Fraunhofer IAO

Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO

Innovation im Blick

Wertschöpfung durch Data Pooling

Herausforderungen und Erfolgsfaktoren

Vincent Philipp Göbels | Diana Fischer-Pressler | Janika Kutz | Bernd Bienzeisler

Hrsg.: Oliver Riedel | Katharina Hölzle | Wilhelm Bauer | Bernd Bienzeisler

Im Auftrag des

IPAI

Inhalt

1 Einleitung 2

2 Begriffsbestimmung 3
2.1 Data Pool 3
2.2 Data-Sharing-Plattformen zur Nutzung von Data Pools 3

3 Anforderungen an Data-Sharing-Plattformen 4
3.1 Technische Anforderungen 5
3.2 Organisatorische Anforderungen 6
3.3 Sicherheitsrelevante Anforderungen 7
3.4 Regulatorische Anforderungen 8

4 Geschäftsmodelle von Data-Sharing-Plattformen 9

5 Potenziale von Data Pools für KI 12

6 Data-Sharing-Plattformen in der Praxis 14
6.1 International Data Spaces 14
6.2 Gaia X 15
6.3 Servicemeister 16
6.4 Axoom. 16

7 Diskussion und Handlungsempfehlungen zu Kooperationsformaten 17

8 Zusammenfassung und Ausblick 19

9 Literaturverzeichnis 20

1 Einleitung

Der Umfang und die Vielfalt verfügbarer Daten haben aufgrund fortschreitender Digitalisierung exponentiell zugenommen. Dies hat zu einer bedeutsamen Verschiebung in der Art und Weise geführt, wie Unternehmen, Institutionen und Individuen Daten nutzen und verwerten. Ein wesentliches Kennzeichen der Digitalisierung ist, dass ehemals organisatorisch und räumlich getrennte Prozesse über digitale Technologien nahtlos miteinander verbunden werden können. Damit einher gehen neue digitale Service-Produkte und Veränderungen in den industriellen Wertschöpfungsketten. Gleichzeitig steigt die Anzahl der verfügbaren Daten entlang von organisatorischen Leistungsprozessen exponentiell.

Trotz dieser Fülle an Daten und der Potenziale, die mit einer unternehmensübergreifenden Bereitstellung von Daten verbunden sind, stehen der effektiven Nutzung und dem reibungslosen Austausch von Daten – im Sinne eines »Data-Pooling« – in der Praxis zahlreiche Herausforderungen entgegen. Eine zentrale Herausforderung betrifft die Verwaltung, Zugänglichkeit und Verwertung von Daten. Dass Daten als eine der wertvollsten Ressourcen des 21. Jahrhunderts gelten, führt unmittelbar zu der Frage, wie diese Ressourcen effizient genutzt werden können. Besonders wichtig ist dabei das Zusammenlegen von Daten aus verschiedenen Quellen und das Regeln der Zugriffsrechte. In dieser Hinsicht stellt sich die Frage, wie es möglich ist, den Zugriff auf eine Vielzahl von Datenquellen zu erleichtern und gleichzeitig die Integrität und Sicherheit dieser Daten zu gewährleisten.

Data Pooling und Data-Sharing-Plattformen treten als vielversprechende Lösungsansätze auf, um diese Herausforderungen zu bewältigen. Diese Ansätze zielen darauf ab, eine Plattform zu schaffen, in der Daten aus verschiedenen Quellen gesammelt, konsolidiert und gemeinsam genutzt werden können. Dies eröffnet neue Möglichkeiten für datenbasierte Innovationen, eine ganzheitliche Prozessdigitalisierung, die Entwicklung neuer digitaler Service-Produkte, aber auch neue wissenschaftliche Erkenntnisse und unternehmerisches Wachstum.

Jedoch ergibt sich in diesem Kontext die kritische Frage nach der Wirksamkeit und Tragfähigkeit solcher Plattformen. Welche Geschäftsmodelle sind nachhaltig genug, um Anreize für Data-Sharing zu schaffen und gleichzeitig die notwendige Infrastruktur aufrechtzuerhalten? Welche Rolle spielen etablierte Marktteilnehmer und Hyperscaler bei der Gestaltung von Datenökosystemen? Inwiefern werden regulatorische und ethische Anliegen die Zukunft des Data Pooling beeinflussen?

Das vorliegende Whitepaper gibt einen Überblick über das Thema Data Pooling und betrachtet Anwendungen und Einsatzmöglichkeiten von künstlicher Intelligenz (KI) für und auf Data Pools. In diesem Zusammenhang werden die wichtigsten Anforderungen, Bestandteile und Wertversprechen zum Hosting einer Data-Sharing-Plattform dargestellt und hinsichtlich ihrer Eignung für potenzielle Interessenten, insbesondere Unternehmen, beschrieben. Des Weiteren wird erläutert, welche Merkmale von Data-Sharing-Plattformen für Gruppen oder Organisationen, die einen breiten Datenbestand besitzen und diesen monetarisieren möchten, von besonderer Relevanz sein könnten.

2 Begriffsbestimmung

2.1 Data Pool

Laut der Website von IBM (2023) ist ein Data Pool ein zentraler Speicher oder eine Sammlung von Daten, in der Akteure Daten in einem standardisierten Format erfassen, speichern und austauschen können. Diese Daten können aus verschiedenen Datenquellen stammen wie Sensoren, Endgeräten, Softwareanwendungen oder Onlinereviews. Data Pools ermöglichen es also, große Mengen von Daten verschiedener Quellen zusammenzuführen und zu aggregieren, um eine umfassende Sicht auf bestimmte Datenbereiche zu erhalten.

Ein ähnliches Konzept zur Speicherung von Daten beschreibt der Begriff »Data Lake«. Nach John und Misra (2017) besteht die Abgrenzung hier vor allem darin, dass in einem Data Lake auch unstrukturierte Daten enthalten sein können. In der Praxis werden unternehmensinterne Speicher oftmals als Data Lake bezeichnet. Ein Data Pool kann Daten aus diesem Data Lake enthalten, die mit anderen geteilt werden sollen.

2.2 Data-Sharing-Plattformen zur Nutzung von Data Pools

Eine Data-Sharing-Plattform ist die technische und organisatorische Lösung für den Zugriff und den Austausch von Daten. Sie erlaubt es den Nutzenden, Daten zu kaufen und zu verkaufen, Datensätze zu finden und Händler zu ermitteln (Meisel und Spiekermann 2019; Fricker und Maksimov 2017).

Im Gegensatz zu Ansätzen, in denen Daten zentral gesammelt und verwaltet werden, ermöglicht eine Data-Sharing-Plattform den dezentralen Zugriff auf die Daten durch die Nutzenden. Dabei stellt die Plattform sicher, dass der Datenaustausch in einem sicheren Rahmen erfolgt und alle relevanten gesetzlichen und regulatorischen Anforderungen erfüllt werden. Die Plattform kann auch zusätzliche Funktionalitäten wie bspw. eine Zugangskontrolle, Analysetools und eine Integration von KI-Anwendungen bieten.

Data Pools bilden in der Praxis die Grundlage für Data-Sharing-Plattformen, die durch den Austausch von Daten Wertschöpfung erzielen. Eine Data-Sharing-Plattform ermöglicht es, Daten auf eine kooperative Art und Weise zu teilen und durch die Austauschbeziehung von den Daten anderer Partner zu profitieren. Sie eröffnet auch die Möglichkeit, neue Geschäftsmodelle zu schaffen oder bestehende Leistungsangebote auszubauen, die auf einer gemeinsamen Datennutzung basieren. Ebenso können existierende Prozesse optimiert und fundiertere Entscheidungen getroffen werden. Dabei können die Partner entweder kostenfrei oder gegen Bezahlung Daten austauschen. Das Bereitstellen von Daten kann dadurch auch zu einem Geschäftsmodell werden. Neben den vielen Möglichkeiten, Daten in einem Data Pool zu sammeln, gibt es bei der Nutzung der Plattformen auch einige Anforderungen. Datenschutz- und Sicherheitsbedenken, Vertraulichkeitsfragen sowie Fragen der Datenqualität und der Datenintegration sind häufig genannte Probleme. Diese werden in Kapitel 3 vertieft. Beispiele für Data-Sharing-Plattformen sind Catena X oder Manufacturing X. Beide sind Plattformen, die im Rahmen der europäischen Initiative Gaia X entstanden sind. Ihr Ziel ist es, den sicheren Austausch von Daten entlang der Wertschöpfungskette in der Automobilindustrie bzw. zwischen Unternehmen in der Fertigungsindustrie zu fördern. Diese und andere Data-Sharing-Plattformen werden in Kapitel 6 vorgestellt.

3 Anforderungen an Data-Sharing-Plattformen

In Abbildung 1 sind die wichtigsten Anforderungen an Data-Sharing-Plattformen in den vier Bereichen Technologie, Organisation, Sicherheit und Regulatorik zusammengefasst und in den folgenden Abschnitten detaillierter beleuchtet.[1]

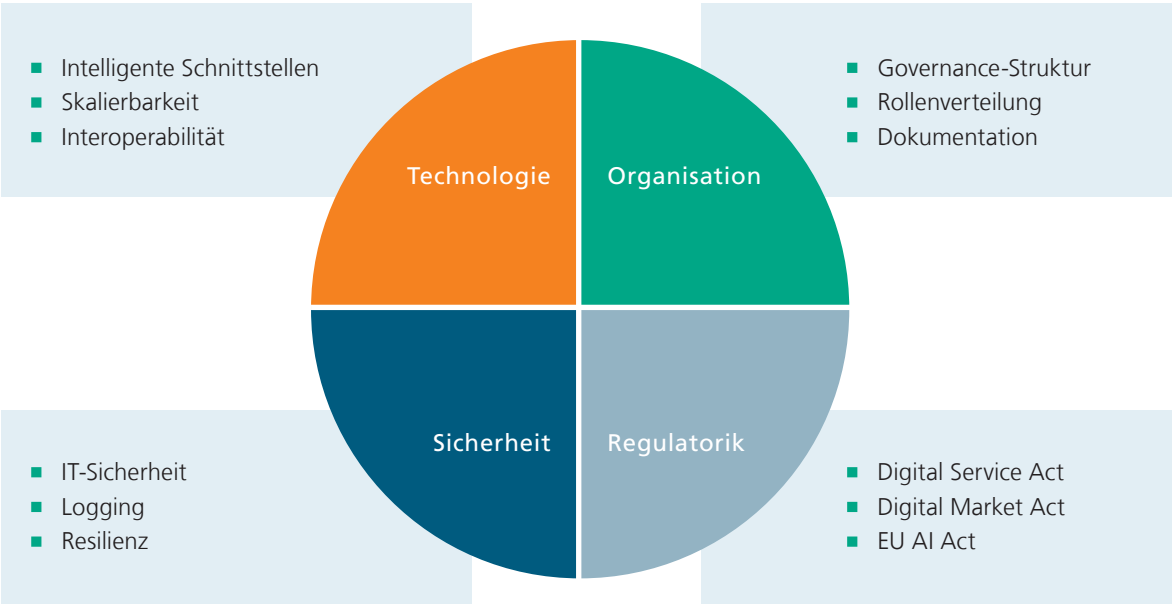

Abbildung 1: Anforderungen an eine Data-Sharing-Plattform im Überblick

1 Für eine tiefergehende Auseinandersetzung mit dem Aufbau von Data-Sharing-Plattformen wird auf Lindner et al. (2021); Otto et al. (2022) und Curry et al. (2022), verwiesen.

