

 **DIO**
Data Intelligence
Offensive

WHITEPAPER

DATENRÄUME UND DATENKREISE

IM FOKUS DER DATA INTELLIGENCE



Autor*innen

Prof. Dr. iur. Peter A. Bruck, PhD. MA
Geschäftsführendes Vorstandsmitglied/ Generalsekretär der Data Intelligence Offensive
und Gründer der Research Studio Austria FG, Vorsitzender des Management Boards des
Leitprojektes „Data Market Austria“ (2016-2019)

Univ. Prof. Dr. Allan Hanbury
Vorstandsmitglied der Data Intelligence Offensive
und Leiter des Forschungsbereichs E-Commerce der Fakultät für Informatik, TU Wien

Mag.^a Nicole Kirowitz
Koordination & Kommunikation in der Data Intelligence Offensive

DI Mag. Günther Tschabuschnig
Präsident der Data Intelligence Offensive
und Chief Information Officer der Zentralanstalt für Meteorologie und Geodynamik

Juli 2021

Data Intelligence Offensive
Leopoldskronstraße 30
5020 Salzburg, Österreich

Tel. +43 662 834 602 0
Internet: www.dataintelligence.at
E-Mail: office@dataintelligence.at
Titelblatt: © STOCK.ADOBE.COM/YOUR123

Dieses Whitepaper gibt einen Überblick über die Begriffe, Konzepte und Struktur von Datenräumen und Datenkreisen.

Zudem wird die Data Intelligence Offensive (DIO) mit einigen Use Cases vorgestellt und die europäische Initiative von GAIA-X in Kontext gesetzt.

Das Whitepaper richtet sich dabei an Datenexpert*innen und Unternehmer*innen sowie Wissenschaftler*innen aus dem Bereich der Data Science und Datenwirtschaft.

Inhaltsverzeichnis

| | |
|--|----|
| Executive Summary | 3 |
| 1 Einleitung | 4 |
| 2 Begriffsdefinitionen | 5 |
| Data Assets | 5 |
| Data Governance | 5 |
| Dateninfrastruktur | 5 |
| Datenökosystem | 5 |
| Datensouveränität/ Datenhoheit | 5 |
| Metadaten | 5 |
| Semantisches Regelwerk | 5 |
| GAIA-X | 5 |
| IDSA | 6 |
| DIO Arbeitsgruppe (AG) | 6 |
| Datenräume und Datenkreise | 6 |
| Datenräume/ Data spaces (DR) | 6 |
| Datenkreise/ Data circles (DK) | 7 |
| 3 Stakeholder*innen der Datenräume und Datenkreise | 8 |
| 4 Struktur der Datenräume und Datenkreise | 8 |
| Technische Struktur der Datenräume und Datenkreise | 9 |
| 5 Die vier Prinzipien von Datenräumen | 11 |
| 1. Datenhoheit | 11 |
| 2. Datenaustausch und Datennutzung | 11 |
| 3. Dezentrale Infrastruktur | 11 |
| 4. Datenverwaltungskonzept (Data Governance) | 12 |
| 6 Die vier Prinzipien von Datenkreisen | 12 |
| 1. Vertrauenswürdigkeit (Trust) | 12 |
| 2. Wirtschaftlichkeit | 13 |
| 3. Interoperabilität | 13 |
| 4. Flexibilität | 13 |
| 7 Die Implementierung von Datenkreisen und Datenräumen | 13 |
| 8 Kompatibilität und Überschneidung zur GAIA-X Architektur | 14 |
| 9 IDS-Architektur | 15 |
| 10 DIO Charta | 15 |
| Referenzen | 17 |

Executive Summary

Viele Unternehmen nutzen heutzutage ihre verfügbaren Daten noch nicht. Dabei kann ein sicherer sowie vertrauenswürdiger Datenaustausch und eine Datennutzung Innovationen vorantreiben, die Wertschöpfung erhöhen und die Wettbewerbsfähigkeit entscheidend verbessern.

Die Data Intelligence Offensive (DIO) entwickelt dazu aufbauend auf dem österreichischen Leitprojekt „Data Markt Austria“ das Konzept der Datenräume (data spaces) und Datenkreise (data circles), welche einen wertschöpfenden Datenaustausch ermöglichen.

Ziel dieses Whitepapers ist es, den Diskurs zur Thematik zu vertiefen, die Begriffe zu klären und aufzuzeigen, worin der Unterschied zwischen DIO-Arbeitsgruppen, -Datenräumen als auch -Datenkreisen besteht und wie diese entwickelt werden können.

DIO-Arbeitsgruppen ermöglichen den Austausch verschiedener Expert*innen zu relevanten Querschnittsthemen sowie die Ausarbeitung von Stellungnahmen.

DIO-Datenräume fokussieren sich auf übergeordnete Domänen, auf welchen Datenkreise – das heißt Teilbereiche von Datenräumen – aufbauen können.

Eine technische Struktur ist essenziell für das Funktionieren der Datenräume und Datenkreise. Dabei gilt es, die Charakteristiken von Datenräumen und Datenkreisen, darunter Wahrung der Datenhoheit, Datenaustausch und gemeinsame Datennutzung, eine dezentrale Infrastruktur, Data Governance, Vertrauenswürdigkeit, Wirtschaftlichkeit, Interoperabilität und Flexibilität zu beachten.

