriften und Werten der EU“²⁵ freisetzt.

2.1.4 Besonderheiten des Gesundheitssektors

Während die zuvor aufgezeigten Potenziale und Hürden des Datenaustausches zwischen Unternehmen über viele Branchen hinweg gelten, herrschen in einzelnen Sektoren wiederum besondere Bedingungen vor. Der Gesundheitssektor spielt hier eine spezielle Rolle, da seit jeher ökonomische, gesellschaftliche und technische Perspektiven miteinander vereinbart werden müssen. Entsprechend stehen hier auf der einen Seite enorme Potenziale für die Verbesserung der medizinischen Versorgung aller Bürger:innen sowie auf der anderen Seite historisch gewachsene Hürden auf regulatorischer und technischer Ebene.

2.1.4.1 Potenziale

Mit dem digitalen Wandel im Gesundheitssektor werden große Erwartungen an die Verbesserung der Gesundheitsversorgung verbunden. So verspricht er eine engere Vernetzung aller Leistungserbringer und damit u. a. die verbesserte Anamnese und Direktion der Patient:innen im komplexen Gesundheitssystem, die Vermeidung von Doppeluntersuchungen und Fehlmedikationen, standortunabhängige Versorgung und Leistungserbringung oder individualisierte Prävention und Therapie, Betreuung und Nachverfolgung. Auch Reaktionszeiten auf gesundheitliche Krisensituationen wie die aktuelle COVID-19-Pandemie werden sich verkürzen lassen. Daraus resultiert eine Verbesserung der individuellen Versorgung und eine Anhebung des Gesundheitsniveaus der gesamten Bevölkerung. Aus ökonomischer Perspektive ergeben sich ebenfalls „große Innovations- und Wertschöpfungspotenziale“, etwa durch die effizientere Gestaltung von Prozessen und die damit einhergehende Optimierung der Ressourcenallokation.²⁶ Bis zu 34 Milliarden Euro sollen sich so im deutschen Gesundheitswesen einsparen lassen.²⁷ Für Unternehmen der Privatwirtschaft entstehen neue Geschäftsfelder, etwa im B2P-Bereich („Business-to-Patient“). Digitale Lösungen zur Therapiebegleitung, dem Management von chronischen Krankheiten oder zum Monitoring des individuellen Gesundheitszustands sind nur einige Beispiele dafür, welche neuen Möglichkeiten gerade aufkommen.²⁸ Grundlage für all diese Lösungen ist die Erhebung, Sammlung und Verarbeitung von Gesundheitsdaten aus dem Versorgungsalltag, die für Versorgung (primäre Nutzung) und Forschung und Entwicklung (sekundäre Nutzung) genutzt werden können. So lassen sich z. B. „Daten in der medizinischen Forschung dazu verwen[n]de[n] [...], vergleichende Wirksamkeitsstudien durchzuführen sowie den Verlauf von Therapien zu verfolgen“.²⁹

Die notwendigen Daten entstehen zum einen durch die Vernetzung der Leistungserbringer untereinander und zum anderen durch die Vernetzung zwischen einzelnen Geräten im medizinischen Alltag. So können Chirurg:innen im Operationssaal in Zukunft auf eine Vielzahl von Sensoren zurückgreifen, etwa in den Instrumenten, dem OP-Tisch oder den eingesetzten Operationsrobotern. Die Verbindung dieser medizinischen Geräte durch ein Internet of Medical Things (IoMT) ermöglicht es, dem Operationsteam alle gesammelten Informationen zentral zur Verfügung zu stellen und ein präzises Bild über den Zustand der betroffenen Person abzubilden. Aus der Herstellerperspektive liegen die Vorteile darin, dass Schäden an teuren Geräten rechtzeitig erkannt und Fehlfunktionen oder eine falsche Bedienung vermieden werden können. Per Remote Maintenance können Techniker:innen dies zentral von ihrem Standort erledigen und müssen nicht mehr in jedes einzelne Krankenhaus fahren.³⁰ Je vernetzter die Geräte sind und desto mehr Daten gesammelt werden, desto besser werden auch die Grundlagen für das Training von algorithmischen Systemen (bspw. Künstlicher Intelligenz) in der

²⁵ Europäische Kommission (2022c)

²⁶ EFI (2022), S.96

²⁷ McKinsey Digital (2018)

²⁸ Deloitte (2019)

²⁹ EFI (2022), S.102

³⁰ Institute for Digital Transformation in Healthcare (2021)

Medizin. Schon heute können entsprechende Programme in der Radiologie zur Bildklassifizierung, Läsionserkennung und Segmentierung genutzt werden, um Radiolog:innen bei der Entscheidungsfindung zu unterstützen.³¹ Die Potenziale von algorithmischen Systemen für alle weiteren Bereiche der medizinischen Versorgung sind ebenfalls groß³² – vorausgesetzt, die Vernetzung von Daten im Gesundheitssektor erreicht ein ausreichendes Niveau. Um zukünftig die Potenziale vollständig freisetzen zu können, müssen jedoch noch einige Hürden überwunden werden.

2.1.4.2 Hürden

Die Potenziale der Datennutzung im Gesundheitssektor werden noch durch unterschiedliche Hürden limitiert. Diese lassen sich auf drei Ebenen beschreiben: Gesetzeslage, komplexe Stakeholder-Situation und Qualität & Standards.

Gesetzeslage

So geht der digitale Wandel im Gesundheitswesen allgemein nur langsam voran. Mit zahlreichen Gesetzen in den letzten Jahren, angefangen mit dem "Gesetz für sichere digitale Kommunikation und Anwendungen im Gesundheitswesen" (E-Health-Gesetz) von 2015, wurde in Deutschland politisch versucht, die Digitalisierung verschiedener Elemente der medizinischen Versorgung voranzubringen. Bisher ist der Nutzen in der Praxis jedoch gering. Die elektronische Patientenakte (ePA), als zentrales Element für den Austausch von Patientendaten Anfang 2021 gestartet, wird laut einer Umfrage des Bitkom e. V. bisher nur von 0,5 Prozent der Deutschen genutzt.³³ Hinzu kommt, dass die im deutschen Gesundheitswesen bestehende technologische Infrastruktur, die einen sicheren und vertrauensvollen Austausch von Gesundheitsdaten zwischen den verschiedenen Stakeholdern des Gesundheitssektors ermöglichen soll, ein geschlossenes System ist. Für den Austausch primärer Behandlungsdaten steht Leistungserbringern (Ärzt:innen, Heilberufe), Leistungsempfängern (Patient:innen) und Kostenträgern (Krankenkassen) in Deutschland die Telematikinfrastruktur (TI) zur Verfügung. Nach jahrelanger Planung sind hier mittlerweile immerhin 89 Prozent der Arztpraxen und 77 Prozent der psychotherapeutischen Praxen angeschlossen.³⁴ Bisher wird allerdings lediglich das Versichertenstammdatenmanagement regelmäßig genutzt.

³¹ Moawad et al. (2022)

³² Plattform Lernende Systeme (2019)

³³ Bitkom (2021)

³⁴ KBV (2022)

① Primäre und sekundäre Datennutzung

Im Gesundheitssektor lassen sich zwei Nutzungsarten voneinander unterscheiden. Daten, die im Rahmen der Versorgung von Patient:innen erhoben werden, etwa über den Blutdruck oder die Herzfrequenz, werden als Primärdaten bezeichnet. Sie werden genutzt, um die individuelle Behandlung des Betroffenen zu gewährleisten (primäre Datennutzung). Werden diese Primärdaten vieler verschiedener betroffener Personen bspw. für die Forschung aggregiert und analysiert, können neue Erkenntnisse gewonnen werden, die über den konkreten Versorgungskontext hinaus gehen (sekundäre Datennutzung). Durch diese Weiterverarbeitung der Primärdaten entstehen Sekundärdaten, bei denen der direkte Bezug zum konkreten Versorgungskontext nicht mehr hergestellt werden kann. „Neben den weiterverarbeiteten Primärdaten gehören zu den Sekundärdaten auch administrative Daten wie z. B. Daten der amtlichen Statistik, Daten aus pharmakoepidemiologischen Datenbanken oder Daten aus (krankheitsbezogenen) Registern. Sekundärdaten bezeichnen somit Daten, die bereits vorhanden sind und deren Datenfelder sich aus ihrem primären Zweck ergeben, die also nicht mehr infolge einer empirischen Heuristik oder einer forschenden Fragestellung de novo angelegt werden können“.³⁵

Die Nutzung von Gesundheitsdaten unterliegt ohnehin regulatorischen Beschränkungen. Nach Art. 9 DSGVO handelt es sich bei Gesundheitsdaten um eine besondere Kategorie personenbezogener Daten. Deren Verarbeitung ist grundsätzlich untersagt, es sei denn, der Verarbeitung für einen bestimmten Zweck wurde ausdrücklich zugestimmt oder die Verarbeitung erfolgt „für Zwecke der Gesundheitsvorsorge oder der Arbeitsmedizin, [...] für die medizinische Diagnostik, die Versorgung oder Behandlung im Gesundheits- oder Sozialbereich oder für die Verwaltung von Systemen und Diensten im Gesundheits- oder Sozialbereich“.³⁶ Genauso ist eine „Verarbeitung [...] aus Gründen des öffentlichen Interesses im Bereich der öffentlichen Gesundheit“ vorgesehen. Mit dem deutschen Patientendatenschutzgesetz (PDSG) § 363 SGB V ist zudem ab 2025 die Verarbeitung von anonymisierten Daten der elektronischen Patientenakte zu öffentlichen Forschungszwecken erlaubt. Hierzu muss ein Antrag bei einem Forschungsdatenzentrum gestellt werden. Unternehmen der Privatwirtschaft gehören jedoch nicht zu den Antragsberechtigten. Hinzu kommt, dass die Landesdatenschutzbeauftragten die DSGVO ohnehin unterschiedlich interpretieren. Verfahren zur Anonymisierung von Daten in Universitätsklinik A werden in Bundesland X als ausreichend bewertet. Dasselbe Verfahren in Universitätsklinik B in Bundesland Y wird als nicht ausreichend bewertet, sodass die DSGVO umfassend greift. Für Unternehmen ist also selbst die Anonymisierung kein Garant für die rechtssichere Verarbeitung von Gesundheitsdaten.

Komplexe Stakeholder-Situation

Ein weiterer Faktor, der den digitalen Wandel und damit die Nutzung von Daten im Gesundheitssektor erschwert, ist die komplexe Stakeholder-Situation. Ärzt:innen und andere Heilberufe erbringen Leistungen an Patient:innen und werden dafür von den Krankenkassen vergütet. In dieser Dreierkonstellation haben alle Beteiligten unterschiedliche Einstellungen zum Thema Digitalisierung und Datenteilen. Niedergelassene Ärzt:innen sehen z. B. noch wenig Nutzen von „[d]igitale[n] Anwendungen zur Sammlung medizinischer Daten“ und sorgen sich um „Sicherheitslücken in der EDV“ und um ein aktuell „[u]ngünstiges Kosten-Nutzen-Verhältnis“.³⁷ Patient:innen hingegen sind zwar grundsätzlich bereit, Daten in Form einer Datenspende für die öffentliche Forschung bereitzustellen³⁸, dies gilt jedoch nicht für das Teilen von Gesundheitsdaten mit „Wissenschaftlern aus der Industrie und von privaten Unternehmen“.³⁹ Gesetzliche Krankenkassen sind beim Thema Datennutzung bereits weiter: Hier

³⁵ Gothe, Ihle, Swart (2021), S.65

³⁶ DSGVO Art. 9

³⁷ KBV (2022), S.4

³⁸ Bitkom (2021)

³⁹ Lesch et al. (2022), S.220

kommen immer häufiger Big-Data-Anwendungen zum Einsatz.⁴⁰ Die Ziele dieser Anwendungen beziehen sich jedoch auf interne Abläufe zur Prozessoptimierung oder verbesserten Kundenansprache mit dem Ziel der Differenzierung vom Wettbewerb. Anwendungen in der medizinischen Versorgung (bspw. Medikationspläne o. ä.) werden aktuell in der Regel nicht angeboten.

Qualität & Standards

Doch selbst wenn die Bereitschaft zum Teilen von Daten bei Patient:innen und Ärzt:innen steigen sollte, stellt die aktuelle Verfügbarkeit von Gesundheitsdaten eine weitere Hürde dar. Die Daten sind fragmentiert und auf unterschiedliche Stakeholder des ersten Gesundheitsmarktes wie Arztpraxen, Krankenhäusern, Psycholog:innen, bei Heil- und Pflegediensten, Apotheken oder Krankenkassen verteilt. Hinzu kommen Datenquellen aus dem zweiten Gesundheitsmarkt, die per se von denen des ersten Gesundheitsmarktes getrennt sind. Mangelnde Standards und Interoperabilität in beiden Märkten führen zu einer niedrigen Datenqualität und verhindern, dass Systeme verschiedener Hersteller miteinander kommunizieren können. „So können patientenbezogene Daten aus technischen und aus datenschutzrechtlichen Gründen nicht über die Grenzen von Standorten hinweg verknüpft werden.“⁴¹ Zwar gibt es bestehende Standards wie HL7, FHIR oder SNOMED CT⁴², diese beziehen sich jedoch jeweils auf spezielle Systeme, wie z. B. Terminologie-Standards und sind in Anbetracht der riesigen Menge an potenziellen Datenquellen noch zu wenig verbreitet. Bevor Gesundheitsdaten in größerem Umfang geteilt werden können, bedarf es einer gemeinsamen „Sprache“ für gesundheitsrelevante Datenbestände. Im Rahmen der 2018 ins Leben gerufenen Medizininformatik-Initiative des BMBF werden solche Fragen nach den Rahmenbedingungen für den Datenzugang und die Datennutzung sowie der Interoperabilität erforscht und für Universitätsmedizin-Standorte bereits umgesetzt. Entsprechende Ergebnisse sind daher in Zukunft im Blick zu behalten.

Medizininformatik-Initiative

Die Medizininformatik-Initiative (MII) ist ein mehrphasiges Förderprogramm des Bundesministeriums für Bildung und Forschung (BMBF), in dem an den deutschen Universitätskliniken und Partnereinrichtungen Datenintegrationszentren aufgebaut und vernetzt werden, um Forschungs- und Versorgungsdaten standortübergreifend verknüpfen zu können. Für konkrete medizinische Anwendungen werden auf dieser Basis IT-Lösungen entwickelt, die die Möglichkeiten digitaler Dienstleistungen und Infrastrukturen im Gesundheitsbereich zeigen sollen. Während einzelne Projekte und Anwendungsfälle in geförderten Konsortien aus Forschungsinstituten, Hochschulen, Unternehmen oder nicht-universitären Krankenhäusern bearbeitet werden, gibt es konsortienübergreifende Arbeitsgruppen zu Consent, Data Sharing, Interoperabilität und Kommunikation, die grundlegende Aspekte der Vernetzung von Daten aus Forschung und Patientenversorgung definieren.

2.2 Marktanalyse

Nachdem die Potenziale und Hürden des Teilens von Daten zwischen Unternehmen im Allgemeinen und speziell zwischen den Stakeholdern im Gesundheitswesen aufgezeigt wurden, soll nun betrachtet werden, welche Wege es für das Teilen von Daten aktuell gibt. Hier lässt sich eine Reihe von Initiativen, Ökosystemen, Projekten und kommerziellen Angeboten identifizieren. Grundsätzlich können zwei Arten unterschieden werden (siehe Abbildung 1): Zum einen gibt es solche, die grundlegende vorwettbewerbliche Konzepte für die technische Durchführung und die Governance zum Teilen von Daten erarbeiten, die aber keine eigene Lösung kommerziell betreiben. Die entsprechenden Konzepte sind für interessierte Unternehmen zugänglich und können als Basis für konkrete Use Cases genutzt werden. Solche Initiativen sollen im Folgenden daher als „Framework Initiativen“ bezeichnet werden.

⁴⁰ Radić et al. (2016)

⁴¹ Lesch et al. (2022), S.212

⁴² Seit dem 01.01.2021 ist Deutschland Mitglied von SNOMED International.

Zum anderen gibt es einzelne Unternehmen sowie Konsortien, die selbst Lösungen für den Austausch von Daten aufbauen und betreiben. Diese „Data Sharing Ökosysteme“ setzen zum Teil die oben genannten Frameworks in der Praxis um, verwenden aber auch eigene Ansätze. Im Folgenden werden die zwei Arten anhand von Beispielen genauer vorgestellt, um im Anschluss ihre verbindenden Merkmale herauszuarbeiten.

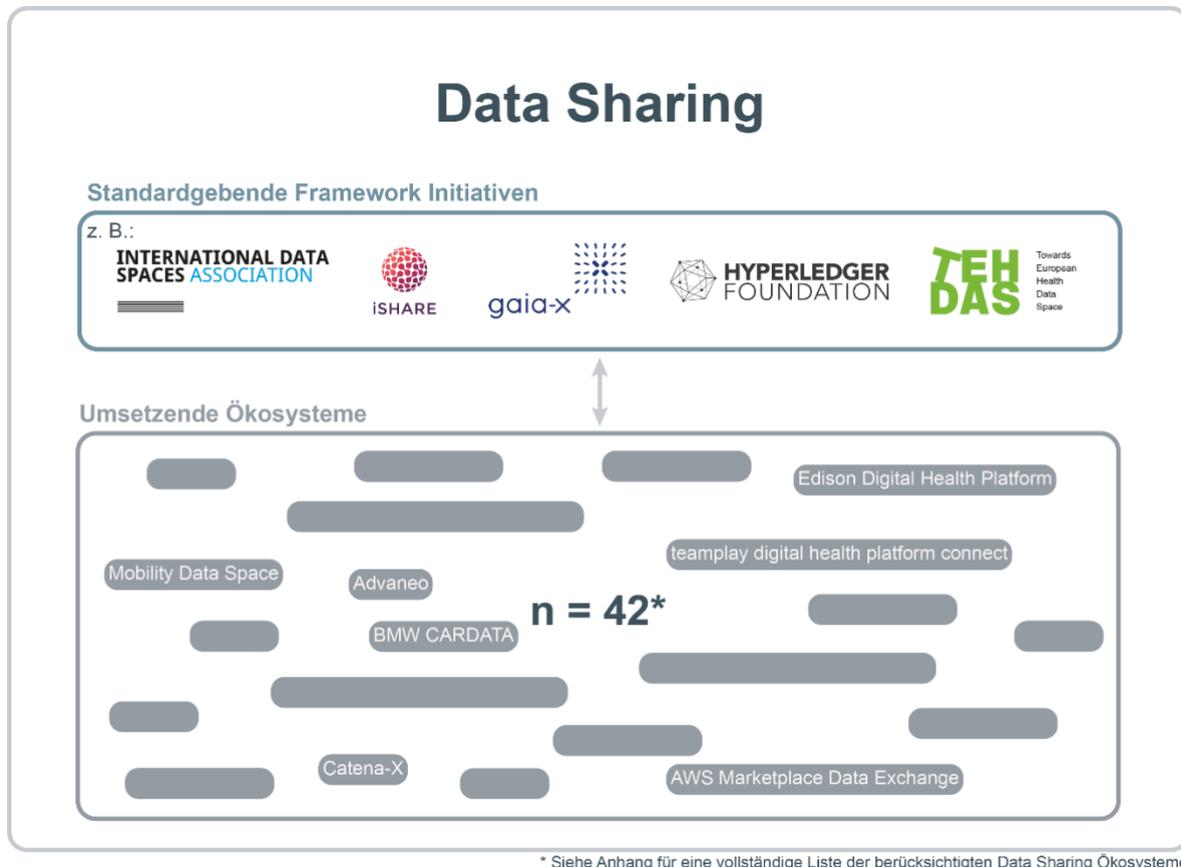


Abbildung 1: Verhältnis von Framework Initiativen und Ökosystemen

2.2.1 Framework Initiativen

Zunächst sollen exemplarisch fünf Framework Initiativen vorgestellt werden, um die unterschiedlichen technischen und organisatorischen Möglichkeiten für das Teilen von Daten zu beleuchten.⁴³ Die ausgewählten Initiativen sind bereits seit längerem aktiv und werden von maßgeblichen Stakeholdern aus Wirtschaft und Forschung unterstützt. Es wurden vor allem europäische Initiativen ausgewählt, da diese die rechtlichen Rahmenbedingungen der DSGVO explizit berücksichtigen. Mit der „Hyperledger Foundation“ ist allerdings auch eine US-amerikanische Initiative vertreten, um das Potenzial der Blockchain-Technologie abseits von Crypto-Währungen aufzuzeigen.

International Data Spaces Association (IDSA)

Die International Data Spaces Association (IDSA) verfolgt das Ziel, den sicheren und standardisierten Austausch und die einfache Verknüpfung von Daten in domänenübergreifenden Datenräumen zu ermöglichen. Dazu wurde der International Data Spaces (IDS)-Standard in Form eines Referenzarchitekturmodells geschaffen, das technische Architektur und Governance-Modelle kombiniert. Er ermöglicht es Unternehmen und Einzelpersonen selbst zu bestimmen, wie, wann und zu welchem Preis ihre Daten über die Wertschöpfungskette hinweg genutzt werden und sorgt so für Datensouveränität.⁴⁴ Hauptmerkmal des IDS-Standards ist die dezentrale Datenhaltung, unter

⁴³ Zu weiteren relevanten Initiativen zählen u. a. noch „FIWARE Foundation“ oder „BVDA – Big Data Value Association“

⁴⁴ IDSA (2022a)

Respektierung proprietärer Systeme, und die Authentisierung der teilnehmenden Organisationen über zertifizierte Konnektoren.⁴⁵ Mit der Festlegung von Nutzungsrichtlinien und deren Durchsetzung in der Laufzeitumgebung dieser Konnektoren kann Datensouveränität technologisch umgesetzt werden. Ursprünglich wurde die Initiative als Forschungsprojekt unter dem Namen „Industrial Data Spaces“ vom Bundesministerium für Bildung und Forschung gefördert, mittlerweile gehören dem eingetragenen Verein über 130 Unternehmen, Forscher und Industrieverbände aus über 20 Ländern an. Die IDSA sieht den IDS als offene Architektur, die mit anderen Frameworks kompatibel ist und sucht die Zusammenarbeit mit anderen Initiativen zum Data Sharing. Daher ist sie u. a. ebenfalls Gründungsmitglied der Initiative Gaia-X AISBL, die im Folgenden vorgestellt wird.

Gaia-X European Association for Data and Cloud AISBL

Das Projekt Gaia-X wurde 2019 unter der Federführung Deutschlands und Frankreichs ins Leben gerufen, um eine Dateninfrastruktur für ein digitales Ökosystem in Europa zu schaffen, das nach europäischen Werten und Datenschutzstandards aufgebaut wird. Ähnlich wie der IDS-Standard setzt auch der Gaia-X-Standard auf eine dezentrale Datenhaltung. In Bezug auf die Anforderungen der verwendeten Architektur ergeben sich daher zwischen Gaia-X und IDS einige Überschneidungen.

📘 Zum Verhältnis der Standards von IDS und Gaia-X

Ein im Januar 2021 veröffentlichtes Positionspapier der IDSA⁴⁶ beschreibt das Verhältnis von GAIA-X und IDS wie folgt:

“Aus der konzeptionellen Zuordnung ergibt sich die folgende übergeordnete Beziehung:

- Der föderierte Katalog umfasst den IDS-Broker, den Vokabularanbieter und das Informationsmodell
- Der Federation Service of Sovereign Data Exchange stützt sich auf die IDS-Nutzungskontrolle und die Clearingstelle
- Der Gaia-X Federation Service of Identity & Trust profitiert vom IDS Identity Provider und Dynamic Attribute Provisioning Service (DAPS)
- Gaia-X Nodes sind mit IDS Konnektoren als Gateways kompatibel
- Das Gaia-X Data Ecosystem ist der Ort, an dem IDS Data Provider und Data Consumer konzeptionell angesiedelt sind
- Der IDS Service Provider, der IDS App Store Provider und der App Provider sind im Gaia-X Infrastructure Ecosystem angesiedelt. [Übers. d. Verf.]⁴⁷

Beide Frameworks teilen dasselbe Ziel eines auf Datensouveränität und Vertrauen basierenden Ökosystems zum Austausch von Daten und können daher potenziell aufeinander abgestimmt werden.