3.1 Technische Anforderungen

Jede Data-Sharing-Plattform benötigt eine IT-Infrastruktur, die die technischen Fähigkeiten der Software abbilden kann. Unter anderem basieren Interoperabilität, Bedienbarkeit, Sicherheit und Vertrauenswürdigkeit der Plattform auf der eingesetzten Technologie (Nagel und Lycklama 2022).

Um die Bereitstellung der Daten so unkompliziert wie möglich zu gestalten, sind **intelligente Schnittstellen** erforderlich. Diese ermöglichen eine automatische Klassifizierung von diversen Datenformaten und das Einfügen von Daten in bestehende Strukturen. Bei der Entwicklung einer neuen Plattform sollte die Interoperabilität mit bestehenden Systemen und Plattformen berücksichtigt werden. Im Idealfall können Datenbestände aus verschiedenen Quellen direkt migriert werden, womit die Zusammenarbeit mit anderen Organisationen erleichtert wird. Zudem sollte es möglich sein, periodische Updates der Datenbestände mit einer einfachen Logik durchzuführen. Die Datenlieferanten sollten jederzeit eine Übersicht über ihre Daten sowie die Möglichkeit haben, den Zugriff auf ihre Daten zu verwalten.

Die **Qualität der Daten** bestimmt ihren Wert. Data-Sharing-Plattformen sollten bereinigte, fehlerfreie, aktuelle und vollständige Datensätze bereitstellen (Hazen et al. 2014). Laut Konduru (2017) sind Detailgrad, Zusammenhänge und Vollständigkeit wichtige Einflussfaktoren für präzise Prognosen. KI-gestützte Algorithmen können dabei helfen, fehlerhafte Werte auszusortieren oder durch Interpolation zu ersetzen, Datenformate anzupassen und automatisiert Beschreibungen hinzuzufügen.

Neben qualitativ hochwertigen Daten ist ein schneller und zuverlässiger Zugang zum Data Pool erforderlich. Dafür ist eine **performante Infrastruktur** notwendig. Je nach Nutzungsintention und Geschäftsmodell sind verschiedene Anforderungen an die Plattform umzusetzen. Verfolgt die Data-Sharing-Plattform ein Analytics-as-a-Service-Konzept, wenn also Analysen angeboten werden sollen, muss der Plattformanbieter die entsprechende Rechenleistung bereitstellen. Falls nur die Daten angeboten werden, ist hauptsächlich Speicherplatz erforderlich, während die Schnittstellen für die Datenannahme eine vergleichsweise geringere Rechenleistung benötigen. Zusätzlich zu den bereits genannten Anforderungen an die Infrastruktur muss die Skalierbarkeit der Plattform gewährleistet sein. Auf der Website von VMware (o.D.) wird argumentiert, dass die Möglichkeit, die Rechenleistung und den Speicherbedarf bei Bedarf zu erhöhen, wichtig ist, um auch bei wachsendem Datenvolumen und steigender Anzahl der Nutzenden eine hohe Leistungsfähigkeit sicherzustellen.

Eine weitere wichtige Anforderung ist die **Flexibilität der Plattform**. Insbesondere sollte die Plattform an spezifische Anforderungen anpassbar bzw. erweiterbar sein, um die Integration von neuen Technologien und Datenformaten zu ermöglichen. Open-Source-Technologien können hier besonders vorteilhaft sein, da sie eine hohe Kompatibilität untereinander bieten. Zum Beispiel werden KI-Modelle auf Huggingface in Open-Source-Formaten bereitgestellt (z. B. h5-Format oder Protobuf), so dass diese von möglichst vielen Softwareumgebungen verarbeitet werden können. Dies ermöglicht den Wechsel zwischen Frameworks wie Tensorflow oder PyTorch oder das Einlesen von einem mit Python entwickeltem Modell in der Java-basierten Umgebung Maven.

3.2 Organisatorische Anforderungen

Angemessene Regeln und Richtlinien sind wichtig, um das volle Potenzial aus datengesteuerten Geschäftsökosystemen zu schöpfen (Nagel und Lycklama 2022). Ein zentraler Aspekt ist dabei sicherzustellen, dass Daten fair und transparent geteilt werden. Eine klare **Governance-Struktur** ist hierbei entscheidend, um die Einhaltung von Datenschutz- und Sicherheitsvorschriften sowie ethischen Standards zu gewährleisten. Dies bedeutet, dass die Plattform klare Regeln für den Zugriff auf Daten, ihre Nutzung und ihre Weitergabe festlegen sollte, um sicherzustellen, dass die Daten von allen Nutzenden verantwortungsvoll und angemessen verwendet werden. In der Praxis sind unterschiedliche Systeme im Einsatz.Beispielsweise werden frei verfügbare Daten üblicherweise unter einer Creative-Commons-Lizenz veröffentlicht. Data-Sharing-Plattformen, die den Richtlinien von Gaia X folgen, sehen vor, dass Datenlieferant und Datennutzende einen Datennutzungsvertrag innerhalb eines festgelegten rechtlichen Rahmens aushandeln.[2] Darüber hinaus können Plattformen auch individuelle Verträge mit Mitgliedern aufsetzen, in denen bestimmte Nutzungsrechte eingeräumt werden. Klar definierte Governance-Strukturen sollen sicherstellen, dass die Plattform eine nachvollziehbare und konsistente Richtlinie für den Umgang mit Daten hat, an die sich Nutzende halten müssen. Dies soll das Vertrauen in die Plattform stärken und garantieren, dass die Daten auf eine Weise verwendet werden, die für alle Beteiligten fair und angemessen ist. Eine transparente Kommunikationsstruktur erleichtert zudem die Zusammenarbeit und ermöglicht eine effektive Entscheidungsfindung. Ein Beispiel für klar definierte und umfangreiche Governance-Strukturen sind die Richtlinien von Catena X (Focken 2023).

Um die interne und externe Zusammenarbeit zu fördern, bedarf es darüber hinaus einer klar definierten **Rollenverteilung** und Verantwortlichkeiten. Da dennoch Konflikte, insbesondere bei der Zusammenarbeit mit verschiedenen Stakeholdern, auftreten können, ist es wichtig, ein Konfliktmanagement-System zu etablieren, um Konflikte schnell und effektiv zu lösen.

Eine Data-Sharing-Plattform erfordert zudem eine **verständliche Dokumentation**, um bei der Suche und Nutzung von Daten zu unterstützen. Dazu sollten Metadaten sowie klare Anweisungen für die Nutzung der Plattform bereitgestellt werden.

2 Die Arbeit von Jung und Dörr (2022) gibt einen Einblick in die Konstruktion dieses rechtlichen Rahmens.

3.3 Sicherheitsrelevante Anforderungen

Sicherheit ist ein kritischer Faktor für jede Data-Sharing-Plattform, da sie dazu beiträgt, die Integrität, Vertraulichkeit und Verfügbarkeit der Daten sicherzustellen. Daher ist es wichtig, sicherheitsrelevante Anforderungen zu definieren und umzusetzen, um ein hohes Maß an Datensicherheit zu gewährleisten und das Vertrauen der Nutzenden in die Plattform zu fördern.

Eine Data-Sharing-Plattform sollte ein robustes **Zugangs- und Authentifizierungsmanagement** aufweisen, um sicherzustellen, dass nur Autorisierte Zugang zu den Daten haben. Hierzu gehören auch Passwortrichtlinien, Mehr-Faktor-Authentifizierung und das regelmäßige Überprüfen von Zugriffsrechten. Weiterhin muss die IT der Data-Sharing-Plattform Maßnahmen implementieren, um sich gegen Malware und Cyberangriffe zu schützen. Hierzu gehören Firewall-Schutz, regelmäßige Überprüfung auf Schwachstellen und die Implementierung von Sicherheitsupdates.

Zusätzlich sollte die Plattform über ein **Monitoring- und Logging-System** verfügen, das verdächtige Aktivitäten erkennt und aufzeichnet. Dabei ist es wichtig, dass das System nicht nur auf bereits bekannte Bedrohungen reagiert, sondern auch auf neue oder unerwartete Angriffe. Es bedarf klarer Regeln, welche Aktivitäten auf der Plattform überwacht werden sollen und wie auf verdächtige Aktivitäten reagiert wird. Logging ermöglicht es, sämtliche Aktivitäten auf der Plattform aufzuzeichnen, was bei der Nachverfolgung von Vorfällen hilfreich sein kann. Beim Design des Logging-Systems sollten Aspekte wie Datenschutz und Speicherkapazität berücksichtigt werden.

Die Plattform sollte ferner über Mechanismen verfügen, die sicherstellen, dass die Daten geschützt und nur für autorisierte Zwecke verwendet werden. Dazu gehören Maßnahmen wie Verschlüsselung, Anonymisierung und Pseudonymisierung. Weiterhin muss garantiert sein, dass die Daten ohne Verletzung von Datenschutzrechten erhoben wurden.

Um sowohl technische als auch organisatorische Lücken in der Absicherung einer Data-Sharing-Plattform zu finden, müssen **Audits** durchgeführt werden. Diese sollten in regelmäßigen Zeitabständen stattfinden, um sicherzustellen, dass die Sicherheitsmaßnahmen wirksam sind, und um unautorisierte Zugriffsversuche oder andere Sicherheitsprobleme schnell zu erkennen. Die Auditierung von Zugriffen und Aktivitäten ist ebenfalls wichtig, um die Einhaltung von Compliance-Anforderungen sicherzustellen.

Eine weitere sicherheitsrelevante Anforderung ist die **Resilienz** einer Data-Sharing-Plattform. Diese ermöglicht im Falle von Infrastrukturausfällen, bspw. durch Schäden nach Naturkatastrophen, eine Wiederherstellung der Daten. Mechanismen wie das regelmäßige Erstellen von Backups und die Implementierung von Notfallwiederherstellungsplänen sind dazu unerlässlich.

Abschließend ist die Aufklärung und Sensibilisierung der Nutzenden über die Bedeutung von Sicherheit der Data-Sharing-Plattform zu nennen. Hierzu können Schulungen und Schulungsmaterialien zur Verfügung gestellt werden, um über bewährte Sicherheitspraktiken und -verfahren aufzuklären.

3.4 Regulatorische Anforderungen

Von Seiten des Gesetzgebers werden ebenfalls Anforderungen an Data-Sharing-Plattformen gestellt. Im GAIA-X-Projekt werden Leitbilder geschaffen, daneben existieren aber auch Gesetze zur Regulierung von Plattformen. Der Digital Services Act (DSA) und der Digital Markets Act (DMA) sind zwei Gesetzentwürfe der Europäischen Union, die im Juli 2022 verabschiedet wurden. Ziel der Gesetze ist es, die Regeln für digitale Dienstleistungen und digitale Märkte zu aktualisieren und zu verbessern, um die Rechte der Nutzerinnen und Nutzer zu stärken und den Wettbewerb zu fördern.

