International Green Data Spaces

Orientierungsbericht zu technischen Herausforderungen inklusive Lösungsansätzen

8.2.2024

Förderung durch:

Bundesministerium
Klimaschutz, Umwelt,
Energie, Mobilität,
Innovation und Technologie

FFG
Forschung wirkt.

DIO

Data Intelligence Offensive

Hintere Zollamtsstr. 1/13. OG

1030 Wien

Inhalt

HINWEIS FÜR LESER:INNEN

Dieser technische Orientierungsbericht bietet einen aktuellen Überblick über spezifische technische Aspekte im Kontext von Data Spaces. Die Informationen, die in diesem Bericht präsentiert werden, erheben nicht den Anspruch, sämtliche Aspekte von Data Spaces in ihrer Vollständigkeit zu erschöpfen.

Die Erstellung dieses Berichts erfolgte unter Einbezug einer Vielzahl von Quellen aus dem Fachgebiet. Der Bericht ist als Leitfaden für technische Überlegungen und zu allgemeinen Informationszwecken gedacht, sollte aber nicht als Ersatz für eine eingehende Fachberatung betrachtet werden.

1 Projekt

Das vorliegende Dokument ist ein Orientierungsbericht zu technischen Herausforderungen und Lösungsansätzen im Kontext des folgenden Förderprojekts:

- International Green Data Spaces (InGDS), Green Data Hub
- FFG-Projektnummer: FO999897492
- eCall Nummer: 45648918

Die Autor:innen dieses Berichtes sind:

- Lisa Nußbaumer, Data Space Architect (DIO)
- Sarah Stryeck, Data Steward (DIO)

1.1 Hintergrund

Der Green Data Hub ist ein vom Bundesministerium für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie (BMK) über zwei FFG F&E-Dienstleistungsverträge finanziertes Projekt des non-profit Vereins Data Intelligence Offensive (im Folgenden „DIO", ZVR-Nummer 1683545785, UID-Nr. ATU76576867), das sich zum Ziel gesetzt hat, ein nachhaltiges Daten-Service-Ökosystem aufzubauen, das Unternehmen, Organisationen und öffentliche Verwaltung miteinander vernetzt, um gemeinsam Innovation in Richtung einer nachhaltigen, europäischen Datenwirtschaft zu ermöglichen und somit zur Erreichung der Klimaziele beizutragen. Im Mittelpunkt steht dabei die Vernetzung, die Datenbereitstellung sowie die Zusammenarbeit unter den relevanten Stakeholdern im Daten-Service-Ökosystem. Innerhalb dieses Ökosystems werden bestimmte Data Spaces als Themenbereiche definiert, wie z.B. Energiewende, Mobilitätswende, Kreislaufwirtschaft und Digitaler Klimazwilling, die wiederum als Rahmen für konkrete Use Cases dienen, bei denen Daten ausgetauscht und genutzt werden, um datengetriebene Lösungen mit sozio-ökologischem sowie nachhaltigem Fokus zu entwickeln.

1.2 Ziel des Projektes

Auch wenn die DIO selbst nicht unmittelbar die technische Infrastruktur für den Datenaustausch in einem Data Space bereitstellt, verfolgt sie das übergeordnete Ziel, diesen Austausch zwischen

Datenanbieter:innen und Datenkonsument:innen zu erleichtern. Die Vielzahl der Akteur:innen in einem Data Space führt zu spezifischen technischen Herausforderungen bei der Vernetzung und im Umgang mit Daten. Der vorliegende Orientierungsbericht konzentriert sich darauf, die technischen Rahmenbedingungen und Potenziale für einen effizienten und sicheren Datenaustausch über Domänen- und Ländergrenzen hinweg abzubilden.

2 Data Spaces im Fokus: Überblick und Zusammenhänge

Angesichts der wachsenden Bedeutung von Daten in der modernen Wirtschaft und Industrie rückt das Konzept der Data Spaces in den Mittelpunkt. In den frühen 2000er-Jahren wurde das Konzept der Data Spaces erstmals in der Informatik vorgestellt[1]. Dabei wurde die Idee verfolgt, zentrale Datenbanken durch eine dezentrale Speicherung von Daten an ihrer jeweiligen Quelle zu ersetzen. Unter dieser Prämisse versteht man unter einem Data Space eine verteilte Dateninfrastruktur, die durch einen Governance-Rahmen definiert ist und sichere und vertrauenswürdige Datentransaktionen unter Wahrung der Datensouveränität ermöglicht. Eine Data Space-Architektur besteht aus verschiedenen technischen Komponenten, die gemeinsam sicherstellen, dass Daten innerhalb eines Data Spaces transparent, sicher und zuverlässig ausgetauscht oder gehandelt werden können. Die Gestaltung eines Data Spaces ist nicht auf eine universal anwendbare Lösung beschränkt. Vielmehr ermöglicht ein modularer Ansatz eine flexible Integration unterschiedlicher technischer Komponenten.

Die dynamische Entwicklung der Data Space Technologien und der dafür notwendigen technischen Infrastruktur ist in erster Linie ein kollektiver Prozess, der durch die Zusammenarbeit verschiedener Akteur:innen und Organisationen geprägt ist. Als Grundlage für die Konzeption von Daten-Ökosystemen, die europäischen Standards und Werten entsprechen, dienen die europäische Datenstrategie[2], sowie die normativen Richtlinien und regulatorischen Vorgaben der Europäischen Union in Bezug auf Datenschutz, Datensicherheit und Transparenz (z.B. DGSVO[3], DGA[4], DA, AIA). Darauf aufbauend analysieren unterschiedliche Organisationen die (technischen) Anforderungen an Data Spaces, spezifizieren und operationalisieren diese durch die Entwicklung von ausführbaren Diensten. Einen wesentlichen Beitrag zur länderübergreifenden Identifikation von Anforderungen und der Umsetzung von Data Spaces leisten:

- **Data Space Support Center (DSSC),** ein von der EU finanziertes Projekt, verfolgt das Ziel, gemeinsame Anforderungen und bewährte Praktiken zu definieren, um die Einrichtung souveräner Data Spaces zu beschleunigen. Durch die Überführung dieser Anforderungen in Bausteine, einschließlich der Definition von wichtigen Komponenten, bietet das DSSC ein Forum für die Ausgestaltung der Spezifikationen im Bereich Data Spaces.

[1] (Franklin, Halevy, & Maier, 2005)

[2] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0066

[3] https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=celex:32016R0679

[4] https://eur-lex.europa.eu/eli/reg/2022/868/oj

- **International Data Space Association (IDSA)** ist auf die (Weiter-)Entwicklung von Standards für internationale Data Spaces fokussiert. Das schließt die Definition einer Referenzarchitektur sowie von Protokollen ein, die einen sicheren und interoperablen Datenaustausch gewährleisten. Im Zentrum steht das übergeordnete Ziel der IDSA: die Etablierung von global anerkannten Standards, um die technische Interoperabilität zu fördern und effiziente Datenaustauschmethoden zu unterstützen.
- **Gaia-X** ist eine europäische Initiative und konzentriert sich auf die Entwicklung von Trust & Compliance Komponenten für Data Spaces. Das primäre Ziel besteht darin, eine sichere Umgebung für den Datenaustausch in Data Spaces zu schaffen. In diesem Zusammenhang übernimmt Gaia-X eine führende Rolle bei der Definition von Anforderungen, die das Vertrauen innerhalb eines Data Space gewährleisten sollen.
- **Eclipse Foundation**: dient als Plattform für kollaborative Open-Source-Projekte, die von Entwicklern unterschiedlicher Unternehmen und Privatpersonen vorangetrieben werden. Insbesondere konzentriert sich eine zentrale Arbeitsgruppe der Eclipse Foundation auf die technische Umsetzung von Data Spaces, unter Einsatz der Bausteine aus dem Data Spaces Support Centre (DSSC). Diese gezielte Zusammenarbeit minimiert Redundanzen und fördert die Effizienz bei der Entwicklung von Data Spaces.

Die hier erwähnten Akteur:innen repräsentieren nur einen Ausschnitt der vielfältigen Initiativen, die an der technologischen Gestaltung von Data Spaces arbeiten. Zahlreiche weitere Initiativen und Organisationen leisten ebenfalls bedeutende Beiträge zu der Entwicklung von Data Space Infrastruktur, auch wenn sie im Rahmen dieses Berichts nicht explizit erwähnt werden können. Diese umfassende Beteiligung unterstreicht aber das kollektive Bestreben, ein robustes Datenökosystem für die Zukunft zu formen.

2.1 Arten von Data Spaces

- **Öffentlicher Data Space**

Jede:r darf dem Data Space beitreten – der Data Space ist für jeden sichtbar. Es gibt keine Einschränkungen.

- **Eingeschränkter Data Space**

Der Beitritt ist eingeschränkt – der Data Space ist für jeden sichtbar. Die Teilnehmer:innen geben eine Richtlinie für den Beitritt vor.