Der Hauptgedanke des föderierten Systems von Gaia-X beinhaltet, dass auf diesem Weg nicht eine einzelne Cloud entsteht, sondern ein vernetztes System, das viele Anbieter von Cloud-Diensten miteinander verbindet.⁴⁸ Mit der Gründung der Gaia-X European Association for Data and Cloud AISBL, in der Vertreter:innen aus Wirtschaft, Politik und Wissenschaft aus der ganzen Welt zusammenarbeiten, wurden die Ambitionen des Vorhabens unterstrichen. Die Gaia-X Association übernimmt verschiedene organisatorische Aufgaben, um die Entwicklung des Gaia-X-Standards voranzubringen. So gehört neben der Erarbeitung der rechtlichen Rahmenbedingungen auch die Stärkung der Kooperation innerhalb der Gaia-X-Community zu ihrer Rolle. Nach außen vertritt sie die Mitglieder von Gaia-X und fördert die internationale Zusammenarbeit.⁴⁹ Auch von staatlicher Seite erhält der Gaia-X-Standard

⁴⁵ Siehe Kapitel 2.3 für eine detaillierte Beschreibung der Funktionsweise eines Datenraums nach IDS-Standard

⁴⁶ Otto, B. et al. (2021)

⁴⁷ ebd., S.4

⁴⁸ Gaia-X-Hub Germany (2022)

⁴⁹ Gaia-X (2021)

Unterstützung: Im Rahmen des Förderwettbewerbs „Innovative und praxisnahe Anwendungen und Datenräume im digitalen Ökosystem Gaia-X“ der Bundesnetzagentur werden branchenspezifische Forschungsprojekte zur praktischen Umsetzung der Gaia-X-Prinzipien gefördert.

TEHDAS

Das durch das EU-Gesundheitsprogramm finanzierte Projekt „Towards European Health Data Space (TEHDAS)“ hat zum Ziel, die grenzüberschreitende Sekundärnutzung von Gesundheitsdaten in Europa mit der Entwicklung von Grundsätzen für den Aufbau eines europäischen Gesundheitsdatenraums zu fördern. Die (Zwischen-)Ergebnisse der von Januar 2021 bis August 2023 laufenden Projektphase werden Einfluss auf den finalen Gesetzesvorschlag der EU-Kommission zum „European Health Data Space“ (siehe 2.1.3.5) nehmen. Koordiniert wird das Projekt, an dem zahlreiche Partner aus 21 EU-Mitgliedstaaten und vier weiteren europäischen Ländern beteiligt sind, vom finnischen Innovationsfonds Sitra. In drei organisatorischen und fünf inhaltlichen Arbeitspaketen werden sowohl technische, regulatorische als auch Governance-bezogene Optionen für die grenzüberschreitende Zusammenarbeit bei der Sekundärnutzung von Gesundheitsdaten zwischen europäischen Ländern entwickelt. Die Klärung der Rolle der einzelnen Bürger:innen und der Dialog mit anderen europäischen Projekten und politischen Entscheidungsträgern gehören ebenfalls zu den Zielen von TEHDAS. Regelmäßig veröffentlichte Reports der einzelnen Arbeitspakete geben einen Einblick in den aktuellen Stand der Entwicklung.⁵⁰ So wurden im April 2022 bspw. Vorschläge für die grundlegenden Softwareelemente, etwa für den Abruf, die Integration sowie die Anonymisierung und Pseudonymisierung von sekundären Gesundheitsdaten im EHDS vorgestellt.⁵¹ Darauf aufbauend soll in Zukunft eine detaillierte Beschreibung der Architektur und der Infrastrukturoptionen der EHDS-Dienste folgen. Hier ist bereits abzusehen, dass der EHDS als dezentrales Peer-to-Peer-Netzwerk gedacht wird, dessen Zugang über einzelne Knotenpunkte und zentralisierte Dienste geregelt wird. Einzelne Elemente erinnern dabei an das Referenzarchitekturmodell der IDSA. Weitere Zwischenergebnisse umfassen z. B. Leitlinien für den Aufbau grenzüberschreitender Partnerschaften zur gemeinsamen Nutzung von Gesundheitsdaten für sekundäre Zwecke sowie die Identifizierung von relevanten Standards und Datenmodellen für die semantische Harmonisierung. Auch wenn TEHDAS zunächst Optionen aufzeigt und keine Aussage über die endgültige Gestaltung des EHDS getroffen werden kann, sollten die Projektergebnisse weiterverfolgt werden.

iSHARE Foundation

Die niederländische „iSHARE Foundation“ hat sich ebenfalls auf die Fahne geschrieben, ein dezentrales Netzwerk zum Austausch von Daten zu schaffen. Der besondere Fokus der gemeinnützigen Organisation liegt auf der Schaffung von Vertrauen zwischen den einzelnen Teilnehmern ihres Netzwerkes. Dieses wird durch die Zugehörigkeit zum iSHARE-Netzwerk erworben, was eine Vereinbarung mit der iSHARE Foundation erfordert. Für die Umsetzung wurde das sog. „iSHARE Trust Framework“ entwickelt, das zunächst für den Transport- und Logistiksektor in den Niederlanden konzipiert wurde und die „Themen Identifizierung, Authentifizierung und Autorisierung von Geschäftsdatenattributen“⁵² adressiert. Das Framework ist dabei eine Kombination aus funktionalen, technischen, betrieblichen und rechtlichen Vereinbarungen, an die sich die beteiligten Parteien halten müssen. Die iSHARE Foundation fungiert als Dachorganisation, die die Durchsetzung der Vereinbarungen, die Verwaltung der mit dem Beitritt zum Netzwerk verbundenen Prozesse und die Erleichterung von Änderungen am System garantiert. Diese Vereinbarungen ermöglichen es den Beteiligten, Logistikdaten mit anderen Parteien auszutauschen – auch mit neuen und bisher unbekanntem Partnern. Der Datengeber behält dabei immer die endgültige Entscheidung darüber, wer auf welche Daten wie zugreifen darf. Genau wie die IDSA hat so auch die iSHARE Foundation zum Ziel, Datensouveränität beim Teilen von Daten zwischen Unternehmen zu gewährleisten. Seit 2019 existiert daher eine Kooperation zwischen den beiden Organisationen, in Folge derer die iSHARE-

⁵⁰ TEHDAS (2022a)

⁵¹ TEHDAS (2022b)

⁵² iSHARE (2021)

Vereinbarungen über die gemeinsame Nutzung von Daten in die IDS-Architektur übernommen wurden.⁵³

Hyperledger Foundation

Das Projekt „Hyperledger“ der Linux Foundation gehört neben Ethereum zu den größten Vorhaben im Bereich der Blockchain-Technologie.⁵⁴ Es setzt sich aus einer Reihe von Open-Source-Tools und Unterprojekten zusammen und ermöglicht seinen Mitgliedern, Blockchain-Netzwerke oder kommerzielle Lösungen zu entwickeln und einzusetzen.⁵⁵ Zu den über 170 Mitgliedern⁵⁶ zählen u. a. IBM, Accenture, Lenovo oder die Deutsche Telekom. Bei der Hyperledger Foundation handelt sich um ein gemeinnütziges Konsortium, das Unternehmenslösungen und universelle Richtlinien für die Blockchain-Implementierung bereitstellt, ohne selbst Blockchain-Netzwerke zu betreiben. Im Kern steht das Softwaremodul „Hyperledger Fabric“, das als Framework für die Speicherung der Transaktionen im Netzwerk sorgt. Die Ziele von Hyperledger liegen im Allgemeinen in der Schaffung von hochskalierenden Industrieanwendungen⁵⁷ – das Teilen von Daten über private Blockchain-Netzwerke mit authentifizierten Mitgliedern ist dabei nur eine Möglichkeit von vielen. Damit unterscheidet es sich von den anderen Framework Initiativen. Hyperledger bietet ähnliche Vorteile (Schaffung von Vertrauen, Verantwortung und Transparenz) wie die bereits vorgestellten Initiativen, setzt aber mit der Nutzung von Blockchain-Netzwerken auf eine andere technologische Basis, um diese Effekte zu erreichen.

Beobachtung: Wie ein Blick auf die vorgestellten Framework Initiativen verdeutlicht, sind diese branchenagnostisch und (europäisch) international aufgestellt. Eine Ausnahme bildet das Projekt TEHDAS, welches sich explizit mit dem Aufbau des European Health Data Space befasst. Die Frameworks stehen nicht in Konkurrenz zueinander, sondern können sich an unterschiedlichen Stellen ergänzen. Einzelne Bausteine sind ähnlich oder miteinander kompatibel und lassen sich daher verbinden. Dies zeigt sich beispielsweise an der Data Spaces Business Alliance⁵⁸, in der sich die Gaia-X European Association for Data and Cloud AISBL, Big Data Value Association (BVDA), FIWARE Foundation und International Data Spaces Association (IDSA) zusammengeschlossen haben, um gemeinsam auf das Thema Data Sharing aufmerksam zu machen, Technologien weiterzuentwickeln und Standards zu formen. Wie die Beispiele der IDSA und iSHARE zeigen, werden solche Verbindungen zuweilen durch bilaterale Kooperationen formalisiert. Darüber hinaus gibt es große Überschneidungen bei den involvierten Stakeholdern; so sind bspw. die Unternehmen Bosch und Deutsche Telekom sowohl in der IDSA als auch in Gaia-X und der Hyperledger Foundation als Mitglieder aktiv. Die Orientierung oder das Mitwirken an einer oder mehrerer Framework Initiativen ist daher sowohl üblich als auch förderlich, wenn es um den Aufbau eines eigenen Data Sharing-Projekts geht. Insbesondere für den Gesundheitssektor werden die Ergebnisse von TEHDAS einen relevanten Einfluss auf zukünftige Entwicklungen haben.

2.2.2 Data Sharing Ökosysteme

Im Gegensatz zu Framework Initiativen konzentrieren sich Data Sharing Ökosysteme auf den Betrieb einer Infrastruktur zum Teilen von Daten, hauptsächlich in bestimmten Branchen. Um die Gemeinsamkeiten und Unterschiede solcher Ökosysteme zu analysieren und die Implikationen für den Aufbau zukünftiger Data Sharing Ökosysteme abzuleiten, wurde ein Sample aus Forschungsprojekten, kommerziellen Angeboten und Konsortien für das Teilen von Daten identifiziert. Ein Teil der Auswahl von Beispielen wurde durch den Austausch mit Expert:innen aus Framework Initiativen, insbesondere von „IDSA“ und „Gaia-X“, getroffen. Ergänzend wurde eine eigene Internet- und Literaturrecherche unter den Begriffen „Data Sharing“, „Plattform“ und „Datenmarktplatz“ sowie verwandten Begriffen

⁵³ IDSA (2019)

⁵⁴ Aggarwal & Kumar (2021)

⁵⁵ Hyperledger Foundation (2022)

⁵⁶ Unternehmen, die Teil der Hyperledger Foundation und damit der Entwicklercommunity werden möchten, können aus zwei Mitgliedschaftsstufen (Premier \$250.000 & General ab \$5000) wählen.

⁵⁷ Aggarwal & Kumar (2021)

⁵⁸ Data Spaces Business Alliance (2022)

durchgeführt. So entstand eine Liste mit knapp 100 Data Sharing-Lösungen aus unterschiedlichsten Bereichen. Im Rahmen der Finalisierung des Samples wurde dann explizit nach gesundheitsbezogenen Ökosystemen gefiltert, ohne jedoch einzelne Branchen komplett auszuschließen (siehe Abbildung 2). Der Fokus bei der Auswahl lag zudem allgemein auf europäischen Stakeholdern, trotzdem sind auch einige US-amerikanische Beispiele enthalten, da diese gerade im Gesundheitssektor aufgrund der dort vorherrschenden Rahmenbedingungen über einen Vorsprung gegenüber Europa verfügen. Zudem sollten große globale Player wie Amazon Web Services bei einer Analyse nicht ausgeblendet werden. Auf Basis der genannten Kriterien wurden schließlich 42 Data Sharing Ökosysteme für das finale Sample ausgewählt (siehe Anhang für vollständige Liste). Die Auswertung brachte dann verschiedene Klassifikationen (siehe Abbildung 3) mit unterschiedlichen Implikationen zu Tage, die im Folgenden beschrieben werden.

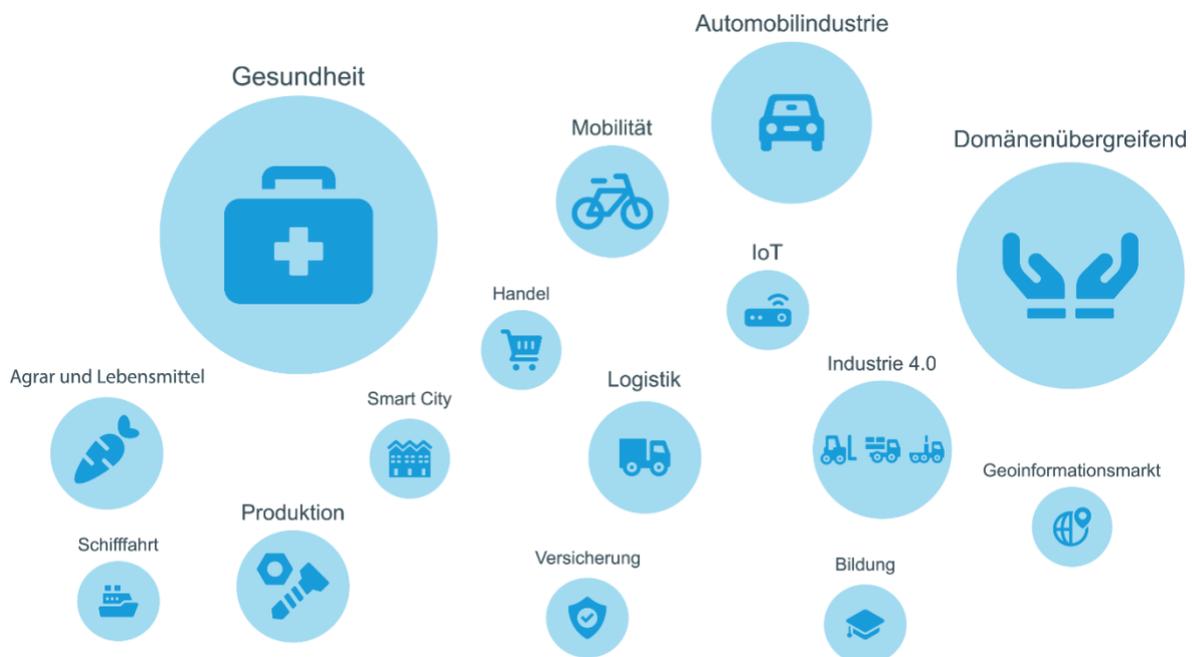


Abbildung 2: Verteilung der Branchen im Sample



Abbildung 3: Klassifikation von Data Sharing Ökosystemen

Forschungsprojekt

Zunächst ist auffällig, dass Data Sharing Ökosysteme zu einem gewissen Anteil (in der Stichprobe ca. 28 Prozent) im Rahmen von Forschungsprojekten, die von der EU-Kommission (Förderprogramme Horizon 2020 bzw. Horizon Europe) oder landesspezifischen Ministerien gefördert werden, entwickelt und erprobt werden. Dabei fällt auf, dass entsprechende Ökosysteme in der Regel auf den IDS-Standard als technische Basis setzen (siehe Absatz „Technische Basis“). Hier schließen sich gewinnorientierte / private und öffentliche / gemeinnützige Träger zu Konsortien zusammen, um gemeinsam an branchenspezifischen Data Sharing-Lösungen zu arbeiten. In der Regel haben diese Forschungs- und Entwicklungsvorhaben eine begrenzte Laufzeit (~ 3 Jahre) und fokussieren sich auf wenige Use Cases. Zum Abschluss entsteht oftmals ein Proof-of-Concept, aber keine fertiggestellte und im Anschluss aktive Data Sharing-Lösung. Nach Projektende lösen sich die Konsortien in der Regel wieder auf, auch wenn hier bisweilen einzelne Projekte den Übergang zu marktreifen Lösungen schaffen und sich als langfristige Lösungen etablieren können. Förderungen werden allerdings nicht nur für Forschungsprojekte, sondern z. B. auch zum Aufbau von Innovationsclustern vergeben. So wurden und werden die Data Sharing Ökosysteme „Catena-X“ und „Mobility Data Space“ im Rahmen der „Konzertierten Aktion Mobilität“⁵⁹ über mehrere Phasen hinweg durch das BMWK gefördert, um einen nachhaltigen Betrieb zu sichern.

⁵⁹ BMWK (2021)

Beispiel für Forschungsprojekte im Gesundheitssektor: „HEALTH-X dataLOFT“

Das Forschungsprojekt HEALTH-X dataLOFT unter der Leitung des Zentrums für Digitale Gesundheit der Charité - Universitätsmedizin Berlin zielt darauf ab, Bürger:innen ins Zentrum der Bereitstellung, Nutzung sowie Kontrolle der eigenen Gesundheitsdaten aus den primären und sekundären Gesundheitsmärkten zu stellen. Zahlreiche Partner aus den Bereichen Privatwirtschaft, Versorgung und Forschung arbeiten daran, ein Ökosystem von Gesundheitsanwendungen zu schaffen, welches über GAIA-X-Standards zugänglich sein soll. Gefördert wird das Projekt im Rahmen der Ausschreibung „Innovative und praxisnahe Anwendungen und Datenräume im digitalen Ökosystem GAIA-X“ durch das BMWK.

Beispiel für Forschungsprojekte in der verarbeitenden Industrie: „MARKET4.0“

MARKET 4.0 entwickelt eine offene digitale Peer-to-Peer-Plattform, die es Produktionsanlagen und Dienstleistern ermöglichen soll, sich mit produzierenden Unternehmen zu vernetzen und zusammenzuarbeiten. Die Lösungen werden im Kontext von drei verschiedenen Bereichen validiert: Metallverarbeitung, Kunststoffverarbeitung und Hightech-Produktionsanlagen. Beteiligt sind zahlreiche private und öffentliche Organisationen aus ganz Europa. Die Förderung des Projekts findet im Rahmen des Forschungs- und Innovationsprogramms Horizon 2020 der Europäischen Union statt.

Beobachtung: Der hohe Anteil an geförderten Forschungsprojekten lässt auf ein großes öffentliches Interesse an der Förderung des Datenaustausches zwischen Unternehmen schließen. Gerade die EU hat mit ihrer „European Data Strategy“ und entsprechenden Gesetzesvorschlägen wie dem „Data Act“ die Schaffung von interoperablen Datenräumen auf der Agenda stehen. Das Thema „Gesundheit“ gehört explizit zu den präferierten Branchen. In Deutschland ist vor allem das BMWK als fördernde Institution aktiv.

Branche

Bei der Branche und damit auch der Art der Daten, die in bestehenden Data Sharing Ökosystemen geteilt werden, lassen sich spezifische und domänenübergreifende Beispiele finden. In der Regel haben diese einen Fokus auf bestimmte Branchen wie Industrie 4.0, Gesundheit, Mobilität, Handel etc. Hierzu zählen bspw. der „Mobility Data Space“ (Mobilität), das Smart „Smart Connected Supplier Network – Market 4.0“ (Industrie 4.0) oder „teamply digital health platform connect“ (Gesundheit). Es gibt aber auch universelle Plattformen, die offen für jegliche Art von Daten sind, wie der „Data Intelligence Hub“ oder „ADVANE0“. Diese domänenübergreifenden Data Sharing Ökosysteme fungieren oft als Datenmarktplätze, bei denen die Teilnehmer über Art und Preis der Daten bestimmen.

Beispiel für Gesundheitsökosysteme: „teampay digital health platform connect“

teampay digital health platform connect ist eine offene und interoperable Lösung als „Platform-as-a-Service“-Angebot, mit der sich die Akteure im Gesundheitswesen vernetzen können. Das offene Konzept ermöglicht den Leistungserbringern über eine einzige Anbindung ein breites Spektrum digitaler Aktensysteme zu nutzen sowie am Datenaustausch zwischen Leistungserbringern in instituts- und sektorübergreifenden Behandlungsszenarien teilzunehmen. Dabei kommen international etablierte Standards (z. B. IHE, HL7 FHIR, HL7 v3 / v3 / CDA) zum Einsatz. Durch die Plattform sollen Leistungserbringer zukünftig zudem nicht nur von erfolgreichen Lösungen profitieren können, sondern auch von Angeboten Dritter, sogenannten Mehrwertdienste.

Beispiel für domänenübergreifende Ökosysteme: „ADVANEO Data Marketplace“

Der ADVANEO Data Marketplace ist ein Kollaborationsportal, um Daten branchenübergreifend zwischen Unternehmen auszutauschen und so datengetriebene Anwendungen entwickeln zu können. Durch den Rückgriff auf den IDS-Standard bleibt die Datensouveränität der beteiligten Unternehmen erhalten: ADVANEO verwaltet lediglich Metadaten, die Rohdaten verbleiben bei den Datenanbieter und werden im Falle einer Transaktion direkt an den Datenkonsumenten übertragen. Über den reinen Datenaustausch hinaus bietet ADVANEO zudem integrierte KI-Werkzeuge, Datenmodelle und Applikationen sowie den freien Zugang zu offenen Daten (Open Data).

Beobachtung: Branchenspezifische Data Sharing Ökosysteme sind in der Überzahl (ca. 80 Prozent). Dafür kann es eine Vielzahl von Gründen geben. Primär ist hier zu nennen, dass sie relevante Stakeholder bündeln und so ein qualitativ hochwertiges Angebot von Daten innerhalb ihrer spezifischen Domäne schaffen können. Außerdem verfügt jede Branche über industriespezifische Hürden, Regulierungen und weiteren Rahmenbedingungen, denen so gezielt Rechnung getragen werden kann.

Organisationsform

Die Betreiber bzw. Initiatoren bestehender Data Sharing Ökosysteme können in drei Kategorien unterschieden werden. Der Großteil der betrachteten Beispiele wird von gewinnorientierten Unternehmen (ca. 57 Prozent) verantwortet. Unter den Beispielen finden sich aber auch vereinzelt Non-Profit-Organisationen (NPO), die als Betreiber bzw. Initiatoren agieren (ca. 10 Prozent). Diese werden – wie im Fall des Mobility Data Space – von kooperierenden Unternehmen gegründet, um den Datenaustausch unter neutraler Kontrolle zu halten. Die NPO dient zudem dazu, die Kollaboration vieler unterschiedlicher Organisationen auf Governance-Ebene zu orchestrieren. Das Ziel ist hier die Ermöglichung eines vorwettbewerblichen Austauschs von Daten zwischen den teilnehmenden Unternehmen. In einem der untersuchten Beispiele wurde zusätzlich ein Verein gegründet, der den Mitgliedern die Möglichkeit zur Mitbestimmung z. B. in Standardisierungsfragen gibt (siehe Beispiel „Catena-X Automotive Network e.V.“).

Neben profit- und nicht-profit-orientierten Unternehmen gibt es vor allem im Rahmen von geförderten Forschungsprojekten gemischte Konsortien, die aus Vertretern beider Kategorien bestehen (können) und sich für die Laufzeit eines Projekts zusammenschließen. Die finanzielle und organisationale Eigenständigkeit der Unternehmen bleibt bestehen. Ein Projektpartner übernimmt dabei die Konsortialführung und tritt nach außen im Namen des Konsortiums auf. Nach Ende der Projektlaufzeit endet in der Regel auch die Zusammenarbeit und das Konsortium wird aufgelöst.