1 Einleitung

Die Data Intelligence Offensive (DIO) ist ein Verein zur Förderung der Datenwirtschaft und der Optimierung der Datentechnologien. Sie ist ein „Spin-Off“ des österreichischen Leitprojektes „Data Market Austria“¹. Die Konsortiumsmitglieder wollten über die Forschung und Entwicklung eines Proof of Technology-Prototypen hinaus mit der Initiierung der „Data Intelligence Offensive“ zur Stärkung des Daten-Innovationsumfeldes beitragen. In DIO finden sich die Akteur*innen zum Aufbau und Pflege entsprechender Data Communities unter Einbindung relevanter Stakeholder (KMUs, Startups, Großunternehmen, Forschung, Verwaltung) zusammen.

Die Data Intelligence Offensive wurde Ende 2018 als Kooperations-Plattform von Personen, Unternehmen, Organisationen und öffentlichen Stellen zur Förderung der Datenwirtschaft und des optimierten Einsatzes von Technologien gestartet, um die Entwicklung von digitalen Märkten in Österreich zu forcieren und einen vertrauenswürdigen und sicheren Handel mit Datenprodukten im Zusammenhang mit Big Data zu stärken.

DIO baut auf folgenden Einsichten auf:

- Funktionierende Datenmärkte sind die essenzielle Voraussetzung, damit Daten als strategische Ressource für Innovation und Wertschöpfung genutzt werden können.
- Digitale Technologien und Prozesse ermöglichen in allen Branchen und gesellschaftlichen Sektoren die Erfassung und Verarbeitung von Daten. Deren Wertschöpfungsmöglichkeiten mit Hilfe von Data Science, Künstlicher Intelligenz (KI) und maschinellem Lernen werden durch einen effizienten Datenhandel wesentlich erhöht.
- Das Innovationspotenzial der Digitalisierung erreicht sowohl einzel- wie gesamtwirtschaftlich eine zusätzliche Ebene der Wertschöpfung, wenn die in/ von einer Organisation/ Institution/ Unternehmung erfassten Daten mit anderen zusammengeführt und auch gehandelt werden.
- Der Einsatz sowie die Wertschöpfung und Steigerung der Wettbewerbsfähigkeit durch Data Science und künstliche Intelligenz hängt wesentlich vom Umfang und der Qualität der verfügbaren Daten ab.

2021 forciert DIO die Zusammenarbeit mit öffentlichen Stellen, privaten Unternehmen sowie gemeinnützigen Organisationen und setzt verschiedene Instrumente ein, um die Transformation der Datenwirtschaft zu ermöglichen.

DIO führt dabei als essenzieller nationaler Player Stakeholder*innen der Datenökonomie an Datenökosysteme, Datenräume sowie Datenkreise heran.

¹ Das Data Market Austria Projekt erhielt seine Förderung aus dem Programm „IKT der Zukunft“ des österreichischen Bundesministeriums für Verkehr, Innovation und Technologie der Österreichischen Forschungsförderung Gesellschaft (FFG) (Projektnummer 855404). Siehe auch: <https://datamarket.at/>

Zudem trägt DIO auf internationaler Ebene durch Unterstützung der Vorhaben wie GAIA-X und IDSA dazu bei, die Datensouveränität Europas zu sichern, seine digitale Souveränität dauerhaft aufzubauen und zu erhalten.

2 Begriffsdefinitionen

Data Assets

Data Assets stellen Datenbestände beziehungsweise die Gesamtheit vorhandener Daten eines Unternehmens dar.

Data Governance

Data Governance (Datenverwaltungskonzept) definiert Prozesse, Rollen, Richtlinien und Funktionen für die Nutzung von Daten in Unternehmen.

Dateninfrastruktur

Eine Dateninfrastruktur stellt eine technische Infrastruktur für den Zugang zu Daten, deren Speicherung sowie Austausch und Nutzung dar.

Datenökosystem

Ein Datenökosystem ist ein Zusammenspiel verschiedenster Akteur*innen aus dem Bereich der Datenwirtschaft. Ziel ist es Daten auszutauschen, um Innovationen voranzutreiben.

Datensouveränität/ Datenhoheit

Datensouveränität beziehungsweise Datenhoheit bezeichnet die größtmögliche Kontrolle und Herrschaft über eigene Daten.

Metadaten

Metadaten sind strukturierte Daten, welche weitere Informationen enthalten und die eigentlichen Daten beschreiben (Bsp.: Dateiname, Zugriffsrechte)

Semantisches Regelwerk

Durch eine einheitliche Sprache können Daten möglichst vieler Akteur*innen besser miteinander verknüpft und ausgewertet werden. Eine damit verbundene Selbstbeschreibung enthält Informationen zum Speicherort, zu verwendeten Technologien, zu Rechen- und Speicherleistung sowie zur bereitgestellten Funktionalität.

GAIA-X

Die Data Intelligence Offensive zielt auf die Projektentwicklung in Bezug auf die Standardisierung des Datenaustausches beziehungsweise der Datennutzung nach dem deutsch-französischen Leitprojekt GAIA-X in Österreich ab. Mit

GAIA-X soll eine **sichere und vernetzte europäische Dateninfrastruktur** geschaffen werden, die den höchsten Ansprüchen an digitaler Souveränität gebührt und Innovationen fördert.