- **Privater Data Space**

Der Beitritt ist eingeschränkt – der Data Space ist nicht öffentlich ersichtlich. Die Teilnehmer:innen geben eine Richtlinie für den Beitritt vor

2.2 Teilnehmer:innen in einem Data Space

Die nachfolgende Einteilung von Teilnehmern in einem Data Space wurde teilweise vom Reference Architecture Model[5] (RAM) der International Data Space Association (IDSA) übernommen und für

[5] https://internationaldataspaces.org/publications/ids-ram/

diesen Bericht zusammengefasst. Die aktiven Teilnehmer:innen eines Data Spaces werden in drei Kategorien eingeteilt:

- Zentrale Teilnehmer:innen
- Intermediäre
- Software/Service Anbieter:innen

2.2.1 Zentrale Teilnehmer:innen

Data Space Authority

Eine Data Space Authority definiert, verwaltet und gewährleistet die Grundsätze eines Data Spaces.

Dateninhaber:in (Data Owner)

Im RAM 4.0 werden Dateninhaber:innen als juristische oder natürliche Person definiert, die Daten erstellt und/oder Kontrolle darüber ausüben. Dadurch ist es Dateninhaber:innen möglich, Datennutzungssrichtlinien festzulegen und den Zugriff auf die Daten bereitzustellen. Die Dateninhaberschaft umfasst mindestens zwei wesentliche Konzepte[6]:

- Die (technischen) Möglichkeiten und die Verantwortung, Nutzungsverträge und -richtlinien festzulegen und den Zugriff auf Daten zu ermöglichen.
- Die (technischen) Möglichkeiten und die Verantwortung, das Zahlungsmodell zu definieren, einschließlich des Modells für die Wiederverwendung von Daten durch Dritte.

In der Regel übernehmen Teilnehmende eines Data Spaces, die als Dateninhaber:in fungieren, automatisch auch die Rolle eine:r Datenanbieter:in. Es kann jedoch Fälle geben, in denen ein:e Datenanbieter:in nicht gleich Dateninhaber:in ist. Dies kann der Fall sein, wenn die Daten technisch von einer anderen Einheit als dem Dateninhaber verwaltet werden, z.B. wenn ein Unternehmen einen externen IT-Dienstleister für die Datenverwaltung nutzt, oder wenn die Datenverwaltung an einen Datenverwalter übertragen wird.[7]

Wenn Dateninhaber:innen nicht gleichzeitig als Datenanbieter:innen fungieren, besteht die einzige Aufgabe von Dateninhaber:innen darin, einem Datenanbieter:innen zu autorisieren, die Daten für einen Datenverbraucher:innen verfügbar zu machen.[8]

Datenanbieter:in (Data Provider[9])

Datenanbieter:innen erleichtern den Datenaustausch zwischen Dateninhaber:in und Datenverbraucher:in, indem sie den Zugang zu den Daten bereitstellen. In den meisten Fällen ist der Datenanbieter auch der Dateninhaber, jedoch ist dies nicht zwingend erforderlich. Die zentrale

[6] (Pettenpohl, Spiekermann, & Both, 2022)

[7] (Otto, Steinbuss, Teuscher, & Lohmann, 2019)

[8] (Otto, Steinbuss, Teuscher, & Lohmann, 2019)

[9] Äquivalenter Begriff zu Big data provider in ISO/IEC 20547-3 (vgl. https://www.iso.org/obp/ui/#iso:std:iso-iec:20547:-3:ed-1:v1:en)

Funktion des Datenanbieters besteht darin, Daten von Dateninhaber:innen den Datenkonsument:innen zugänglich zu machen.

Datenverbraucher:in	
Datenkonsument:in (Data Consumer[10])	**Datenbenutzer:in (Data User)[11]**
Datenkonsument:innen fordern Daten an und verwenden die vom Datenanbieter bereitgestellten Daten.	Datenbenutzer:innen sind berechtigt Daten gemäß der festgelegten Nutzungsvereinbarungen zu verwenden.

2.2.2 Intermediäre

Intermediäre sind Organisationen, die den Datenaustausch zwischen Datenanbieter:innen und Datenkonsument:innen erleichtern. Sie spielen eine entscheidende Rolle, um einen fairen Zugang zu Daten für Teilnehmer:innen aller Größen sicherzustellen, insbesondere für Start-ups und kleine und mittlere Unternehmen (KMU), die Daten teilen oder nutzen möchten. Dienste von Intermediären unterstützen sowohl den bilateralen als auch den multilateralen Austausch von Daten sowie die Schaffung von Plattformen zur gemeinsamen Nutzung von Daten in Daten-Ökosystemen.

Broker

Ein Broker-Dienstleister speichert und verwaltet Informationen über verfügbare Datenquellen in Data Spaces. Dabei können mehrere Broker-Dienstleister gleichzeitig existieren, um unterschiedliche Anwendungsbereiche abzudecken. Organisationen, die Broker-Dienste anbieten, können auch andere Vermittlerrollen übernehmen. Der Schwerpunkt der Aktivitäten eines Broker-Dienstleisters liegt auf dem Empfang und der Bereitstellung von Metadaten. Über die bereitgestellte Schnittstelle können Datenanbieter:innen ihre Metadaten an den Broker-Dienstleister senden. Diese Metadaten werden in einem internen Repository gespeichert und können von Datenverbraucher:innen strukturiert abgefragt werden.[12]

Clearingstelle (Clearing House)

Die Clearingstelle fungiert als Vermittler für Clearing- und Abwicklungsdienstleistungen im Rahmen von Finanz- und Datenaustauschtransaktionen. In Data Spaces erfolgt eine klare Trennung zwischen den Clearing-Aktivitäten und den Broker-Diensten, da sie technisch verschiedene Aspekte umfassen – Clearing bezieht sich auf den Transfer von Vermögenswerten, während Broker-Dienste sich auf die Verwaltung eines Metadaten-Repositories konzentrieren. Dennoch besteht die Möglichkeit, dass dieselbe Organisation sowohl die Rolle der Clearingstelle als auch des Broker-Dienstleisters übernimmt, da beide als vertrauenswürdige Vermittler zwischen Datenanbieter:innen und Datenverbraucher:innen agieren. Die Clearingstelle dokumentiert sämtliche Aktivitäten im Rahmen des Datenaustauschs. Nach Abschluss eines Datenaustauschs bestätigen sowohl der

[10] Äquivalenter Begriff zu Big data consumer in ISO/IEC 20547-3 (vgl. https://www.iso.org/obp/ui/#iso:std:iso-iec:20547:-3:ed-1:v1:en)

[11] Sollten im Text Datenkonsument:innen und Datenbenutzer:innen gemeint sein, werden diese unter dem Begriff Datenverbraucher:innen subsumiert.

[12] (Pettenpohl, Spiekermann, & Both, 2022)

Datenanbieter als auch der Datenverbraucher den erfolgreichen Übertragungsprozess durch die Protokollierung der Transaktionsdetails bei der Clearingstelle. Diese Protokollierungsinformationen dienen der Abrechnung und können ebenso zur Konfliktlösung herangezogen werden, beispielsweise zur Überprüfung, ob ein Datenpaket vom Datenverbraucher empfangen wurde oder nicht. Darüber hinaus erstellt die Clearingstelle Berichte über durchgeführte Transaktionen, die für Abrechnungszwecke, Konfliktlösung usw. verwendet werden können.[13]

Identitätsanbieter (Identity Provider)

Identitätsanbieter:innen sind verantwortlich für die Bereitstellung eines Dienstes, der die Erstellung, Pflege, Verwaltung, Überwachung und Validierung von Identitätsinformationen für die Teilnehmer:innen des Data Spaces gewährleistet. Diese Funktionalitäten sind von essenzieller Bedeutung, um einen sicheren Betrieb von Data Spaces sicherzustellen und unbefugten Zugriff auf Daten zu verhindern.

2.2.3 Software/Service Anbieter:innen

In der Kategorie der Software- und Serviceanbieter:innen finden sich IT-Unternehmen, die den Teilnehmer:innen eines Data Spaces, Software und/oder Dienstleistungen bereitstellen, beispielsweise im Rahmen eines Software-as-a-Service-Modells.

Data Space Provider

Ein Data Space Provider stellt Technologie und gemeinsam genutzte Dienste für den Betrieb eines Data Spaces bereit und verwaltet diese.

Serviceanbieter:innen

Wenn Teilnehmer:innen eines Data Spaces nicht über die erforderliche technische Infrastruktur verfügen, besteht die Möglichkeit, die Daten an einen Serviceanbieter zu übertragen, der die benötigte Infrastruktur für andere Organisationen bereitstellt. Diese Rolle schließt auch Anbieter:innen ein, die zusätzliche Datendienste bereitstellen, um die Qualität der im Data Space ausgetauschten Daten zu verbessern, beispielsweise Datenanalyse, Datenintegration, Datenbereinigung oder semantische Anreicherung. Ein solcher Serviceanbieter kann sowohl als Datenanbieter als auch als Datenkonsument betrachtet werden, da er Daten von einem Anbieter empfangen, seinen spezifischen Service anbieten und anschließend selbst Daten im Data Space bereitstellen kann.[14]

Softwareanbieter:innen

Softwareanbieter:innen implementieren die Funktionalitäten eines Data Spaces durch die Bereitstellung von Softwarekomponenten gemäß den festgelegten Anforderungen. Die Nutzung

[13] (Otto, Steinbuss, Teuscher, & Lohmann, 2019)

[14] (Otto, Steinbuss, Teuscher, & Lohmann, 2019)

dieser Software erfolgt auf Grundlage individueller Vereinbarungen zwischen dem Softwareanbieter und dem Benutzer, sei es als Datenkonsument:in, Datenanbieter:in oder Broker-Dienstleister.[15]

2.3 Interoperabilität beim Datenaustausch/-handel

In einem Data Space, einer föderierten und offenen Dateninfrastruktur, ist Interoperabilität von entscheidender Bedeutung für den nahtlosen Datenaustausch zwischen technischen Komponenten. Dies schließt gemeinsame Datenstandards und -formate, definierte Kommunikationsprotokolle, Schnittstellenspezifikationen, Zugriffs- und Nutzungsrechte, Sicherheitsmechanismen, Datenintegration und -präsentation sowie Identitätsmanagement und Authentifizierung mit ein. Die aktuelle Aufnahme erster Softwarekomponenten in den ISO-Zertifizierungsprozess[16] betont die zunehmende Relevanz von technischen Standards in Data Spaces. Diese Entwicklung stellt sicher, dass technische Elemente gemäß anerkannten Normen entwickelt werde. Damit wird eine Basis geschaffen, die reibungslose Interoperabilität, erhöhte Sicherheit und gesteigerte Effizienz in Data Spaces ermöglicht.