Beispiel für Non-Profit-Organisationsform: „Catena-X Automotive Network e.V.“

Die ursprünglich als „Automotive Alliance“ gestartete Initiative „Catena-X“ ist ein Partnernetzwerk aus Stakeholdern der deutschen Automobilbranche. Seit Mai 2021 werden die Aktivitäten rund um den Aufbau eines Datenökosystems für alle Teilnehmer der automobilen Wertschöpfungskette in dem Verein „Catena-X Automotive Network e.V.“ gebündelt. Dieser umfasst 84 Mitglieder, u. a. mit Branchengrößen wie BMW, Daimler, Volkswagen, Bosch, SAP und Siemens (Stand: 15.02.2022).⁶⁰ Das gemeinsame Ziel ist, eine Umgebung für den Aufbau, den Betrieb und die kollaborative Nutzung durchgängiger Datenketten entlang der gesamten automobilen Wertschöpfungskette zu schaffen. Die Grundsätze zur Datensouveränität, -sicherheit und -interoperabilität werden dabei auf Basis von Gaia-X und Konzepten der IDSA umgesetzt. Für die Zukunft ist die Gründung einer eigenen Betreibergesellschaft vorgesehen.

Beobachtung: Wenn es um einen gemeinschaftlichen Aufbau einer Infrastruktur für den Austausch von Daten geht, sind Zusammenschlüsse mehrerer Unternehmen zu beobachten. Die bürokratischen und regulatorischen Prozesse für die Gründung einer entsprechenden Rechtseinheit haben dabei unterschiedliche Ausprägungen, sodass zum Teil auch mehrere Rechtseinheiten mit unterschiedlichen Rollen gegründet werden (siehe Beispiel „Catena-X Automotive Network e.V.“). Die unterschiedlichen Ansätze der Ökosysteme im Sample zeigen, dass es aktuell noch keine klare „Blaupause“ dafür gibt, wie eine entsprechende Rechtseinheit bzw. Kollaborationsstruktur aussehen sollte.

Zielsetzung

Allgemein kann festgehalten werden, dass Data Sharing Ökosysteme die Vernetzung bestimmter Akteure, entweder einer bestimmten Branche oder über Sektorengrenzen hinweg, fördern wollen. Die konkrete Zielsetzung ist dann eng mit der jeweiligen Organisationsform verbunden. Einzelne gewinnorientierte Betreiber wie bspw. Amazon Web Services geht es um den Aufbau von Geschäftsfeldern und damit um die Monetarisierung von Data Sharing-Dienstleistungen.

Data Sharing Ökosysteme, die von Non-Profit-Organisationen geführt werden, verfolgen in der Regel das Ziel einer vorwettbewerblichen Verfügbarmachung von Daten bspw. einer bestimmten Branche. Auf diese Weise soll eine gemeinsame Datenbasis geschaffen werden, von der alle Teilnehmer profitieren. Der Wettbewerb erfolgt dann bspw. auf der Ebene der beim jeweiligen Marktteilnehmer entwickelten Anwendungen.

Forschungskonsortien haben in der Regel durch die jeweilige Ausschreibung festgelegte Zielsetzungen. Meist geht es um die Entwicklung und Erprobung von bestimmten Konzepten, aber auch die Verwertung der während der Projektlaufzeit erarbeiteten Ergebnisse in weiterführenden Vorhaben wird angestrebt.

Mobility Data Space

Der Mobility Data Space versteht sich als Data Sharing Community für Mobilitätsanbieter: Fahrzeughersteller, Mitfahrdienste, Betreiber des öffentlichen Verkehrs, Softwarehersteller für Navigationssysteme, Forschungseinrichtungen, Bike-Sharing-Anbieter etc. Sie soll den Wettbewerb um innovative, umweltfreundliche und nutzerfreundliche Mobilitätskonzepte anreizen, indem sie allen Nutzern gleichberechtigt und transparent Zugang zu Daten verschafft. An der Konzeption haben mehr als 200 Stakeholder der deutschen Mobilitätslandschaft aus Wissenschaft, Wirtschaft und öffentlicher Verwaltung mitgewirkt. Als Trägergesellschaft fungiert die Non-Profit-Organisation „DRM Datenraum Mobilität GmbH“.

⁶⁰ Catena-X (2022)

AVIATION DataHub

Die Lufthansa gründete 2019 den AVIATION DataHub als neutrale Plattform für den Austausch von luftfahrtbezogenen Daten. Damit richtet er sich an Fluggesellschaften, Unternehmen für Wartung, Reparatur und Betrieb, OEMs, Datenanbieter sowie Anwendungsentwickler. Ziel ist es, die Luftfahrtbranche und hier insbesondere die Flugzeugwartung zu digitalisieren und den Austausch von Daten ohne Restriktionen gegenüber Wettbewerbern zu ermöglichen. Dabei sollen auch die verwandten Bereiche „Flugbetrieb“ und „Bodenabfertigung“ einbezogen werden. Die Teilnehmer erhalten volle Kontrolle darüber, welche Daten sie mit wem teilen möchten. Während Lufthansa Technik derzeit den Hauptanteil an dem Unternehmen hält, sollen in naher Zukunft weitere Gesellschafter und strategische Partner aus der Luftfahrtindustrie hinzukommen.

Beobachtung: Bei der Verfügbarmachung von Daten haben Data Sharing Ökosysteme nicht nur kurzfristige Gewinne, sondern auch die Verbesserung der gesellschaftlichen Rahmenbedingungen zum Ziel. Damit wird in der Regel einem öffentlichen Interesse Rechnung getragen. Bestehende Prozesse einzelner Branchen sollen so bspw. nachhaltiger gestaltet werden, um etwa Mobilität für eine klimafreundlichere Zukunft anzupassen oder eine nachfrageorientierte Produktion und ein gesteigertes Bewusstsein für Lebensmittel und deren Wert zu schaffen. Bei Ökosystemen aus dem Gesundheitssektor steht aktuell vor allem die Optimierung der medizinischen Forschung durch Vereinfachung des Zugangs zu klinischen Daten im Mittelpunkt. Dies betrifft bspw. die Bekämpfung von Volkskrankheiten wie Diabetes mellitus Typ 2, Herz-Kreislauf-Erkrankungen oder Krebs. Data Sharing Ökosysteme aus dem Gesundheitssektor, die als Forschungsprojekte klassifiziert werden können, zielen hingegen vor allem auf den medizinischen Fortschritt hin zu einer personalisierten Gesundheitsversorgung ab.

Beteiligte Stakeholder

Die Beispiele bestehender Data Sharing Ökosysteme zeigen, dass an diesen sowohl private als auch öffentliche Stakeholder beteiligt sind. Private Betreiber sind am häufigsten vertreten (64 Prozent) und bauen Data Sharing-Lösungen auf, um mit Geschäftsmodellen wie Pay-per-Use, Vermittlungsgebühren oder Abonnements Umsätze zu erzielen. Bei den wenigen öffentlichen Stakeholdern (knapp 7 Prozent) wird in der Regel Open Data zur kostenfreien Nutzung zur Verfügung gestellt. Hier steht der Datenaustausch im Vordergrund. Gerade im Kontext von Forschungsprojekten lassen sich auch Kooperationen zwischen dem privaten und öffentlichen Sektor feststellen. Solche öffentlich-privaten Partnerschaften (eng.: Public-Private-Partnership, kurz PPP) zeigen das Interesse von öffentlicher Seite an der Etablierung entsprechender Strukturen.

Beispiel für öffentliche Stakeholder: „Mobilitäts Daten Marktplatz“

Der „Mobilitäts Daten Marktplatz“ ist seit 2017 Deutschlands nationaler Zugangspunkt für Verkehrsdaten nach Vorgabe der IVS-Direktive der EU. Er wird von der Bundesanstalt für Straßenwesen betrieben und bündelt Mobilitätsdaten über verschiedene Verkehrsmittel, Netzelemente und Akteure hinweg. Dazu werden u. a. Daten von weiteren deutschen Open-Data-Portalen integriert. Straßenbetreiber sowie Unternehmen und Forschungseinrichtungen können Daten wie z. B. Echtzeit-Verkehrsdaten, Fahrpreis- und Vertriebsinformationen oder Informationen über Parkplätze und Rastanlagen anbieten, recherchieren und abonnieren.

Beispiel für öffentlich-private Partnerschaften: PAIRS

Das Forschungsprojekt „Privacy-Aware, intelligent and Resilient Crisis Management“ (kurz: PAIRS) wird im Rahmen des „Innovationswettbewerb Künstliche Intelligenz“ des BMWK gefördert. Das Konsortium besteht neben dem Initiator „Advaneo GmbH“ u. a. aus weiteren Unternehmen wie IBM Deutschland GmbH und der SICK AG, es sind aber auch öffentliche Stakeholder wie die Bundesanstalt Technisches Hilfswerk (THW) oder die Universität des Saarlandes beteiligt. Gemeinsam soll eine serviceorientierte, offene Dateninfrastruktur entwickelt werden, mit der die Auswirkungen von Krisensituationen prognostiziert werden können. Die KI-Hybrid-Technologie von PAIRS sollen in einem domänenübergreifenden Datenraum sowohl das initiale Krisenereignis als auch die Reaktionen verschiedener Akteure antizipieren, um auf dieser Basis zielgerichtete Handlungsempfehlungen zu generieren.

Beobachtung: Abgesehen von den wenigen rein öffentlichen Ökosystemen spielen vor allem zwei Arten von Ökosystemen eine Rolle: (1) Vollständig privatwirtschaftliche Ökosysteme und (2) solche, die einen kollaborativen Ansatz wählen und auf Privat Public Partnerships (PPP) setzen. Während private Ökosysteme ein Geschäftsmodell darin sehen, Datenaustausch zu erleichtern und sich so zukünftige Geschäftsmöglichkeiten sichern möchten, liegt der Fokus bei PPP eher auf den Anwendungsfällen und den Mehrwerten, die der Austausch von Daten ermöglicht. Für PPPs geht es vor allem um den harmonisierten Aufbau einer Infrastruktur, die von jedem teilnehmenden Unternehmen dann für unterschiedliche Zwecke genutzt werden kann. Diese Infrastruktur bedient auch ein öffentliches Interesse, etwa an der Verbesserung der Versorgung im Gesundheitswesen. Die Involvierung von öffentlichen Stakeholdern, die sich ohnehin bereits mit dieser Thematik auseinandersetzen, ist daher naheliegend. Die häufige Involvierung von öffentlichen Stakeholdern lässt zudem darauf schließen, dass Data Sharing Ökosysteme politisch aktiv gewollt sind und entsprechend auch gefördert werden.

Use Case

Die betrachteten Data Sharing Ökosysteme unterscheiden sich darin, ob sie eine begrenzte Anzahl spezifischer Use Cases fokussieren oder Raum für multiple Use Cases schaffen. Spezifische Use Cases sind seltener (bei ca. 21 Prozent der Ökosysteme) und werden vor allem in Forschungsprojekten behandelt, was der kurzen Laufzeit der Projekte sowie der limitierten Fördersumme geschuldet ist. Zudem gehört die Erprobung von erarbeiteten Konzepten anhand weniger spezieller Use Cases in der Regel zu den Förderbedingungen der jeweiligen Ausschreibung. Kommerzielle Data Sharing Ökosysteme, die von privaten Betreibern initiiert werden, sind hingegen offener gegenüber möglicher Use Cases, da hier vor allem die Infrastruktur für das Teilen von Daten zur Verfügung gestellt wird. Für welche konkreten Projekte die teilnehmenden Unternehmen diese Infrastruktur nutzen, bleibt ihnen überlassen.

Beispiel für spezielle Use Cases: TEAM-X – Trusted Ecosystem of Applied Medical Data eXchange

Mit Trusted Ecosystem of Applied Medical Data eXchange sollen die Domänen Gesundheit, Pflege und Smart Living auf Basis des GAIA-X-Frameworks sinnvoll zusammengeführt werden. In diesem Rahmen werden Ansätze erprobt, um die Kompetenz der Bürger:innen im Umgang mit ihren Daten zu erhöhen und die Datensouveränität zu fördern. Das Konzept von TEAM X wird anhand der Use Cases Frauengesundheit (Onkologie, Schwangerschaft) und digitale Pflegeplattformen (ambulant und stationär) evaluiert. Mit Blick auf die besonders schutzwürdigen Gesundheitsdaten sollen Weiterentwicklungsbedarfe im regulatorischen Rahmenwerk herausgearbeitet werden.

Beispiel für multiple Use Cases: Vendia Share Platform

Vendia ist ein US-amerikanisches Startup, das 2020 gegründet wurde. Mit der Vendia Share Plattform wird eine serverlose Technologie angeboten, die Blockchain als sichere Basis für den Austausch von Daten und die Zusammenarbeit von Unternehmen nutzt. Die „Software-as-a-Service“ (SaaS)-Lösung ist u. a. in der Lage, Daten von unterschiedlichen Cloud-Anbietern miteinander zu verbinden. Vendia unterstützt unterschiedliche Use Cases für Unternehmen, die sensible Daten austauschen müssen, sei es von der Bestandsverfolgung über die gesamte Lieferkette bis hin zur Abwicklung von Transaktionen mit mehreren Parteien.

Beobachtung: Es lassen sich zwei Herangehensweisen für den Aufbau von Data Sharing Ökosystemen beobachten: Entweder es wird eine Infrastruktur aufgebaut, um diese als Service anzubieten (multiple Use Cases) oder es wird zunächst ein bestimmtes Problem identifiziert und in einem oder wenigen spezifischen Use Cases bearbeitet. Beide Herangehensweisen haben jeweils eigene Vor- und Nachteile. So bringt der Aufbau einer reinen Infrastruktur ein hohes Maß an Skalierbarkeit, lässt allerdings einen Fokus und damit ein Alleinstellungsmerkmal vermissen. Bei spezifischen Use Cases verhält es sich hingegen genau andersherum, denn hier wird ein bestimmtes Thema als Alleinstellungsmerkmal fokussiert, welches schnelle praktische Ergebnisse liefert, das jedoch zunächst nicht oder nur schwer skalierbar ist.

Zugang

Wie die Analyse von bestehenden Data Sharing Ökosystemen zeigt, kann der Zugang zum entsprechenden Angebot offen, halboffen oder geschlossen sein. Offene Lösungen erlauben die Registrierung jeglicher Teilnehmer und setzen keine Grenzen. Hierdurch wird die Kontrolle sowohl der Identität des Teilnehmers als auch der Qualität der bereitgestellten Daten erschwert. Bei halboffenen Angeboten wird eine Auswahl der berechtigten Teilnehmer durch den jeweiligen Betreiber getroffen. Neue Teilnehmer müssen bestimmte Bedingungen erfüllen, um partizipieren zu dürfen. Dies können etwa die Zertifizierung durch eine Zertifizierungsstelle oder eine Vereinsmitgliedschaft sein. Bei geschlossenen Data Sharing-Lösungen bestimmt der Betreiber, welche Teilnehmer zugelassen werden. Dies kann unterschiedliche Gründe haben, etwa um sicherzustellen, dass nur qualitativ hochwertige Daten getauscht oder dass nur Daten aus bestimmten Branchen gehandelt werden. Unter den untersuchten Data Sharing Ökosystemen finden sich Data Sharing-Lösungen, die aufgrund ihres Status als Forschungsprojekte als geschlossene Lösungen bezeichnet werden können. Da hier vor allem Proof-of-Concepts entwickelt werden, ist der Austausch von Daten zunächst auf das jeweilige Konsortium und evtl. assoziierte Partner beschränkt.

Beispiel für offene Ökosysteme: City Dataspace

Der City Dataspace ist eine Plattform, die es unterschiedlichen Stakeholdern ermöglicht, relevante Daten für die Entwicklung zu einer Smart City (Open & Urban Data) zu veröffentlichen, zu finden und zu nutzen. Alle Daten, die im Kontext von Smart City interessant sein können, werden dabei unterstützt. Jeder kann sich beteiligen und Daten hochladen: Von der Kommune über Unternehmen bis hin zu Bürger:innen als Privatpersonen. Der City Dataspace wird im Projekt Bergisch.Smart_Mobility unter der Leitung des Lehrstuhls für Technologien und Management der Digitalen Transformation zusammen mit den Städten Wuppertal, Solingen und Remscheid entwickelt.

Beispiel für halboffene Ökosysteme: Dawex

Mit dem Global Data Marketplace und der Data Exchange Platform bietet das französische Unternehmen Dawex Systems SAS zwei Lösungen für den Austausch von Daten an. Während Erstere als offener Marktplatz gestaltet ist, auf dem alle Arten von Daten gehandelt werden können, bietet Letztere die Möglichkeit, den Austausch und die Auswertung von Daten intern und innerhalb bestimmter Gruppen durchzuführen. Hier können Unternehmen den Zugang bspw. auf bestimmte Kunden, Tochtergesellschaften oder Lieferanten einschränken. Auch die Teilnahme am Marktplatz unterliegt einer Bedingung: Die Identität aller auf dem Marktplatz agierenden Stakeholder muss zunächst durch Dawex validiert werden.

Beobachtung: Alle Formen des Zugangs bieten gewisse Vorteile. Offene Data Sharing-Lösungen können schneller skalieren, da sie in kürzester Zeit eine Vielzahl von neuen Teilnehmern vernetzen können. Die damit einhergehende Vielfalt an Datenquellen ermöglicht wiederum eine große Anzahl von Use Cases. Im Gegensatz dazu haben halboffene und geschlossene Angebote Vorteile bei der Sicherstellung der Datenqualität und der Vertrauenswürdigkeit der Teilnehmer, da hier selektiert wird.

Architektur

Die bestehenden Data Sharing Ökosysteme lassen sich in Bezug auf ihre Architektur nach einer zentralen oder dezentralen Ausrichtung unterscheiden. Beides ist in etwa gleich oft vertreten, aber stellenweise gibt es auch hybride Lösungen, die beide Ansätze kombinieren. Bei zentral aufgebauten Angeboten werden die zu tauschenden Daten an einer Stelle gesammelt (durch z. B. den Upload in die Cloud), um von dort aus zur Verfügung gestellt zu werden. Dies ist einfacher für die Teilnehmer, da keine besonderen technischen Voraussetzungen benötigt werden. Gleichzeitig muss dem Betreiber vertraut werden, dass dieser die Daten ausschließlich zum Zwecke des Datenaustausches verarbeitet. Zudem wird die Kontrolle über die Daten abgegeben, sodass zentrale Architekturen weniger Datensouveränität für die Teilnehmer bieten. Anders ist es bei dezentral aufgebauten Lösungen. Hier verbleiben die Daten bei den einzelnen Teilnehmern und werden erst bei Bedarf direkt an den Empfänger der Daten übertragen. Auf diese Weise bleibt die Kontrolle über die Daten beim jeweiligen Teilnehmer. Dazu benötigen alle Teilnehmer eine entsprechende technische Komponente (bspw. den IDS Konnektor, siehe Kapitel 2.3.1), deren Einrichtung und Nutzung mit mehr Aufwand verbunden sind als ein einfacher Datenupload. Bei hybriden Data Sharing-Lösungen liegen die Daten bei den einzelnen Teilnehmern und werden von dort aus mit anderen geteilt, aber einzelne Services werden über eine zentrale Infrastruktur zur Verfügung gestellt. Dies kann die Transformation von Daten oder die Anbahnung des Datenaustausches umfassen.

Beispiel für eine zentrale Architektur: BMW CarData

CarData der BMW Group ist eine Plattform für Fahrzeugtelematikdaten der Marken BMW und MINI, die sich an Unternehmen wie Versicherungen, Werkstätten oder Mobilitätsdienstleister richtet. Die Telematikdaten werden dazu zentral auf BMW-Servern gesammelt und können von dort aus abgerufen werden – vorausgesetzt, die das Fahrzeug führende Person stimmt der Weitergabe der Daten zu. Die Kosten für die Datenübertragung werden über ein Pay-per-Use-Modell abgerechnet.

Beispiel für eine dezentrale Architektur: Datapace

Datapace ist ein dezentraler, Blockchain-basierter Datenmarktplatz für die sichere und automatisierte Monetarisierung und den Austausch von IoT-Sensoren und digitalen Assets mit technischer und richtlinienbasierter Datenverifizierung und dem Zugang zu einer weltweiten Netzwerkinfrastruktur von Sensoren. Die Data Sharing-Lösung kann zum Streamen beliebiger Daten aus unterschiedlichen Quellen wie IoT-Geräten, physischen Assets, autonom fahrenden Autos, Drohnen und vielen mehr verwendet werden. Es ermöglicht, Daten von Drittanbietern zu integrieren und sie über den Marktplatz zu monetarisieren.

Beobachtung: Wird die verwendete Architektur nach Branche betrachtet, sind dezentrale oder hybride Ansätze bei Data Sharing Ökosystemen aus dem Gesundheitssektor mit ca. 64 Prozent in der Überzahl. Dies lässt sich insbesondere mit den besonders schützenswerten Daten im Gesundheitssektor erklären. Gerade weil Unternehmen ohnehin Bedenken beim Zugang auf Daten durch Dritte haben, sind vor allem dezentrale oder hybride Architekturen geeignet, da bei diesen Vertrauen und Datensouveränität systemimmanente Eigenschaften sind. Hier verbleiben die Daten bei den einzelnen Teilnehmern und es wird keine Kontrolle über die Daten abgegeben.

Technische Basis

Die betrachteten Data Sharing Ökosysteme bauen auf unterschiedlichen Standards für den Datenaustausch auf. Es lässt sich eine Vielzahl an Projekten mit IDS und / oder GAIA X Basis finden, die nach den in Kapitel 2.3 beschriebenen Prinzipien aufgebaut sind. Entsprechend müssen die Teilnehmer solcher „Datenräume“ zertifiziert sein und über einen Konnektor verfügen, der zur Authentifizierung der Teilnehmer und zur Durchsetzung von Nutzungsrichtlinien dient. Zudem existieren ein Identitätsmanagement und Zertifizierungsstellen. Daneben gibt es auch Lösungen, die auf Blockchain-Technologie zur Datenfreigabe und Transaktionsabwicklung setzen. Blockchain-Technologie kommt auch für den Abschluss von Lizenzverträgen zwischen Data Provider und Data Consumer zum Einsatz. Beide Standards ermöglichen eine dezentrale oder hybride Datenhaltung (siehe Absatz „Architektur“). Neben der Nutzung dieser offenen Standards kommen proprietäre Systeme zum Einsatz, deren Aufbau von außen nicht immer ersichtlich ist. Hier wird meist ein zentraler Datenhaltungsansatz gewählt, sodass die Abhängigkeit vom Betreiber der Data Sharing-Lösung groß ist.

Beispiel für Blockchain: Data Broker Global

Databroker ist ein Blockchain-basierter Peer-to-Peer-Marktplatz für Daten aus Belgien, der seit 2020 vom Unternehmen SettleMint kommerziell betrieben wird. Die gekauften Daten werden direkt zwischen den beiden Parteien übertragen, ohne dass die Databroker-Plattform Zugriff auf die Daten erlangt. Aktuell werden Daten aus den Bereichen Geografie, Umwelt, Menschen, Landwirtschaft, Energie, Verkehr, Wirtschaft und Lieferkette gehandelt. Zudem wird das System als White-Label-Produkt einer „Platform-as-a-service“ (PaaS)-Lösung angeboten, sodass Unternehmen mit großen Datenbeständen eigene Datenaustauschplattformen betreiben können.

Beispiel für IDS: Data Intelligence Hub

Der Data Intelligence Hub der Deutschen Telekom (DIH) ist eine Plattform für den sicheren Datenaustausch zwischen Unternehmen. Er verwendet die Komponenten IDS Konnektor, IDS Broker und IDS App Store der IDS-Referenzarchitektur der International Data Spaces Association, um Teilnehmer:innen zu verifizieren und den direkten Austausch von Daten zu ermöglichen. Die Plattform bietet den Nutzer:innen neben der Erfassung, dem Austausch und der Verarbeitung von Daten auch Analysewerkzeuge an. Der Zugang zum DIH ist kostenlos, es werden aber Gebühren für optionale Leistungen wie Speicherung, Rechenleistung und Transaktionen erhoben.