Die Data Intelligence Offensive unterstützt als offizielles **Mitglied der gemeinnützigen Vereinigung GAIA-X AISBL** das Vorhaben von GAIA-X und trägt somit auf nationaler sowie internationaler Ebene dazu bei, die Datensouveränität Europas zu sichern, seine digitale Souveränität dauerhaft aufzubauen und zu erhalten.

IDSA

GAIA-X liefert die Infrastruktur, die einen sicheren sowie vertrauenswürdigen Datenaustausch (beziehungsweise Datennutzungen) ermöglicht und dessen Einhaltung überwacht. Die GAIA-X-Architektur baut auf dem International Data Spaces (IDS)-Standard auf, welcher einen offenen, transparenten und selbstbestimmten Austausch von Daten beziehungsweise Datennutzung ermöglicht.

Der IDS-Standard wurde von der International Data Spaces Association (IDSA)², einer Initiative branchenübergreifender Unternehmen und Forschungsorganisationen, zum Aufbau einer Infrastruktur für die gemeinsame Datennutzung entwickelt. Ziel der IDSA ist es, internationale Datenräume zu konzipieren, da Datenräume der Schlüssel für eine globale und digitale Wirtschaft sind.

DIO-Arbeitsgruppe (AG)

Arbeitsgruppen der Data Intelligence Offensive ermöglichen den Austausch verschiedener Expert*innen zu relevanten Querschnittsthemen der Datenwirtschaft sowie die Erarbeitung von **Lösungsansätzen** in Form von Workshops und Stellungnahmen. In Arbeitsgruppen werden **Querschnittsthematiken** wie beispielsweise Daten und Recht, Data Governance, Daten und Startups, Blockchain, etc. behandelt. Zudem arbeiten Arbeitsgruppen an „Blueprints“, die Lösungsvorschläge für Herausforderungen beim Aufbau von Datenräumen und Datenkreisen, bei der Gestaltung und Vereinbarung von Geschäftsmodellen, bei der Regelung interner und externer Beziehungen der Datenräume und Datenkreise sowie bei der Nutzung, dem Betrieb und bei der Außerbetriebsetzung (Lebenszyklusmodell) bieten.

Die Strategie der Arbeitsgruppen zielt auf die Etablierung von Datenräumen und in weiterer Folge fokussierter Datenkreise ab, die Use Cases umsetzen. Diese umfasst die Abfolge von Datenakquirierung über Analyse bis hin zu Datennutzung.

² Siehe auch www.internationaldataspaces.org

Datenräume und Datenkreise

Datenräume/ Data Spaces (DR)

Datenräume fokussieren sich auf **übergeordnete Domänen** (Wirtschaftsbereiche, Industriesektoren oder sonstige fachliche Anwendungsfelder), mit einer dezentralen (verteilten) Dateninfrastruktur, auf der Datenkreise aufbauen

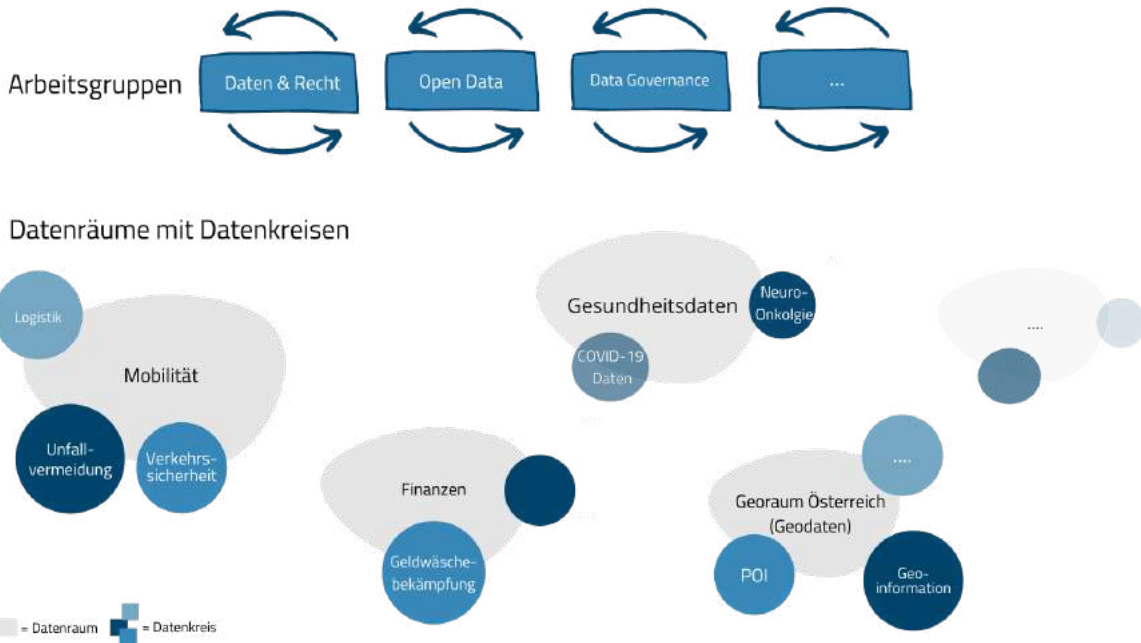
können. In einem Datenraum werden Daten unter Wahrung der **Datensouveränität/ Datenhoheit**, das heißt der größtmöglichen Kontrolle und Herrschaft über die eigenen Daten, für potenzielle innovative Dienste verfügbar gemacht. Zentral ist dabei, dass **unterschiedliche Akteur*innen** eines Datenökosystems auf Daten zugreifen und diese nutzen, um das volle Innovationspotenzial von Daten auszuschöpfen. Domänenspezifische Datenräume können sich auch mit anderen Datenräumen verbinden, wie beispielsweise der Datenraum Mobilität mit dem Datenraum Tourismus.