Das Dataspace Protocol[17] der International Data Spaces Association (IDSA) bildet die technische Grundlage für Interoperabilität in und zwischen Data Spaces. Das Dataspace Protocol definiert Strukturen und Protokolle, um den technischen Datenaustausch zu orchestrieren. Dies umfasst die Spezifikation von Regeln und Vereinbarungen, die die Kommunikation, Authentifizierung, Autorisierung und Datenübertragung zwischen autonomen Einheiten (=Teilnehmer:innen) innerhalb eines Data Spaces ermöglichen.

2.3.1 Herausforderungen

Interoperabilität von Altsystemen

Durch die Entwicklung von Anwendungen und Informationssystemen als (domänen-)spezifische und lokale Lösungen (bottom-up Lösung), sind in einigen Fällen fragmentierte IKT-Inseln entstanden (z.B.: öffentliche Verwaltung), die schwer miteinander interoperieren.

Interoperabilität von Metadaten und Daten

Gemeinsame Vokabulare zur Standardisierung von Spezifikationen für die Beschreibung von Datensätzen und Service schaffen eine einheitliche Basis in einem Data Space. Derzeit sind Gaia-X-Kataloge für Services, Organisationen und Datensätze als Standards verfügbar.

Interoperabilität von Konnektoren

Interaktionen in einem Data Space erfolgen meist über Konnektoren. Da eine Vielzahl an Konnektoren sowohl als open-source, geschützte als auch closed-source Ausführungen verfügbar sind, ist Interoperabilität nicht bei allen Kombinationen gegeben.

Internationale Standards für Datensätze

---

[15] (Pettenpohl, Spiekermann, & Both, 2022)

[16] https://www.iso.org/standard/86589.html

[17] https://docs.internationaldataspaces.org/ids-knowledgebase/v/dataspace-protocol/overview/readme

Um Fragmentierung, hohe Integrationskosten und die Bildung von isolierten Datensilos zu vermeiden, rät die EU-Kommission im Hinblick auf gemeinsame europäische Data Spaces, die Datensätze auf internationalen Standards wie INSPIRE[18] (für räumliche Daten) und den FAIR-Prinzipien[19] aufzubauen.

2.3.2 Lösungsansätze

- **Verwendung von offenen Spezifikationen**, um technische Interoperabilität sicherzustellen.
- Mindestinteroperabilitätsmechanismen definieren, die sowohl funktionale als auch technische Fragestellungen abdecken
- Universelle Plugins zur Integration von Lösungen
- W3C Verifiable Credentials-Standard
- Interoperabilität von digital wallets
- Standardisierte Vokabulare (Semantik, Datenmodelle, APIs): JSON-LD für Skill-Datenmodelle, semantische Übersetzer, GraphQL-API und Open Badges v3 (OBv3).

2.4 Datensouveränität/Datenhoheit

In der Literatur werden häufig Begriffe wie "Datenhoheit" und "Datensouveränität" verwendet, wobei ihre Bedeutung nicht immer eindeutig ist. Der Begriff "Datensouveränität" bezieht sich auf das Konzept, dass Daten den gesetzlichen Bestimmungen und Governance-Strukturen des Landes oder der Region unterliegen, in denen sie sich befinden[20]. Das impliziert, dass Einzelpersonen und Organisationen die Kontrolle über ihre eigenen Daten besitzen und darüber entscheiden können, wie diese gesammelt, verarbeitet, gespeichert und weitergegeben werden. Die Thematik der Datensouveränität ist insbesondere im Kontext des globalen Datenverkehrs von hoher Relevanz, da verschiedene Länder unterschiedliche Gesetze und Regelungen zum Schutz der Datenprivatsphäre und -sicherheit haben. Ein Beispiel hierfür ist die Datenschutz-Grundverordnung (DSGVO) der Europäischen Union, die strenge Richtlinien zur Sammlung und Verarbeitung personenbezogener Daten innerhalb der EU festlegt.

Wichtige Faktoren:

- **Transparenz und Kontrolle:** Dateninhaber:innen sollten transparente Einblicke in die Verwendung ihrer Daten haben und die Kontrolle darüber behalten können, wie ihre Daten genutzt werden. Mechanismen zur Verfolgung und Überwachung von Datenflüssen sowie zur einfachen Verwaltung von Zugriffsrechten können dazu beitragen, die Datensouveränität zu gewährleisten.

[18] INSPIRE EU-weit: https://inspire.ec.europa.eu/inspire-principles/9; INSPIRE für Österreich: https://www.inspire.gv.at/Allgemein/Standards-und-Normen.html

[19] FAIR: **F**indable, **A**ccessible, **I**nteroperable und **R**e-usable (siehe https://www.go-fair.org/fair-principles/)

[20] (Hummel, Braun, Tretter, & Dabrock, 2021)

- **Standards und Interoperabilität**: Die Einhaltung von branchenweiten Standards und die Förderung von Interoperabilität zwischen verschiedenen Datenplattformen können die Datensouveränität stärken, indem sie die Portabilität von Daten und die Auswahlmöglichkeiten für Benutzer:innen erhöhen.
- **Technologische Souveränität:** Die Fähigkeit, technologische Entscheidungen unabhängig zu treffen und technologische Lösungen selbst zu kontrollieren, ist ein weiterer wichtiger Aspekt der Datensouveränität. Die Nutzung von Open-Source-Software und die Förderung von technologischer Vielfalt können dazu beitragen, die Abhängigkeit von einzelnen Anbieter:innen zu verringern und die Souveränität über die eigenen Daten zu stärken.

3 Technische Infrastruktur von Data Spaces

In diesem Kapitel wird eine lösungsneutrale und modulare Struktur für die Gestaltung von Data Spaces anhand von unterschiedlichen technischen Bausteinen (= Komponenten) vorgestellt, die je nach individuellen Anforderungen an die Funktionsweise eines Data Spaces kombinierbar sind. Als Referenzpunkte für die hier vorgestellten Bausteine dienen die Technical Building Blocks[21] empfohlen vom DSSC.

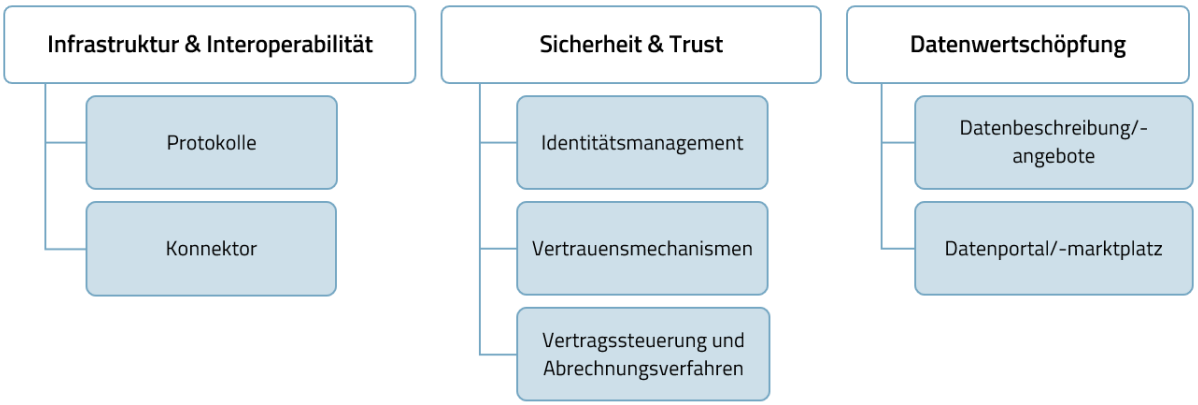

Abbildung 1: Technische Bausteine für die Gestaltung eines Data Spaces

3.1 Infrastruktur und Interoperabilität

Vor der Betrachtung der technischen Bausteine der Infrastruktur ist es wichtig, das IDSA-Referenzarchitekturmodell (IDS-RAM)[22] (aktuelle Version 4.0) zu erwähnen. Obwohl das IDS-RAM kein eigenständiger technischer Baustein ist, bietet es einen grundlegenden Rahmen für die Gestaltung sicherer und interoperabler Datensysteme. Es definiert eine technische Spezifikation, die eine standardisierte Architektur als Richtlinie bereitstellt für bewährte Praktiken und Standards, denen Organisationen folgen können, um ihre technologische Infrastruktur zu konzipieren und umzusetzen. Sein primäres Ziel besteht darin, Organisationen die nahtlose gemeinsame Nutzung von Daten- und KI-Ressourcen über diverse Unternehmensstandorte zu ermöglichen, wobei jedoch

[21] https://dssc.eu/space/BBE/178422228/Technical+Building+Blocks (version: October, 9.2023)

[22] https://internationaldataspaces.org/offers/reference-architecture/

stets die Datensouveränität gewährleistet und die Konformität mit rechtlichen Vorschriften und regulatorischen Anforderungen sichergestellt wird.