Der **DSA** ist ein Gesetzentwurf der Europäischen Union, welcher darauf abzielt, die Verantwortlichkeit von Online-Dienstleistern, einschließlich sozialer Netzwerke, Marktplätze und Suchmaschinen, in Bezug auf die Verbreitung von Inhalten auf ihren Plattformen zu erhöhen. Insbesondere sollen die Plattformen verpflichtet werden, illegale Inhalte effektiver zu entfernen. Des Weiteren sollen die Plattformbetreiber für mehr Transparenz sorgen und Nutzenden mehr Kontrolle über ihre Daten geben.

Der **DMA** zielt darauf ab sicherzustellen, dass große digitale Plattformen wie Google, Amazon oder Facebook den Wettbewerb nicht behindern und ein fairer Wettbewerb ermöglicht wird. Zu diesem Zweck sollen große Digitalunternehmen bspw. verpflichtet werden, Daten mit kleineren Unternehmen zu teilen und Konkurrenten nicht zu benachteiligen.

Online-Plattformen mussten bis zum 17. Februar 2023 ihre Anzahl aktiver Nutzerinnen und Nutzer veröffentlichen. Plattformen mit mehr als 45 Millionen aktiven Nutzenden gelten für den Gesetzgeber als sehr große Plattformen und haben vier Monate Zeit, den Verpflichtungen des **DSA** nachzukommen. Für alle weiteren Plattformen treten die Anforderungen des **DSA** ab dem 17. Februar 2024 in Kraft. Die EU-Mitgliedsstaaten müssen bis zu diesem Zeitpunkt Koordinatoren für digitale Dienste ernennen. Der **DMA** trat am 1. November 2022 in Kraft und Unternehmen müssen der Kommission bis zum 3. Juli 2023 Informationen über ihre Anzahl an registrierten Nutzerinnen und Nutzern übermitteln, damit die Kommission »Gatekeeper« identifizieren kann. Gatekeeper haben bis März 2024 Zeit, um den Verpflichtungen des **DMA** nachzukommen.

Die beiden Gesetzentwürfe haben das Potenzial, das digitale Ökosystem in Europa grundlegend zu verändern, indem sie die Regeln für digitale Dienstleistungen und digitale Märkte aktualisieren und verbessern. Die Umsetzung dieser Gesetze könnte jedoch zu einer erhöhten Regulierung und Überwachung sowie zu technischen Herausforderungen für Plattformbetreiber führen. Die bisherigen Unterkapitel zeigen jedoch bereits auf, dass eine technisch skalierbare Lösung mit intelligenten Schnittstellen und Sicherstellung von Privatsphäre sowie Datensicherheit eine hohe Relevanz für Data-Sharing-Plattformen hat.

Werden bei einer Data-Sharing-Plattform KI-Anwendungen eingesetzt, z. B. zur automatisierten Datenanalyse, ist künftig die Konformität mit dem **EU AI Act** zu gewährleisten (Europäische Kommission 2020). Ende April 2024 soll der Gesetzesentwurf förmlich verabschiedet werden und die darin formulierten rechtlichen Anforderungen an KI-Anwendungen, die innerhalb der EU eingesetzt werden sollen ab 2026 gelten (Europäischer Rat 2023; Thogmartin und Orth 2023). Der EU AI Act sieht vor KI-Anwendungen nach dem Ausmaß ihrer potenziellen Gefahren in Risikoklassen einzustufen. Entsprechend der Risikoklassifizierung einer KI gelten unterschiedliche Anforderungen: Die Mehrheit der KI-Anwendungen wird nach aktueller Betrachtung als »low risk« eingestuft und unterliegt daher nur geringer Regulierung. Hingegen gelten für „high risk" Anwendungen, bei denen sensible Daten verarbeitet werden oder die KI sicherheitsrelevante Entscheidungen treffen kann, spezifische Anforderungen an Fairness, Autonomie und Kontrolle, Transparenz, Verlässlichkeit, Sicherheit und Datenschutz. Darüber hinaus existieren KI-Anwendungen die gänzlich innerhalb der EU verboten werden. Die Publikation „KI-Zertifizierung und Absicherung im Kontext des EU AI Act" (Kutz et al. 2023) liefert einen tiefergehenden Diskurs des EU AI Acts.

4 Geschäftsmodelle von Data-Sharing-Plattformen

Data-Sharing-Plattformen ermöglichen es Datenanbietern, ihre Datenbestände zu monetarisieren und bieten Datenkonsumenten eine bequeme und effiziente Möglichkeit auf relevante Daten zuzugreifen. Zu den Hauptaktivitäten von Data-Sharing-Plattformen gehören der Aufbau und die Wartung der Plattforminfrastruktur, der Erwerb und die Pflege von Datensätzen, die Gewährleistung der Datenqualität und -sicherheit, die Entwicklung und Verbesserung von Benutzeroberflächen, der Aufbau von Partnerschaften mit Datenanbietern, die Vermarktung und Werbung für die Plattform sowie die Erleichterung von Transaktionen zwischen Datenkäufern und -verkäufern.

Abbildung 2 zeigt sieben typische Dimensionen und deren mögliche Ausprägungen für Geschäftsmodelle von Data-Sharing-Plattformen. Die nachfolgenden Erläuterungen orientieren sich entlang dieser Dimensionen.

Data-Sharing-Plattformen bieten ein **Wertversprechen**, indem sie Zugang zu einer breiten Palette von Datensätzen und/oder datennahen Dienstleistungen bieten. Generell können zwei Arten unterschieden werden: transaktionsorientierte oder datenorientierte Plattformen. Transaktionsorientierte Plattformen treten als Vermittler von Datengütern auf, entweder durch die Bereitstellung der notwendigen Infrastruktur oder durch direkte Vermittlung der Daten. Datenorientierte Plattformen bieten neben der Infrastruktur auch Werkzeuge zur Datenanalyse und -aufbereitung an und treten als Anbieter von datennahen Dienstleistungen auf (wie z. B. »Analytics-as-a-Service« Plattform Servicemeister (Timmermann o.D.)).

Die **Beteiligungsart** gibt Aufschluss über die Unabhängigkeit der Plattform. Datenmarktplätze können von denselben Akteuren betrieben werden, die auch im direkten Datenhandel tätig sind, oder von neutralen Akteuren, die weder auf der Anbieter- noch auf der Nachfrageseite stehen. In der Regel betreiben vor allem größere Unternehmen ihre eigenen Datenplattformen, um den regelmäßigen Datenaustausch mit anderen Akteuren zu managen. Kleinere Unternehmen nutzen aus Kostengründen hingegen eher neutrale Plattformen.

Dimension	**Ausprägung**			
Wertversprechen	Transaktionsorientiert	Datenorientiert		
Beteiligungsart	Datenlieferant	Neutral		
Marktzugang	Geschlossen	Hybrid	Offen	
Datenintegration	Domänenspezifisch	Domänenunspezifisch		
Verarbeitungsgrad	Rohdaten	Aufbereitete Daten		
Architektur	Zentral	Hybrid	Dezenztral	
Erlösmodel	Abonnoment Progressiver Preis	Freemium Mitgliedschaften /freie Nutzung, Tausch)	Pay-per-Use etc.	Paketpreis

Abbildung 2: Dimensionen zur Klassifikation von Data-Sharing-Plattformen (eigene Darstellung in Anlehnung an Spiekermann (2019); Engelhardt et al. (2017) und Lindner et al. (2021))

Der **Marktzugang** definiert den Grad der Offenheit einer Plattform. Während eine geschlossene Plattform möglicherweise höhere Sicherheits- und Vertrauensaspekte für die Nutzenden bietet, können offene Datenplattformen eine größere Vielfalt an Datenquellen einbeziehen und somit potenziell einen höheren Mehrwert für die Nutzenden der Plattform generieren. Dies erhöht zwar die Aktivitäten auf der Plattform durch eine größere Zielgruppe, reduziert gleichzeitig aber auch die Kontrolle über die Qualität und Nutzung der Datenprodukte. Mischformen ermöglichen den Austausch zwischen ausgewählten Akteuren und öffnen gleichzeitig einen Teil ihrer Plattform für Neueinsteiger (Spiekermann 2019).

Eine weitere Dimension bezieht sich auf die **Datenintegration**. Diese bestimmt die Dateninhalte, die auf der Plattform verfügbar sind. Plattformen können zwischen domänenspezifisch und -unspezifisch unterschieden werden. Domänenspezifische Plattformen bieten vor allem Daten, die aus einem Anwendungsbereich oder einer speziellen Branche stammen. Die inhaltliche Fokussierung erleichtert die Ansprache bestimmter Zielgruppen. Die Zielkunden von Data-Sharing-Plattformen können z. B. je nach Schwerpunkt variieren. Sie können sich an Unternehmen richten, die bestimmte Datensätze suchen, an Forschende, die Daten für Analysen benötigen, oder an Entwicklungsteams, die Daten für die Erstellung von Anwendungen benötigen. Domänenunspezifische Plattformen sind inhaltlich offen, stellen ein breites Spektrum an Daten bereit und beziehen Daten aus verschiedenen Wirtschafts- und Anwendungsbereichen mit ein. Die Plattform ist somit durch die Diversität der Daten und die große Anzahl an Datensätzen für Kunden interessant. Datenplattformen können als domänenspezifische Plattform gegründet werden und organisch zu einer domänenunspezifischen Plattform wachsen (Otto 2022).

Beim **Verarbeitungsgrad** der Daten wird zwischen Rohdaten und aufbereiteten Daten unterschieden. Rohdaten beziehen sich auf die Weitergabe von Datenpaketen in unbearbeiteter Form, so wie sie bspw. von einem Sensor ausgegeben werden. Aufbereitete Daten hingegeben wurden in verschiedenen Abstufungen schon bearbeitet. So können solche Daten rudimentär bearbeitet sein, indem sie mit einem standardisierten Datenmodell abgeglichen und in ein einheitliches Format überführt werden. Der Verarbeitungsgrad kann aber noch deutlich höher sein, indem Daten inhaltlich überprüft, aggregiert sowie in logischen Paketen organisiert werden, um dann in Datensätzen für weitere Analysen zur Verfügung zu stehen (Spiekermann 2019; Lindner et al. 2021).