Datenkreise/ Data Circles (DK)

Datenkreise sind **themenspezifische Plattformen innerhalb eines Datenraumes**. Sie fokussieren sich auf einen Teilbereich der jeweiligen Domäne und ermöglichen den **Austausch beziehungsweise die Nutzung und den Handel von Daten** in einem klar abgegrenzten Teilgebiet. Voraussetzungen für Datenkreise sind in erster Linie geänderte Grundeinstellungen hinsichtlich des Handels mit Daten: So sollen Daten auf einem virtuellen Datenmarkt gehandelt und schließlich gemeinsam genutzt werden.

Ziel von Datenkreisen ist es, den Austausch, Handel und die gemeinsame Nutzung von anonymisierten und pseudoanonymisierten Daten für Forschung und Technologieentwicklung zu ermöglichen und zu beschleunigen. Da Datenkreise eine kollaborative Zusammenführung von Datenanbieter*innen, Umsetzer*innen und Datenkonsument*innen ermöglichen sollen, liegt das Hauptaugenmerk auf der **Vertrauenswürdigkeit des Datenaustausches beziehungsweise der Datennutzung**.

Beispiele für Arbeitsgruppen, Datenräume und Datenkreise:



Grafik 1: Beispielhafte Darstellung einzelner Arbeitsgruppen, Datenräume und Datenkreise

Veranschaulichung von DR/ DK anhand des Beispiels „Mobilität“:

Die Data Intelligence Offensive hat zum Ziel, Unternehmen, die für eine reibungslose Mobilität in Österreich sorgen, untereinander bundesweit zu vernetzen.

Dazu werden **Datenräume** gebildet, die auf den ersten Blick fachlich weit voneinander entfernte Unternehmen mit verschiedenen Daten (z. B.: Mobilitätsdaten der Mobilfunk Firma 1, Berichte über Unfälle und Stau durch Radiosender, Wetterdaten durch Wetterdienst, Mobilitätsdaten der Mobilfunk Firma 2) für konkrete Use Cases zusammenschließen.

In weiterer Folge setzen die beteiligten Unternehmen in einem **Datenkreis** (beispielsweise Unfallvermeidung oder Verkehrssicherheit) einen konkreten Datenaustausch beziehungsweise Datennutzung unter Wahrung der Datensouveränität um. Dadurch können nachhaltige Services (Stauvorhersagen, Verkehrsflussvorhersagen, Lieferzeitvorhersagen, Simulation von Straßensperrungen) aufgebaut und kommerziell als auch nicht-kommerziell erschlossen werden.

3 Stakeholder*innen der Datenräume und Datenkreise

Für die Gestaltung von Datenräumen müssen relevante Stakeholder*innen und deren Anliegen identifiziert werden:

- Datenkonsument*innen: Sie greifen auf Datenräume zu, um Daten zu verwenden.
- Datenanbieter*innen: Sie sammeln und verwalten Daten und stellen sie in Datenräumen zur Verfügung.
- Datenproduzent*innen: Sie erstellen Daten (z. B.: in der Industrie durch automatisierte Fertigung; Handel).
- Dateneigentümer*innen: Sie haben das Recht, Bedingungen für den Zugriff und die Nutzung von Daten zu gewähren oder zu widerrufen.
- Datenanwendungsanbieter*innen: Sie stellen Anwendungen bereit, die Daten transformieren, verarbeiten oder visualisieren.
- Datenmarktanbieter*innen: Sie bieten Funktionen, die den Betrieb von Datenmarktplätzen ermöglichen.
- Identitätsanbieter*innen: Sie bieten Funktionen zur Identifizierung von Parteien.
- Datenbroker: Sie sammeln Daten, um sie in weiterer Folge an Dritte zu verkaufen.

4 Struktur der Datenräume und Datenkreise

Für einen effektiven Datenaustausch, Datenhandel sowie Datennutzung muss einerseits die Infrastruktur durch den Aufbau einer Daten-Community, als auch die technische Struktur (Engineering) rund um Daten sichergestellt werden.

Der Datenaustausch, Datenhandel sowie die Datennutzung erfolgen über **virtuelle Datenmarktplätze**, auf denen Datenanbieter*innen Daten bereitstellen und Datenkonsument*innen Daten anfordern können.

Dabei ist die Gewährleistung der Datenhoheit/ Datensouveränität essenziell. Jede*r Beteiligte kann Entscheidungen darüber treffen, wie digitale Prozesse, Infrastrukturen und Datenflüsse strukturiert, aufgebaut und verwaltet werden sollen. Als Grundlage dient ein geeignetes **Governance-Schema**, das die Festlegung von Begriffen, Akteur*innen und Bedingungen ermöglicht.

Die **Wahrung der Datensouveränität** stellt eine notwendige Bedingung für den wettbewerbsfähigen und sicheren Aufbau einer europäischen Dateninfrastruktur im Rahmen von GAIA-X und für die Schaffung gleicher Wettbewerbsbedingungen dar.