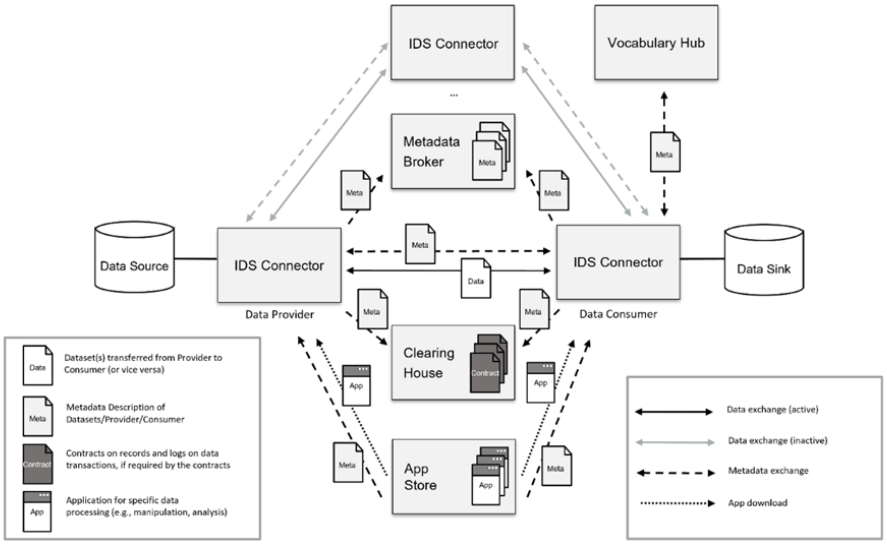

Abbildung 2: Systemdiagramm des IDSA RAM, aus Abschnitt 3.5: Systemschicht des IDS RAM 4.0[23]

Die Kernpunkte des IDS-RAM beinhalten:

- **IDS-Architektur**: Die IDS-Architektur basiert auf Konnektoren, die als sichere Gateways für den Datenaustausch zwischen Teilnehmer:innen dienen. Diese Konnektoren sind in eine Steuerungsebene (für Identitäts- und Zugriffsmanagement, Vertragsbearbeitung, Protokollierung usw.) und eine Datenebene (für die eigentliche Datenübertragung) unterteilt.
- **IDS-RAM-Komponenten**: Die IDS-RAM-Komponenten umfassen Identitätsanbieter, Metadaten-Broker, Clearing House und einen App Store für zusätzliche Datendienste. Diese Dienste sind modular und interoperabel, was eine nahtlose Integration mit anderen Systemen ermöglicht.
- **Data Space-Komponenten**: Ein Data Space setzt sich aus Asset-Anbietern und -Konsumenten, Konformitätsdiensten, Identitätsdiensten, einem Katalog für Vermögenswerte und Datenaustauschdiensten zusammen. Diese Komponenten gewährleisten einen sicheren und effizienten Datenaustausch innerhalb des Systems.
- **Datenaustausch**: Der Datenaustausch umfasst Datenübertragung, Vertragsabschlüsse, Nutzungsrichtlinien und Protokollierung. Die Komplexität dieser Dienste variiert je nach den Anforderungen und dem Kontext des Datenraums.
- **Katalog**: Ein zentraler Katalog enthält standardisierte Beschreibungen von Diensten und ermöglicht es den Teilnehmer:innen, gezielt nach relevanten Diensten zu suchen und sie auszuwählen.

[23] https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0/blob/main/documentation/3_Layers_of_the_Reference_Architecture_Model/3_5_System_Layer/3_5_0_System_Layer.md#system-layer

- **Identitätsdienste**: Diese Dienste umfassen die Verwaltung von Identitäten und Berechtigungen, häufig basierend auf selbstverwalteten Identitäten (SSI) und Verifiable Credentials. Sie stellen sicher, dass nur autorisierte Benutzer auf die Daten zugreifen können.
- **Regelkonformität**: Spezialisierte Dienste gewährleisten die Einhaltung von Vorschriften und Standards, oft durch Zertifikate wie im Gaia-X Trust Framework. Dies stellt sicher, dass der Datenaustausch den geltenden gesetzlichen Anforderungen entspricht.
- **Zusätzliche Komponenten**: Dazu können Portale für den Zugang zum Data Space, Orchestrierungsdienste und Plugins für spezifische Funktionen gehören. Diese Komponenten erweitern die Funktionalität des Systems und bieten zusätzliche Möglichkeiten zur Anpassung an spezifische Anforderungen.

3.1.1 Protokolle

Im Kontext von Datentransaktionen innerhalb eines Data Spaces spielen Kommunikationsprotokolle eine entscheidende Rolle bei der Gewährleistung eines sicheren und effizienten Datenaustauschs. Abbildung 3 beschreibt die Interaktion von Teilnehmer:innen in einem Data Space schematisch. Die Interaktion erfolgt über Konnektoren, die verschiedene Protokolle implementieren (Catalog Protocol, Contract Negotiation Protocol, Transfer Process Protocol etc.). Die meisten Interaktionen finden zwischen den Konnektoren statt, aber für einen erfolgreichen Datenaustausch sind auch Interaktionen mit anderen Systemen erforderlich.

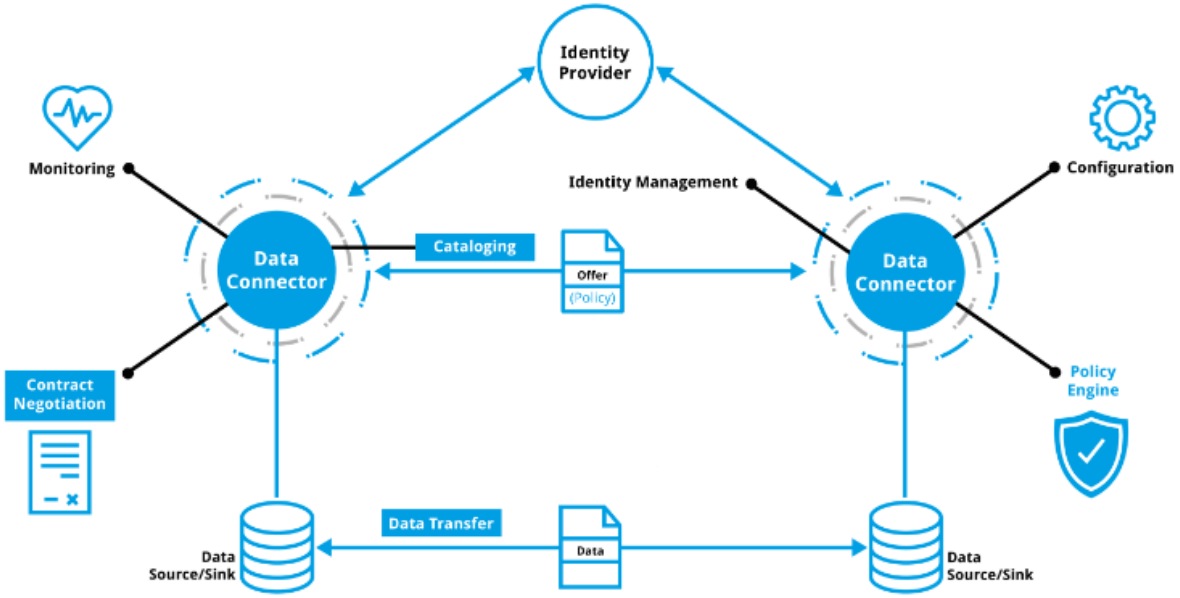

Abbildung 3: Schematische Abbildung zu technischer Interoperabilität in Data Spaces[24]

[24] übernommen von: https://docs.internationaldataspaces.org/ids-knowledgebase/v/dataspace-protocol/overview/readme)

Lösungen:

- **IDS-Protocol**: Das IDS (International Data Spaces) Protokoll[25] ist darauf ausgelegt, eine sichere und vertrauenswürdige Kommunikation innerhalb von Datenräumen zu ermöglichen. Es umfasst Unterprotokolle wie Multipart oder IDSCP, die jeweils verschiedene Aspekte des Datentransfers wie Nachrichtenformatierung, Verschlüsselung und Authentifizierung behandeln. Das IDS-Protokoll betont die Datensouveränität, indem es strenge Richtlinien zur Datenverwendung durchsetzt und sicherstellt, dass die Daten unter der Kontrolle des Datenanbieters bleiben. Durch die Nutzung des IDS-Protokolls kann der Datenaustausch innerhalb von Datenräumen auf vertrauenswürdige und konforme Weise erfolgen, was die Interoperabilität und das Vertrauen zwischen den Teilnehmern fördert.
- **Dataspace Protocol**: Das Dataspace Protokoll[26] repräsentiert durch HTTPS (Hypertext Transfer Protocol Secure), ist ein weit verbreitetes Protokoll für den sicheren Datentransfer über das Internet. HTTPS verwendet robuste Verschlüsselungsmechanismen während der Datenübertragung, um die Vertraulichkeit und Integrität der ausgetauschten Informationen zu gewährleisten. Es gewährleistet die sichere Übertragung von Daten zwischen verschiedenen Datenbanken oder Systemen innerhalb des Data Spaces und schützt sie so vor unbefugtem Zugriff oder Manipulation. Durch die Nutzung des Dataspace Protokolls können Organisationen sichere Verbindungen für den Datenaustausch herstellen und damit die Risiken im Zusammenhang mit Datenlecks oder unbefugtem Datenzugriff minimieren.
- **Kommunikationsprotokolle**: Verschiedene Kommunikationsprotokolle wie MQTT (Message Queuing Telemetry Transport) und AMQP (Advanced Message Queuing Protocol) werden für den Echtzeitdatenaustausch und die Nachrichtenübermittlung innerhalb von Data Spaces verwendet[27]. Diese Protokolle ermöglichen eine effiziente und zuverlässige Kommunikation zwischen Geräten, Sensoren und Systemen in verteilten Datenumgebungen und unterstützen die Integration von IoT (Internet der Dinge) und Sensordaten.
- **Übertragungsprotokolle**: Übertragungsprotokolle wie FTP (File Transfer Protocol), SFTP (Secure File Transfer Protocol) und REST (Representational State Transfer) werden häufig zum Datenaustausch zwischen Systemen und Anwendungen in Data Spaces verwendet[28]. Diese Protokolle definieren Regeln und Mechanismen für den Datentransfer, um die

[25] (Pampus, Jahnke, & Quensel, 2022)

[26] (Olms, et al., 2020)

[27] (Gieß, Hupperz, Schoormann, & Möller, 2024)

[28] (Gieß, Hupperz, Schoormann, & Möller, 2024)

Datenintegrität, Sicherheit und Interoperabilität in Datenaustauschprozessen sicherzustellen.

Insgesamt sind Protokolle wesentliche Bestandteile des Datentransfers innerhalb von Data Spaces und ermöglichen eine sichere, vertrauenswürdige und konforme Kommunikation zwischen Datenquellen und -nutzern[29]. Die Einhaltung dieser Protokolle gewährleistet die Vertraulichkeit, Integrität und Authentizität der Datenaustausche und trägt somit zur Datensouveränität und Interoperabilität innerhalb eines Data Space-Ökosystems bei.

3.1.2 Konnektor

Ein Data Space Konnektor ist eine essenzielle Komponente, die Verbindungen zu Datenquellen herstellt, Metadaten und Nutzungsbedingungen verwaltet und den sicheren Austausch und das vertrauenswürdige Teilen von Daten zwischen verschiedenen Entitäten erleichtert. Konnektoren spielen eine entscheidende Rolle bei der Sicherstellung der Datensouveränität, indem sie die Einhaltung von Datennutzungsrichtlinien durchsetzen und die Interoperabilität beim Datenaustausch innerhalb von Data Spaces unterstützen[30]. Sie können auf verschiedene Arten implementiert werden, wie beispielsweise als Software as a Service, Platform as a Service oder Self-Service-Optionen. Funktional fungiert ein Konnektor als Brücke zwischen Datenanbieter:in und -nutzer:in, indem er die Schnittstellen (=Datenendpunkte) in den jeweiligen Systemen oder Plattformen dieser Entitäten bereitstellt oder implementiert und es somit ermöglicht, dass Teilnehmer:innen eines Data Spaces, Daten austauschen, verarbeiten und empfangen können, während die Datensouveranität gewahrt bleibt. Darüber hinaus können Konnektoren je nach Kontext plattformagnostisch oder spezifisch gestaltet sein und verschiedene Kommunikations- und Übertragungsprotokolle unterstützen, um die Interoperabilität und Flexibilität während einer Datentransaktion zu verbessern[31].

Lösungen:

In regelmäßigen Abständen veröffentlicht die IDSA einen Data Space Connector Report, der einen umfassenden Überblick über das aktuelle Angebot an und die Entwicklungen von Data Space Konnektoren (inklusive Open-Source- und proprietärer Software) bietet. Der aktuelle Connector Report[32] (Stand: Jänner 2024) umfasst 30 Data Space Konnektoren, darunter 13 Open-Source-, zwei teilweise Open-Source- und 15 Closed-Source-Konnektoren. Er identifiziert vier Hauptentwicklungslinien für Data Space Konnektoren: Eclipse Dataspace Components, Dataspace Connector, TNO Security Gateway und TRUE Connector (siehe Abbildung 4).

[29] (Gieß, Hupperz, Schoormann, & Möller, 2024)

[30] (Hellmeier & von Scherenberg, 2023)

[31] (Gieß, Hupperz, Schoormann, & Möller, 2024)

[32] https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-Data-Connector-Report-89-No-11-January-2024.pdf

Der **Eclipse Data Space Connector (EDC)**[33] hat sich als der de-facto **offene** Standard für Konnektoren innerhalb der aktuellen Data Space Landschaft etabliert. In (Gieß, Hupperz, Schoormann, & Möller, 2024) wird der EDC als Beispiel für die Vielseitigkeit von Konnektoren hervorgehoben. Der EDC zeichnet sich durch seine plattformagnostische Natur aus, was bedeutet, dass er problemlos in verschiedene Cloud-Infrastrukturen integriert werden kann.

Wesentliche technische Merkmale des EDC:

- Modulare Architektur: Das EDC-Framework folgt einer modularen Architektur, die es Entwicklern ermöglicht, die Funktionalität von Data Space Konnektoren basierend auf spezifischen Anforderungen anzupassen und zu erweitern.
- Protokollunterstützung: Der EDC unterstützt verschiedene Data Space Protokolle, einschließlich des IDS-Protokolls, HTTPS und anderer Kommunikationsprotokolle, um Kompatibilität mit verschiedenen Data Space Umgebungen sicherzustellen.
- Sicherheit und Compliance: Das EDC-Framework integriert Sicherheitsfunktionen und Compliance-Kontrollen, um einen sicheren Datenaustausch, die Datenhoheit und die Einhaltung von Data Governance Richtlinien innerhalb von Data Spaces zu gewährleisten.

[33] https://github.com/eclipse-edc/Connector

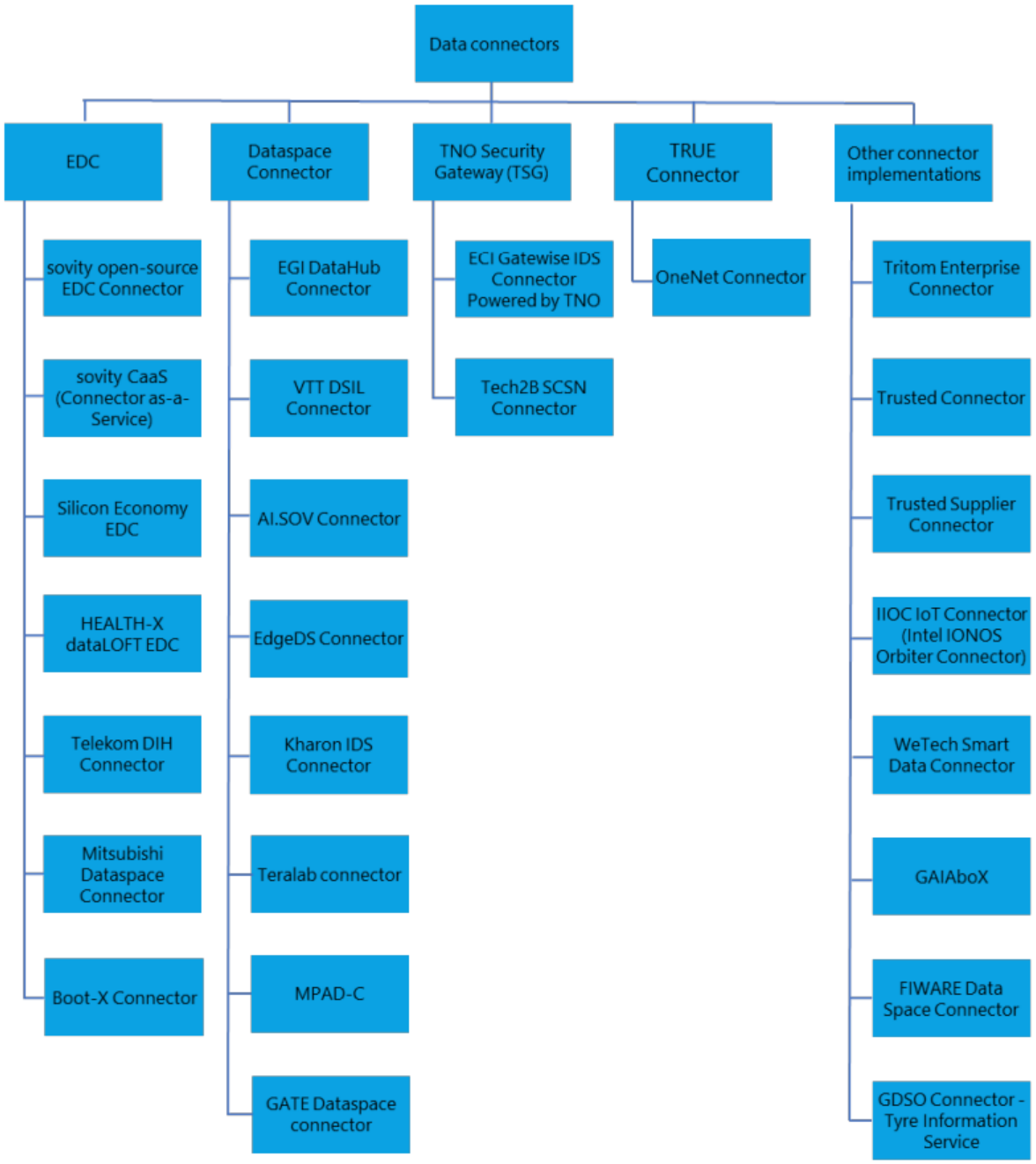

Abbildung 4: Data Space Konnektoren Entwicklungslinien (vgl. Fig. 5 in IDSA Connector Report 01/2024)