Eine weitere Dimension bezieht sich auf die **Plattformarchitektur**, die wiederum in zentral, hybrid und dezentral unterschieden werden kann. Bei zentralen Architekturen werden die Daten von verschiedenen Anbietern über eine zentrale Stelle wie bspw. eine Cloud-Infrastruktur angeboten. Dies ermöglicht eine erleichterte Zugänglichkeit durch den zentralen Speicher und ein erleichtertes Zugriffsmanagement. Zudem muss weder der Datenlieferant noch der Kunde für den Datentransfer eine eigene komplexe technische Infrastruktur besitzen. Nichtsdestotrotz könnte mangelndes Vertrauen der

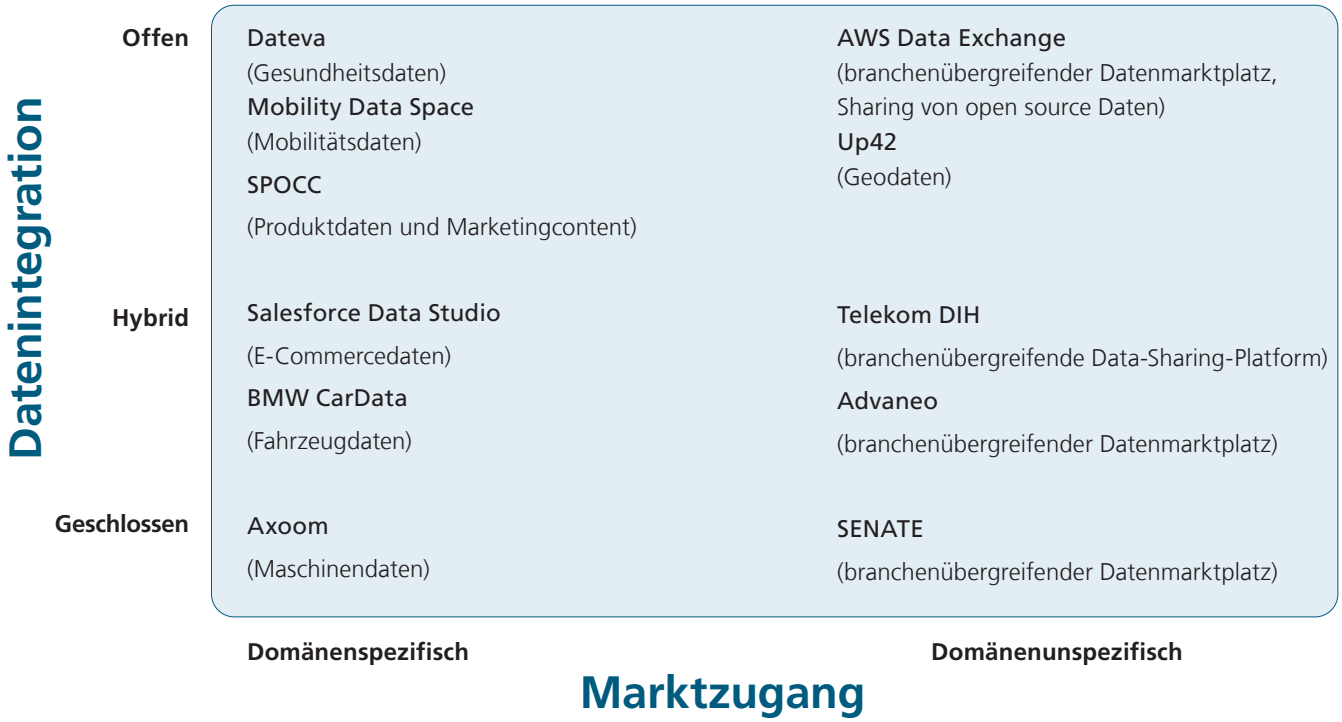

Abbildung 3: Einordnung von Praxisbeispielen nach Marktzugang und Datenintegration gemäß Lindner et al. (2021) (S. 76f.)

Parteien zu Hemmnissen in der Nutzung der Plattform führen. Bei dezentralen Plattformarchitekturen verbleiben die Daten und somit die Datenhoheit beim Datengebenden. Hier werden nur erforderliche Informationen zu den Daten bereitgestellt, während für Transaktionen die Daten direkt vom Datengebenden an den Kunden geschickt werden. Es gibt darüber hinaus Ansätze, Daten mithilfe der Blockchaintechnologie dezentral zu speichern. Lieferanten müssten also nicht den Speicherplatz für ihre Daten aufbringen (Liang et al. 2020). Dabei würden Daten fragmentiert an verschiedenen Speicherorten in einem dezentralen Netzwerk verschlüsselt abgelegt. Wenn ein Kunde Daten erwirbt, könnten ihm die Speicherorte und die Entschlüsselungsmethode übermittelt werden. Ein dezentrales Netzwerk erfordert jedoch genügend Nutzende, die dafür Rechenleistung bereitstellen. Hybride Formen kombinieren beide Ansätze, indem sie einen dezentralen Datentransfer verfolgen, jedoch durch die Bereitstellung zusätzlichen technischer Infrastruktur einen erleichterten Austausch ermöglichen (Spiekermann 2019).

Data-Sharing-Plattformen generieren Einnahmen über verschiedene **Erlösmodelle**, die sie anbieten (Stahl et al. 2015; Spiekermann 2019). Dazu können Abonnementgebühren für einen Premium-Zugang, Transaktionsgebühren für den Verkauf von Daten auf dem Marktplatz, Lizenzgebühren für die Datennutzung, API-Nutzungsgebühren, Werbeeinnahmen oder Gebühren für Mehrwertdienste wie Datenanalysen oder benutzerdefinierte Berichte gehören. Im Folgenden werden übliche Preismodelle etwas genauer vorgestellt.

- Abonnement-Modell: Ein Preis- und Zahlungsmodell, bei dem die Nutzendenden in regelmäßigen Abständen, in der Regel monatlich oder jährlich, eine wiederkehrende Gebühr für den Zugang zu den Diensten oder Ressourcen der Plattform und deren Nutzung zahlen. Dabei handelt es sich um ein laufendes Abonnement, das den Kunden kontinuierlichen Zugang zu den Merkmalen, Funktionen und Datenfreigabefunktionen der Plattform gewährt (Spiekermann 2019).
- Freemium-Modell: In diesem Modell stehen den Nutzerinnen und Nutzern neben der kostenpflichtigen Version auch kostenlose Dienste zur Verfügung. Dieser Ansatz ermöglicht den Nutzenden den kostenlosen Zugang zu grundlegenden Funktionen und bietet gleichzeitig die Möglichkeit, ein kostenpflichtiges Abonnement für zusätzliche Vorteile und erweiterte Funktionen abzuschließen (Spiekermann 2019).
- Pay-per-Use-Modell: Ein Preismodell, bei dem die Nutzendenden auf der Grundlage der tatsächlichen Nutzung oder des Gebrauchs der Dienste oder Ressourcen der Plattform abgerechnet werden. Dieses Modell ermöglicht es den Kunden, nur für die spezifischen Ressourcen zu zahlen, die sie nutzen, und bietet somit Flexibilität und Kosteneffizienz (Stahl et al. 2015).
- Paketpreise/Tiered Pricing: In diesem Modell werden dem Kunden unterschiedlich gestaffelte Pakete (bspw. API-Aufrufe, Anzahl Datenquellen) zu einem festen Preis und für eine bestimmte Nutzungsdauer angeboten. Im Unterschied zum Festpreis wird neben der Zeit auch die Menge festgelegt. In der Regel sind größere Pakete insgesamt teurer, aber relativ pro Einheit günstiger (Stahl et al. 2015).
- Progressiver Preis: Hier richtet sich der Preis eines Datensatzes nach dem Zeitpunkt des Erwerbs. Der Preis erhöht sich, je mehr Kunden eine nicht verkäufliche Lizenz zur Nutzung des Datensatzes erwerben. Dieses Preismodell wird vor allem angewandt, wenn die Verbreitung limitiert werden soll (u. a. auch bei künstlerischen Werken). Praktiziert wird dies bspw. von Dreamstime über sogenannte Preis-Level (Stahl et al. 2015).
- Mitgliedschaften: Die Mitgliedschaft als kostenlose Nutzung oder Tauschgeschäft ermöglicht den Nutzendenden den Zugang zu den Diensten und Daten der Plattform, ohne direkt Geld zu zahlen. Stattdessen wird die Mitgliedschaft im Austausch für die Bereitstellung von Daten oder anderen wertvollen Ressourcen gewährt. So können Anwendende ihre eigenen Daten oder Ressourcen als Währung anbieten und beide Seiten profitieren von der wechselseitigen Austauschbeziehung (Stahl et al. 2015).

Die einzelnen Erlösmodelle können unterschiedlich kombiniert werden. Daneben sind auch maßgeschneiderte Modelle für einzelne Nutzergruppen möglich.

In Abbildung 3 werden Praxisbeispiele in die Dimensionen Marktzugang und Datenintegration eingeordnet.

5 Potenziale von Data Pools für KI

Große Datenmengen, die über Data Pools verfügbar gemacht werden, bieten ein enormes Einsatzpotenzial für KI. Durch die Zusammenführung und Bereitstellung größerer und repräsentativerer Datenmengen können genauere und robuster trainierte KI-Modelle als mit einseitigen Daten entwickelt werden. Mithilfe einer größeren Datenbasis können KI-Algorithmen komplexere Muster und Zusammenhänge erkennen, was zu präziseren Vorhersagen, besseren Entscheidungen und insgesamt verbesserten Leistungen führt.

Darüber hinaus ermöglicht Data Pooling auch den Zugang zu vielfältigeren Datenquellen, einschließlich unterschiedlicher Datenformate, Domänen und Kontexte. Durch die Einbeziehung von Daten aus verschiedenen Quellen und Bevölkerungsgruppen reduzieren sich Verzerrungen und die Voreingenommenheit (Bias) in Datensätzen. Dies eröffnet die Möglichkeit, verschiedene Aspekte eines Problems abzudecken und unterschiedliche Perspektiven einzubeziehen, was zu faireren, umfassenderen und ganzheitlicheren Lösungen führen kann.

Zusätzlich beschleunigt Data Pooling die Entwicklung und den Einsatz von KI-Modellen, indem es den Zugriff auf bereits vorhandene Datensätze erleichtert. Anstatt Daten von Grund auf neu sammeln oder generieren zu müssen, können Entwicklerinnen und Entwickler oder Forscherinnen und Forscher auf bereits vorhandene Datenressourcen zurückgreifen, was Zeit, Ressourcen und Kosten spart.

Die Vorteile von Data Pooling beschränken sich nicht nur auf die bloße Bereitstellung von Daten. Auch die Qualität, die Datenbereinigung und die Gewährleistung von Datenschutz spielen eine wichtige Rolle. Durch eine sorgfältige Auswahl, Aufbereitung und Standardisierung der Daten können potenzielle Fehler oder Verzerrungen minimiert und die Datenqualität maximiert werden, was letztendlich zu besseren Ergebnissen der KI-Anwendungen führt.

Im Folgenden werden Datenerhebung, -verarbeitung und Training von KI-Modellen in Netzwerken betrachtet. Netzwerke bestehen aus Knotenpunkten und Verbindungen untereinander. Abbildung 4 zeigt die drei Netzwerktypen, die in diesem Kapitel betrachtet werden. Die Knotenpunkte stellen einzelne Geräte bzw. Nutzendende dar, die sowohl Daten erheben als auch Rechenleistung übernehmen können.