Die Herstellung von **gleichen Wettbewerbsbedingungen** aller Teilnehmer*innen des Datenraums ist notwendig, um einer monopolistischen Struktur des Datenökosystems entgegenzuwirken. Somit sind neue Marktteilnehmer*innen bei der Beantragung der Zulassung zu einem Datenraum keinen unüberwindbaren Hindernissen ausgesetzt.

Technische Struktur der Datenräume und Datenkreise

Die den Datenräumen und Datenkreisen zugrundeliegende **Software** soll es den teilnehmenden Organisationen ermöglichen, eine Sandboxing-Umgebung aufzubauen, auf der in einem geschützten und experimentellen Rahmen die pilothafte Umsetzung von Data Assets ausprobiert werden kann. Dabei werden Daten

aus mehreren unterschiedlichen Quellen in einer gemeinsamen einheitlichen Datenstruktur zusammengeführt.

Datenquellen können so durch die Teilnehmer*innen gemeinsam genutzt werden, wobei die meisten Daten vorzugsweise dezentral gehalten werden sollen. Dies wird durch den **Aufbau eines gemeinsamen dezentralen Metadatenkatalogs bei dezentraler souveräner Datenhaltung** unterstützt.

Um eine Anbindung an andere Datenplattformen zu gewährleisten sollte die Software in bestehende nationale und internationale Rahmen passen. Die Software soll die **IDS-Referenzarchitektur** verstehen und IDS-zertifiziert sein, weil davon ausgegangen werden kann, dass sich der IDS-Ansatz als de-facto Standard für den Datenaustausch – wie auch im Rahmen der europäischen Initiative GAIA-X geplant – durchsetzen wird.

Wichtige Aspekte bei der Generierung von Datenkreisen:

- **Datensouveränität:** Kontrolle und Herrschaft über eigene Daten durch Aufbau eines sicheren Datenraums mittels IDS-Referenzarchitektur
- **Trust Plattform:** Möglichkeit, Schnittstellen anzubinden
- **IDS-Referenzarchitektur:** Grundlagen-Architekturmodell zur Schaffung von Datenräumen
- **Identitäten:** Nutzung öffentlich zugänglicher Zertifizierungs-/ Identitätsmanagementsysteme (wie etwa der WKO, GAIA-X, Ö-Cloud oder andere)
- **Rechte- und Rollen-System:** Möglichkeit, abgestufte Nutzungsrechte nach Rolle und Partner*in zu vergeben (spezifizierbare Zugriffsberechtigung auch innerhalb des Datenkreises möglich, sodass Datenkreisteilnehmer*innen sehr granulär steuern können, wer welche Daten im Datenkreis sieht – Beispiel: ein*e Datenkreisteilnehmer*in erhält als Wissenschaftler*in Zugang zu allen Daten des Kreises, ein*e Industriepartner*in hingegen nicht zu den Daten des/ der anderen Industriepartner*in.)
- **Metadaten:** Nutzung virtueller oder kaskadierender (stufenförmig) Metadatenkataloge für die Datenkreise (dezentrale Datenhaltung und dezentrales Metadatenmanagement)
- **Anpassbarkeit:** Möglichkeit, Datenkreise domainspezifisch und spezifizierbar auf die Bedürfnisse des jeweiligen Datenkreises/ Industrie/ Thema anzupassen
- **Nutzbarkeit:** User Experience und Niederschwelligkeit als wichtige Komponente
- **Interoperabilität:** Andockmöglichkeit an bestehende Systeme und solche, die im Aufbau sind (wie GAIA-X)
- **Offenheit:** Plattformunabhängigkeit und Anbindung an bestehende (Daten-)Systeme der Teilnehmer*innen der Datenkreise
- **Konnektoren:** Möglichst geringer Aufwand für die Anbindung der Datenkreisteilnehmer*innen an die auszuwählende Software („One API fits all“)
- **Contracting:** Möglichkeit der Nutzung bestehender und neuer Smart-Contracting-Systeme und -Frameworks

5 Die vier Prinzipien von Datenräumen

1. Datenhoheit

Datenhoheit ist die Fähigkeit eines/ einer Dateneigentümer*in zur ausschließlichen **Selbstbestimmung in Bezug auf eigene Daten** als Wirtschaftsgut. Dies ist eines der zentralen Konzepte, das Datenräumen zugrunde liegt. Für Teilnehmende in Datenräumen bedeutet Datenhoheit die Möglichkeit zum Anzeigen, Verarbeiten, Verwalten und Sichern ihrer Daten sowie essenzielle Kontrolle über ihre eigenen Daten, auch, wenn diese anderen Marktteilnehmer*innen zugänglich gemacht werden.

2. Datenaustausch und Datennutzung

Gleiche Wettbewerbsbedingungen für den Datenaustausch, Datenhandel und die Datennutzung können entstehen, wenn ein solches Datenökosystem vor dem Hintergrund der **kollaborativen Zusammenarbeit** anstelle des Wettbewerbs beherrscht wird. Dies kann durch ein solides Design und eine gründliche Wartung der den Datenräumen zugrundeliegenden föderierten Infrastruktur erreicht werden.

Die API-Programmierung macht 60 bis 80 Prozent der Gesamtbetriebskosten aller Datenprojekte aus. Die API-Technologie eines Datenraums kann Datenschnittstellen analysieren und präsentiert diese über eine generalisierte REST-API. Die Funktionen dieser API werden durch die Daten selbst definiert.