3.2 Sicherheit und Trust

3.2.1 Identitätsmanagament

Das Identitätsmanagement ermöglicht die Identifizierung, Authentifizierung und Autorisierung von Teilnehmer:innen innerhalb eines Data Spaces. Es gewährleistet, dass Organisationen, Einzelpersonen, Maschinen und andere Akteur:innen digitale Identitäten erhalten und diese Identitäten authentifiziert und überprüft werden können. Zusätzlich ermöglicht es die Bereitstellung von zusätzlichen Informationen, die von Autorisierungsmechanismen genutzt werden, um Zugriffs- und Nutzungskontrolle zu ermöglichen.

In der analogen Welt kann die Identität oft durch physische Objekte wie Personalausweis oder Reisepass nachgewiesen werden. In der digitalen Welt gestaltet sich diese Identifizierungsformen jedoch schwieriger. Im Kontext des Identitätsmanagements wird eine Identität als die Darstellung einer Entität in einem spezifischen Umfeld verstanden[34]. Diese Identität setzt sich einerseits aus Identifiern zusammen, also eindeutigen Kennungen zur Identifizierung einer Organisation, eines Objekts oder einer Person. Andererseits umfasst sie auch die Attribute dieser Entität, wie beispielsweise Passwörter oder demografische Informationen. Identitätsanbieter (IdP oder IDP) speichern und verwalten die digitalen Identitäten von Nutzer:innen und ermöglichen die Authentifizierung von Nutzer:innen für Single Sign-On Dienste (SSO) und die Autorisierung des Zugriffs auf Ressourcen durch spezielle APIs[35]. Die Kommunikation zwischen dem IDP und den Service Providern erfolgt über Sicherheitsprotokolle wie SAML, OpenID oder OAuth.

Eine Self-Sovereign Identity (SSI) ermöglicht die Erzeugung digitaler Identitäten ohne Abhängigkeit von Vermittlern oder zentralen Parteien. Das Ziel von SSI besteht darin, die individuelle Kontrolle, Sicherheit und Portabilität digitaler Identitäten über verschiedene Dienste hinweg zu gewährleisten. Ein zentraler Baustein dabei sind digitale Zertifikate (Credentials). Diese Zertifikate können sowohl selbstattestierte Identitätsattribute als auch von Dritten bestätigte Attribute enthalten. Die von Dritten bestätigten Credentials werden als Verifiable Credentials (VCs) bezeichnet, für die bereits ein Standard existiert, der die Struktur eines solchen Zertifikats festlegt[36]. VCs sind entscheidende Elemente zur Nachweisführung von Identitätsattributen zwischen den verschiedenen Rollen in einer SSI-Lösung, wie sie schematisch Abbildung 5 dargestellt sind.

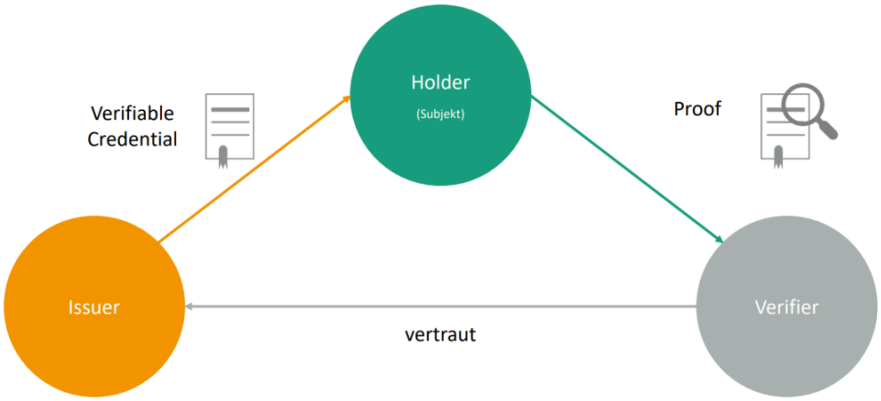

Abbildung 5: Rollen in einem SSI-System (vgl. Abb.5 in (Strüker, et al., 2021))

Um eine sichere Kommunikation und den Schutz der Privatsphäre zu gewährleisten, ermöglicht der DID-Standard (Decentralized Identifiers) den Parteien oder Teilnehmer:innen eines Data Spaces, eine Ende-zu-Ende-verschlüsselte bilaterale Kommunikation auf verschiedenen Infrastrukturen einzurichten49. Dadurch wird eine verlässliche und geschützte Übertragung von Informationen

[34] (Strüker, et al., 2021)

[35] (Pohlmann, 2019)

[36] https://www.w3.org/TR/vc-data-model/Overview.html

gewährleistet. Ein DID besteht dabei aus dem URL-Schema DID, gefolgt von einer DID-Methode und einem DID-methodenspezifischen Identifier (siehe Abbildung 6)[37].

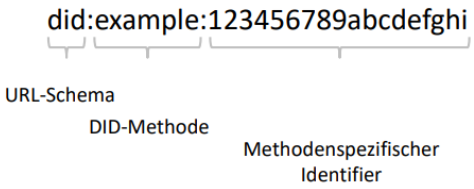

Abbildung 6: Bestandteile eines DID (vgl. Abb. 6 in (Strüker, et al., 2021))

Die Einrichtung von Identitätsverwaltungssystemen spielt eine entscheidende Rolle für einen sicheren und transparenten Datenaustausch zwischen den Teilnehmenden. Diese Systeme bieten die Möglichkeit, Berechtigungen zu überprüfen, ohne dabei sensible Informationen offenzulegen. Bei der Umsetzung des Identitätsmanagements können bestehende Plattformen genutzt werden, die bereits Teile der erforderlichen Funktionalität bereitstellen. Zu den Beispielen für Open-Source-Lösungen zählen die Apache Syncope-Identitätsmanagement-Plattform[38], das FIWARE-Identitätsmanagement-Framework[39], die KeyCloak-Infrastruktur[40] sowie die Open-Source-Identitätsmanagement-Plattform des Shibboleth-Konsortiums[41]. Zusätzlich könnte die Integration nationaler e-IDs wie der österreichischen ID Austria[42] und einer europäischen digitalen Identität EUid[43] in Identitätsverwaltungssysteme die Teilnahme an Data Spaces erleichtern. Diese digitalen Identitäten ermöglichen eine nahtlose Authentifizierung und Autorisierung für den Zugriff auf Data Spaces, was zu einer verbesserten Interoperabilität und Benutzerfreundlichkeit führt.

Neben dem Identitätsmanagement kann auch ein rollen/-attributbasierter Zugriff innerhalb eines Data Spaces von Bedeutung sein. Eine rollenbasierte Zugriffsverwaltung ermöglicht es Teilnehmenden in einem Data Space, verschiedene Rollen einzunehmen, darunter beispielsweise Administrator, Benutzergruppe: Forschung, Benutzergruppe: Wirtschaft, einfaches Mitglied und Externe. Je nach zugewiesener Rolle könnten unterschiedliche Zugriffsoptionen für die verfügbaren Daten im Data Space definiert werden. Beispielsweise könnte die Benutzergruppe "Forschung" uneingeschränkten Zugriff auf alle Daten haben, um Forschungsprojekte durchzuführen. Im Gegensatz dazu könnte die Benutzergruppe "Wirtschaft" nur beschränkten Zugriff haben, der auf bestimmte Datensätze oder -bereiche beschränkt ist, um die Vertraulichkeit von geschäftskritischen Informationen zu wahren. Dieses differenzierte Zugriffsschema ermöglicht eine

[37] Detaillierte Informationen zur DID-Syntax: https://www.w3.org/TR/did-core/Overview.html#did-syntax
[38] https://syncope.apache.org/
[39] https://github.com/FIWARE/catalogue/tree/master/security
[40] https://www.keycloak.org/
[41] https://shibboleth.atlassian.net/wiki/spaces/IDP4/overview
[42] https://www.oesterreich.gv.at/id-austria.html
[43] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-digital-identity_de

gezielte Steuerung des Datenzugriffs entsprechend den Anforderungen und Rollen der Teilnehmenden im Data Space.

3.2.2 Vertrauensmechanismen

Dieser Baustein bezieht sich auf Vertrauensmechanismen, die einen zuverlässigen Datenaustausch zwischen Teilnehmer:innen erleichtern, indem er die Identität und die Einhaltung von Regeln durch andere Teilnehmer innerhalb einer Datenaustauschtransaktion sicherstellt. Dies kann durch organisatorische Maßnahmen wie Zertifizierung oder technische Maßnahmen wie Remote-Attestation erfolgen.