Beobachtung: Die Merkmale „technische Basis“ und „Architektur“ haben viele Überschneidungen. Wird eine dezentrale oder hybride Architektur gewählt, wird in der Regel auf offene Standards wie IDS und Blockchain zurückgegriffen. Mit 44 Prozent liegen IDS-basierte Lösungen hier vor Blockchain-Lösungen, die 33 Prozent der Data Sharing Ökosysteme mit dezentraler oder hybrider Architektur ausmachen. Beide ermöglichen einen nachvollziehbaren Datenaustausch, der nach klaren Regeln strukturiert ist. Zudem bieten sie eine potenzielle Interoperabilität zu anderen Data Sharing Ökosystemen.

2.2.3 Typisierung von Data Sharing Ökosystemen

Die Betrachtung der Merkmale aktueller Data Sharing Ökosysteme offenbart eine Reihe an Überschneidungen und Abhängigkeiten, anhand derer sich zwei zentrale Typen von Ökosystemen für das Teilen von Daten identifizieren lassen: Datenplattformen und Datenräume (siehe Abbildung 4). Die zwei identifizierten Typen können vor allem anhand der Merkmale „Organisationsform“, „Architektur“ und „technische Basis“ voneinander differenziert werden. Zudem lassen sich Auswirkungen auf die Zielsetzung und die beteiligten Stakeholder festmachen. Dabei ist festzuhalten, dass manche Beispiele Merkmale beider Typen vereinen und daher als Hybriden anzusehen sind.

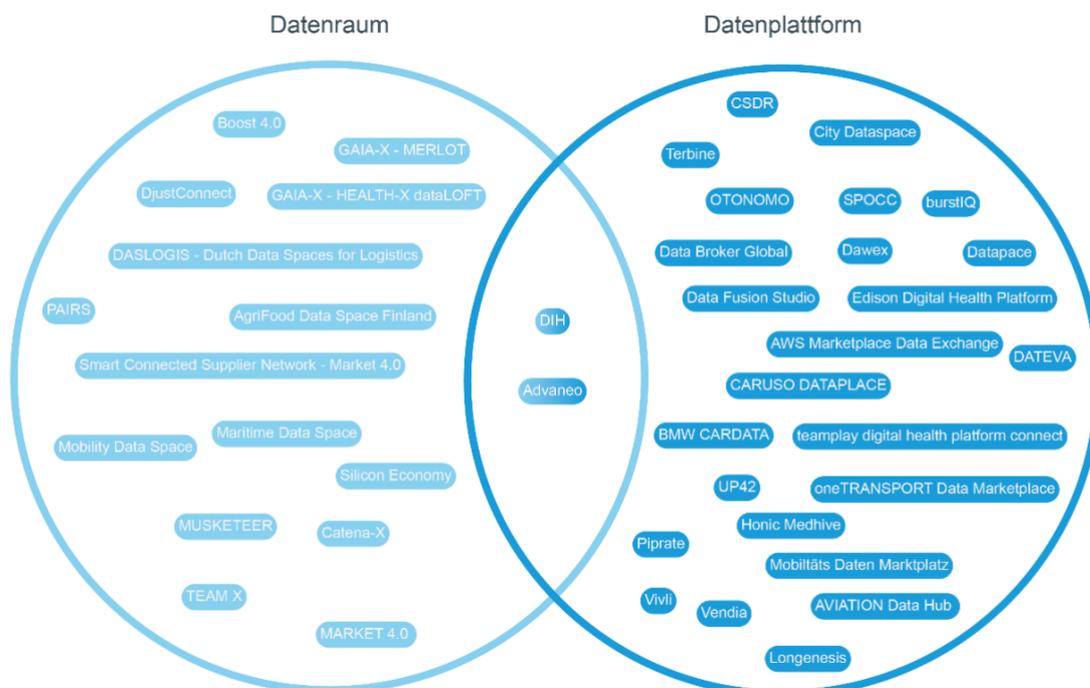


Abbildung 4: Datenraum vs. Datenplattform

Durch die Differenzierung innerhalb der Data Sharing Ökosysteme ergibt sich ein detaillierteres Bild des Verhältnisses von standardgebenden Framework Initiativen und umsetzenden Ökosystemen (siehe Abbildung 5).

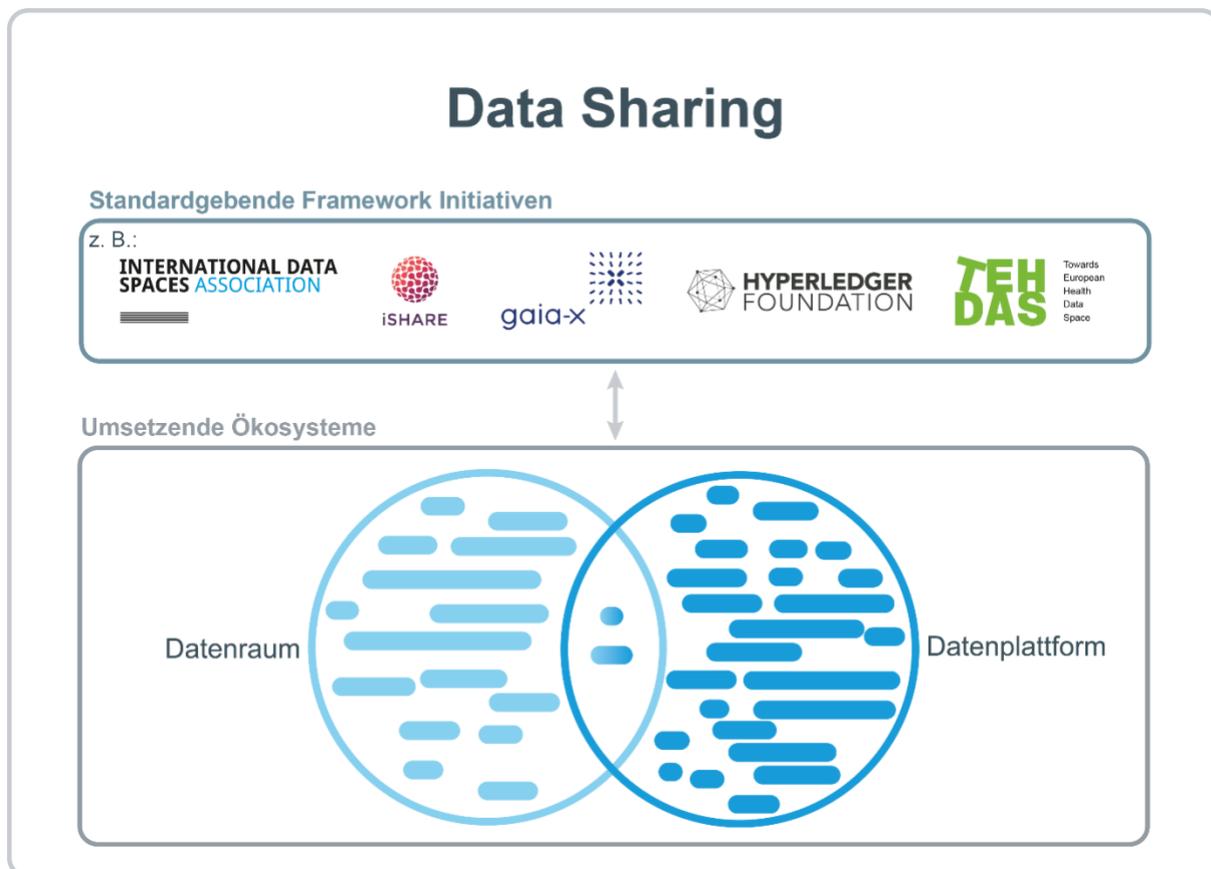


Abbildung 5: Typisierung von Data Sharing Ökosystemen im Verhältnis zu Framework Initiativen

Im Folgenden sollen beide Typen von Data Sharing Ökosystemen genauer beschrieben werden, um die jeweiligen Unterschiede und Potenziale für das Teilen von Daten aufzuzeigen. Zunächst liegt der Fokus auf Datenplattformen, da diese am häufigsten zu beobachten sind.

Datenplattformen

Datenplattformen machen mit ca. 60 Prozent den Großteil des untersuchten Samples von Data Sharing Ökosystemen aus. Unternehmen müssen sich auf diesen Plattformen als Teilnehmer registrieren, um anschließend als Datengeber und/oder Datenkonsument agieren zu können. Der Zugang ist bei der Mehrheit (ca. 70 Prozent) der untersuchten Datenplattformen offen gestaltet, sodass keine Zugangsvoraussetzungen gefordert werden.

Aus technologischer Sicht handelt es sich bei Datenplattformen um Cloudangebote mit zentraler Architektur. Das bedeutet, dass Daten auf Servern des Plattformbetreibers gespeichert werden. Unternehmen, die über Datenplattformen Daten teilen möchten, sind so gezwungen, die Daten auf die entsprechenden Server hochzuladen. Dem Betreiber der jeweiligen Datenplattform muss dabei vertraut werden, da dieser die Daten nicht nur speichert, sondern auch verwaltet. Als technische Basis dienen vor allem proprietäre Lösungen; nur vier Datenplattformen im Sample setzen auf Komponenten des IDS-Standards, sieben auf eine Blockchain-Lösung für einzelne Elemente wie die Generierung von Lizenzverträgen. Zudem ist der Datenaustausch nur zwischen den registrierten Mitgliedern der Plattform möglich, sodass Unternehmen gegebenenfalls auf mehreren Datenplattformen registriert sein müssen, um Daten mit allen für sie relevanten Partnern austauschen zu können.

Datenplattformen verbindet, dass sie in den meisten Fällen jeweils von einzelnen Unternehmen und nicht von Konsortien bzw. kooperierenden Partnern betrieben werden. Bei lediglich 16 Prozent der Beispiele verhält es sich anders. Die Datenaustauschplattform für klinische Studien „Vivli“ wird von einer Non-Profit-Organisation betrieben, die sich aus Mitgliedsbeiträgen der institutionellen Mitglieder und Stiftungen finanziert. Die ebenfalls auf klinische Daten spezialisierte Plattform „Clinical Study Data Request“ bündelt das Angebot von sog. „Datensponsoren“ aus der Pharmabranche. Daneben entstand die „teampay digital health platform connect“ von Siemens Healthineers in Kooperation mit IBM. Zu guter Letzt stehen hinter der deutschen Gesundheitsdatenplattform „honic“ neben der verantwortlichen Health Data Technologies GmbH weitere Partner für die Umsetzung der Cloud-Infrastruktur, der Rolle des Datentreuhänders und der Softwareentwicklung.

Zu den primären Zielen von Datenplattformen zählen sowohl die Schaffung von Netzwerken für das Teilen von Daten als auch die Monetarisierung eben dieser Dienstleistung. Da es sich bei den untersuchten Datenplattformen fast ausschließlich um kommerzielle Angebote handelt, werden hier verschiedene Geschäftsmodelle wie Freemium, Pay-per-Use- oder Mitgliedschaftsmodelle umgesetzt. So erhalten registrierte Kund:innen bspw. im Gegenzug für eine jährliche Mitgliedschaft Zugang zu den Daten auf der Datenplattform.

Neben bekannten Größen wie Amazon Web Services, Siemens Healthineers oder BMW findet sich eine Vielzahl kleinerer Player, die Datenplattformen für bestimmten Branchen anbieten. Insbesondere im Gesundheitssektor dominieren dabei US-amerikanische Anbieter (5 von 8), was auf die libertär-ökonomische Gestaltung des US-amerikanischen Gesundheitswesens zurückgeführt werden kann. Wie bereits in Kapitel 2.1.4 deutlich wurde, sind die Bedingungen im deutschen Gesundheitswesen erheblich komplexer. Abgesehen von privaten Anbietern gibt es im Sample mit dem „Mobilitäts Daten Marktplatz“ der Bundesanstalt für Straßenwesen und dem „City Dataspace“ der Bergischen Universität Wuppertal auch zwei Datenplattformen, die von öffentlicher Hand geführt werden. Weitere Beteiligungen von oder Kooperationen zwischen öffentlichen und privaten Stakeholdern konnten für Datenplattformen nicht identifiziert werden.

Neben den bereits erwähnten branchenspezifischen Datenplattformen ist ein gewisser Teil (ca. 24 Prozent) der Beispiele aus der untersuchten Stichprobe domänenübergreifend aufgebaut, sodass jegliche Art von Daten geteilt werden kann. Dies gilt vor allem, wenn sie wie die Beispiele „Data Broker Global“ oder „AWS Data Exchange“ als Marktplätze konzipiert wurden, auf denen „die Preisfindung für die verfügbaren Datensätze direkt durch Angebot und Nachfrage der Marktteilnehmenden statt[findet] und der Plattformbetreiber [...] nicht involviert [ist]“.⁶¹

Der Reifegrad der untersuchten Datenplattformen ist hoch. Fast alle sind bereits aktiv und können von Unternehmen genutzt werden. Geförderte Forschungsprojekte befinden sich nicht unter den als Datenplattform typisierten Data Sharing Ökosystemen.

Datenräume

Der Anteil von Data Sharing Ökosystemen, die sich als Datenraum bezeichnen lassen, liegt im vorhandenen Sample bei ca. 40 Prozent. „Datenräume sind gemeinsame und vertrauenswürdige Transaktionsräume, über die Daten von verschiedenen Akteursgruppen dezentral bereitgestellt und gemeinsam genutzt werden.“⁶² Im Gegensatz zu Datenplattformen wird für diese Transaktionsräume kein zentraler Betreiber benötigt, vielmehr ermöglicht das föderierte System, dass jeder Teilnehmer, der über einen Zugangspunkt (z. B. IDS Konnektor) verfügt, ein oder mehrere Rollen im Datenraum erfüllen kann. Ein Datenraum lässt sich dann als Summe aller seiner Teilnehmer definieren. Durch

⁶¹ Lindner, Straub, Kühne (2021), S.31

⁶² EFI (2022)

Zertifizierungen und ein Identitätsmanagement erlaubt ein solcher Datenraum den sicheren und vor allem souveränen Austausch von Daten zwischen zwei oder mehr Parteien.⁶³

Da Datenräume als föderiertes System gedacht werden, verwenden sie eine dezentrale Architektur. Zwar wäre es technisch auch denkbar, zu diesem Zweck die Blockchain-Technologie einzusetzen, allerdings zeigt das Sample, dass hier primär auf den IDS-Standard gesetzt wird. Vereinzelt sind auch Projekte auf Gaia-X-Basis vertreten, wobei auch in solchen Fällen der IDS-Standard eine Rolle spielen kann, da einzelne Elemente von Gaia-X mit dem IDS-Standard umgesetzt werden können (aber nicht müssen). So implementiert bspw. das Forschungsprojekt „HEALTH-X dataLOFT“ die „Gaia-X Federation Services“, setzt aber für die Grundlage des Datenaustausches auf IDS Konnektoren. Diese Modularität zeichnet Datenräume aus, denn hier werden wiederverwendbare, generische „Bausteine“ verwendet, die bereichs- und branchenübergreifend eingesetzt werden können. Einzelne „Bausteine“ zur Sicherstellung von Interoperabilität und Vertrauen sowie der Wertschöpfung auf der Grundlage von Daten und der Governance eines Datenraumes lassen sich je nach Bedarf miteinander kombinieren (siehe Kapitel 2.3).⁶⁴

Das vornehmliche Ziel von Datenräumen ist die gemeinsame Nutzung von Daten, auf Basis derer individuelle Geschäftsmodelle entwickelt werden können. Dazu stellt ein Datenraum die Infrastruktur zur digitalen Vernetzung zwischen Unternehmen zur Verfügung. Ein Beispiel dafür ist der „Mobility Data Space“, der die dezentrale Systemarchitektur der IDSA für den Aufbau eines Ökosystems im Mobilitätssektor nutzt. Mit der Verknüpfung von öffentlichen und privatwirtschaftlichen Daten ist der Mobility Data Space ein digitaler Vertriebskanal für datengetriebene Geschäftsmodelle.

Auch wenn kein zentraler Betreiber für einen Datenraum benötigt wird, so zeigen die Beispiele „Catena-X“ und „Mobility Data Space“, dass dies in der Praxis auch anders gehandhabt wird. Beide Projekte haben eine Betreibergesellschaft gegründet bzw. haben dies vor. Catena-X verfügt überdies auch über eine Vereinsstruktur, durch die interessierte Unternehmen im Rahmen einer Mitgliedschaft den Aufbau und die Weiterentwicklung des Datenraums unterstützen können.

Der Data Intelligence Hub (DIH) der T-Systems International GmbH zeigt zudem, dass die Grenzen zwischen Datenraum und Datenplattform fließend sein und einzelne zentrale Merkmale zu Hybriden verschmelzen können. Der DIH setzt ebenfalls auf dezentrale Datenhaltung und Teilnehmerzertifizierung nach dem IDS-Standard, um Unternehmen den Austausch von Daten zu ermöglichen. Darauf aufbauend wurde zudem ein branchen- und industrieübergreifender Datenmarktplatz entwickelt, über den Preise für Datenpakete zwischen den Teilnehmern ausgehandelt werden können. Der Zugang zum DIH ist kostenlos, es werden aber Gebühren für optionale Leistungen wie Speicherung, Rechenleistung und Transaktionen erhoben. Aufgrund der Möglichkeit bei Bedarf Daten auch zentral abspeichern zu können sowie der Monetarisierung der Infrastruktur und des Betriebs durch ein Einzelunternehmen stellt der DIH einen Hybrid aus Datenraum und Datenplattform dar.

Der Reifegrade von aktuellen Datenräumen im Sample ist noch gering, da es sich bei dem Großteil (ca. 76 Prozent) um geförderte Forschungsprojekte handelt, die teilweise noch am Anfang ihrer Projektlaufzeit stehen. Dies gilt dementsprechend auch für Datenräume, die sich dem Gesundheitssektor zuordnen lassen. Zwar gibt es stellenweise Angebote wie „honic“, die sich selbst als Datenraum bezeichnen, deren Funktionsweise aber auf eine Datenplattform (siehe oben) schließen lässt. Zudem handelt es sich bei allen Gesundheitsdatenräumen des Samples um Forschungsprojekte, die einen spezifischen Use Case verfolgen und sich zudem noch entweder in der anfänglichen Konzeptionsphase oder in einer Pilotphase befinden.

⁶³ Die detaillierte Beschreibung der Funktionsweise eines Datenraums erfolgt in Kapitel 2.3.

⁶⁴ IDSA (2021)

Aufgrund des hohen Anteils an Forschungsprojekten zeichnen sich die Datenräume der Stichprobe in Bezug auf die beteiligten Stakeholder durch eine Zusammenarbeit aus öffentlichen und privaten Playern aus. Im Gesundheitssektor arbeiten bspw. Stakeholder aus Privatwirtschaft, Versorgung und Forschung zusammen, um Daten aus dem primären und sekundären Gesundheitsmarkt zu nutzen.

2.2.4 Fazit

Basierend auf der Typisierung bestehender Data Sharing Ökosysteme im vorgestellten Sample können zwei Ansätze für die Mitglieder der BDI Initiative Gesundheit digital aufgezeigt werden, mit denen sich eine Data Sharing-Lösung im Gesundheitssektor aufbauen ließe. Grundsätzlich kommen sowohl Datenplattformen als auch Datenräume in Betracht, allerdings ist die BDI Initiative Gesundheit digital mit der Idee angetreten, einen offenen, kollaborativen Ansatz bei gleichzeitiger Sicherstellung von Datensouveränität zu verfolgen. Darüber hinaus gilt es, das Vertrauen der Patient:innen in die deutsche und europäische Gesundheitsindustrie zu stärken und bisher vorherrschender öffentlicher Skepsis entgegenzutreten. Eine Expertenbefragung durch das Beratungsunternehmen Roland Berger führte 2020 zu dem Ergebnis, dass Pharma, Medtech & Co sowohl bei Gesunden (im Rahmen der Prävention) als auch bei Erkrankten als der am wenigsten bevorzugte Partner für den Zugang zu gesundheitsbezogenen Leistungen angesehen werden.⁶⁵ Um dies langfristig zu verändern, ist die Schaffung von Vertrauen und Transparenz durch Datensouveränität unabdingbar. Mit einem offenen System, welches alle forschenden Stakeholder miteinander verbindet und für die Interoperabilität von Daten- und Informationsflüssen sorgen kann, ließe sich der größte (kollektive) Mehrwert für Patient:innen schaffen. Unter Berücksichtigung der in 2.2.2 gemachten Beobachtungen kann daher ein Datenraum als angemessene Lösung angesehen werden, da ein solcher Datensouveränität als systemimmanentes Merkmal impliziert. Zudem werden entsprechende Datenräume von der Politik auf nationaler und europäischer Ebene unterstützt, fördern die Interoperabilität, schaffen die Vernetzung von privaten und öffentlichen Stakeholdern, fokussieren Datenqualität und setzen auf eine offene technologische Basis. Des Weiteren steht hier bspw. mit dem IDS-Standard ein offenes und damit anschlussfähiges Referenzarchitekturmodell bereit. Im Folgenden sollen die einzelnen strukturellen und strategischen Vorteile im Detail betrachtet werden.

Sicherung der Datensouveränität

Im Hinblick auf die datenschutzrechtlichen Hürden im Gesundheitssektor bietet ein Datenraum mit seiner föderierten Datenhaltung und der damit verbundenen Souveränität der Teilnehmer über diese Daten die Möglichkeit, zweckgebundene Freigaben von Daten technologisch abzubilden. Die Daten verbleiben zudem am jeweiligen Ort ihrer Entstehung. Damit kann Datensouveränität sowohl auf technischer- als auch auf Governance-Ebene umgesetzt und so ein unberechtigter Zugriff durch Dritte verhindert werden. Das macht einen Datenraum zu einer sicheren Umgebung gerade für besonders schützenswerte Daten, wie auch der in Kapitel 2.2.2 beobachtete höhere Anteil von Lösungen mit dezentraler Architektur im Gesundheitswesen zeigt. Ein weiterer Vorteil des föderierten Systems liegt darin, dass es keinen einzelnen Betreiber geben muss, dem Vertrauen entgegengebracht werden muss. Wie bei der Organisationsform der im Sample betrachteten Beispiele zu beobachten, ist die gemeinsame Gründung einer nicht-gewinnorientierten Betreibergesellschaft als neutrale Instanz trotzdem eine Option. Damit ist ein Datenraum flexibler als eine Datenplattform.

Unterstützung aus der Politik

Wie die Analyse gezeigt hat, werden Datenräume von der Politik sowohl ideell als auch finanziell gefördert. Darauf lässt der hohe Anteil an Förderprojekten im Kontext von Datenräumen innerhalb des untersuchten Samples schließen, der in Kapitel 2.2.2 beobachtet wurde. Beispiele wie der „Mobility Data Space“, der im Rahmen eines Fördervorhabens von 2022 bis 2024 durch das BMDV mit rund 8,5

⁶⁵ Roland Berger (2020), S.11

Millionen Euro unterstützt wird⁶⁶, oder der Gesundheitsdatenraum HEALTH-X dataLOFT, welcher 13 Millionen Euro vom BMWK erhält, unterstreichen die Relevanz von Datenräumen auf nationaler Ebene.

Die geplante Einführung des in Kapitel 2.1.3.5 vorgestellten „European Health Data Space (EHDS)“ ist zudem ein wichtiges Signal der EU-Kommission, die digitale Souveränität Europas langfristig zu sichern. Der kommende Verordnungsentwurf wird klare Regeln für die primäre und sekundäre Nutzung elektronischer Gesundheitsdaten innerhalb der EU festlegen. Interoperabilität, Souveränität und Sicherheit stehen dabei im Vordergrund. Kompatibilität mit dem EHDS sollte beim Aufbau eines Datenraums Gesundheit entsprechend berücksichtigt werden. Hier wird es zudem weitere Fördermöglichkeiten über das EU4Health-Programm geben. Um anschlussfähig für potenzielle zukünftige Regulierungen oder Datenrauminitiativen aus der Politik zu sein, ist eine Berücksichtigung dieser Entwicklungen zu empfehlen.