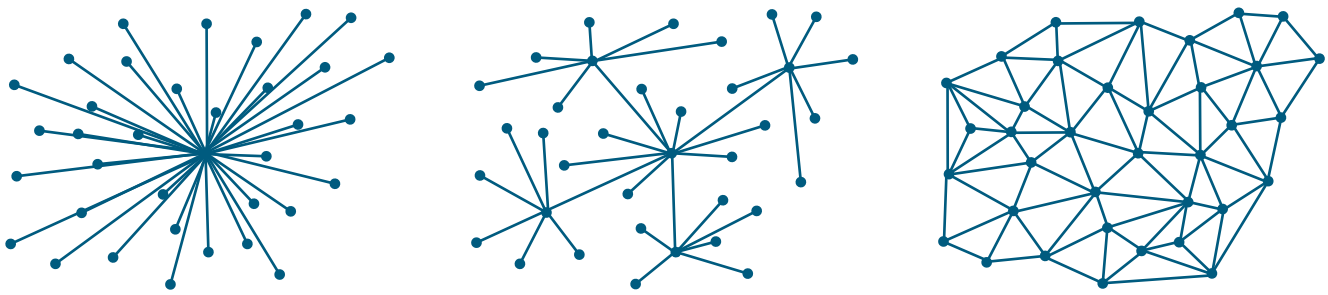

Abbildung 4: Verschiedene Typen von Netzwerken gemäß Demiral (2019)

Das Centralised Learning sieht vor, dass die Daten der Knoten an einer zentralen Stelle aggregiert werden und mit der Gesamtheit an Daten ein zentrales KI-Modell entwickelt wird. Dieser Ansatz entspricht klassischen Anwendungsfällen von KI, bei der (Cloud-)Rechenzentren die Rechenleistung für die Entwicklung der KI stellen, während die Endgeräte lediglich Daten sammeln und übermitteln.

Dabei können gewisse Barrieren auftreten: Einerseits könnte nicht genügend Rechenkapazität zur Verfügung stehen, andererseits kann es vorkommen, dass Individuen ihre Daten nicht frei zugänglich bereitstellen möchten (Ursin et al. 2019). Dank Cloud-Anbieter steht für die meisten KI-Anwendungen eine beträchtliche Rechenkapazität zur Verfügung, was die Möglichkeit bietet, große KI-Modelle über längere Zeiträume zu trainieren (Bian et al. 2022). Dennoch kann es bei Echtzeitanwendungen erforderlich sein, diese großen Modelle in kürzerer Zeit zu trainieren, was die Cloud allein nicht liefern kann.

Beim Decentralised Learning wird zwar auch ein zentraler Datensatz erzeugt, jedoch wird der Trainingsprozess auf die verschiedenen Knoten verteilt. Jeder Knoten berechnet einen Teil des letztendlichen KI-Modells, welches an zentraler Stelle aus den Teilen zusammengesetzt wird. Wenn so zusätzlich zu Rechenzentren die Rechenleistung aktiver Nutzendender eines Services für ein Live-Training eingebunden werden kann, besteht die Möglichkeit, die erforderliche Performance zu erreichen. Gleichzeitig ermöglicht dies den Entwickelnden, Entwicklungskosten fair auf die Anwendenden zu übertragen, ohne auf Abonnementmodelle für ihre Dienste angewiesen zu sein.

Beim Decentralised Learning kann es aber zu Problemen in Bezug auf den Datenschutz kommen. Um den Datenschutz von Verbrauchern zu gewährleisten, schreibt die Datenschutz-Grundverordnung (DSGVO) vor, dass sensible Informationen wie Name, Wohnort oder biometrische Daten nur mit Einwilligung der betroffenen Person verarbeitet werden dürfen (Deutscher Bundestag 27.04.2016). Diese Einwilligung kann jederzeit ohne Angabe von Gründen widerrufen werden, was dazu führt, dass die verarbeitenden Stellen die personenbezogenen Daten vollständig löschen müssen (Deutscher Bundestag 27.04.2016). Im Kontext von KI-Modellen kann dies einen erheblichen Aufwand bedeuten, da möglicherweise ein vollständiges Neutraining erforderlich ist. Die Verarbeitung personenbezogener Daten kann daher für Entwicklerinnen und Entwickler unattraktiv sein. Um dennoch Informationen auswerten zu können, müssen personenbezogene Daten vollständig anonymisiert und aggregiert werden, was die Aussagekraft und den Wert der Daten erheblich einschränkt. In Deutschland ist die Bereitschaft, Daten zu teilen, vergleichsweise gering, selbst wenn mit Vorteilen wie niedrigeren Kosten oder personalisierten Services geworben wird (Brandt 2018).

Ein Lösungsansatz, um KI-Modelle zu trainieren, ohne Daten zentral zu sammeln, ist das Distributed Learning. Entwicklerinnen und Entwickler von KI-Services sind letztendlich nicht an den tatsächlichen Daten der Individuen interessiert, sondern an dem KI-Modell, das auf diesen Daten trainiert wurde. Beim Distributed Learning trainieren Teilnehmende das Modell mit ihren eigenen Daten auf den eigenen Geräten und teilen nur die Verbesserungen des Modells mit einem zentralen Server. Der zentrale Server kombiniert dann diese Verbesserungen und aktualisiert das Modell insgesamt. Dabei bleiben die eigentlichen Daten der Teilnehmenden geschützt und werden nicht direkt weitergegeben. Durch eine verantwortungsvolle Entwicklung nach den Grundsätzen der KI-Absicherung können keine Rückschlüsse auf die zugrunde liegenden Daten aus dem KI-Modell gezogen werden (Poretschkin et al. 2021). Die von Google vorgestellte Technologie des föderierten Lernens liefert einen konkreten Ansatz, wie Distributed Learning umgesetzt werden kann.

Nachfolgend werden Ressourcen zu einer tiefergehenden Auseinandersetzung mit der Anwendung von KI-Technologien in Data Pools aufgelistet:

- Föderiertes Lernen bei Google und im Kontext von Google Assistant (McMahan und Thakurta 2022; McMahan und Ramage 2017; Bellwood und McCloud o.D.)
- Differential Privacy zur algorithmischen Sicherstellung von Datenschutz (Dwork und Roth 2013)
- KI-basierte Programmierungsuntersützung mit StarCoder (BigCode 2023)
- Föderiertes Lernen im Gesundheitssektor und Anwendung zur Covid-19-Bekämpfung (Giuseppi et al. 2022)

6 Data-Sharing-Plattformen in der Praxis

Nachfolgend werden relevante Initiativen rund um die Entwicklung von Data-Sharing-Plattformen sowie Anwendungsbeispiele von Data-Sharing-Plattformen vorgestellt. Die in der EU aktuell größten Initiativen für einen vertrauenswürdigen Austausch von Daten sind International Data Spaces und Gaia X.

6.1 International Data Spaces[3]

Die International Data Spaces (IDS) Initiative wurde als Forschungsprojekt des Bundesministeriums für Bildung und Forschung (BMBF) gemeinsam mit den Fraunhofer Instituten IML, ISST, IAIS und FOKUS als »Industrial Data Space« im Jahr 2015 gestartet. Das Forschungsprojekt entwickelte über drei Jahre eine sichere und vernetzte Dateninfrastruktur für Industrieunternehmen im europäischen und internationalen Raum. Aus der Initiative entstand im Jahr 2016 die gemeinnützige Organisation International Data Spaces Association (IDSA). Diese entwickelt die Architektur im Rahmen des »IDS Reference Architecture Model« (IDS-RAM) fort. Die IDSA-Initiative wird zunehmend als potenzieller Katalysator für die europäische Datenwirtschaft gesehen. Sie soll neue Datenmärkte fördern, indem sie einen sicheren und vertrauenswürdigen Datenaustausch ermöglicht. Insbesondere im privaten Sektor (industrielle Implementierungen) kann der definierte Standard des IDS-RAM die Datenwirtschaft fördern und sektorübergreifende Anwendungen ermöglichen, die bisher schwer zugänglich oder nur durch bilaterale Vereinbarungen möglich waren.

Merkmale der IDSA / IDS-RAM:

- Sichere Umgebung, die die Monetarisierung und den Datenaustausch ermöglicht
- IDS-RAM basierte Data-Sharing-Plattformen sind Vermittler zwischen Datenlieferant und Datennutzende
- Klare rechtliche Bedingungen für Dateneigentümer und Nutzendende
- Datenlieferanten behalten Hoheit über ihre Daten
- Fokus auf proprietäre Daten

Zugang zu der IDSA wird nur nach vorheriger Bewerbung gewährtSteigende Mitgliedszahlen deuten auf Interesse an der Architektur und dem Mitwirken an den Anwendungsfällen, d. h. den verschiedenen Data-Sharing-Plattformen, hin. Im Jahr 2019 hatte die Initiative nach dem IDSA Fact Sheet 100 Mitglieder, im Jahr 2022 waren es 130 Mitglieder (Otto 2022) und für das Jahr 2023 werden auf der Website der International Data Spaces Association 145 Mitglieder angegeben. Mitglieder der IDSA vertreten verschiedene Branchen und entwickeln basierend auf dem Standard Data-Sharing-Plattformen. Ein Anwendungsfall, der von der IDSA unterstützt wird, ist der Mobility Data Space (Data Spaces Overview 2022).

Mobility Data Space[4]

Der Mobility Data Space (MDS) ist ein neutraler Datenmarktplatz, der vom Bundesministerium für Digitales und Verkehr gefördert wird. Ziel ist es, Anbietern von Daten die Vernetzung mit Geschäftspartnern zu ermöglichen, um die Mobilität sicherer, effizienter und umweltfreundlicher zu gestalten. Der MDS bietet den Teilnehmenden Zugang zu einem Netzwerk von Dienstleistern und fördert den Austausch mit Branchenexperten. Der MDS basiert auf einem datenschutzfreundlichen Ansatz und ermöglicht eine dezentrale Datenspeicherung.

3 https://internationaldataspaces.org/

4 https://mobility-dataspace.eu/

6.2 Gaia X[5]

Das Gaia-X-Projekt wurde im Herbst 2019 gestartet (Otto 2021). Das erklärte Ziel ist die Entwicklung eines souveränen digitalen Ökosystems für Europa, das zwar nicht geografisch begrenzt ist, sich aber auf die europäischen Werte bezieht. Dies umfasst die Vermeidung von Anbieterbindungen zur Förderung eines freien Marktes sowie grundlegende Werte wie Transparenz und verantwortliche Datennutzung (Otto 2021).

Merkmale von Gaia X:

- Schaffung einer dezentralen und föderalen Dateninfrastruktur
- Gewährleistung von Datensouveränität und Interoperabilität
- Bereitstellung von Infrastruktur für einen sicheren und vertrauenswürdigen Datenaustausch
- IDS-Standard (International Data Spaces) ist ein zentrales Element der Gaia-X-Architektur

Die größten Anwendungsfälle von Gaia X sind die zwei Data-Sharing-Plattformen, Catena X und Manufacturing X, erstere für die Automobilindustrie und letztere für die Maschinenindustrie. Beide werden überwiegend von deutschen Unternehmen gestützt (Catena X o.D., Vision und Ziele, Plattform Industrie 4.0 o.D., Sauer 2023).