Lösungen:

- Der **Gaia-X Trust Framework**[44] umfasst eine Reihe von Regeln, die die Mindestanforderungen für die Teilnahme in einem Gaia-X-Ökosystem festlegen. Diese Regeln gewährleisten eine gemeinsame Governance und grundlegende Interoperabilitätsniveaus über einzelne Ökosysteme hinweg und sollen Benutzer:innen gleichzeitig die volle Kontrolle über ihre Entscheidungen geben. Im Gaia-X Framework erfolgt der Datenaustauschals Ergebnis von Vertragsverhandlungen, die in die Infrastruktur verankert sind. Das Gaia-X Data Clearing House (GXDCH)[45] bietet Betriebsdienste an, um sicherzustellen, dass der Gaia-X-Trust Framework eingehalten wird und die Interoperabilität beim Datenaustausch gewährleistet ist.
- Das **iSHARE**-Framework ist ein übergreifendes System, das einen sicheren und kontrollierten Datenaustausch zwischen Teilnehmer:innen innerhalb und zwischen Data Spaces erleichtert[46]. Es bietet sowohl technische als auch rechtliche Strukturen zur Unterstützung des vertrauenswürdigen Datenaustauschs. Zu den Schlüsselkomponenten gehören der iSHARE-Satellit[47] für Validierung, Autorisierungsmechanismen und standardisierte Nutzungsbedingungen. Teilnehmer:innen im iSHARE-Ökosystem arbeiten unter denselben Vertrags- und Schema-Regeln, um die Einhaltung von Datenlizenzen und -berechtigungen sicherzustellen. Aus technischer Sicht integriert das Framework Konnektoren wie den IDSA-Connector, um eine Kontrolle über die Datenverwendung und -verarbeitung in sicheren Umgebungen zu ermöglichen.

[44] http://docs.gaia-x.eu/policy-rules-committee/trust-framework/22.10/

[45] https://docs.gaia-x.eu/framework/?tab=clearing-house

[46] Für mehr Informationen zum iSHARE-Framework kann man sich hier das Cookbook herunterladen: https://ishare.eu/inspiration/cookbook-for-data-spaces-eng/?#gf_3

[47] Der iSHARE-Satellit fungiert als zentraler Koordinator und Governance-Kern innerhalb des iSHARE-Frameworks. Er spielt eine entscheidende Rolle bei der Überprüfung der Legitimität der Teilnehmer:innen und der Erleichterung von Operationen innerhalb eines Data Spaces.

3.2.3 Vertragssteuerung und Abrechnungsverfahren

Ein Datenaustausch bzw. -handel erfordert in den meisten Fällen den Abschluss eines Datenliefer- bzw. Datennutzungsvertrages zwischen dem Datenanbieter und dem Datenkonsument (auch bei kostenfreier zur Verfügungstellung der Daten). In Data Spaces haben Datenanbieter die Möglichkeit, vordefinierte Bedingungen für Datentransaktionen in Form von Datennutzungsvereinbarungen, Policies oder Lizenzen festzulegen. Diese Richtlinien fungieren als Leitfaden für den sicheren und kontrollierten Austausch von Daten. Datenanbieter können dabei die Bedingungen für die Nutzung, den Zugriff und die Weitergabe ihrer Daten festlegen, was ihnen erlaubt, ihre (sensiblen) Informationen zu schützen und die Kontrolle über die Datennutzung zu behalten. Diese datenzentrische und souveräne Herangehensweise fördert eine vertrauensvolle Zusammenarbeit zwischen den Teilnehmer:innen im Data Space. Durch die Automatisierung dieses Prozesses können Data Spaces effektive Werkzeuge für den sicheren und effizienten Datenhandel zwischen Vertragspartner:innen sein.

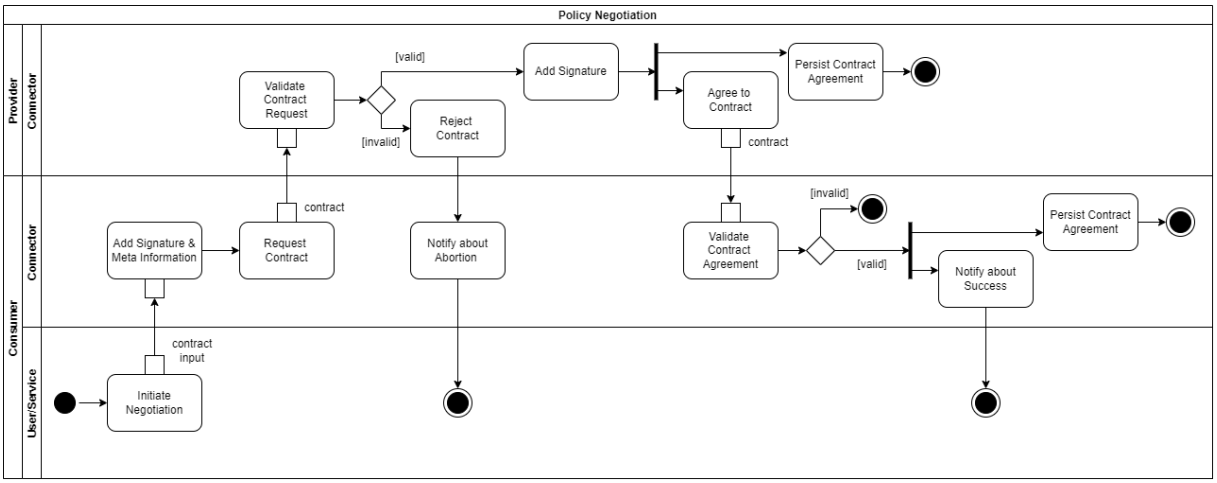

Abbildung 7: Die einfachste Version eines Vertragsverhandlungsprozesses, die gemäß der IDSA RAM Version 4.0 notwendig ist, um eine Vertragsvereinbarung zu erzielen[48].

Abbildung 7 zeigt exemplarisch einen technologieunabhängigen Nachrichtenfluss in einem Vertragsverhandlungsprozess gemäß des IDSA Referenzarchitekturmodells Version 4.0. Der Verhandlungsprozess beginnt, wenn der Konnektor des Datenkonsumenten dem Datenanbieter eine Vertragsanfrage sendet. Die Inhalte dieser Anfrage können vom ursprünglichen Vertragsangebot abweichen oder dieses direkt übernehmen. Die entsprechenden Metainformationen im Vertrag werden daraufhin angepasst (z. B. das Datum, die Laufzeit oder die Unterschrift)[49]. Der Vertragsantrag wird entweder abgelehnt oder akzeptiert. Bei Zustandekommen

[48] https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0/blob/main/documentation/3_Layers_of_the_Reference_Architecture_Model/3_4_Process_Layer/3_4_3_Contract_Negotiation.md#contract-negotiation

[49] Die Anpassung der Metadaten während des Vertragsverhandlungsprozesses stellt eine Idealkonzeptualisierung dar. Eine praktische Implementierung kann von den hier skizzierten theoretischen Ablauf abweichen, da sie von verschiedenen Faktoren und technischen Gegebenheiten beeinflusst wird.

eines Vertrags wird dieser auch vom Konnektor des Datenanbieters unterzeichnet, und der Datenkonsument wird über den Vertragsabschluss informiert. Sobald ein Vertragsabschluss erreicht wurde, wird dieser in beiden Konnektoren instanziiert und bereitgestellt. Das bedeutet, dass er auf beiden Seiten gespeichert wird.

Neben Datennutzungsverträgen, Policies und Lizenzen, die zur digitalen Schließung von Datenverträgen herangezogen und meist unter dem Begriff „digital contracts" subsumiert werden, gibt es weitere Möglichkeiten, die Lieferung und Nutzung von Daten vertraglich unter den Data Space Teilnehmer:innen festzuhalten. Es ist jedoch wichtig zu betonen, dass für einige dieser Optionen noch rechtliche und/oder technische Unsicherheiten bestehen.

- **Externe Verträge**: Mit externen Verträgen sind (zumeist schriftliche) Verträge gemeint, die außerhalb der technischen Infrastruktur bilateral zwischen Datenanbieter und Datenkonsument geschlossen werden. Diese Verträge regeln auch die Bedingungen für den Austausch und die Nutzung von Daten und enthalten Informationen über die beteiligten Parteien, den Zweck der Datennutzung, den Umfang der Daten, Datenschutz, Sicherheitsmaßnahmen, Laufzeit und Haftungsregelungen.
- **Ad-hoc Verträge**: Ein Ad-hoc-Vertrag ist eine informelle, kurzfristige und spontane Vereinbarung zwischen Parteien, die ohne vorherige detaillierte Verhandlungen oder formelle Dokumentation geschlossen wird[50]. Am Beispiel eines Data Spaces könnten ad-hoc Verträge zu dem Zeitpunkt der Datenkonsumation automatisiert geschlossen werden, inklusive eines Bezahlmechanismus der daraufhin angestoßen werden kann. Ad-hoc Verträge sind jedoch oft weniger umfangreich und enthalten weniger rechtliche Absicherungen als ausführlichere Vereinbarungen. Deshalb besteht für diese Art Vertrag noch Rechtsunsicherheit.
- **Smart Contracts**: sind selbstausführende, codebasierte Verträge, bei denen die Bedingungen der Vereinbarung direkt in Codezeilen geschrieben sind[51]. Sie werden in der Regel auf Blockchain-Technologie implementiert und ermöglichen eine sichere und transparente Ausführung von Transaktionen ohne die Notwendigkeit von Vermittlern.