Förderung der Interoperabilität

Wie bereits in Kapitel 2.1.4.2 ausgeführt, stellt die Interoperabilität von Daten gerade im Gesundheitssektor aktuell noch eine große Hürde dar. Die Bedingungen innerhalb eines Datenraums bieten das Potenzial, diese Interoperabilität zu fördern. Zum einen ist er in der Lage, auch unterschiedliche Stakeholder mit unterschiedlicher Datenhaltung miteinander zu vernetzen, da die minimale Zugangsvoraussetzung zunächst lediglich aus der Zertifizierung eines Konnektors besteht. Dies zeigt auch die in Kapitel 2.2.2 beobachtete Involvierung von öffentlichen Stakeholdern in entsprechenden Data Sharing Ökosystemen. Zum anderen definiert die Rolle des Vocabulary Providers (siehe Kapitel 2.3.3) die gemeinsame Semantik des Datenraums, was eine weitere Grundvoraussetzung für Interoperabilität ist. Auf diese Weise bietet ein Datenraum seinen Teilnehmern die Chance, Interoperabilität gemeinsam nachhaltig herzustellen und zu standardisieren. In diesem Zusammenhang ist perspektivisch die Anbindung bestehender Initiativen wie der in Kapitel 2.1.4.2 erwähnten Medizininformatik-Initiative (MII) zu prüfen, die in ihren Datenintegrationszentren bereits Interoperabilitätsstandards festgelegt und als MII Kerndatensatz beschrieben hat. Datenräume sind zudem untereinander sowie theoretisch mit weiteren Datenplattformen kompatibel, sodass spätere Möglichkeiten des Datenteilens etwa innerhalb des EHDS sichergestellt sind.

Fokussierung auf Datenqualität

Wie in Kapitel 2.1.4 aufgezeigt, gelten im Gesundheitssektor spezifische Bedingungen, die u. a. ein besonderes Augenmerk auf Datenqualität, Sicherheit und Vertrauen erfordern. Entsprechend ist der Fokus auf ein branchenspezifisches Data Sharing Ökosystem zu legen; eine Praxis, die auch schon in Kapitel 2.2.2 zu beobachten war. Datenräume sind für die Umsetzung eines solchen Ökosystems prädestiniert, da sie aufgrund ihres Aufbaus über Konnektoren und Zertifizierungen einen mindestens halboffenen Zugang erfordern. Wie ebenfalls in Kapitel 2.2.2 gesehen, hat der eingeschränkte Zugang zu Data Sharing Ökosystemen eben jenen Vorzug, Datenqualität und Vertrauenswürdigkeit der Teilnehmer sicherzustellen.

Vernetzung von privaten und öffentlichen Stakeholdern möglich

Im Kontext der an Data Sharing Ökosystemen beteiligten Stakeholder konnte in Kapitel 2.2.2 beobachtet werden, dass auch öffentliche Stakeholder integriert werden. Dies geschieht primär bei den Ökosystemen, die sich als Datenräume typisieren lassen. Gerade für den regulierten Gesundheitssektor ist dies von großer Bedeutung. Primärdaten entstehen in Arztpraxen, Krankenhäusern und bei Krankenkassen, die wiederum an die Telematikinfrastruktur der gematik angeschlossen sind.

⁶⁶ BMDV (2021)

i gematik

Die gematik GmbH verantwortet die Telematikinfrastruktur (TI) – die zentrale Plattform für die Vernetzung des ersten Gesundheitsmarktes innerhalb des deutschen Gesundheitswesens. Mehrheitsgesellschafter in der gematik ist mit 51% das Bundesministerium für Gesundheit (BMG), weitere Gesellschafter sind unter anderem die Kassenärztliche Bundesvereinigung und der GKV-Spitzenverband. Die gematik übernimmt die Definition und Durchsetzung verbindlicher Standards für Dienste, Komponenten und Anwendungen in der TI.

Die Data Sharing-Lösung, die auf der Grundlage dieser Studie entwickelt werden kann, soll technologisch offen für einen Anschluss an diese bestehende Infrastruktur sein. Die Architektur eines Datenraums ermöglicht die Anbindung neuer Data Provider über Konnektoren, wodurch theoretisch auch eine Verbindung zwischen Telematikinfrastruktur und Datenraum geschaffen werden könnte. Zudem ist auch eine potenzielle Integration der Daten des Forschungsdatenzentrums (als Data Provider mit Konnektor) theoretisch denkbar.

i Forschungsdatenzentrum

Das Forschungsdatenzentrum (FDZ) am Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) befindet sich derzeit noch im Aufbau. Zukünftig sollen hier Abrechnungsdaten aller gesetzlich Krankenversicherten in Deutschland für Forschungszwecke nutzbar gemacht werden. Ab Januar 2023 kommen Versorgungsdaten aus der elektronischen Patientenakte (ePA) hinzu. Forscher:innen wissenschaftlicher Institutionen können diese Daten auf Antrag für Forschungsvorhaben verwenden.

Allerdings gibt es (noch) regulatorische Hürden für eine potenzielle Vernetzung von Telematikinfrastruktur und Forschungsdatenzentrum mit einem Datenraum Gesundheitswesen, die in einem nächsten Schritt detailliert geprüft werden müssen. So liegt etwa nach § 303e Absatz 1 SGB V keine Antragsberechtigung am FDZ für Privatunternehmen vor; zudem ist nach § 306 Absatz 1 SGB V kein Anschluss an die Telematikinfrastruktur für Privatunternehmen vorgesehen.

Offene technologische Basis

Der Aufbau der Infrastruktur zum Teilen von Daten im B2B-Bereich besitzt bestimmte technologische Anforderungen. Mit dem IDS-Standard steht ein offenes Referenzarchitekturmodell für Datenräume zur Verfügung, welche die notwendigen Elemente auf den Ebenen Business, Funktion, Prozess, Information und System unter den Perspektiven Sicherheit, Zertifizierung und Governance beschreibt. In Kapitel 2.2.2 konnte bereits beobachtet werden, dass die Mehrheit von Data Sharing Ökosystemen auf Basis einer dezentralen oder hybriden Architektur auf den IDS-Standard zurückgreift. Durch den modularen Aufbau lassen sich sowohl das komplette Referenzarchitekturmodell als auch einzelne Elemente wie Konnektor, Informationsmodell oder Meta Data Broker übernehmen. Da viele Module zudem open source sind, wird die Weiterentwicklung auch durch die Community stetig vorangetrieben.⁶⁷ Über die Kompatibilität des IDS-Standards mit GAIA-X kann eine nachhaltige Entwicklung des Datenraums auch auf europäischer Ebene sichergestellt werden. Dies gilt auf lange Sicht auch für die Interoperabilität mit dem EHDS.

2.3 Aufbau & Funktionen eines Datenraums

In Kapitel 2.2.3 wurden bereits die grundlegenden Merkmale von Datenräumen beschrieben. Daran anknüpfend soll nun die mögliche Ausgestaltung eines Datenraums inklusive der entsprechenden Rollen und Module sowie deren Zusammenspiel erläutert werden. Dazu wird exemplarisch ein

⁶⁷ IDSA (2022b)



Abbildung 7: Data Owner & Data Provider

Data Owner & Data Provider

Der Data Owner ist im Kontext eines Datenraums ein Unternehmen oder eine Privatperson, die Daten erzeugt und/oder die Kontrolle über diese ausübt. Bei ihr entstehen die Daten, die innerhalb des Datenraumes ausgetauscht werden. Entsprechend kann sie bestimmen, wer Zugang zu ihren Daten bekommen soll und welche Richtlinien für die Verwendung ihrer Daten gelten sollen. Allerdings muss sie auch über die technischen Mittel verfügen, um diese Bestimmungen umsetzen zu können. Ist dies der Fall, ist ein Data Owner in der Regel gleichzeitig auch ein Data Provider, der die Daten selbst im Datenraum anbietet. Es gibt allerdings auch Szenarien, in denen der Data Owner und der Data Provider unterschiedliche Akteure sind. Dies ist bspw. der Fall, wenn eine zu behandelnde Person ihre Vitaldaten über einen Fitnesstracker erhebt. Sie selbst ist (technisch) nicht in der Lage, die über sie erhobenen Daten im Datenraum anzubieten. Dazu benötigt sie etwa den Hersteller des Fitnesstrackers, der die Daten für die Person in einer Datenbank speichert und diese von dort aus im Datenraum bereitstellen könnte. Damit wäre der Fitnesstracker-Hersteller der Data Provider. Die Rolle des Data Owners beschränkt sich in diesem Fall darauf, den Data Provider zur Weitergabe seiner Daten zu autorisieren. Dasselbe würde gelten, wenn etwa ein Unternehmen einen externen IT-Dienstleister mit der Datenverwaltung beauftragt. Das auftraggebende Unternehmen wäre der Data Owner, während der IT-Dienstleister als Data Provider fungieren würde.

Der Data Provider ist folglich dafür zuständig, die Daten des Data Owners für den Austausch mit dem Data Consumer zur Verfügung zu stellen. Wie bereits zuvor dargestellt, kann es sich bei Data Provider und Data Owner um denselben Stakeholder handeln, dies ist jedoch nicht zwangsläufig der Fall. Im Zusammenspiel mit zusätzlichen Rollen, die ein Datenraum enthalten kann, die aber nicht für die grundlegende Funktionalität notwendig sind (siehe Kapitel 2.2.4), entstehen weitere Handlungsmöglichkeiten für den Data Provider. Neben den eigentlichen Daten überträgt er auch Metadaten an einen optionalen Broker (siehe Kapitel 2.3.3). Zudem übernimmt der Data Provider nach dem Datentransfer auch die Aufgabe, das Protokoll über den erfolgreichen oder missglückten Datenaustausch an eine Clearingstelle zu senden, um die Abrechnung oder eine Konfliktlösung zu ermöglichen. Des Weiteren kann der Data Provider auf Daten-Apps zurückgreifen, um die Daten während des Transfers anzureichern oder ihre Qualität zu verbessern.



Abbildung 8: Data Consumer & Data User

Data Consumer & Data User

Der Data Consumer empfängt die Daten, die ein Data Provider im Datenraum bereitstellt. Er stellt diese Daten dann dem Data User zur Verfügung. Letzterer ist derjenige, der den rechtlichen Anspruch auf die Verwendung der Daten gemäß der vereinbarten Nutzungsrichtlinie besitzt. Er ist damit der Vertragspartner des Data Owner. Analog zum Verhältnis von Data Provider und Data Owner kann es sich bei Data Consumer und Data User um denselben Akteur handeln, aber auch hier kann es Szenarien geben, in denen dies nicht zutrifft. Angenommen, eine betroffene Person teilt ihre Fitnesstracker-Daten über ein Webportal mit ihrem Fitnesstrainer:in, dann handelt es sich bei dem

Webportal um den Data Consumer, während der Fitnessstrainer als Data User agiert. Er ist derjenige, der die Daten zur Analyse und Interpretation nutzt und derjenige, der dafür vom Data Owner beauftragt wurde. Ähnlich verhält es sich im Beispiel von IT-Dienstleister und auftraggebendem Unternehmen, in dem der IT-Dienstleister die Rolle des Data Consumers übernimmt und das Unternehmen als Data User fungiert.

Da der Data Consumer als empfangendes Pendant des Data Providers zu sehen ist, kann er über den Datentransfer hinaus ähnliche Handlungen im Datenraum vollziehen. Vorausgesetzt, dass ein oder mehrere der optionalen Rollen im Datenraum aktiv sind, kann der Data Consumer mit diesen interagieren. Er kann mit Hilfe eines optionalen Brokers nach Anbietern von bestimmten Daten suchen, die Clearingstelle über das Ergebnis des Datentransfers informieren oder Daten-Apps zur Anreicherung, Umwandlung usw. der empfangenen Daten verwenden.



Abbildung 9: Identity Provider

Identity Provider

Der Identity Provider übernimmt das Identitätsmanagement innerhalb des Datenraums. Nur wenn sichergestellt wird, dass alle Teilnehmer auch diejenigen sind, für die sie sich ausgeben, kann ein sicherer und vertrauenswürdiger Datenaustausch stattfinden. Dazu übernimmt der Identity Provider die Aufgabe, Identitätsinformationen von und für Datenraumteilnehmer zu erstellen, zu pflegen, zu verwalten, zu überwachen und zu validieren. Indem die Identitäten der einzelnen Teilnehmer geprüft werden, wird sichergestellt, dass nur berechnigte Teilnehmer am Datenraum partizipieren und Zugang zu Daten erhalten können. Bei „berechtigten Teilnehmern“ handelt es sich um Organisationen, die durch die Zertifizierungsstelle des Datenraums zertifiziert wurden (siehe Absatz „Zertifizierungsstelle und Prüfstellen“).

Bei einem IDS-basierten Datenraum umfasst der Identity Provider drei Elemente.

1. Die Zertifizierungsinstanz, die digitale Zertifikate (nach X.509 Standard) für die Teilnehmer des Datenraums verwaltet.
2. Der Dynamic Attribute Provisioning Service (DAPS), der die dynamischen Attribute und Zugangstokens der Teilnehmer verwaltet.
3. Der Participant Information Service (ParIS), der allgemeinen Informationen (z. B. Umsatzsteuer-Identifikationsnummer, gesetzliche Vertreter oder Organisationsstruktur) über alle Teilnehmer des Datenraums bereithält.



Abbildung 10: Zertifizierungsstelle

Zertifizierungsstelle und Prüfstellen

Die letzte Rolle, die für ein minimal funktionsfähigen Datenraum benötigt wird, ist die der Zertifizierungsstelle. Für die Teilnahme am Datenraum ist es essenziell, dass sowohl Konnektoren als

auch die teilnehmenden Organisationen auf ihre Konformität mit den Vorgaben des Datenraumes anhand von vordefinierten Kriterienkatalogen überprüft werden. Mit Hilfe von ausgewählten Prüfstellen werden diese Vorgaben evaluiert und Zertifizierungen durchgeführt. Dabei bieten Prüfstellen auch Beratung und Testverfahren für Softwarekomponenten vor der Zertifizierung an. Die Evaluierung eines Teilnehmers oder eines Konnektors erfolgt auf Antrag eines Teilnehmers und stützt sich auf den Vertrag zwischen dem Teilnehmer und der Prüfstelle. In diesem Prozess überwacht die Zertifizierungsstelle die Maßnahmen und Entscheidungen der einzelnen Prüfstellen. Damit handelt es sich bei der Zertifizierungsstelle um ein zentrales Governance-Organ eines Datenraums.



Abbildung 11: Konnektor

Konnektor

Der Konnektor dient bei einem IDS-basierten Datenraum dazu, eine authentifizierte und verschlüsselte Verbindung zwischen Data Provider bzw. Data Consumer und dem Datenraum herzustellen. Es handelt sich um eine Software, die über eine eindeutige Kennung und ein gültiges Zertifikat verfügt und somit den Teilnehmer ausweisen kann. Damit ist der Konnektor Ausweis und Schlüssel zugleich. Der Konnektor kann auf zwei Arten genutzt werden:

1. Datenbanken von Data Providern und Softwaresysteme von Data Consumern werden an den Konnektor angeschlossen, sodass er als Brücke zwischen beiden Teilnehmern verwendet wird. (Siehe 2.3.2 Beispiel 1: Rechtliche Durchsetzung)
2. Datenbank und Softwaresystem werden innerhalb der Laufzeitumgebung eines Konnektors bereitgestellt. (Siehe 2.3.2 Beispiel 2: Technische Durchsetzung)

Da die Daten so bei den einzelnen Teilnehmern gespeichert bleiben, ermöglicht der Konnektor, dass der Datenraum mit einer dezentralen Datenhaltung funktioniert. Darüber hinaus ist der Konnektor für die Kontrolle und Durchsetzung von Nutzungsrichtlinien zuständig. Beides – dezentrale Datenspeicherung und die Durchsetzung von Nutzungsrichtlinien – unterstützen den Gedanken der Datensouveränität. Damit ist der Konnektor die zentrale technische Komponente für den Datenaustausch und die Datenverarbeitung in einem IDS-basierten Datenraum.

Eclipse Dataspace Connector

Der IDS Konnektor ist nur eine Möglichkeit, eine sichere Verbindung in einem Datenraum herzustellen. In dem Open-Source-Projekt „Eclipse Dataspace Connector“ der Eclipse Foundation wird derzeit an einer Alternative gearbeitet, die zunächst den IDS Standard sowie relevante Protokolle und Anforderungen im Zusammenhang mit Gaia-X integrieren wird.⁶⁹ Darüber hinaus kann dieser so erweitert werden, dass er mit weiteren Referenzarchitekturmodellen kompatibel sein wird.

2.3.2 Grundlegender Datenfluss

Mit den zuvor vorgestellten Rollen und Komponenten ist ein direkter Datenaustausch zwischen einem Data Provider und einem Data Consumer in einem Datenraum möglich. Beide Parteien müssen bereits miteinander bekannt sein, da es in dieser minimal funktionsfähigen Version noch keine Möglichkeit gibt, nach bestimmten Datensätzen zu suchen (siehe 1.3.3 Optionale Rollen). Abbildung 12 zeigt einen solchen Fall.

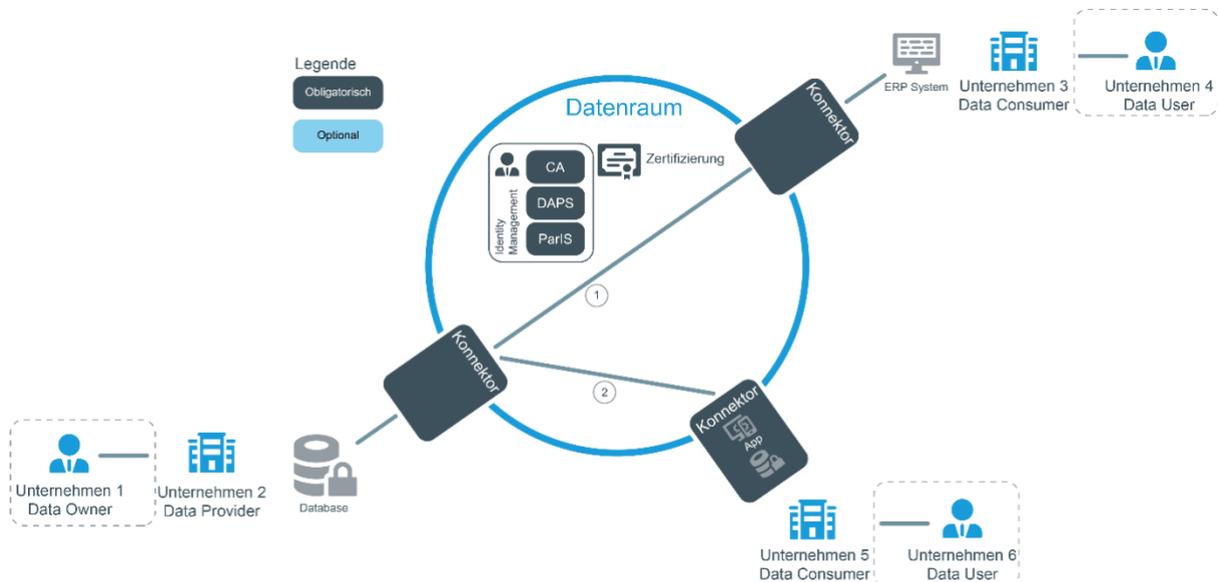


Abbildung 12: Einfacher Datenfluss

Zunächst müssen sowohl Data Provider als auch Data Consumer durch die Zertifizierungsstelle des Datenraums geprüft und durch den Identity Provider mit digitalen Zertifikaten ausgestattet sein. Der Datenfluss beginnt damit, dass ein Data Consumer einen Datensatz von einem ihm bekannten Data Provider anfordert. Daraufhin wird ein Vertragsverhandlungsprozess gestartet, in dem die Regeln und Bedingungen (Nutzungsrichtlinien) für einen Datentransfer ausgehandelt werden.

Es obliegt dem Data Owner, die endgültigen Nutzungsrichtlinien festzulegen. Ist der Data Provider gleichzeitig auch der Data Owner, legt er diese Regeln selbst fest, ansonsten wird er vom Data Owner lediglich zur Durchführung des Datentransfers unter den gewünschten Nutzungsrichtlinien autorisiert. Solche Nutzungsrichtlinien umfassen Szenarien wie z. B. die Beschränkung der Datennutzung auf eine bestimmte Gruppe von Teilnehmern, die Beschränkung der Nutzung auf bestimmte Zwecke oder eine bestimmte Dauer, die Nutzung von Daten nicht öfter als n-mal, usw. Die Einhaltung dieser Regeln wird über den Konnektor, der Data Provider und Data Consumer mit dem Datenraum verbindet,

⁶⁹ Eclipse Foundation (2022)

gewährleistet. Dabei gibt es zwei Wege, die Einhaltung durchzusetzen, je nachdem, wie der Konnektor eingebunden ist.

Beispiel (1): Rechtliche Durchsetzung

Abbildung 12 zeigt einen Datenfluss (1) zwischen Unternehmen 2 und Unternehmen 3. Der Konnektor des Datenkonsumenten (Unternehmen 3) ist in diesem Fall mit einem ERP-System verbunden. Daher muss der Data Provider (Unternehmen 2) bzw. der Data Owner (Unternehmen 1) zustimmen, dass die Daten außerhalb des Konnektors verarbeitet werden. Hier kommt ein digitaler Vertrag zwischen den beiden Parteien zustande, der für beide Parteien rechtsverbindlich ist, auch wenn der Data Provider die technische Kontrolle verliert. Hier muss darauf vertraut werden, dass der Data Consumer bzw. Data User den Vertrag einhält.

Beispiel (2): Technische Durchsetzung

Abbildung 12 zeigt ebenfalls einen Datenfluss (2) zwischen Unternehmen 2 und Unternehmen 5. Hier wird eine Anwendung, die direkt im Konnektor des Data Consumers (Unternehmen 5) bereitgestellt wird, zur Weiterverarbeitung der empfangenen Daten verwendet. Die Daten verlassen den Konnektor und damit den Datenraum also nicht. Wie in Beispiel 1 wird ein digitaler Vertrag zwischen beiden Parteien geschlossen. Hinzu kommt aber, dass Anwendungen, die in der Laufzeitumgebung eines Konnektors bereitgestellt werden, Nutzungsrichtlinien auf technische Weise durchsetzen können, wie z. B. das Löschen von Daten nach 5 Tagen. Damit wird den geschlossenen digitalen Verträgen eine weitere Vertrauensebene hinzugefügt.

Unabhängig von der Art und Weise der Nutzungsrichtliniendurchsetzung zeigen beide Beispiele lediglich das Grundprinzip des Datenaustausches innerhalb eines Datenraums. Die mehrwertstiftenden Funktionen eines Datenraums, wie etwa die hohe Qualität und Auffindbarkeit von Daten sowie ein einheitliches Vokabular umfassen, werden durch weitere, optionale Rollen geschaffen.

2.3.3 Optionale Rollen

Während der reine Austausch von Daten in einem Datenraum bereits mit den zuvor beschriebenen Rollen möglich wird, gibt es weitere Rollen, die die Attraktivität des Datenraums für weitere Teilnehmer steigert. Bei allen in der Folge beschriebenen Rollen handelt es sich um Intermediäre, die im Datenraum agieren und dafür sorgen, dass Vertrauen geschaffen, Metadaten bereitgestellt und Geschäftsmodelle entwickelt werden können.