Die schnelle und einfache Eingliederung von Schnittstellen (interfaces) ist essenziell beim Datenaustausch (data connection) und der Datennutzung. Dies setzt voraus, dass Standards und organisatorische Tools dafür geschaffen sind. Zudem besteht die Möglichkeit mehrere Datenplattformen miteinander zu verbinden (Plattformverbund/ federated platform), wobei jede Plattform die Kontrolle über ihre eigenen Operationen behält.

3. Dezentrale Infrastruktur

Die Infrastruktur von Datenräumen darf weder eine monolithische, noch zentralisierte IT-Infrastruktur sein. Stattdessen wird die Gesamtheit der interoperablen Implementierungen von Datenräumen berücksichtigt. Dies entspricht einer Reihe von Vereinbarungen in Bezug auf funktionale, technische, betriebliche und rechtliche Aspekte. Eine solche „strukturelle Infrastruktur“ ist für Teilnehmer*innen des Datenraums unsichtbar. Dies beinhaltet Anforderungen in Bezug auf Interoperabilität, Portabilität (Verwendbarkeit derselben Software in verschiedenen Umgebungen), Auffindbarkeit, Sicherheit, Datenschutz und Vertrauenswürdigkeit.

Aus technischer Sicht wird eine **föderierte Infrastruktur** als eine Sammlung interoperabler, API-basierter IT-Plattformen angesehen, auf denen Benutzer*innen den Datenfluss über fortschrittliche Mechanismen des Identitäts- und

Einwilligungsmanagements steuern. Das Design der Infrastruktur umfasst Mechanismen von Datenverbindungen.

4. Datenverwaltungskonzept (Data Governance)

Ein solides **Datenverwaltungskonzept** ist für die Gestaltung, Erstellung und Aufrechterhaltung von Ökosystemen in Bezug auf datengetriebene Entwicklung von entscheidender Bedeutung. Im Rahmen der Data Governance sollen die Rollen und Verantwortlichkeiten **aller Stakeholder*innen** definiert werden. Dazu zählen Unternehmen, Einzelpersonen und Regierungen, die als Datennutzer*innen oder Datenanbieter*innen fungieren sowie deren Technologiepartner*innen und IT-Expert*innen. Zentral bei der Einrichtung von Datenräumen ist die Vertretung öffentlicher Werte und Interessen als auch privater Interessen.

Eine solche Data Governance hat ebenso die Aufgabe, eine **breite Akzeptanz europäischer Datenräume** zu fördern. Dabei sind langfristige Anstrengungen erforderlich, um alle Beteiligten an Bord zu holen.

Die Data Governance ist ein wichtiger Faktor, um das volle Potenzial von Daten zu erschließen. Dazu gehören Tools wie Data Lifecycle Management (Datenverwaltung während ihres Lebenszyklus), Datenrisikomanagement und Datenqualitätsmanagement. Dabei ist besonders in Datenräumen auf Aspekte der Data Literacy, das heißt Datenkompetenz, beziehungsweise den kompetenten Umgang mit Daten hinzuweisen.

6 Die vier Prinzipien von Datenkreisen

1. Vertrauenswürdigkeit (Trust)

Ein vertrauenswürdiger Datenaustausch impliziert, dass die Datenkreise den erwarteten Anforderungen entsprechen. Dies bedeutet, dass die Entwicklung von Daten-

Anwendungen oder Daten-Analysen **in einem sicheren Umfeld** erfolgen muss:

- **Security-by-Design**, das heißt Sicherheit von Datenkreisquellen durch eindeutige, nicht widerlegbare Vereinbarungen (z. B. Smart Contracts)
- **Privacy-by-Design**, das heißt Integration von Datenschutzbedenken in die Entwicklung von Datenplattformen und Datenaustauschanwendungen, ethisches und menschenzentriertes persönliches Datenmanagement
- **Assurance-by-Design**, das heißt Integration von Sicherheits- und Datenschutzerfordernungen in die Entwicklung von Datenplattformen und Datenfreigabeanwendungen

Ein solches vertrauenswürdige Framework umfasst fünf Konzepte: Identifizieren, Schützen, Erkennen, Reagieren und Wiederherstellen. Zu behandelnde Probleme umfassen Zugriffskontrolle, Nutzungskontrolle, Vertrauens- und Identitätsverwaltung. Diese Parameter müssen jedenfalls von einem Datenkreis und dem darüber liegenden Datenraum konzipiert sein.

Da die Datenplattformen dezentral und somit föderiert sind, werden auch Schutzmechanismen verbunden. Das **föderierte Datenschutzmanagement** besteht aus dem

Datenschutzmanagement einzelner Datenbereiche, das durch eine föderierte Zusammenarbeit mit dem globalen Datenschutzmanagement verbunden ist. Dazu ist ein gemeinsamer Rahmen erforderlich, dessen Aufbau auf einer Erweiterung von ISO / IEC 27101 (z. B. NIST-Datenschutzrahmen¹⁹) basieren kann, in der fünf verwandte Konzepte identifiziert werden: Datenschutz identifizieren, Datenschutz regeln, Datenschutz kontrollieren, Datenschutz kommunizieren und Privatsphäre schützen.