Gaia X steht unter der Kritik, nicht mit den Entwicklungen von Unternehmen wie Microsoft und Google mithalten zu können (Scheuer 2023). Dies liegt daran, dass aufgrund der Vielzahl an beteiligten Akteuren die Entwicklung von Standards lange dauert. Zudem wurde noch keine Lösung gefunden, ein globales IT-Infrastrukturnetz aufzubauen. Gaia X ist als Initiative notwendig, um Standards und Verantwortung für Datenmanagement zu schaffen, stellt derzeit aber noch keine Alternative zu etablierten Hyperscalern, wie bspw. Amazon AWS, Microsoft Azure und Google Cloudproject, dar, sondern wird im Gegenteil von diesen mitgestaltet (Finkenzeller 2021).

Catena X[6]

Catena X ist eine Initiative der Automobilindustrie, die es ermöglichen soll, den digitalen Informationsfluss über die gesamte Lieferkette hinweg zu gewährleisten. Die Vision von Catena X ist eine Zukunft, in der alle Akteure in durchgängigen Wertschöpfungsketten vernetzt sind, in denen alle Partner gleichberechtigt sind, über ihre Daten souverän verfügen und keine Lock-In-Effekte entstehen. Das Ziel von Catena X ist es für die Digitalisierung der Lieferketten, insbesondere auch für mittelständische und kleine Unternehmen, eine nachhaltige Lösung bereitzustellen und die Kooperation und Zusammenarbeit von Marktteilnehmern und Wettbewerbern zu unterstützen.

Manufacturing X[7]

Manufacturing X ist eine Plattform zur Förderung der digitalen Transformati-on in der Fertigungsindustrie. Die Plattform zielt darauf ab, durch den multilateralen Datenaustausch im Wertschöpfungsnetzwerk der Maschinenindustrie neue Potenziale zu erschließen. Insbesondere soll die digitale Transformation des Mittelstands vorangetrieben werden. Dazu werden relevante industrielle Anwendungsfälle durch den Einsatz existierender offener Standards der IDS oder Gaia X umgesetzt. Die Plattform fördert die Zusammenarbeit mit Partnern und schafft die Basis für Skalierung, während Standardisierung, Internationalisierung und Governance von Anfang an mitgedacht werden sollten.

5 https://gaia-x.eu/

6 https://catena-x.net/de/

7 https://www.plattform-i40.de/IP/Navigation/DE/Manufacturing-X/Initiative/initiative-manufacturing-x.html

6.3 Servicemeister[8]

Servicemeister ist eine KI-Plattform, die sich auf Serviceprozesse in der Fertigungsindustrie spezialisiert hat. Ursprünglich als Forschungsprojekt des BMWK im Jahr 2020 gestartet, bietet das Konsortium aus Maschinenbauern und Froschungseinrichtungen nun datennahe Dienstleistungen wie die Diagnose von Störungen auf der Grundlage von Maschinendaten für Mitgliedsunternehmen an. Mittelständischen Unternehmen fehlen oft sowohl personelle als auch technische Ressourcen, um digitale Prozesse und KI-Anwendungsfälle umzusetzen. Servicemeister soll diese Unternehmen unterstützen und ihnen die Wettbewerbsfähigkeit gegenüber Großunternehmen, die solche Anwendungen intern abbilden können, ermöglichen. Die Daten, die an Servicemeister übermittelt werden, sollen zusätzlich unternehmensübergreifend ausgewertet werden, um bspw. Fachwissen in Form von KI-basierten Service-Chatbots anbieten zu können (Timmermann o.D., Bundesministerium für Wirtschaft und Klimaschutz o.D.b.).

Um solche Zwecke zu erfüllen, kann die Datenhoheit nicht bei den Datenlieferanten verbleiben, sondern die Daten müssen für die Betreiber der Plattform zugänglich sein. Hier unterscheidet sich Servicemeister von Initiativen, die auf dem IDS-Standard basieren, indem die Daten nicht gegen Bezahlung mit anderen Unternehmen auf der Plattform gehandelt werden. Der direkte Mehrwert für die Unternehmen besteht in der Analyse ihrer eigenen Daten sowie in dem langfristigen Versprechen, dass der Service für ihre Maschinen mithilfe der bereitgestellten Daten verbessert werden kann.

6.4 Axoom[9]

Axoom ist eine digitale Plattform und Tochtergesellschaft des deutschen Maschinenbauunternehmens Trumpf. Die Plattform wurde gegründet, um Lösungen für die digitale Transformation von Fertigungsprozessen anzubieten. Seit 2019 führt das IT-Unternehmen GFT die Plattform in enger Kooperation mit Trumpf. GFT hat über Axoom Partnerschaften mit verschiedenen Unternehmen der Fertigungsindustrie aufgebaut und unterstützt Kunden bei der Umsetzung ihrer digitalen Strategien (vgl. Artikel in Automationspraxis 2019).

Axoom ermöglicht es Fertigungsunternehmen, Prozesse zu optimieren, die Produktivität zu steigern und neue Geschäftsmodelle zu entwickeln, indem sie datengestützte Dienstleistungen und Produkte anbietet (Bundesministerium für Wirtschaft und Klimaschutz o.D.a.). Beispiele sind eine IIoT-Plattform (Industrial Internet of Things), Datenanalyse, Cloud-Services und digitale Lösungen für die vernetzte Produktion, Echtzeitüberwachung, vorausschauende Wartung und die Integration von Maschinen und Anlagen in digitale Ökosysteme.

Zum Zeitpunkt der Veröffentlichung des Whitepapers besteht Unklarheit darüber, ob GFT noch Projekte unter dem Namen Axoom durchführt. Dennoch bleibt GFT weiterhin aktiv in der Branche tätig.

8 https://www.servicemeister.org/

9 https://www.plattform-i40.de/IP/Redaktion/DE/Anwendungsbeispiele/491-axoom/beitrag-axoom.html

7 Diskussion und Handlungsempfehlungen zu Kooperationsformaten

Herausforderungen bei der Umsetzung von Data-Sharing-Plattformen gibt es in verschiedenen Bereichen: Technik, Betrieb und Organisation, Rechtskonformität sowie nationale und regionale Bedingungen (Curry et al. 2022). Häufig diskutiert werden Schwierigkeiten im Bereich der Datenakquise, Herausforderungen in Bezug auf Vertrauen und Akzeptanz von Sharing-Plattformen, Konkurrenz sowie sicherheitsrelevante Fragestellungen. Nachfolgend werden diese Herausforderungen aufgegriffen und im Kontext bestehender Data-Sharing-Plattformen diskutiert.

Datenakquise
Wie können nachhaltig qualitativ hochwertige und aktuelle Daten für die Plattform akquiriert werden?

Der intrinsische Wert einer Data-Sharing-Plattform manifestiert sich in dem Bestand der darin hinterlegten Daten. Nur wenn genügend relevante Daten verfügbar sind, sind Akteure bereit, einer Plattform beizutreten und im Gegenzug ihre Daten zur Verfügung zu stellen. Viele Initiativen, die eine weitreichende Diversifizierung an Industrien anvisieren, erhöhen zwar damit das Potenzial an relevanten Daten, jedoch sind sie zugleich in der Gefahr, in einer Anfangsphase mit nur wenigen Datenkontingenten pro Industrie zu stagnieren. Bspw. konnten weder der IDS noch Gaia X trotz enormer Fördersummen eine branchenübergreifende, zusammenhängende Datenplattform hervorbringen. Im Gegenteil: Erfolgreiche Produkte der Forschungsprojekte sind branchenspezifische Data-Sharing-Plattformen.

In diesem Kontext haben Data-Sharing-Plattformen für Forschungszwecke, wie Huggingface oder EUDAT, einen maßgeblichen Vorzug aufgrund der Tatsache, dass ein substanzieller Anteil ihrer vorhandenen Daten aus wissenschaftlichen Publikationen stammt und bereits per se frei zugänglich ist. Zusätzlich haben für viele Forschende ihre Daten bzw. Informationen besonders dann einen Wert, wenn sie möglichst weit verbreitet werden und einem großen Nutzendenkreis zur Verfügung stehen, d. h. Data-Sharing ist im Eigeninteresse der Datenlieferanten (Fulterer 2022; Gentzsch et al. 2014). In der Industrie ist Data-Sharing noch stärker mit Zweifeln verbunden.

Vertrauen und Akzeptanz
Wie können sich Betreiber einer Data-Sharing-Plattform als vertrauenswürdig etablieren? Wie kann eine Data-Sharing-Plattform zum Standard werden?

Wenn eine Plattform etabliert ist und von Marktführern adoptiert wurde, steigt der Druck für andere Akteure, Teil der Plattform zu werden. Entsprechende Zweifel und potenzielle Herausforderungen werden eher überwunden, wenn der Anreiz mitzuhalten groß genug ist. Doch dafür muss eine Plattform zunächst die Marktführer überzeugen. Daneben ist das Vertrauen in die Betreiber der Plattform ein kritischer Faktor.

Initiativen, die auf EU-Ebene gefördert und von Non-Profit-Organisationen geführt werden, können einen gewissen Vertrauensvorschuss genießen, leiden aber oftmals unter komplexeren Abstimmungsprozessen und damit langsamerer Entwicklung, was die Etablierung als Standard erschwert. Data-Sharing-Plattformen nach dem IDS-Modell versuchen, zusätzliches Vertrauen zu gewinnen, indem sie nur den Raum für den Datenaustausch bieten, nicht aber die Daten selbst speichern. Wenn Datenlieferanten ihre Datenhoheit behalten und genau auswählen können, wer ihre Daten erhält, nützt diese Kontrolle, um die Sorge des ungewollten Sharing zu mildern.

Auch wenn die Gründe für die Auslagerung von Axoom aus Trumpf an GFT nicht vollständig bekannt sind, könnte das Vertrauen von potenziellen Mitgliedern in die Plattform ein Grund sein. Direkte Konkurrenten von Trumpf sind vielleicht nicht einverstanden, ihre Daten mit Trumpf zu teilen, aber offen dafür, ihre Daten mit einem IT-Dienstleister zu teilen, dessen Interesse nicht in den Inhalten der Daten liegt.

Konkurrenz
Wie können sich Data-Sharing-Plattformen gegen Hyperscaler durchsetzen?

Die Aussicht auf einen erfolgreichen Marktzugang für neu entstehende Data-Sharing-Plattformen wird durch die Präsenz etablierter Plattformen oder Hyperscaler signifikant erschwert. Gaia X konnte bislang keine übergreifende Cloud-Plattform bereitstellen, da Unternehmen wie Microsoft, Amazon und Google eine dominierende Marktstellung haben (Stiens und Kerkmann 2021). Für Unternehmen, die bereits Lösungen von Hyperscalern im Einsatz haben, ist es finanziell unattraktiv, eine Data-Sharing-Plattform von einem neuen Anbieter zu nutzen, wenn der Hyperscaler ein vergleichbares Angebot liefert. Der Digital Markets Act versucht dieser Thematik entgegenzuwirken. Gleichwohl kristallisiert sich nach wie vor ein erhebliches Erfordernis für distinktive Alleinstellungsmerkmale seitens neuer Plattformen heraus, um sich erfolgreich am Markt zu etablieren.