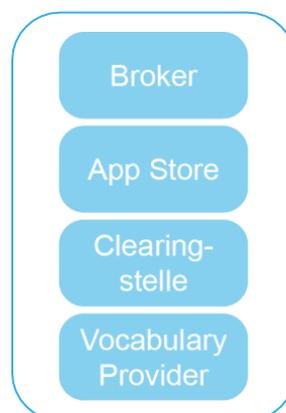


Abbildung 13: Optionale Rollen

Broker

Die Rolle des Brokers ist es, Informationen über die von den Teilnehmern des Datenraums angebotenen Datenendpunkte zu speichern und zu verwalten. Er dient als Anlaufstelle für Data Consumer, die auf der Suche nach bestimmten Daten sind und die in Frage kommenden Data Provider nicht kennen. Ein Hauptmerkmal des Brokers ist es, dass er die Datensätze nicht selbst speichert, sondern lediglich mit Metainformationen agiert. Dazu stellt er dem Data Provider eine Schnittstelle für die Übermittlung der Metadaten zur Verfügung. Nach Erhalt der Metadaten werden diese in einem internen Repository gespeichert, damit sie von Data Consumern in strukturierter Weise abgefragt werden können.

Ein entsprechender Datenfluss mit Beteiligung des Brokers lässt sich wie folgt darstellen:

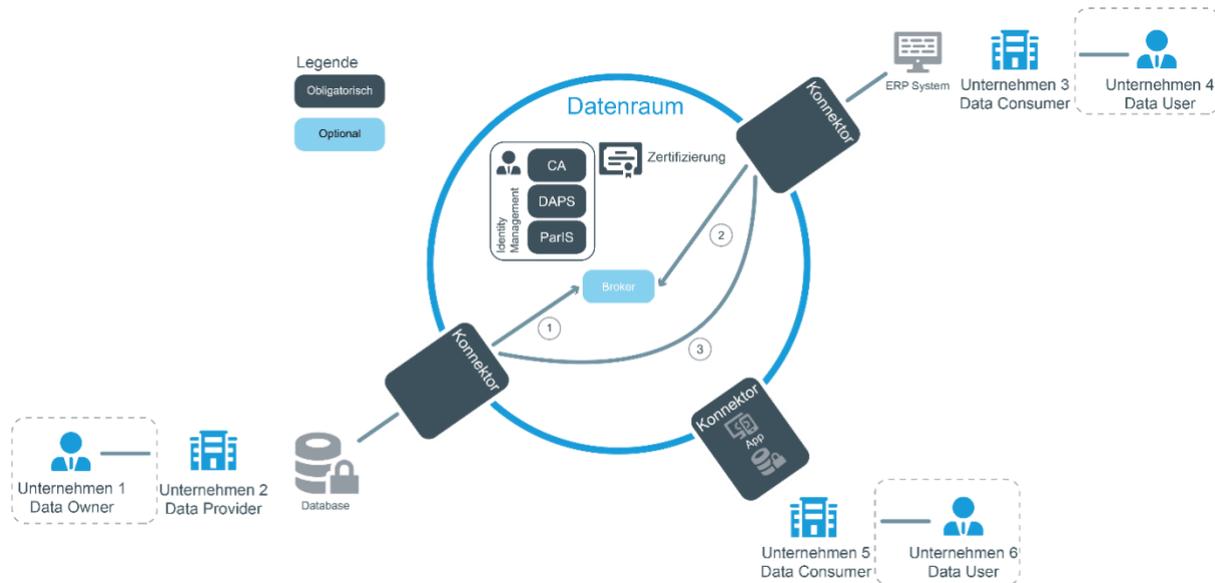


Abbildung 14: Datenfluss mit Broker

(1) Der Data Provider registriert seine Datensätze beim Broker und übermittelt die notwendigen Metadaten. Diese enthalten z. B. Informationen über den Data Provider selbst aber auch Informationen über die Syntax, Serialisierung und Semantik der Datenquelle sowie deren Anwendungsbereich. (2) Ein Data Consumer, der auf der Suche nach entsprechenden Daten ist, sendet eine entsprechende Suchanfrage an den Broker und erhält die relevanten Metadaten über Datenset und Data Provider. Mit diesen Informationen ist er in der Lage, in Vertragsverhandlungen mit dem Data Provider zu gehen. Nachdem der Broker dem Data Consumer die Metadaten über einen bestimmten Data Provider zur Verfügung gestellt hat, ist seine Aufgabe abgeschlossen. (3) Der nachfolgende Datentransfer findet gemäß dem bereits beschriebenen grundlegenden Datenfluss direkt zwischen Data Provider und Data Consumer statt.

App Providers & App Store Provider

Vor oder nach dem eigentlichen Datenaustausch müssen die Daten möglicherweise verarbeitet oder umgewandelt werden. App Provider entwickeln Daten-Apps, die diese Datenverarbeitungsprozesse bei Data Consumern oder Data Providern optimieren können. Sie können in der Konnektor-Umgebung installiert werden, um zusätzliche Datenverarbeitungsfunktionen zu implementieren. Dazu zählen etwa die Umwandlung von Daten aus einem Eingabeformat in ein anderes Ausgabeformat oder komplexere Datenanalysefunktionen. Jede dieser Daten-Apps muss im App Store des Datenraums veröffentlicht werden. Nur so können sie von Data Providern und Data Consumern gesucht, abgerufen und genutzt werden. Zur besseren Auffindbarkeit wird jede Daten-App mit Hilfe von Metadaten in Bezug auf ihre Semantik, Funktionalität, Schnittstellen usw. beschrieben. Der App Store Provider ist wiederum für die

Ein Datenfluss unter Einbindung einer Clearingstelle lässt sich wie folgt darstellen:

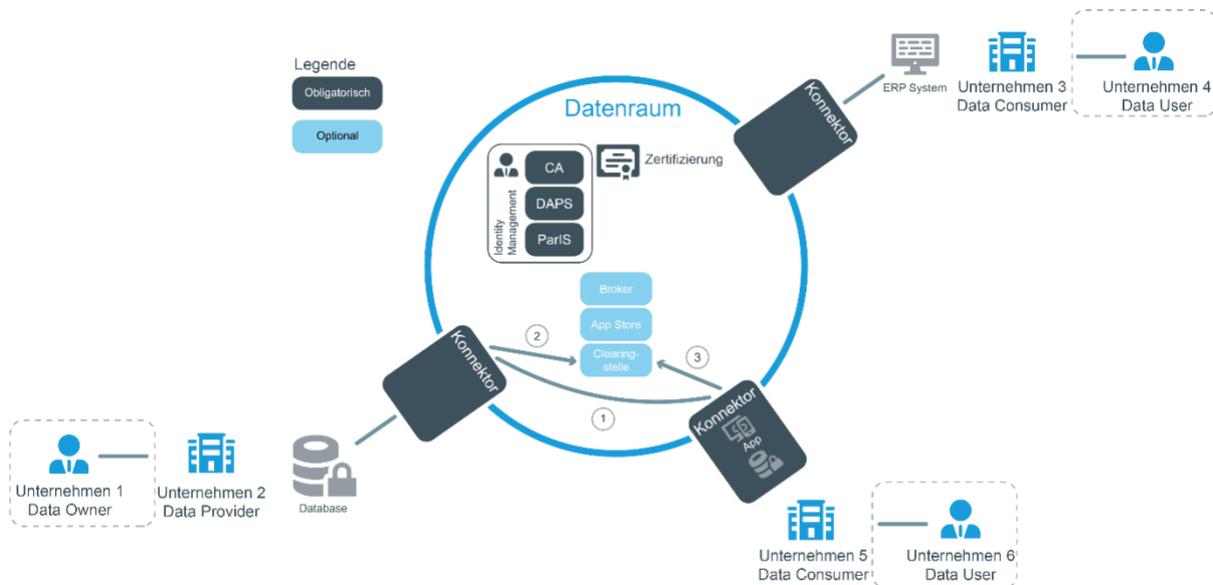


Abbildung 16: Datenfluss mit Clearingstelle

(1) Zwei Teilnehmer des Datenraums tauschen nach dem bekannten Prozess Daten aus. Nach der Transaktion melden sowohl der (2) Data Provider als auch der (3) Data Consumer Informationen über den Verlauf des Austausches an die Clearingstelle. Dies geschieht sowohl bei einem erfolgreichen als auch bei einem missglückten Datentransfer. Da der Datenfluss so durch einen neutralen Dritten nachgewiesen werden kann, erhalten beide Teilnehmer die Sicherheit, dass der Transfer korrekt abgerechnet werden kann. Zudem lassen sich potenzielle Konflikte nach anhand der Protokolle durch die Clearingstelle lösen.

Vocabulary Provider

Der Vocabulary Provider verwaltet und bietet Wörterbücher (d. h. Ontologien, Referenzdatenmodelle oder Metadatenelemente) an, die zur Annotation und Beschreibung von Datensätzen verwendet werden können. Insbesondere stellt der Vocabulary Provider das Informationsmodell des Datenraums bereit, welches die Grundlage für die Beschreibung von Datenquellen bildet. Darüber hinaus können auch andere domänenspezifische Vokabulare bereitgestellt werden. Auf diese Weise kann sichergestellt werden, dass alle Datenraumteilnehmer das gleiche Vokabular für ihre Datensätze verwenden und somit eine gemeinsame „Sprache sprechen“.

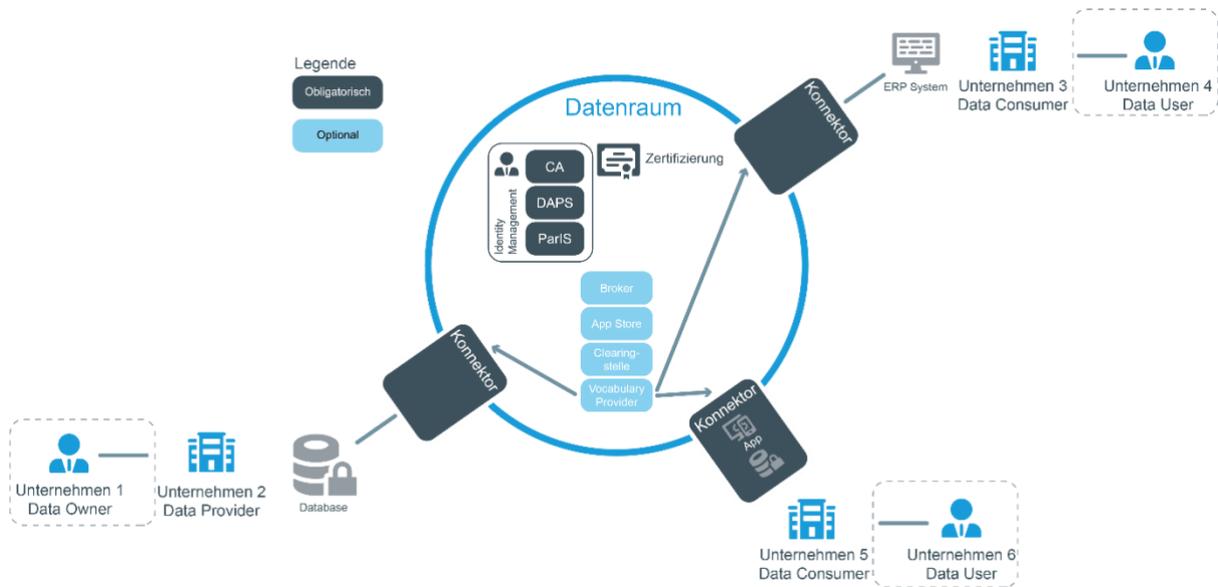


Abbildung 17: Funktion des Vocabulary Providers

Damit wurden nun die grundlegende Funktionsweise eines Datenraums nach IDS-Standard sowie die involvierten Rollen beschrieben. Es zeigt sich, dass nur wenige Rollen und Komponenten gebraucht werden, um einen einfachen Datenaustausch durchzuführen, dass aber der Mehrwert eines Datenraums vor allem durch die Integration der optionalen Rollen geschaffen wird. Im folgenden Kapitel soll nun ein Konzept für einen kollaborativen Datenraum Gesundheitswesen vorgestellt werden, der auf der hier verkürzt gezeigten Architektur basieren könnte.

3. Konzept für einen kollaborativen Datenraum Gesundheitswesen

Nachdem in Kapitel 2 die Ausgangslage und die Hürden für den effektiven Austausch von Daten in Deutschland und Europa diskutiert, ein Blick auf bestehende Ökosysteme zum Austausch von Daten geworfen und der IDS als exemplarisches Referenzarchitekturmodell vorgestellt wurde, werden in Kapitel 3 nun konkrete Vorschläge und Ideen für einen Datenraum Gesundheit der BDI Initiative Gesundheit digital skizziert.

3.1 Vision & Strategie

Aus zahlreichen strukturierten Interviews mit Mitgliedern der BDI Initiative Gesundheit digital sowie weiteren Vertreter:innen von Verbänden, Selbstverwaltungsorganen und der Industrie ergab sich eine gemeinschaftlich geteilte Vision für einen Datenraum Gesundheitswesen. Dieser soll einen durchgängigen Datenaustausch für alle an der Gesundheitsindustrie Teilnehmenden ermöglichen. Dafür ist ein gemeinsames und kollaboratives Vorgehen notwendig. Nur durch Kollaboration zwischen unterschiedlichen Stakeholdern können Innovationen für eine optimierte und personalisierte Versorgungslandschaft entstehen. Der Datenraum soll das Netzwerk und die technologische Basis dafür bereitstellen und damit ein leistungsfähiges Gesamtsystem für alle partizipierenden Akteure bieten. Im Vordergrund steht die zukünftige Wirtschaftlichkeit aller Netzwerkpartner und des deutschen und europäischen Gesundheitssystems. Daher kommt der Anschlussfähigkeit des kollaborativen Datenraums Gesundheitswesen an den EHDS und dessen Data Governance große Bedeutung zu. Ein nachhaltiger Aufbau des Datenraums ist nur im Einklang mit den Aktivitäten der EU-Kommission denkbar.

Um die skizzierte Vision zu erreichen, braucht es eine Strategie und ein strukturiertes Vorgehen. Schließlich bringt der Datenaustausch zwischen verschiedenen Stakeholdern unterschiedliche Grade der Komplexität mit sich – sowohl was Data Provider und Data Consumer als auch die entsprechenden Datenquellen betrifft. Dies hat Auswirkungen auf die Anwendungsfälle, die in der Anfangsphase des kollaborativen Datenraums Gesundheitswesen realisiert werden können, denn besonders weitreichende und damit komplexe Anwendungsfälle lassen sich nicht kurzfristig umsetzen. Wie bereits in Kapitel 2.1.4 beschrieben wurde, gibt es hier aktuell bestimmte Hürden, die den Austausch von Daten im Gesundheitswesen erschweren. So ist der Zugang zu Primär- und Sekundärdaten aus der medizinischen Versorgung für Unternehmen in Deutschland beschränkt und bspw. nur in genau definierten Szenarien wie prospektiven Studien oder dem Post Market Clinical Follow-Up möglich. Anträge beim Forschungsdatenzentrum Gesundheit können von Unternehmen der Privatwirtschaft jedoch nicht gestellt werden. Da der erste Gesundheitsmarkt, in dem die meisten relevanten Daten entstehen, ein geschlossener Markt ist, der nur für Leistungsempfänger, Leistungserbringer und Kostenträger zugänglich ist, ist es für Unternehmen zudem schwer, Vertreter der Selbstverwaltung zu involvieren. Formate des direkten Austausches sind bisher nicht flächendeckend gegeben. Aus Perspektive der Industrie müssen hier Veränderungen der regulatorischen Rahmenbedingungen stattfinden. Die EU-Kommission sorgt mit ihrem Vorschlag für den European Health Data Space zwar für Bewegung bei diesem Thema, es kann jedoch nicht damit gerechnet werden, dass dieser bereits in naher Zukunft entsteht. Es werden noch einige Jahre vergehen, bis eine Einigung erzielt und entsprechende Verordnungen in Deutschland umgesetzt werden. In dieser Situation ist es sinnvoll voranzugehen und einen kollaborativen Datenraum Gesundheitswesen sukzessive aufzubauen. Schließlich es gibt auch Anwendungsfälle, die in einem weniger komplexen Setup umsetzbar sind und trotzdem auf die Zielsetzung einer personalisierten und optimierten Gesundheitsversorgung einzahlen. Mit diesen können Datenaustauschkapazitäten Stück für Stück aufgebaut und sukzessive weitere Teilnehmer im Netzwerk aufgenommen werden. Abbildung 18 zeigt anhand von drei Phasen, wie ein solches schrittweises Vorgehen aussehen kann.

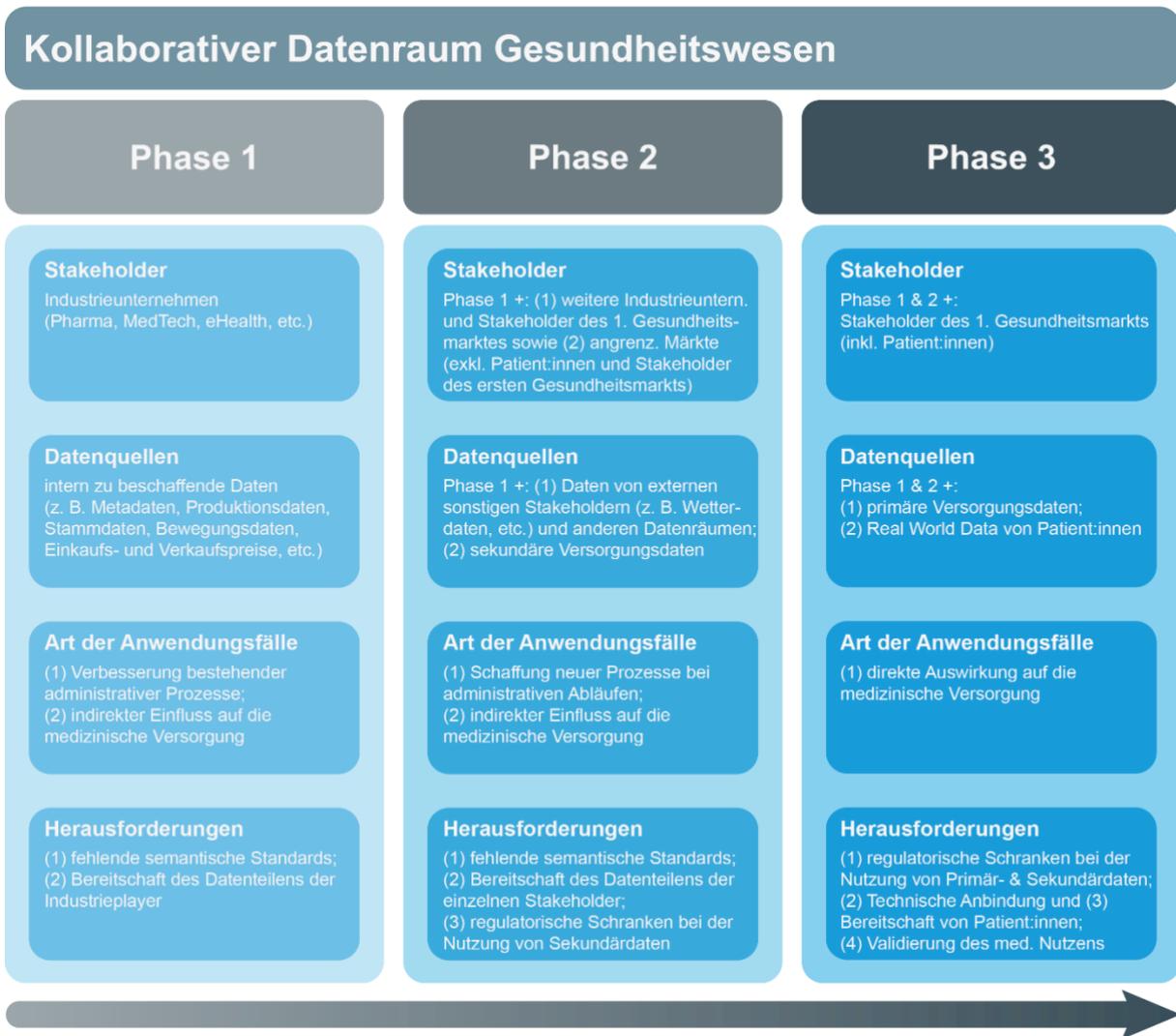


Abbildung 18: Drei Phasen als strategisches Leitbild des kollaborativen Datenraums Gesundheitswesen

In **Phase 1** sind zunächst Industrieunternehmen des Gesundheitssektors als Teilnehmer des kollaborativen Datenraums Gesundheitswesen involviert. Pharma-, Medizintechnik-, eHealth-Unternehmen u. a. bringen ihre intern zu beschaffenden Daten unterschiedlicher Sensibilitätsgrade ein. Dazu zählen Metadaten, Produktionsdaten, Stammdaten, Bewegungsdaten, Einkaufs- und Verkaufspreise etc. Die darauf aufbauenden möglichen Anwendungsfälle tragen zur Verbesserung bestehender administrativer Prozesse bei. Damit haben sie in dieser Phase zunächst einen indirekten Effekt auf die medizinische Versorgung. Zu den Herausforderungen in dieser Phase zählen das Fehlen semantischer Standards sowie die fehlende gemeinsame Infrastruktur. Darüber hinaus benötigt es die grundsätzliche Bereitschaft der Industrie-Player, Daten miteinander zu teilen.

Mit **Phase 2** wird das Teilnehmerfeld um weitere Industrieunternehmen und Stakeholder des zweiten Gesundheitsmarktes und angrenzender Märkte erweitert. Zu den internen Datenquellen aus Phase 1 kommen nun die Daten von externen Stakeholdern (Wetterdaten, Mobilitätsdaten etc.) und anderen Datenräumen hinzu. Zudem sollte das Ziel sein, sekundäre Versorgungsdaten wie Abrechnungsdaten in den kollaborativen Datenraum Gesundheitswesen zu integrieren. So können Anwendungsfälle, die neue Prozesse in administrativen Abläufen schaffen, realisiert werden. Der Einfluss auf die medizinische Versorgung bleibt wie in Phase 1 indirekt. Auch in Phase 2 bestehen die größten zu lösenden Herausforderungen aus dem Fehlen semantischer Standards und der mangelnden Bereitschaft des Datenteilens der einzelnen Stakeholder. Zudem sind regulatorische Hürden im Hinblick auf die Nutzung von sekundären Versorgungsdaten durch Unternehmen zu überwinden.

In **Phase 3** kommen zu den bestehenden Stakeholdern nun auch Stakeholder des ersten Gesundheitsmarktes (inkl. Patient:innen) hinzu. Damit können auch primäre und sekundäre Versorgungsdaten sowie Real World Data der zu behandelnden Personen ausgetauscht werden. Aus den neuen Datenquellen ergeben sich direkte Auswirkungen der möglichen Anwendungsfälle auf die medizinische Versorgung. Zu den Herausforderungen der ersten beiden Phasen kommen nun noch weitere regulatorische Einschränkungen bzgl. der Nutzung von Gesundheitsdaten hinzu. Auch die technische Anbindung von Einzelpersonen an den Datenraum Gesundheitswesen ist zu klären, da die direkte Anbindung von Patient:innen oder Bürger:innen in aktuellen Datenraummodellen noch nicht vorgesehen ist. Zudem zählt die benötigte Bereitschaft von Patient:innen, Gesundheitsdaten zu teilen, zu den Herausforderungen der Phase 3. Zu guter Letzt müssen Anwendungsfälle in der medizinischen Versorgung ihren medizinischen Nutzen erst noch beweisen. Gelingt es, diese Hürden zu überwinden, sind die Voraussetzungen geschaffen, die Vision einer personalisierten Gesundheitsversorgung auf Basis europäischer Werte Wirklichkeit werden zu lassen.