Darüber hinaus können Leitlinien aus aktuellen und künftigen Datenschutzstandards verwendet werden. Das **Federated Assurance Management** besteht aus dem Assurance Management einzelner Datenbereiche, das mit einer föderierten Zusammenarbeit beim globalen Sicherheits- und Datenschutz-Assurance-Management verbunden ist.

Zunächst sollte an der Vereinbarung über eine einheitliche individuelle Sicherheit gearbeitet werden, bevor mit der Definition der föderierten Sicherheit begonnen wird.

Bereits seit mehreren Jahren beteiligt sich DIO an verschiedenen Initiativen wie beispielsweise TRUSTS³, um die intelligente Nutzung von Daten voranzutreiben. Das Horizon 2020-Projekt **TRUSTS – Trusted Secure Data Sharing Space** in Kooperation mit der DIO hat zum Ziel, eine Datenaustauschplattform für einen **sicheren, vertrauenswürdigen und DSGVO-konformen Datenaustausch** zu entwickeln. Vertrauen innerhalb eines Datenmarktes ist eine Voraussetzung dafür, dass sich die Teilnehmenden sicher fühlen, auch persönliche Daten zu teilen. Dies ist nicht nur eine Frage der technologischen Entwicklung – auch eine solide rechtliche und ethische Grundlage ist notwendig.

TRUSTS wird die Integration und Einführung zukünftiger Plattformen in verschiedenen Rechtsordnungen ermöglichen.

2. Wirtschaftlichkeit

Die Gesetzmäßigkeit der Wirtschaftlichkeit basiert auf der Annahme, dass Voraussetzungen für den Datenaustausch und -handel geschaffen und konkrete wirtschaftliche Aspekte aus dem Datenaustausch und der datengetriebenen Geschäftsentwicklung gezogen werden müssen.

Es kann zwischen nicht-finanziellen und finanziellen Anreizmechanismen, einschließlich Modelle zur Monetarisierung von Daten und Methoden zur Bestimmung des

³ <https://www.trusts-data.eu/>

Datenwertes, unterschieden werden. Dies bedingt hingegen eine ökonomische Betrachtung des Zielsystems.

3. Interoperabilität

Interoperabilität ist ein entscheidender Faktor für einen störungsfreien Datenaustausch. Dabei soll allen Anwendungen in Datenräumen die Möglichkeit gegeben werden, Daten zu erstellen, zu verwenden, zu übertragen und effektiv auszutauschen. Dies erfordert die **Definition eigener Standards** und die Nutzung bestehender Standards für den Datenaustausch.

Semantische Interoperabilität dient der Sicherstellung der Bedeutung des Datenmodells im Kontext eines Themenbereiches.

4. Flexibilität

Es soll die Möglichkeit geben, **angepasste Funktionen** in Datenverarbeitungsanwendungen und Datenplattformen hinzuzufügen. Aber auch die **Flexibilität in Bezug auf Interoperabilität** soll gegeben sein, das heißt die Erweiterung von Datenräumen mit spezifischen Interoperabilitätsfähigkeiten, sodass diese mit anderen Datenkreisen interagieren können (z. B.: Datenkreis Wettervorhersage mit dem Datenkreis Logistikanwendungen).

In der DIO-Community werden Lösungen für den Aufbau sowie die Verwendung von Datenräumen und Datenkreisen forciert.

7 Die Implementierung von Datenkreisen und Datenräumen

Für die Umsetzung von Datenräumen sowie Datenkreisen bedarf es einer Softwarelösung, welche einen strukturierten und geregelten Datenaustausch erlaubt und somit einen neuartigen digitalen Marktplatz ermöglicht, der Datenanbieter*innen und Datensuchende unter Einhaltung der rechtlichen Voraussetzungen in einem Ökosystem zusammenbringt.

Das österreichische Klimaschutzministerium (BMK), das für Innovation und Technologie zuständig ist, erstellt eine solche innovative Softwarelösung in Kooperation mit einem Technologieunternehmen. Ziel ist es, einen dezentralen und funktionierenden Datenmarkt in Österreich zu errichten.

8 Kompatibilität und Überschneidung zur GAIA-X Architektur

Die Software, die für den Datenaustausch konzipiert wird, sollte **Konformität** zu bestehenden nationalen und internationalen Standards wie GAIA-X, IDSA mit

IDS-Zertifizierung aufweisen, damit eine Interoperabilität mit anderen Lösungen gegeben ist.

9 IDS-Architektur

Die dezentrale **IDS- Referenzarchitektur** stellt die Infrastruktur für einen privaten und öffentlichen Datenaustausch zur Verfügung, mit dem Ziel, vertrauenswürdige Geschäftsbeziehungen zwischen Datenbesitzer*innen, Datenproduzent*innen und Datenkonsument*innen herzustellen. Dabei gewährleistet die Referenzarchitektur Interoperabilität und ermöglicht größtmögliche Transparenz über den Datenaustausch und die Datennutzung.