Sicherheit
Wie kann der Datenschutz sichergestellt werden, ohne die Datenhoheit zu verletzen? Wie können Sicherheitslücken präventiv verhindert werden?

Im Zuge des Digital Service Acts müssen Betreiber von Plattformen sicherstellen, dass ihre Inhalte rechtskonform sind und damit u. a. auch den Ansprüchen der Datenschutz-Grundverordnung (bzw. General Data Protection Regulation auf europäischer Ebene) genügen. Für eine umfangreiche Prüfung müssten Datenanbieter der Plattform vollständigen Zugriff auf ihre Daten liefern und würden effektiv ihre Datenhoheit verlieren. Weiterhin könnten korrupte Daten oder gar Schadsoftware in den Daten versteckt werden. Sicherheitsmechanismen von Plattformen müssten diese Szenarien präventiv und vollumfänglich ausschließen. Im Rahmen der IDS-Architektur wurden Algorithmen für einen sicheren Austausch zwischen Lieferanten und Nutzendenden entwickelt, dabei tritt die Plattform jedoch nur als Vermittler auf und bietet keine datenbezogenen Dienstleistungen. Bei Gaia X sollen auch derartige Dienstleistungen abgebildet werden, für die kryptographische Verfahren entwickelt werden. Aktuell sind allerdings nur ausgewählte Pilot-Use-Cases möglich (Gaia X o.D., »Use cases«; Cloud Computing Report o.D., »Gaia X wird konkreter«).

8 Zusammenfassung und Ausblick

In diesem Whitepaper wurde das Themenfeld der Data Pools sowie das der Data-Sharing-Plattformen behandelt. Letzteres ist wichtig, da Data Pools durch Data-Sharing-Plattformen zugänglich gemacht werden. Die beschriebenen technischen, organisatorischen, sicherheitsrelevanten und regulatorischen Anforderungen an Data-Sharing-Plattformen zeigen auf, dass die Gestaltung und Implementierung solcher Plattformen eine äußerst komplexe und multidisziplinäre Aufgabe darstellt. Darüber hinaus wurden die den Data-Sharing-Plattformen zugrundeliegenden Geschäftsmodelle beschrieben. Die dargelegten Anwendungsfälle illustrieren die Vielseitigkeit der Konzeption und deren Relevanz für datenbasierte Innovationen sowie effiziente Ressourcenallokation. Nichtsdestotrotz unterstreicht das Whitepaper auch die inhärenten Herausforderungen und Risiken bei der Umsetzung erfolgreicher Data-Sharing-Plattformen, insbesondere im Hinblick auf die Sicherung der Datenintegrität und -sicherheit sowie der Etablierung nachhaltiger Geschäftsstrategien. Zudem wurde das Einsatzpotenzial von KI in Bezug auf Data Pools und Data-Sharing-Plattformen beschrieben.

Es zeichnet sich ab, dass die Entwicklung von Data Pools und Data-Sharing-Plattformen bereits von großer Relevanz ist und auch zukünftig sein wird. Auch für regionale KI-Innovationsökosysteme kann der Aufbau von Data Pooling ein wichtiger Schritt für die Entwicklung von regionsspezifischen smarten Services sein. Die zunehmende Interdependenz von Unternehmen und die wachsende Bedeutung datengetriebener Geschäftsmodelle und Entscheidungsprozesse legen nahe, dass der Einsatz solcher Plattformen immer wichtiger wird. Die Integration von fortschrittlichen Technologien wie KI und maschinellem Lernen dürfte die Effizienz und Innovationskraft dieser Plattformen weiter steigern. Hier zeigt sich, wie die Entwicklung innovativer KI-Modelle Sharing-Plattformen zum Teilen von Daten und KI-Modellen zu Wachstum verhilft. Auch die EU unterstützt die Sicherung von internationalem Data Sharing zu fairen Bedingungen. Gleichzeitig werden jedoch auch regulatorische und ethische Fragestellungen in den Vordergrund rücken, die eine ausgewogene Balance zwischen Datenfreigabe und Datenschutz gewährleisten sollen. Data Pools und Data Sharing sind aktive Forschungsfelder und werden durch die Methoden der KI kontinuierlich weiterentwickelt.

9 Literaturverzeichnis

Automationspraxis (Hg.) (2019): Trumpf gibt Axoom an GFT ab. Online verfügbar unter https://automationspraxis.industrie.de/industrie-4-0/trumpf-gibt-axoom-an-gft-ab/, zuletzt aktualisiert am 02.07.2019, zuletzt geprüft am 15.08.2023.

Bellwood, Lucy; McCloud, Scott (o.D.): Federated Learning. Hg. v. Google AI. Online verfügbar unter https://federated.withgoogle.com/#about, zuletzt geprüft am 08.08.2023.

Bian, Yi-jie; Xie, Lu; Li, Jing-qi (2022): Research on influencing factors of artificial intel-ligence multi-cloud scheduling applied talent training based on DEMATEL-TAISM. In: Journal of Cloud Computing 11 (1), S. 35. DOI: 10.1186/s13677-022-00315-4.

BigCode (2023): BigCode Repository bei Huggingface. Huggingface. Online verfüg-bar unter https://huggingface.co/bigcode, zuletzt geprüft am 08.08.2023.

Brandt, Mathias (2018): Deutsche behalten Daten lieber für sich. Hg. v. statista. Onli-ne verfügbar unter https://de.statista.com/infografik/7830/bereitschaft-persoenlichen-daten-zu-teilen/, zuletzt geprüft am 08.08.2023.

Bundesministerium für Wirtschaft und Klimaschutz (Hg.) (o.D.a): AXOOM GmbH - ID4i – Die digitale Produktionsplattform für die Industrie. Online verfügbar unter https://www.plattform-i40.de/IP/Redaktion/DE/Anwendungsbeispiele/491-axoom/beitrag-axoom.html, zuletzt geprüft am 15.08.2023.

Bundesministerium für Wirtschaft und Klimaschutz (Hg.) (o.D.b): Projektbericht zu Servicemeister. Online verfügbar unter https://www.digitale-technolo-gien.de/DT/Redaktion/DE/Standardartikel/KuenstlicheIntelligenzProjekte/KuenstlicheIntelligenz_ErsterFoerderaufruf/ki-projekt_service-meister.html, zuletzt geprüft am 15.08.2023.

Catena X (o.D.): Vision und Ziele. Online verfügbar unter https://catena-x.net/de/vision-ziele, zuletzt geprüft am 11.08.2023.

Cloud Computing Report (Hg.) (o.D.): GAIA-X wird konkreter: Cloud&Heat präsen-tiert Demonstrator. Online verfügbar unter https://www.cloud-computing-report.de/gaia-x-wird-konkreter-cloudheat-praesentiert-demonstrator/29-05-2020/, zuletzt geprüft am 08.08.2023.

Curry, Edward; Scerri, Simon; Tuikka, Tuomo (2022): Data Spaces. Cham: Springer International Publishing.

Data Spaces Overview (2022). Online verfügbar unter https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/220812_Use-Case-Bro_2022_35-MB.pdf, zuletzt aktualisiert am 12.08.2022, zuletzt geprüft am 08.08.2023.

Demiral, Muharrem (2019): Infrastructure Architecture. Hg. v. Medium. Online verfüg-bar unter https://medium.com/@muharrem.dmrl/infrastructure-architectures-9cc414afb4ab, zuletzt geprüft am 21.11.2023.

Deutscher Bundestag (27.04.2016): Datenschutz-Grundverordnung. DSGVO, vom 27.04.2016. Online verfügbar unter https://dsgvo-gesetz.de, zuletzt geprüft am 08.08.2023.

Dwork, Cynthia; Roth, Aaron (2013): The Algorithmic Foundations of Differential Privacy. In: FNT in Theoretical Computer Science 9 (3-4), S. 211 – 407. DOI: 10.1561/0400000042.

Engelhardt, Sebastian von; Wangler, Leo; Wischmann, Steffen (2017): Eigenschaften und Erfolgsfaktoren digitaler Plattformen. Eine Studie im Rahmen der Begleitforschung zum Technologieprogramm AUTONOMIK für Industrie 4.0 des Bundesministeriums für Wirtschaft und Energie. Hg. v. Alfons Botthof, Begleitforschung AUTONOMIK für Industrie 4.0 und Institut für Innovation und Technik (iit) in der VDI / VDE Innovation + Technik GmbH. Online verfügbar unter https://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/autonomik-studie-digitale-plattformen.pdf?__blob=publicationFile&v=6, zuletzt geprüft am 08.08.2023.

Europäische Kommission (2020): Ein europäischer Ansatz für künstliche Intelligenz. Online verfügbar unter https://digital-strategy.ec.europa.eu/de/policies/european-approach-artificial-intelligence, zuletzt aktualisiert am 31.01.2024, zuletzt geprüft am 22.02.2024.

Europäischer Rat (2023): Timeline – Artificial intelligence. Online verfügbar unter https://www.consilium.europa.eu/en/policies/artificial-intelligence/timeline-artificial-intelligence/, zuletzt aktualisiert am 13.12.2023, zuletzt geprüft am 22.02.2024.

Finkenzeller, Karin (2021): „Gaia-X wird keine europäische Cloud schaffen". Hg. v. WirtschaftsWoche. Online verfügbar unter https://www.wiwo.de/unternehmen/it/europaeisches-datenprojekt-gaia-x-wird-keine-europaeische-cloud-schaffen/27922086.html, zuletzt aktualisiert am 27.12.2021, zuletzt geprüft am 08.08.2023.

Focken, Hanno (2023): Governance Framework für Datenraum-Betrieb. Hg. v. Catena X. Online verfügbar unter https://catena-x.net/de/catena-x-einfuehren-umsetzen/governance-framework-fuer-datenraum-betrieb, zuletzt aktualisiert am 31.07.2023, zuletzt geprüft am 10.08.2023.

Fricker, Samuel A.; Maksimov, Yuliyan V. (2017): Pricing of Data Products in Data Marketplaces. In: Arto Ojala, Helena Holmström Olsson und Karl Werder (Hg.): Soft-ware Business, Bd. 304. Cham: Springer International Publishing (Lecture Notes in Business Information Processing), S. 49–66.

Fulterer, Ruth (2022): KI-Forschung nach dem Vorbild Cern: Die ehemalige Spass-Firma Hugging Face fordert Meta und Google heraus. Hg. v. Neue Zürcher Zeitung. Online verfügbar unter https://www.nzz.ch/technologie/huggingface-diese-nach-einem-emoji-benannte-firma-fordert-meta-und-google-heraus-ld.1682435, zuletzt aktualisiert am 17.05.2022, zuletzt geprüft am 15.08.2023.

Gaia X (o.D.): Use cases. Online verfügbar unter https://gaia-x.eu/use-cases/, zuletzt geprüft am 08.08.2023.