Dieses Phasenmodell ermöglicht, mit pragmatischen Anwendungsfällen beginnen zu können, die den anwendenden Personen und Organisationen aufzeigen, dass der Datenraum zum einen funktionsfähig ist und zum anderen schnell erste Mehrwerte liefert. Niederschwellige B2B-Industrieanwendungsfälle können zudem auch Basisbausteine auf dem Weg hin zu einer personalisierten Gesundheitsversorgung sein, da sie grundlegende Mechanismen etablieren, die in späteren Phasen relevant werden. Gleichzeitig haben erfolgreiche Anwendungsfälle das Potenzial, einen Mentalitäts- und Paradigmenwandel in der Industrie voranzutreiben. Ebenso nicht zu vernachlässigen ist die Signalwirkung auf externe Stakeholder und politische Vertreter:innen. Potenzielle Anforderungen an Gesetzesänderungen lassen sich so ebenfalls formulieren. Auch auf die Öffentlichkeit und die potenziellen späteren anwendenden Organisationen aus den Selbstverwaltungsorganen kann ein schrittweises Vorgehen vertrauensbildend wirken. Positive Beispiele wirken überzeugend und es kann bereits ein höherer Reifegrad der Zusammenarbeit, Organisation und technischen Implementierung demonstriert werden. Um das schrittweise Vorgehen effizient zu gestalten, sollte jede der Phasen in Vorbereitung auf die Umsetzung sehr klar beschrieben und voneinander abgegrenzt werden.

3.2 Anwendungsfälle

Mit der Schaffung eines sicheren und souveränen Datenaustausches zwischen unterschiedlichen Stakeholdern ermöglicht der kollaborative Datenraum „Gesundheitswesen“ eine Vielzahl potenzieller Anwendungsfälle. Für den im vorangegangenen Kapitel beschriebenen schrittweisen Aufbau des Datenraums haben vor allem diejenigen Fälle Relevanz, die kein direktes Differenzierungsmerkmal gegenüber dem Wettbewerb für ein bestimmtes Unternehmen darstellen, sondern zunächst allen Teilnehmern gleichermaßen einen Mehrwert bieten. Die Differenzierung und der Wettbewerb entstehen dann darüber, wie die bereitgestellten Daten verwendet werden. In diesem Kapitel soll erörtert werden, ob es solche Anwendungsfälle im Gesundheitswesen gibt und wie diese kategorisiert werden können.

3.2.1 Typisierung

Zunächst wurden möglichst viele potenzielle Anwendungsfälle recherchiert und gesammelt. Einige dieser Beispiele konnten durch den Blick auf bestehende externe Projekte identifiziert werden, andere wurden in Gesprächen mit Mitgliedern der BDI Initiative Gesundheit digital sowie von einem befragten Expertennetzwerk, bestehend aus Vertreter:innen von Verbänden, Selbstverwaltungsorganen sowie der Industrie, eingebracht. Die so erfassten 39 Anwendungsfälle stehen exemplarisch für die große Bandbreite an möglichen Themen, die ein Datenraum Gesundheitswesen abdecken kann. Der Reifegrad der eingereichten Anwendungsfälle bewegt sich zum größten Teil auf der Ebene vorläufiger Ideen, es wurden lediglich drei Fälle durch die einreichenden Organisationen umgesetzt (ohne Datenraum). Das schließt nicht aus, dass ähnliche Anwendungsfälle bereits in anderen Kontexten umgesetzt wurden. Bei der Clusterung der Beispiele ergeben sich verschiedene Typen von Anwendungsfällen, die sich nach Versorgungstiefe und Innovationsgrad unterscheiden und in Quadranten einteilen lassen (siehe Abbildung 19 & 20).



Abbildung 19: Quadranten von Anwendungsfällen nach Versorgungstiefe und Innovationsgrad

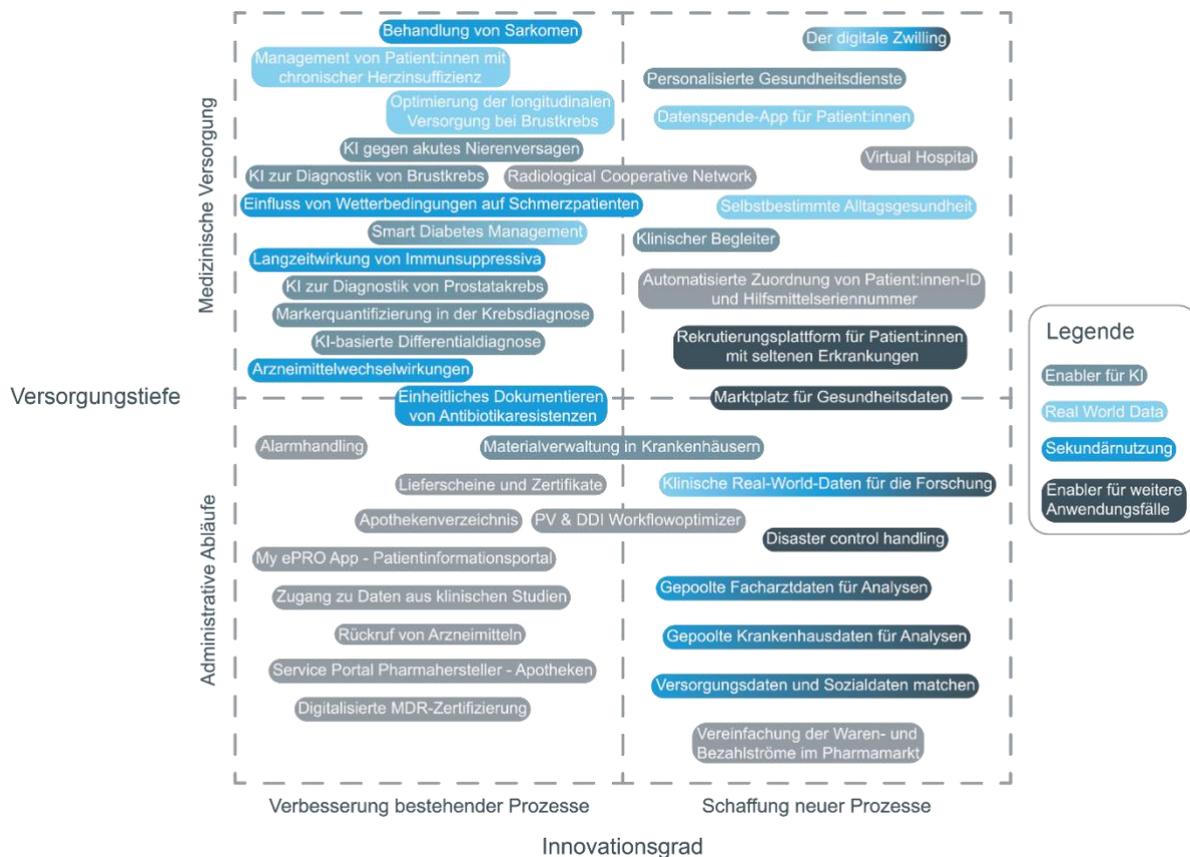


Abbildung 20: Exemplarische Anwendungsfälle nach Quadranten

Quadrant A: Verbesserung bestehender administrativer Prozesse

Ein Teil der Anwendungsfälle konzentriert sich darauf, bestehende administrative Abläufe durch die Vernetzung unterschiedlicher Stakeholder zu verbessern. Hierzu zählen bspw. die Vereinfachung der Waren- und Bezahlströme im Pharmamarkt, die Optimierung der Abläufe beim Rückruf von Arzneimitteln oder die automatisierte Materialverwaltung in Krankenhäusern. Diese Anwendungsfälle haben gemeinsam, dass sie nicht die Patient:innen, sondern die notwendigen organisatorischen Rahmenbedingungen des Gesundheitswesens fokussieren. Die Versorgungstiefe ist daher gering. Mit ihnen werden Prozesse, die heute bereits ohne (oder mit eingeschränkter) Vernetzung und Datenaustausch stattfinden, erheblich verbessert, da sich Möglichkeiten der Automatisierung bieten. Davon profitieren alle potenziellen Teilnehmer des kollaborativen Datenraums Gesundheitswesen gleichermaßen, sodass hier bereits diejenigen Anwendungsfälle zu finden sind, die sich für die erste Phase des Aufbaus des kollaborativen Datenraums Gesundheitswesens eignen.

Quadrant B: Verbesserung bestehender Prozesse in der medizinischen Versorgung

Ein weiterer Teil der Anwendungsfälle zielt ebenfalls auf die Verbesserung bestehender Prozesse ab, hat aber die medizinische Versorgung von Patient:innen im Blick. Die Versorgungstiefe ist daher hoch. Unterstützung bei Diagnose und Therapie sowie bei der Selbstverwaltung der betroffenen Personen im Alltag gehören zu den Beispielen. Dabei fokussieren die Anwendungsfälle in der Regel einzelne weitverbreitete Erkrankungen wie Diabetes, verschiedene Krebsarten oder chronische Herzinsuffizienz, da hier sowohl die Datenmenge als auch die wirtschaftlich relevante Marktgröße gegeben sind. Es gibt jedoch auch Anwendungsfälle, in denen gerade die selteneren Erkrankungen betrachtet werden, da diese aufgrund von kleinen Kohorten von Vernetzung und Datenaustausch in besonderem Maße profitieren.

Beispiel: DeepSarc – Behandlung von Sarkomen (Frankreich)

Sarkome sind seltene und facettenreiche Tumore, die Weichteile, Knochen und Knorpel betreffen. In Frankreich wird sie jedes Jahr bei fast 4.000 Menschen diagnostiziert; die Überlebensprognose ist schlecht. 40 Jahre klinische Forschung haben es ermöglicht, die meisten Behandlungsempfehlungen zu erstellen. Aufgrund der zu geringen Größe der Kohorten, verbunden mit der Seltenheit dieser Krebsarten, konnten bestimmte Fragestellungen mit diesem Ansatz jedoch nicht beantwortet werden. Ein datenbasierter Ansatz, der auch rückblickend die vorhandenen Daten aller Patient:innen mit Sarkom berücksichtigt, die in einer Datenbank (NetSarc) in Frankreich gesammelt wurden, soll neue Erkenntnisse bringen.⁷⁰

Der Datenaustausch über einen kollaborativen Datenraum Gesundheitswesen verbessert auch die Patientensicherheit, etwa wenn umfangreichen Daten zu Arzneimittelwechselwirkungen analysiert oder die Gefahren von Begleiterscheinungen von Krankenhausaufenthalten (z. B. akutes Nierenversagen) schneller erfasst werden können. Da für die Anwendungsfälle in Quadrant B sowohl Primär- als auch Sekundärdaten aus der Versorgung, u. a. personenbezogene Patientendaten, benötigt werden und entsprechende Anpassungen des regulatorischen Rahmens erforderlich sind, sind diese Anwendungsfälle voraussichtlich erst in der letzten Phase des Datenraums Gesundheitswesens umsetzbar.

Quadrant C: Schaffung neuer Prozesse bei administrativen Abläufen

Neben der Verbesserung bestehender Prozesse auf administrativer Ebene und in der medizinischen Versorgung schafft ein weiterer Teil der Anwendungsfälle gänzlich neue Prozesse auf administrativer Ebene. Der Innovationsgrad ist entsprechend hoch. Dabei geht es

⁷⁰ DeepSarc (2019)

um neue Möglichkeiten wie die automatisierte Identifikation von Patienten-ID und Hilfsmitteleriennummer für Hilfsmittelhersteller, die Verknüpfung von Datenquellen zur Sekundärdatennutzung oder die Nutzung des Datenraums Gesundheitswesen für den Katastrophenschutz. Bei den meisten der hier einzuordnenden Anwendungsfälle handelt es sich um den Aufbau grundlegender Strukturen, auf deren Basis wiederum weitere neue Anwendungsfälle kreiert werden können. Ein Marktplatz, der den Austausch und die Monetarisierung von Daten unter Wahrung der Privatsphäre und der Datensouveränität ermöglicht, würde bspw. den Anreiz erhöhen, qualitativ hochwertige Daten zu erheben und anzubieten. Davon können wiederum Anwendungsfälle mit einer größeren Versorgungstiefe profitieren. Für Anwendungsfälle aus diesem Quadranten werden in der Regel keine Primärdaten aus der Versorgung benötigt, jedoch ist der Zugang zu Sekundärdaten wie Kassendaten bspw. über das Forschungsdatenzentrum oder die gesetzlichen Krankenkassen erforderlich. Es ist zu erwarten, dass der Zugang für Unternehmen zu dieser Art von Daten früher realisiert werden kann als zu Primärdaten. Aus diesem Grund kommen solche Anwendungsfälle in der zweiten Phase des kollaborativen Datenraums Gesundheit in Betracht.

Quadrant D: Schaffung neuer Prozesse in der medizinischen Versorgung

Zu guter Letzt konnten Anwendungsfälle identifiziert werden, die neue Prozesse in der medizinischen Versorgung schaffen. Diese haben die größte Versorgungstiefe und den höchsten Innovationsgrad. Hier sind Anwendungsfälle zu finden, die Patient:innen permanent im Alltag begleiten – auch bereits bevor sie den Status als „Patient:in“ besitzen. Der digitale Zwilling als datenbasiertes Abbild der zu behandelnden Person ermöglicht bspw. über die Erfassung multidimensionaler Gesundheitsdaten ein umfassenderes Bild des individuellen Gesundheitszustands zu zeichnen und damit die Durchführung von Präventivmaßnahmen zu beginnen, bevor Krankheiten überhaupt ausbrechen. Andere, KI-basierte Dienste erlauben bspw. die Überprüfung der Verträglichkeit von Medikamenten auf Basis des persönlichen genetischen Hintergrunds. Entsprechend lässt sich sagen, dass Anwendungsfälle in diesem Quadranten ganz konkret die Personalisierung der Medizin durch die Kombination aller verfügbaren Gesundheitsdaten eines einzelnen Individuums über den Datenraum Gesundheitswesen umsetzen können. Damit liegen in diesem Quadranten die Anwendungsfälle, die die in Kapitel 3.1 skizzierte Vision des kollaborativen Datenraums Gesundheitswesens Wirklichkeit werden lassen. Entsprechend ist die Umsetzung erst in der letzten Phase zu erwarten.

Unabhängig von der zuvor vorgenommenen Einteilung in Quadranten können weitere Cluster von Anwendungsfällen beobachtet werden. Diese schließen sich nicht gegenseitig aus, sondern es können auch mehrere dieser Cluster auf einen Anwendungsfall zutreffen.

So werden die Vernetzung und der Austausch von Daten zum Enabler für Anwendungen mit **Künstlicher Intelligenz (KI)**. Dies verwundert nicht, da das Training von Algorithmen eine große Datenbasis voraussetzt. Je einfacher diese Daten ausgetauscht werden können, desto größer und aussagekräftiger wird auch der Trainingsdatensatz. Wie die Anwendungsfälle zeigen, bietet KI vor allem in der medizinischen Versorgung große Potenziale. KI-unterstützte Diagnostik entlastet Ärzt:innen, verringert Fehleinschätzungen und ermöglicht patientenindividuelle Therapien.

Hinzu kommen Anwendungsfälle, die auf bisher kaum genutzte Daten aus dem Alltag der Patient:innen zurückgreifen. Diese „**Real World Data**“ entstehen im Rahmen der medizinischen Versorgung, werden aber auch per Apps und / oder Wearables direkt von den Betroffenen erhoben. Diese Daten haben die Gemeinsamkeit, dass sie im Vergleich zu den kontrollierten Bedingungen in klinischen Studien die tatsächliche Realität im Behandlungsalltag abbilden. Mit der Verbindung und Vernetzung von Langzeit-Selbsttracking im Alltag, medizinproduktespezifisch relevanten Daten aus dem primären Gesundheitsmarkt (z. B. von Implantaten) und ePA-Daten über einen Datenraum Gesundheitswesen

können so z. B. neue Gesundheitsdienste entstehen, die die ärztliche Diagnostik und Therapie verbessern und für ein gesundes Altern im Alltag sorgen.

Beispiel für die Nutzung von „Real World Data“

Etwa 75 Prozent der Menschen mit chronischen Schmerzerkrankungen, wie z. B. Arthritis, glauben, dass das Wetter ihre Schmerzen beeinflusst. Es ist jedoch nicht klar, welche Wetterbedingungen den größten Einfluss haben. Das britische Projekt „Cloudy with a chance of pain“ ging dieser Frage unter der Zuhilfenahme von Real World Daten nach. Die am Projekt teilnehmenden Personen zeichneten ihre tägliche Schmerzintensität in einer Smartphone-App auf. Die GPS-Standorte ihrer Telefone wurden dann mit lokalen Wetterdaten verknüpft. Bei der Analyse von 5,1 Millionen Schmerzberichten verglichen die Forscher:innen bei jeder Person das Wetter an Tagen, an denen eine signifikante Zunahme der Schmerzen auftrat, mit dem Wetter an Tagen, an denen keine solche Zunahme der Schmerzen auftrat. Sie fanden heraus, dass Tage mit höherer Luftfeuchtigkeit, niedrigerem Luftdruck und stärkerem Wind signifikant eher mit starken Schmerzen verbunden sind.⁷¹

Ein weiteres Szenario, welches über einen Datenraum Gesundheitswesen ermöglicht wird und welches die Anwendungsfälle demonstrieren, ist die **Sekundärnutzung von Gesundheitsdaten**. Über die Verknüpfung von unterschiedlichen Datenquellen wie Gesundheitsdaten aus Routineversorgung, Kassendaten, Registerdaten etc. können neue Erkenntnisse über Krankheiten gewonnen werden. In Frankreich ist die Analyse bestehender Datenbestände über den „Health Data Hub“⁷² bereits Realität. Die dort verfügbaren Datenbanken etwa des „National Health Data System“ können für verschiedene Forschungsprojekte genutzt werden. Ein Datenraum Gesundheitswesen kann dies perspektivisch ebenfalls abbilden.

Wie bereits zu Quadrant C angemerkt, gibt es auch Anwendungsfälle, die als **Enabler für weitere Anwendungsfälle** fungieren. Diese nutzen die Gegebenheiten des Datenraums Gesundheitswesen, um Strukturen zu etablieren, auf denen weitere Analysen und daraus resultierende Anwendungsfälle entstehen können. Beispiele sind hier das Pooling von Facharzt- und Krankenhausdaten, um Mengen zu erfahren, Versorgungsabläufe zu erkennen etc. oder eine Plattform für Patientenprofile mit seltenen Erkrankungen, um Probanden für Studien akquirieren zu können.

3.2.2 Merkmale relevanter Anwendungsfälle

Wie sich gezeigt hat, konnten in einer ersten Recherche zahlreiche Anwendungsfälle im Gesundheitswesen identifiziert werden, die sich mit einem kollaborativen Datenraum Gesundheitswesen umsetzen lassen. Dabei ist zu erwarten, dass sich mit einem großangelegten Vorgehen und strukturierten Ideen- und Iterationsphasen in unterschiedlichen Indikationen oder Unternehmensfunktionen ein noch umfassenderes Bild ergeben würde. Auch die letztendliche Einteilung der Anwendungsfälle in die einzelnen Quadranten kann abschließend erst nach einer detaillierten Ausarbeitung des jeweiligen Falls erfolgen. Trotz alledem lässt sich bestätigen, dass unter den recherchierten Anwendungsfällen genügend Fälle zu finden sind, die auf die Vision einer optimierten und personalisierten Gesundheitsversorgung einzahlen. Viele der Anwendungsfälle sind allerdings in der Umsetzung sehr komplex und nicht für jede Phase des kollaborativen Datenraums Gesundheitswesen geeignet (siehe Abbildung 21). Die Komplexität variiert innerhalb aller Quadranten von Fall zu Fall, jedoch ist zu beobachten, dass Anwendungsfälle mit großer Versorgungstiefe (Quadrant B und D) im Allgemeinen komplexer sind als Anwendungsfälle mit geringer Versorgungstiefe (Quadrant A und C). Dies liegt an der Anzahl der benötigten Stakeholder, die in der medizinischen Versorgung neben den industriellen Playern auch Krankenversicherer, Versorger, Patient:innen und die

⁷¹ Dixon et al. (2019)

⁷² Health Data Hub (2022)

gematik einschließt. Zudem werden primäre und sekundäre Daten aus der Versorgung benötigt, um die Anwendungsfälle umsetzen zu können. Diese Komplexität, die nicht nur ein Engagement der beteiligten Akteure, sondern auch veränderte Rahmenbedingungen in der Gesetzeslage durch die Politik erfordert, führt dazu, dass entsprechende Anwendungsfälle (unabhängig vom Innovationsgrad) voraussichtlich verzögert umsetzbar sind.



Abbildung 21: Quadranten in Relation zu den Phasen des Datenraums

Auch Anwendungsfälle mit geringer Versorgungstiefe aber hohem Innovationsgrad (Quadrant C) sind zunächst schwieriger umzusetzen, da ebenfalls Daten von weiteren Stakeholdern wie anderen Datenräumen benötigt werden. Hier sind insbesondere Sekundärdaten erforderlich, um bspw. Anwendungsfälle wie das Pooling von Facharzt- und Krankenhausdaten umsetzen zu können. Da allerdings Versicherer, Versorger, Patient:innen oder die gematik zum Großteil noch nicht oder nur indirekt benötigt werden, sind Anwendungsfälle dieses Quadranten bereits in Phase 2 des Datenraums Gesundheitswesen denkbar.

Daraus ergibt sich, dass Anwendungsfälle mit geringer Versorgungstiefe und niedrigem Innovationsgrad für Phase 1 geeignet sind. Hier werden vor allem Daten, über die die beteiligten Stakeholder Pharma, MedTech und Industriepartner selbst verfügen, benötigt. Zudem sind die Prozesse, die abgebildet werden, bereits aus der Praxis bekannt. Der Einstieg sollte daher hier niederschwellig sein.

3.3 Kollaborations- und Organisationsstruktur

Für den Aufbau des Datenraums Gesundheitswesen und die Umsetzung der Vision und Strategie wird ein organisatorisches Gerüst benötigt. Wie Kapitel 2.3 gezeigt hat, erfordert ein Datenraum nach IDS-

Standard neben den Data Providern und Data Consumern sowie der technischen Komponenten „Konnektor“ einen Identity Provider sowie eine Zertifizierungs- und Prüfstelle, um arbeitsfähig zu sein. Diese beiden Rollen sind essenziell, um Vertrauen in das Ökosystem zu stärken. Es ist daher nicht unerheblich, wer die Zertifizierung der Mitglieder des Datenraums übernimmt und überprüft. Im „IDSA Rule Book“ beschreibt die IDSA hierfür prototypisch eine „IDSA Support Organisation“, die den Betrieb des Datenraums sicherstellen soll.⁷³ Diese agiert nicht nur als Zertifizierungsstelle, sondern organisiert die essenziellen Dienste, die für den Betrieb des Datenraums notwendig sind. Zu den Aufgaben des „Essential Service Provider Managements“ zählen das Verwalten der Geschäftsordnung, Durchführung von RfQ, Auswahl, Zuordnung, On- und Offboarding von Anbietern, Rechnungsstellung sowie die Einrichtung einer Kontaktstelle für allgemeinen Support. Weitere Rollen, die nicht zwingend für einen Datenraum benötigt werden, die aber dessen Anwenderfreundlichkeit erheblich verbessern, sind Broker, App Store, Clearingstelle und Vocabulary Provider (siehe Kapitel 2.3.3). Jede dieser Rollen kann durch die Support Organisation oder durch einzelne Teilnehmer des Datenraums übernommen werden. Die Praxis zeigt jedoch, dass die Komplexität eines Datenraums neben dem reinen Betreiber eine weitere organisatorische Einheit erfordert bzw. dass die genannten Aufgaben auch unterschiedlich verteilt werden können, etwa die Zertifizierung betreffend.