Das Referenzarchitekturmodell der International Data Spaces besteht aus den folgenden Teilkomponenten:

- **Sicherheitsperspektive:** gewährleistet einen sicheren Datenaustausch und Datenschutz
- **Zertifizierungsperspektive:** Der Zutritt zu den Datenräumen ist mittels einer speziellen **IDS-Zertifizierung**, welche einen sicheren und vertrauenswürdigen Datenaustausch gewährleistet, möglich. Zudem müssen die Software-Komponenten wie der IDS Connector zertifiziert werden (technische Compliance, Interoperabilität).
- **Software/ Technologie** für Datenaustausch, Datenhandel und Datennutzung
- **Governance-Perspektive:** zur Definition der Rollen, Funktionen und Prozesse in den Datenräumen (IDS)

10 DIO Charta

Die Promotoren, Mitglieder und Förderer der Data Intelligence Offensive haben ihre Sichtweise auch in der DIO Charta⁴ zusammengefasst und streben eine Umsetzung der nachfolgend angeführten Ziele als konstitutive Erfolgsfaktoren für die Funktionsfähigkeit und die Optimierung des Wertschöpfungspotentials von Datenmärkten an:

VERTRAUEN UND ETHIK

Der Handel mit Daten und das Marktfunktionieren basieren auf dem Vertrauen der Stakeholder*innen in einander und in die Grundlagen ihres Zusammenwirkens. Dazu gehören unter anderem die klare Feststell- und Nachvollziehbarkeit der Identität der einzelnen Akteur*innen, die dokumentierte Autorisierung ihrer Vertreter*innen und die Zertifizierung der Vollmachten und deren automatisierte Überprüfung. Die gesellschaftliche Akzeptanz und der nachhaltige Fortschritt basiert auf einem ethischen

Umgang mit Daten und den Anwendungen von Data Science und KI Methoden.

⁴ https://www.dataintelligence.at/wp-content/uploads/2021/02/DIO_Charta_20210210.pdf

SICHERHEIT UND DATENSOUVERÄNITÄT

Die Transaktion von Daten hängt entscheidend von der fehlerlosen Authentifizierung, Autorisierung und Garantie von Service Levels ab. DIO sieht in der Verpflichtung zu den höchsten Sicherheitsniveaus von Systemen und Komponenten einen entscheidenden Faktor. Dies betrifft beispielsweise die Sicherheit für Datenanbieter*innen, dass Nutzer*innen ihre Policy nicht umgehen können, wie auch technisch-innovative Lösungen, die Datenauswertungen ohne einen direkten Zugriff auf Originaldatensätze ermöglichen.

DATENQUALITÄT UND INTEROPERABILITÄT

Eine nachhaltige Wertschöpfung setzt einerseits eine möglichst hohe Datenqualität voraus und erfordert andererseits, dass Technologieanbieter*innen den Austausch von Daten zwischen technischen Systemen unabhängig vom Herstellenden umfassend gewährleisten.

DOMÄNENÜBERGREIFENDES ÖKOSYSTEM

Die Entwicklung von Datenmärkten wird dann forciert, wenn Technologien domänenunabhängig funktionieren und Daten domänenübergreifend ausgetauscht und genutzt werden können. Einem verteilten Ansatz folgend verbleiben die Daten physisch bei den Eigentümer*innen und vertrauensstarke Broker leisten den Austausch angefangen bei der Suche bis hin zur Ergebnisverwertung.

TECHNOLOGIEN UND INNOVATION

DIO setzt sich für eine optimale Transparenz und Verlässlichkeit in den eingesetzten Technologien und Lösungen ein. Der verstärkte Einsatz von Data Science, Künstlicher Intelligenz und maschinellem Lernen soll in seinen Herausforderungen gemeinsam gemanagt werden.

ZUKUNFTSSTÄRKE DURCH COMMUNITY

Ziel ist es, ein vernetztes Datensystem zu schaffen. Dies gelingt nur mit einer breiten, interdisziplinären Datencommunity, die bereit ist, Daten, Know-How und Vernetzung einzubringen. DIO dient als Plattform, um dieser Community und ihren Stakeholder*innen ein Feld zu geben, damit sie sich auf ein vernetztes, zukunftsorientiertes Datennetzwerk fokussieren zu können.

Referenzen

Bundesministerium für Wirtschaft und Energie (2019). Das Projekt GAIA-X. *Bundesministerium für Wirtschaft und Energie*, S.1-56.

Bruck, P.A. (2019) Datenökonomie braucht Datenhandel: Data Market Austria als technologischer Showcase und Data Intelligence Offensive als Plattform für Österreich. *WKO Wirtschaftspolitischen Blätter* 3/2019, 247-264. <https://www.wko.at/site/WirtschaftspolitischeBlaetter/datenoekonomie-alle-beitraege-autoren.html>

Eggers, G. et al. (2020). GAIA-X: Technical Architecture. *Federal Ministry for Economic Affairs and Energy*, 1-56.

Gaia-X (2021). Policy Rules Document. *Gaia-X European Association for Data and Cloud AISBL*, 1-7.

Gaia-X Technical Committee et al. (2021). Gaia-X Architecture Document. *Gaia-X European Association for Data and Cloud AISBL*, 1-89.

Otto, B. et al. (2016). Industrial Data Space. *Fraunhofer-Gesellschaft*, 1-40.

Otto, B. et al. (2019). Reference Architecture Model. *International Data Spaces Association*, 1-118.

Reed, D. et al. (2021). Decentralized Identifiers (DIDs) v1.0. Core architecture, data model, and representations. *W3C*.

Schuler, M. (2021). Softwarelösung für Datenkreise. Abgerufen 16. Juni 2021, von <https://www.ioeb-innovationsplattform.at/challenges/detail/softwareloesung-fuer-datenkreise/>