Gentzsch, Wolfgang; Lecarpentier, Damien; Wittenburg, Peter (2014): Big Data in Science and the EUDAT Project. In: 2014 Annual SRII Global Conference. 2014 An-nual SRII Global Conference (SRII). San Jose, CA, USA, 23.04.2014 - 25.04.2014: IEEE, S. 191–194.

Giuseppi, Alessandro; Manfredi, Sabato; Menegatti, Danilo; Poli, Cecilia; Pietrabissa, Antonio (2022): Decentralised Federated Learning for Hospital Networks With Appli-cation to COVID-19 Detection. In: IEEE Access 10, S. 92681–92691. DOI: 10.1109/ACCESS.2022.3202922.

Hazen, Benjamin T.; Boone, Christopher A.; Ezell, Jeremy D.; Jones-Farmer, L. Allison (2014): Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and ap-plications. In: International Journal of Production Economics 154, S. 72–80. DOI: 10.1016/j.ijpe.2014.04.018.

IBM (2023): Data pool. Online verfügbar unter https://www.ibm.com/docs/en/imdm/11.6?topic=gds-data-pool, zuletzt aktualisiert am 16.03.2023, zuletzt geprüft am 07.08.2023.

IDSA (Hg.) (2019): INTERNATIONAL DATA SPACES FACT SHEET UND KERNAUSSA-GEN. Online verfügbar unter https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-Fact-sheet-und-Kernaussagen-Deutsch.pdf, zuletzt geprüft am 09.08.2023.

International Data Spaces Association (2023): Website IDSA. Online verfügbar unter https://internationaldataspaces.org/, zuletzt geprüft am 08.08.2023.

John, Tomcy; Misra, Pankaj (2017): Data Lake for enterprises. Leveraging Lambda architecture for building Enterprise Data Lake. Birmingham, UK: Packt Publishing.

Jung, Christian; Dörr, Jörg (2022): Data Usage Control. In: Boris Otto, Michael ten Hompel und Stefan Wrobel (Hg.): Designing Data Spaces. Cham: Springer Interna-tional Publishing, S. 129–146.

Konduru, Veda (2017): 3 ways data quality impacts Predictive Analytics. Hg. v. Towa-rds Data Science. Online verfügbar unter https://towardsdatascience.com/3-ways-data-quality-impacts-predictive-analytics-cc72cdb1043, zuletzt aktualisiert am 02.08.2017, zuletzt geprüft am 08.08.2023.

Kutz, Janika; Göbels, Vincent Philipp; Brajovic, Danilo; Fresz, Benjamin; Renner, Niclas; Omri, Safa et al. (2023): KI-Zertifizierung und Absicherung im Kontext des EU AI Act. Unter Mitarbeit von Fraunhofer-Gesellschaft.

Liang, Wei; Fan, Yongkai; Li, Kuan-Ching; Zhang, Dafang; Gaudiot, Jean-Luc (2020): Secure Data Storage and Recovery in Industrial Blockchain Network Environments. In: IEEE Trans. Ind. Inf. 16 (10), S. 6543–6552. DOI: 10.1109/TII.2020.2966069.

Lindner, Maximilian; Straub, Sebastian; Kühne, Bettina (2021): How to share data? Data-Sharing-Plattformen für Unternehmen. Hg. v. Peter Gabriel. Begleitforschung Smarte Datenwirtschaft; Institut für Innovation und Technik (iit) in der VDI / VDE Inno-vation + Technik GmbH. Online verfügbar unter https://www.digitale-technolo-gien.de/DT/Redaktion/DE/Downloads/Publikation/SDW/2021_04_19_Datasharing_Studie.pdf?__blob=publicationFile&v=28, zuletzt aktualisiert am 19.04.2021, zuletzt geprüft am 08.08.2023.

McMahan, Brendan; Ramage, Daniel (2017): Federated Learning: Collaborative Ma-chine Learning without Centralized Training Data. Hg. v. Google Research. Online verfügbar unter https://ai.googleblog.com/2017/04/federated-learning-collaborative.html, zuletzt aktualisiert am 06.04.2017, zuletzt geprüft am 08.08.2023.

McMahan, Brendan; Thakurta, Abhradeep (2022): Federated Learning with Formal Differential Privacy Guarantees. Hg. v. Google Research. Online verfügbar unter https://ai.googleblog.com/2022/02/federated-learning-with-formal.html, zuletzt aktua-lisiert am 28.02.2022, zuletzt geprüft am 08.08.2023.

Meisel, Lukas; Spiekermann, Markus (2019): Datenmarktplätze. Plattformen für Da-tenaustausch und Datenmonetarisierung in der Data Economy. Hg. v. Prof. Dr.-Ing. Boris Otto, Prof. Dr. Jakob Rehof. Fraunhofer ISST. Online verfügbar unter https://www.isst.fraunhofer.de/content/dam/isst-neu/documents/Publikationen/Datenwirtschaft/2019-2_ISST-Bericht_Datenmarktplaetze-ISSN-0943-1624.pdf, zuletzt geprüft am 08.08.2023.

Nagel, Lars; Lycklama, Douwe (2022): How to Build, Run, and Govern Data Spaces. In: Boris Otto, Michael ten Hompel und Stefan Wrobel (Hg.): Designing Data Spaces. Cham: Springer International Publishing, S. 17–28, zuletzt geprüft am 08.08.2023.

Otto, Boris (2021): GAIA-X and IDS.

Otto, Boris (2022): The Evolution of Data Spaces. In: Boris Otto, Michael ten Hompel und Stefan Wrobel (Hg.): Designing Data Spaces. Cham: Springer International Publishing, S. 3–15.

Otto, Boris; Hompel, Michael ten; Wrobel, Stefan (Hg.) (2022): Designing Data Spac-es. Cham: Springer International Publishing.

Plattform Industrie 4.0 (o.D.): Manufacturing-X. Initiative zur Digitalisierung der Liefer-ketten in der Industrie. Hg. v. Bundesministerium für Wirtschaft und Klimaschutz und Bundesministerium für Bildung und Forschung. Online verfügbar unter https://www.plattform-i40.de/IP/Navigation/DE/Manufacturing-X/Initiative/initiative-manufacturing-x.html, zuletzt geprüft am 11.08.2023.

Poretschkin, Maximilian; Schmitz, Anna; Akila, Maram; Adilova, Linara; Becker, Daniel; B. Cremers, Armin; Hecker, Dirk (2021): Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz. KI-Prüfkatalog. Hg. v. Fraunhofer-Institut für Intelligente Analyse und Informationssysteme IAIS. Online verfügbar unter https://www.iais.fraunhofer.de/s/ki-pruefkatalog/epaper/ausgabe.pdf.

Sauer, Olaf (2023): Was bringt Manufacturing-X und auf welchen Annahmen basiert es? Hg. v. Fraunhofer IOSB. Online verfügbar unter https://blog.iosb.fraunhofer.de/was-bringt-manufacturing-x-und-auf-welchen-annahmen-basiert-es/, zuletzt aktualisiert am 10.05.2023, zuletzt geprüft am 11.08.2023.

Scheuer, Stephan (2023): Europas Versagen im Cloud-Markt rächt sich jetzt beson-ders. Tech-Konzerne in Europa haben den Cloud-Markt ihren Rivalen aus den USA und Asien überlassen. Beim Thema Künstliche Intelligenz wird das zum großen Prob-lem. Hg. v. Handelsblatt. Online verfügbar unter https://www.handelsblatt.com/meinung/kommentare/kommentar-europas-versagen-im-cloud-markt-raecht-sich-jetzt-besonders/29123710.html, zuletzt aktualisiert am 02.05.2023, zuletzt geprüft am 08.08.2023.

Spiekermann, Markus (2019): Data Marketplaces: Trends and Monetisation of Data Goods. In: Intereconomics 54 (4), S. 208–216. DOI: 10.1007/s10272-019-0826-z.

Stahl, Florian; Löser, Alexander; Vossen, Gottfried (2015): Preismodelle für Daten-marktplätze. In: Informatik Spektrum 38 (2), S. 133–141. DOI: 10.1007/s00287-013-0751-7.

Stiens, Teresa; Kerkmann, Christof (2021): Gaia-X-Gipfel in Mailand: Das Cloud-Projekt wird zum Problemfall. Beim europäischen Datenprojekt ist Ernüchterung ein-gekehrt – Verzögerungen und Richtungsstreits machen dem Vorhaben zu schaffen. In diesen Tagen treffen sich Vertreter der Initiative in Mailand. Hg. v. Handelsblatt. Online verfügbar unter https://www.handelsblatt.com/politik/deutschland/datensouveraenitaet-gaia-x-gipfel-in-mailand-das-cloud-projekt-wird-zum-problemfall/27809120.html, zuletzt aktualisiert am 18.11.2021, zuletzt geprüft am 08.08.2023.

Thogmartin, David; Orth, Mosche (2023): EU AI Act: Das KI-Gesetz der Europäischen Union. Einzelheiten und Hintergründe zu den Umsetzungsanforderungen. Deloitte. Online verfügbar unter https://www2.deloitte.com/de/de/pages/data/articles/european-ai-act.html, zuletzt aktualisiert am 12.12.2023, zuletzt geprüft am 22.02.2024.

Timmermann, Hauke (Hg.) (o.D.): Darstellung der Konzeptidee. KI Service - Industrie 4.0. eco - Verband der Internetwirtschaft e.V. Online verfügbar unter https://www.servicemeister.org/darstellung-der-konzeptidee/, zuletzt geprüft am 08.08.2023.

Ursin, Giske; Stenbeck, Magnus; Chang-Claude, Jenny; Gunter, Marc; Kaaks, Rudolf; Kampman, Ellen et al. (2019): Data must be shared-also with researchers outside of Europe. In: Lancet (London, England) 394 (10212), S. 1902–1903. DOI: 10.1016/S0140-6736(19)32633-9.

VMware (o.D.): When to use cloud scalability. Online verfügbar unter https://www.vmware.com/topics/glossary/content/cloud-scalabili-ty.html#:~:text=Scalability%20is%20one%20of%20the%20driving%20reasons%20to%20migrate%20to,to%20increase%20storage%20and%20performance, zuletzt geprüft am 08.08.2023.

Impressum

Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO
Bildungscampus 9
74076 Heilbronn
www.iao.fraunhofer.de

Kontakt
Vincent Philipp Göbels
Tel. +49 152 22543933
vincent-philipp.goebels@iao.fraunhofer.de

Dr. Bernd Bienzeisler
Tel. +49 151 16327691
bernd.bienzeisler@iao.fraunhofer.de

Satz und Layout
NetSyn, Joachim Würger

Titelbild
AKrasov – stock.adobe.com

Fraunhofer-Publica
http://dx.doi.org/10.24406/publica-2856

Alle Rechte vorbehalten
© Fraunhofer IAO, April 2024

Kontakt

Vincent Philipp Göbels
Tel. +49 152 22543933
vincent-philipp.goebels@iao.fraunhofer.de

Dr. Bernd Bienzeisler
Tel. +49 151 16327691
bernd.bienzeisler@iao.fraunhofer.de

Fraunhofer-Institut für Arbeitswirtschaft
und Organisation IAO
Bildungscampus 9
74076 Heilbronn

www.iao.fraunhofer.de