Wie bereits in Kapitel 2.2.2 herausgearbeitet, gibt es unterschiedliche Möglichkeiten einen Datenraum zu orchestrieren. Unter anderem ist hier die Einrichtung einer neutralen Dachorganisation zu beobachten. Im Falle des Mobility Data Space wurde bspw. die „DRM Datenraum Mobilität GmbH“ als neutrale Non-Profit-Organisation gegründet. Sie hat die Aufgabe, den Mobility Data Space weiter aufzubauen und ihn technisch sowie kommerziell aufeinander abzustimmen. Die Arbeit des Datenraums der Automobilindustrie, Catena-X, wird durch den Verein „Catena-X Automotive Network e.V.“ organisiert. Hier soll zusätzlich noch eine Betreibergesellschaft für den kommerziellen Betrieb gegründet werden.⁷⁴ Die Beispiele haben gemeinsam, dass sie die Gründung einer übergeordneten Netzwerkplattform demonstrieren, über die sich beteiligte Stakeholder über die gemeinsamen Ziele austauschen, die Gewinnung weiterer Partner vorantreiben und die Schaffung neuer Rahmenbedingungen fördern können (siehe Abbildung 22). Hier werden auch die Richtlinien für die Zertifizierung neuer Lösungen definiert, sodass das Netzwerk als Zertifizierungsstelle im Sinne des IDSRAM angesehen werden kann. Nach außen fungiert das Netzwerk als zentraler Ansprechpartner und leistet Aufklärungsarbeit, z. B. über eine Website und Social-Media-Aktivitäten.

Die aktuell bestehende Projektgruppe innerhalb der BDI Initiative Gesundheit digital kann ein Ausgangspunkt für eine solche Netzwerkplattform für den kollaborativen Datenraum Gesundheitswesen darstellen. Um auch in der Praxis als Netzwerkplattform agieren zu können, benötigt es allerdings noch weitere Schritte der rechtlichen Formalisierung. Neben der oben angeführten Gründung eines Vereins oder einer GmbH gibt es alternative Möglichkeiten, entsprechenden Netzwerken einen Rahmen zu geben. So sind etwa die Gründung einer gGmbH oder einem loserem, auf Absichtserklärungen der teilnehmenden Unternehmen basierendem Netzwerk ebenfalls legitime Optionen. Wichtig ist, dass für klare Entscheidungswege gesorgt wird, um zeitnah mit der Umsetzung des kollaborativen Datenraums Gesundheitswesen beginnen zu können. Dazu braucht es im besten Fall einen zentralen Steuerungskreis, dem einzelne Arbeitsgruppen untergeordnet werden. Letztere beschäftigen sich mit inhaltlichen Fragen der konkreten Ausgestaltung einzelner Elemente des kollaborativen Datenraums Gesundheitswesen. So werden lange Entscheidungsprozesse, wie sie aktuell bei einzelnen Datenräumen zu beobachten sind, vermieden.

⁷³ Steinbuss et al. (2021)

⁷⁴ Catena-X (2021), S.15

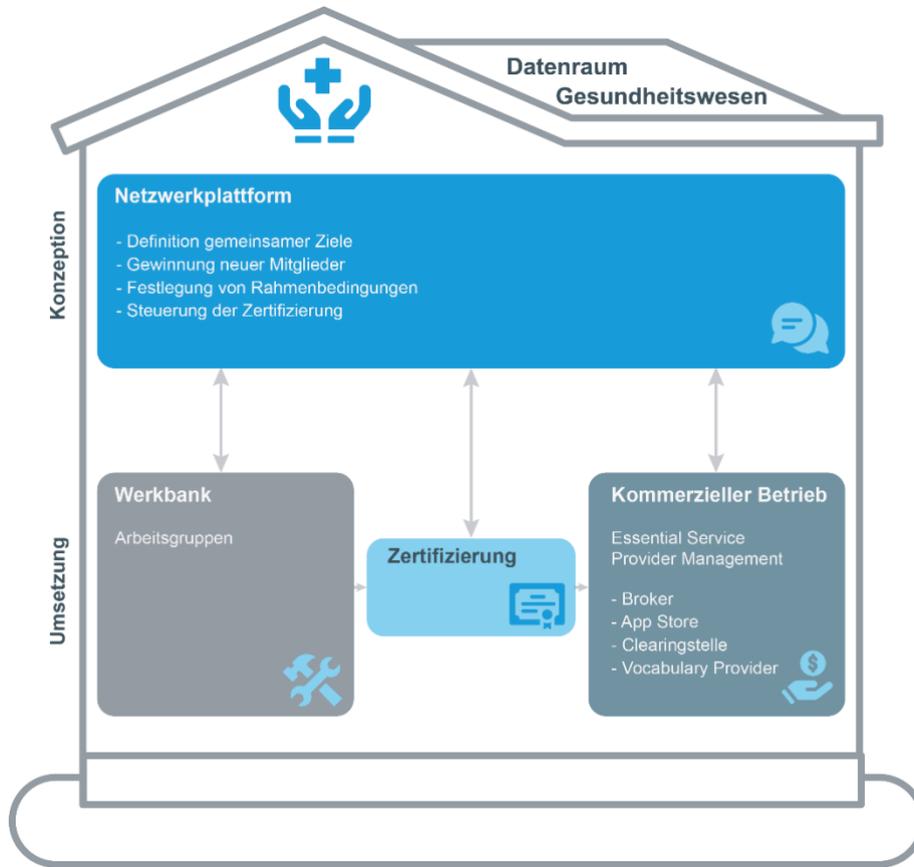


Abbildung 22: Kollaborationsstruktur eines Datenraums Gesundheitswirtschaft

4. Fazit und Ausblick

Wie in Kapitel 2 dargestellt wurde, sind die Potenziale des Data Sharings für Unternehmen enorm. Branchenübergreifend können diese von Effizienzsteigerungen oder der Erschließung neuer Geschäftsfelder profitieren, auch wenn es noch einige Herausforderungen gibt. Speziell im Gesundheitswesen kann der Datenaustausch über alle relevanten Stakeholder hinweg zu einer Optimierung der medizinischen Versorgung bis hin zu einer personalisierten Gesundheitsversorgung führen, wovon sowohl einzelne Patient:innen als auch die Gesellschaft als Ganzes profitieren. Hier gilt es, neben regulatorischen Hemmnissen vor allem Fragen des Vertrauens, der digitalen Verantwortung und der Datensouveränität zu klären. Auf der Suche nach aktuellen Möglichkeiten, Daten zwischen unterschiedlichen Stakeholdern auszutauschen und somit die aufgezeigten Potenziale auszuschöpfen, wurde eine Marktanalyse bestehender Data Sharing-Lösungen durchgeführt. Diese brachte eine Differenzierung in Framework Initiativen und Data Sharing Ökosysteme zu Tage. Während Erstere den technologischen und strukturellen Rahmen für das Teilen von Daten entwickeln, setzen Letztere diesen Rahmen in der Praxis um. Dabei zeigte die Analyse der Merkmale von Data Sharing Ökosystemen, dass diese einer der zwei identifizierten Typen „Datenplattform“ und „Datenraum“ zuzuordnen sind. Datenplattformen bringen selbst einige Einschränkungen mit sich, Datenräume hingegen können bei einer Vielzahl der zuvor identifizierten Herausforderungen eine Lösung darstellen. Neben der Unterstützung durch die Politik fördern diese Interoperabilität, sichern Datensouveränität, können private und öffentliche Stakeholder vernetzen und bleiben aufgrund ihrer offenen technologischen Basis anpassungsfähig. Mit dem Referenzarchitekturmodell der International Data Spaces Association (IDSA) liegt ein umsetzungsbereites Konzept für den Aufbau eines solchen Datenraums vor. Dieses kann die Basis für den „kollaborativen Datenraum Gesundheitswesen“ bilden, der in Kapitel 3 diskutiert wurde. Für diesen konnte in strukturierten Interviews und Gesprächen mit Mitgliedern der BDI Initiative Gesundheit digital sowie weiteren Vertreter:innen von Verbänden, Selbstverwaltungsorganen und der Industrie eine erste Vision und Strategie aufgezeigt werden, die die Optimierung der medizinischen Versorgung bis hin zu einer personalisierten Gesundheitsversorgung zum Ziel hat. Dieses Ziel soll schrittweise anhand von drei Phasen erreicht werden. Eine umfangreiche Sammlung möglicher Anwendungsfälle für den kollaborativen Datenraum Gesundheitswesen zeigte, dass in jeder dieser Phasen andere Use Cases möglich werden. Um die Vision und Strategie anhand erster Anwendungsfälle umzusetzen, benötigt es jedoch eine Kollaborations- und Organisationsstruktur, die den Aufbau einer gemeinsamen Netzwerkplattform für Stakeholder des Gesundheitswesens beinhaltet. Trotz konzeptionellem Rahmen bleiben jedoch noch einige Detailfragen in Bezug auf die Umsetzung eines kollaborativen Datenraum Gesundheit offen, die im Zuge einer Wirtschaftsplanung erörtert werden müssen. Zudem gilt es, Öffentlichkeit und Politik von der Vision des kollaborativen Datenraums Gesundheitswesen zu überzeugen. Wie schnell sich aktuelle Rahmenbedingungen auch zu Gunsten eines kollaborativen Datenraums Gesundheitswesen verändern können, zeigt ein Blick in das Vereinte Königreich: Das staatliche Gesundheitssystem Englands, der National Health Service (NHS), kündigte im April 2022 mit einer Ausschreibung Pläne für eine föderierte Datenplattform an.⁷⁵ Dieses wird sich auf die fünf Hauptanwendungsfälle Volksgesundheit, Koordinierung der Pflege, elektive Genesung, Impfstoffe und Immunisierung sowie Lieferketten konzentrieren. Für die Ausschreibung werden insgesamt 240 Millionen Pfund zur Verfügung gestellt.

⁷⁵ Healthcare IT News (2022)

5. Literaturverzeichnis

- Aggarwal, S.; Kumar, N. (2021): Chapter Sixteen - Hyperledger☆☆Working model. In: Aggarwal, S.; Kumar, N. und Raj, P. (Hrsg.): Advances in Computers: The Blockchain Technology for Secure and Smart Applications across Industry Verticals, Bd. 121: Elsevier, S. 323–343. URL: <https://www.sciencedirect.com/science/article/pii/S0065245820300711>.
- Ankowitsch, E. (2022): EU-Gesundheitsdatenraum: Mega-Projekt läuft an. URL: <https://table.media/europe/analyse/eu-gesundheitsdatenraum-mega-projekt-laeuft-an/> (abgerufen am 14.03.2022)
- Bitkom (2021): Drei Viertel der Deutschen wollen elektronische Patientenakte nutzen. URL: <https://www.bitkom.org/Presse/Presseinformation/Drei-Viertel-wollen-elektronische-Patientenakte>
- BMDV (2021): BMDV fördert nächste Phase des „Mobility Data Space“. URL: <https://www.BMDV.de/SharedDocs/DE/Pressemitteilungen/2021/132-scheuer-BMDV-foerdert-naechste-phase-des-mobility-data-space.html> (abgerufen am 31.03.2022)
- BMWK (2020): Plattform Industrie 4.0, Diskussionspapier, Anonymisierung im Datenschutz als Chance für Wirtschaft und Innovationen.
- BMWK (2021): Eine Milliarde Euro für die Zukunft der Automobilindustrie - Expertenausschuss legt Förderempfehlungen zum Zukunftsfonds Automobilindustrie vor. URL: <https://www.bmwk.de/Redaktion/DE/Pressemitteilungen/2021/08/20210816-1-mrd-euro-fur-die-zukunftsthemen-der-automobilindustrie-expertenausschuss-legt-seine-forderempfehlungen-vor.html>
- Catena-X (2021): Catena-X Automotive Network - Building the First Operating System for a Data Driven Value Chain -. URL: https://catena-x.net/fileadmin/user_upload/intro_praesentationen/catena-x_im_ueberblick.pdf (abgerufen am 22.04.2022)
- Catena-X (2022): Mitgliederliste. URL: https://catena-x.net/fileadmin/user_upload/downloads/20220221_catena-x_aktuelle_mitgliederliste_de.pdf (abgerufen am 18.04.2022)
- Companies Market Cap (2022): Largest Companies by Market Cap. URL: <https://companiesmarketcap.com/> (abgerufen am 07.06.2022)
- Data Spaces Business Alliance (2022): The Data Spaces Business Alliance – Unleashing the European Data Economy. URL: <https://data-spaces-business-alliance.eu/> (abgerufen am 07.06.2022)
- DeepSarc (2019): DeepSarc. URL: <https://www.health-data-hub.fr/partenariats/deepsarc> (abgerufen am 18.04.2022)
- Deloitte (2019): Digitalisierung des Gesundheitsmarktes
- Dixon, W.G.; Beukenhorst, A.L.; Yimer, B.B. et al. (2019): How the weather affects the pain of citizen scientists using a smartphone app. npj Digit. Med. 2, 105. <https://doi.org/10.1038/s41746-019-0180-3>
- Eclipse Foundation (2022): Eclipse Dataspace Connector. URL: <https://projects.eclipse.org/projects/technology.dataspaceconnector> (abgerufen am 07.06.2022)
- EFI - Expertenkommission Forschung und Innovation (2022): Gutachten zu Forschung, Innovation und technologischer Leistungsfähigkeit Deutschlands. Gutachten No. 2022. Expertenkommission Forschung und Innovation (EFI), Berlin
- Europäische Kommission (2022a): Data Act Fact Sheet. URL: <https://ec.europa.eu/newsroom/dae/redirection/document/83564> (abgerufen am 11.03.2022)
- Europäische Kommission (2022b): Europäischer Raum für Gesundheitsdaten (EHDS). URL: https://ec.europa.eu/health/ehealth-digital-health-and-care/european-health-data-space_de
- Europäische Kommission (2022c): Datengesetz: Kommission schlägt Maßnahmen für eine faire und innovative Datenwirtschaft vor. Pressemitteilung vom 23. Februar 2022. URL: https://ec.europa.eu/commission/presscorner/detail/de/ip_22_1113 (abgerufen am 11.03.2022)
- European Commission (2019): Towards European Media Sovereignty. An Industrial Media Strategy to leverage Data, Algorithms and Artificial Intelligence. URL: https://ec.europa.eu/info/sites/default/files/gk_report_final.pdf (abgerufen am 02.05.2022)

- European Commission (2020): Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European data governance (Data Governance Act)
- European Commission (2022): The EU values. URL: <https://ec.europa.eu/component-library/eu/about/eu-values/> (abgerufen am 02.05.2022)
- Fortuna, G. & Bertuzzi, L. (2022): LEAK: Die Pläne der EU-Kommission für einen Gesundheitsdatenraum. URL: <https://www.euractiv.de/section/innovation/news/leak-die-plaene-der-eu-kommission-fuer-einen-gesundheitsdatenraum/> (abgerufen am 14.03.2022)
- Gaia-X (2021): Vision & Strategy. URL: <https://gaia-x.eu/wp-content/uploads/2021/12/Vision-Strategy.pdf> (abgerufen am 06.04.2022)
- Gothe, H.; Ihle, P.; Swart, E. (2021): Was verstehen wir unter Sekundärdaten? – Ein Grundsatzbeitrag zur terminologischen Einordnung und Definition. In: Gesundheitswesen 2021; 83(S 02): S64-S68, Georg Thieme Verlag, Stuttgart. DOI: 10.1055/a-1686-8936
- Health and Food Safety Directorate General (2022): EU Health Policy Platform Holds Annual Meeting. URL: <https://ec.europa.eu/newsroom/sante/newsletter-archives/38163> (abgerufen am 25.04.2022)
- Health Data Hub (2022): Nos offres de service. URL: <https://www.health-data-hub.fr/> (abgerufen am 19.04.2022)
- Healthcare IT News (2022): NHS England announces plans for £240M federated data platform. URL: <https://www.healthcareitnews.com/news/emea/nhs-england-announces-plans-240m-federated-data-platform> (abgerufen am 28.04.2022)
- Hyperledger Foundation (2022): What We Do. URL: <https://www.hyperledger.org/> (abgerufen am 06.04.2022)
- IDSA (2019): IDSA and iSHARE join forces to achieve data sovereignty. URL: <https://internationaldataspaces.org/idsa-and-ishare-join-forces-to-achieve-data-sovereignty/> (abgerufen am 06.04.2022)
- IDSA (2021): The Magic of Data Spaces Now. URL: <https://internationaldataspaces.org/the-magic-of-data-spaces-now/> (abgerufen am 24.04.2022)
- IDSA (2022a): Data spaces overview. Let's build the future data economy! URL: <https://internationaldataspaces.org/download/31819/>
- IDSA (2022b): The International Data Spaces Association on GitHub. URL: <https://github.com/International-Data-Spaces-Association/idsa> (abgerufen am 26.04.2022)
- INFORM (2017): Data Sharing als wichtiger Faktor für die Supply Chain. URL: <https://www.inform-software.de/news/news-details/news/data-sharing-als-wichtiger-erfolgskfaktor-fuer-die-supply-chain>
- Institute for Digital Transformation in Healthcare (2021): "Beam me up, Scotty!" Digitale Transformation der Chirurgie
- iSHARE (2021): iSHARE Scheme. URL: <https://ishareworks.atlassian.net/wiki/spaces/IS/pages/70222191/iSHARE+Scheme>
- KBV (2022): Praxisbarometer Digitalisierung – Schlussfolgerung. URL: https://www.kbv.de/media/sp/Schlussfolgerungen_Praxisbarometer_2021.pdf (abgerufen am 05.05.2022)
- Kolain, M & Molavi, R. (2019): Zukunft Gesundheitsdaten. Wegweiser zu einer forschungskompatiblen elektronischen Patientenakte. Bundesdruckerei GmbH, Berlin
- Lauf, F.; Scheider, S.; Meister, S.; Radic, M.; Herrmann, P.; Schulze, M.; Nemat, A. T.; Becker, S. J.; Rebbert, M.; Abate, C. Konrad, R.; Bartsch, J.; Dehling, T.; Sunyaev, A. (2021): Data Sovereignty and Data Economy - Two Repulsive Forces?. Position Paper. DOI: 10.24406/iss-n-634865
- Lesch W.; Richter G.; Semler S.C. (2022): Daten teilen für die Forschung: Einstellungen und Perspektiven zur Datenspende in Deutschland. In: Richter G.; Loh W.; Buyx A.; Graf von Kielmansegg S. (Hrsg.): Datenreiche Medizin und das Problem der Einwilligung. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-62987-1_11
- Lindner, M; Straub, S.; Kühne, B. (2021): How to share data? Data-Sharing-Plattformen für Unternehmen.

- McKinsey Digital (Hrsg.) (2018): Digitalisierung im Gesundheitswesen: die Chancen für Deutschland.
- Moawad, Ahmed W. MD, Fuentes, David T. PhD; ElBanan, Mohamed G. MD; Shalaby, Ahmed S. MD; Guccione, Jeffrey MD; Kamel, Serageldin MD; Jensen, Corey T. MD; Elsayes, Khaled M. MD (2022): Artificial Intelligence in Diagnostic Radiology: Where Do We Stand, Challenges, and Opportunities, *Journal of Computer Assisted Tomography*: 1/2 2022 - Volume 46 - Issue 1 - p 78-90. doi: 10.1097/RCT.0000000000001247
- Otto, B. (2016): Digitale Souveränität: Beitrag des Industrial Data Space. Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., Munich (2016). doi:10.13140/RG.2.2.35125.68321
- Otto, B.; Steinbuss, S.; Teuscher, A.; Lohmann, S. et al. (2019): Reference Architecture Model. International Data Spaces Association. URL: <https://internationaldataspaces.org/download/16630/>
- Otto, B. et al. (2021): GAIA-X and IDS. International Data Spaces. <https://doi.org/10.5281/zenodo.5675897>
- Otto, B., Burmann, A. Europäische Dateninfrastrukturen. *Informatik Spektrum* 44, 283–291 (2021). <https://doi.org/10.1007/s00287-021-01386-4>
- Paal, B. & Hennemann, M. (2017): "Big Data im Recht – Wettbewerbs- und daten(schutz)rechtliche Herausforderungen," *NJW* (70:24), pp. 1697-1701.
- Plattform Lernende Systeme (Hrsg.): *Lernende Systeme im Gesundheitswesen – Bericht der Arbeitsgruppe Gesundheit, Medizintechnik Pflege*, München 2019.
- Radić, D.; Radić, M.; Metzger, N.; Pohl, A.-S.; Schuldt, N. (2016): Big Data im Krankenversicherungsmarkt. Relevanz, Anwendungen, Chancen und Hindernisse. Fraunhofer-Zentrum für Internationales Management und Wissensökonomie IMW, Leipzig.
- Röhl, K.-H., Bolwin, L., Hüttl, P. (2021): *Datenwirtschaft in Deutschland. Wo stehen die Unternehmen in der Datennutzung und was sind ihre größten Hemmnisse? Gutachten im Auftrag des Bundesverbands der Deutschen Industrie e.V. (BDI), Köln*
- Roland Berger (2020): Future of health 2 – The rise of healthcare platforms. URL: <https://www.rolandberger.com/en/Insights/Publications/Future-of-Health-The-rise-of-healthcare-platforms.html> (abgerufen am 29.04.2022)
- Steinbuss S., Ottradovetz K., Langkau J., Punter M. et al. (2021): IDSA Rule Book. International Data Spaces Association. <https://doi.org/10.5281/zenodo.5658294>
- TEHDAS (2022a): Results. URL: <https://tehdas.eu/results/> (abgerufen am 06.04.2022)
- TEHDAS (2022b): Options for the minimum set of services for secondary use of health data in the EHDS. URL: <https://tehdas.eu/app/uploads/2022/04/tehdas-options-for-the-minimum-set-of-services-for-secondary-use-of-health-data-in-the-ehds.pdf> (abgerufen am 06.04.2022)

Anhang

Advaneo Data Marketplace
AgriFood Data Space Finland
AVIATION Data Hub
AWS Marketplace Data Exchange
BMW CARDATA
Boost 4.0
burstIQ
CARUSO DATAPLACE
Catena-X
City Dataspace
CSDR
DASLOGIS - Dutch Data Spaces for Logistics
Data Broker Global
Data Fusion Studio
Data Intelligence Hub
Datapace
DATEVA
Dawex
DjustConnect
Edison Digital Health Platform
GAIA-X - HEALTH-X dataLOFT
GAIA-X - MERLOT
Honic Medhive
Longgenesis
Maritime Data Space
MARKET 4.0
Mobility Data Space
Mobilitäts Daten Marktplatz
MUSKETEER
oneTRANSPORT Data Marketplace
OTONOMO
PAIRS
Pirate
Silicon Economy
Smart Connected Supplier Network - Market 4.0
SPOCC
TEAM X - Trusted Ecosystem of Applied Medical Data exchange
teampay digital health platform connect
Terbine
UP42
Vendia
Vivli

Tabelle 1: Liste der berücksichtigten Data Sharing Ökosysteme

Danksagung

Das Projekt „Kollaborativer Datenraum Gesundheitswesen – Vorstudie über die konzeptionellen Grundlagen“ entstand unter Förderung und Mitwirken folgender Organisationen:

- Bayer Vital GmbH
- Bitkom e.V.
- Brainlab AG
- Bundesverband Gesundheits-IT - bvitg e.V.
- Bundesverband der Pharmazeutischen Industrie e.V.
- GlaxoSmithKline GmbH & Co. KG
- Johnson & Johnson MEDICAL GmbH
- Lilly Deutschland GmbH
- Medtronic GmbH
- MSD SHARP & DOHME GmbH
- Pfizer Pharma GmbH
- Roche Pharma AG
- Sanofi-Aventis Deutschland GmbH
- Siemens Healthcare GmbH
- Verband Forschender Arzneimittelhersteller e.V.

Ein herzlicher Dank gilt allen Expert:innen, die sich im Rahmen von Gruppen- und Einzelgesprächen an dem Prozess beteiligt haben.

Impressum

idigiT – Institute for Digital Transformation in Healthcare GmbH

Alfred-Herrhausen-Straße 45

58455 Witten

Telefon +49 2302 926 874

www.transforming-healthcare.com

Verantwortlich

Isabel Gadea

Autor:innen

Dr. Sarah J. Becker

Dr. André T. Nemat

Marcel Rebbert

Transparenzhinweis

Dr. André T. Nemat, Managing Partner der Institute for Digital Transformation in Healthcare GmbH, ist seit 2016 ehrenamtliches Vorstandsmitglied des International Data Spaces e. V.