Herausforderungen

- **„Smartness" of Smart Contracts:** Derzeit beschränkt sich ein Smart Contract darauf, eine einfache "Wenn-dann-Beziehung" darzustellen, indem er den ihm übergebenen Input verarbeitet, ohne dabei eigene neue und somit "intelligente" Inhalte zu generieren[52]
- **Sicherheitsaspekte von Smart Contracts:** Die Sicherheit von Smart Contracts ist von entscheidender Bedeutung, da einmal bereitgestellter Code in der Regel nicht mehr

[50] https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2010/wp176_de.pdf

[51] (Zheng, Shaoan, Dai, & Chen, 2020)

[52] (Groß, 2022)

revidiert werden kann. Es gab bereits Angriffe auf Smart Contracts aufgrund von Programmierfehlern und unsicheren Ausführungsumgebungen. Die Gewährleistung der Korrektheit und Sicherheit von Smart Contracts erfordert weitere Forschung und Entwicklung, einschließlich formaler Verifikation von Sprachen und Unterstützung für Entwickler bei der Codevalidierung vor der Aufnahme in die Blockchain.[53]

- **Daten(ver)kauf:** Eine Herausforderung beim Daten(ver)kauf besteht darin, einen effizienten Bezahlungsprozess und die Vereinbarung von Datenlieferverträgen zu realisieren, insbesondere für kostenpflichtige Daten. Die Vereinbarung von Datenlieferverträgen zu bestimmten Bedingungen (z. B. Zugriff, Nutzungsdauer, Datenqualität, Kosten) erfordert klare und präzise Konditionen, die sowohl die Interessen der Datenlieferanten als auch der Datenkäufer berücksichtigen, rechtlich bindend und technisch durchsetzbar sind.
- **Bargaining:** In einem Data Space, in dem der Tausch von Daten weder kostenfrei noch gegen Geld erfolgen muss, liegt die zentrale Schwierigkeit in der technischen Umsetzung alternativer "Währungen". Die Entwicklung geeigneter Mechanismen und Protokolle zur Unterstützung verschiedener Tauschoptionen stellt eine komplexe Aufgabe dar, um den effizienten und vielfältigen Handel mit Daten zu ermöglichen.

3.3 Datenwertschöpfung

3.3.1 Datenbeschreibung/-angebote

Die Beschreibung von Daten ist ein entscheidender Aspekt des Datenaustauschs, der es verschiedenen Systemen und Anwendungen ermöglicht, effektiv miteinander zu kommunizieren und die Daten des jeweils anderen zu verstehen. Zur Datenbeschreibung werden Metadaten benutzt, die kontextuelle Informationen liefern, die zum einen Daten beschreiben und organisieren und zum anderen verfolgen, wie Daten erstellt, verwaltet und genutzt werden. Dafür können unterschiedliche Metadatenschemata verwendet werden, die einen standardisierten Rahmen für die Darstellung und Kommunikation von Daten ermöglichen. Metadatenschemata, wie zum Beispiel das Dublin Core Metadata Element Set (DCMES)[54], bieten vordefinierte Strukturen und Vokabulare zur Beschreibung verschiedener Datentypen, was Systemen erleichtert, Daten korrekt zu interpretieren und zu verarbeiten. Bei der Auswahl eines Metadatenschemas für den Datenaustausch ist es wichtig, Faktoren wie die Art der ausgetauschten Daten, die beabsichtigte Zielgruppe oder Anwendung sowie etwaige regulatorische oder Compliance-Anforderungen zu berücksichtigen. Darüber hinaus können effektive Datenbeschreibungstechniken wie konsistente Benennungskonventionen, ordnungsgemäße Codierung und Formatierung, ausführliche

[53] (Schütte, et al., 2017)

[54] https://www.dublincore.org/specifications/dublin-core/dces/

Dokumentation und Validierungsprüfungen dazu beitragen, einen nahtlosen und zuverlässigen Datenaustausch sicherzustellen.

Daten können als Vermögenswert eines Unternehmens betrachtet werden, weswegen diese auch als sogenannte "Data Assets" bezeichnet werden. Data Assets beinhalten einen Link zu den Daten selbst sowie zu den beschreibenden Metadaten. Über Datennutzungsvereinbarungen kann jedes "Data Asset" individuell zu den eigenen Bedingungen angeboten werden, indem ein "Data Offer" erstellt wird. In Data Spaces, als dezentralen digitalen Räumen, ist es Teilnehmer:innen dann möglich, die "Data Offers" anderer Teilnehmer:innen zu sehen, zu verhandeln und unter Wahrung der zugrundeliegenden Werte zu beziehen (siehe schematische Skizzierung der Inhalte eines Datenaustausches bzw. -handels in Abbildung 8).

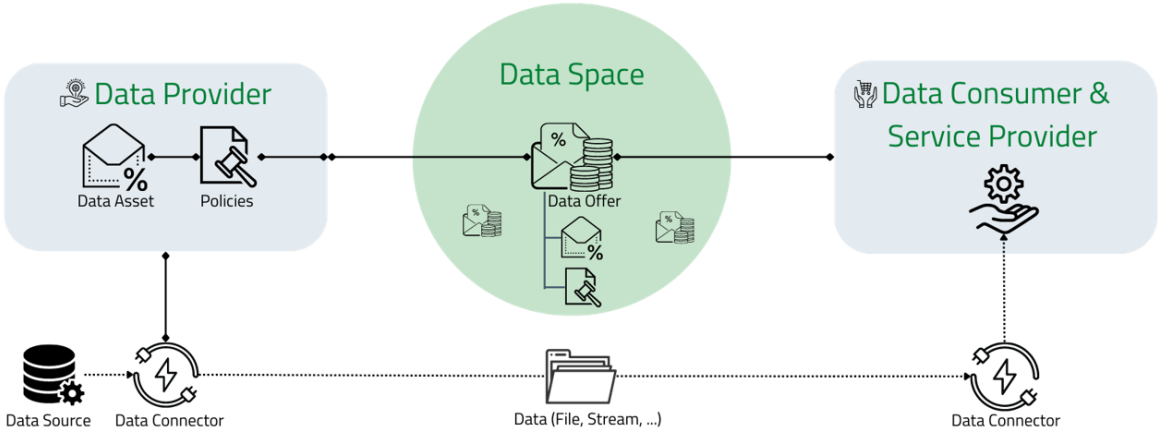

Abbildung 8: Schematische Darstellung einer Datentransaktion in einem Data Space

3.3.2 Datenportal/-marktplatz

Datenportale bzw. Datenmarktplätze in Form von grafischen Benutzeroberflächen (GUIs) spielen eine zentrale Rolle bei der Vereinfachung der Benutzerinteraktion mit Data Space Konnektoren und Plattformen. Im Kontext von Data Spaces bieten GUIs eine visuelle Benutzeroberfläche, die es Benutzer:innen ermöglicht, auf intuitive und benutzerfreundliche Weise mit Data Space Konnektoren zu interagieren. Dabei können Benutzer:innen durch verschiedene Funktionen navigieren, auf Daten zugreifen, Einstellungen konfigurieren und Aufgaben mithilfe grafischer Elemente wie Schaltflächen, Menüs und visuellen Darstellungen von Daten ausführen.

Einer der Hauptvorteile von GUIs in Data Spaces liegt in ihrer Fähigkeit, Benutzer:innenn eine effektive Verwaltung und Beschreibung[55] von Daten zu ermöglichen. GUIs bieten Tools zur Organisation, Analyse und Präsentation von Daten in einem verständlichen Format. Dadurch können Benutzer:innen Datenbestände (= Data Assets) anzeigen, Datenaustauschprozesse überwachen.

Ein weiterer wichtiger Aspekt von GUIs in Data Spaces ist ihre Zugänglichkeit und Benutzerfreundlichkeit. GUIs machen Data Space-Plattformen für Benutzer:innen mit

[55] Die Datenbeschreibung erfolgt meist mittels Metadatenschemata, die in das GUI integriert sind.

unterschiedlichem technischen Know-how zugänglich und ermöglichen es auch nicht-technischen Benutzer:innen, mit Daten zu interagieren und Aufgaben ohne umfangreiche Schulung durchzuführen. Die visuelle Natur von GUIs vereinfacht das Benutzererlebnis und erleichtert es Benutzern, komplexe Data Space-Umgebungen zu navigieren und Data Space-Funktionen zu nutzen.

Schließlich sind GUIs darauf ausgelegt, nahtlos mit der Funktionalität von Data Space Konnektoren zu integrieren, sodass Benutzer:innen mit Datenquellen interagieren, Datenaustauschprozesse konfigurieren und Data Assets innerhalb der GUI-Umgebung verwalten können. GUIs können Funktionen wie Datensuche, Filtermöglichkeiten und Datenaustauschoptionen enthalten, um die Benutzerproduktivität und -effizienz bei Data Space Operationen zu verbessern.

Insgesamt spielen GUIs eine entscheidende Rolle bei der Nutzerinteraktion und der Datenverwaltung in Data Spaces. Durch ihre visuell ansprechende und benutzerfreundliche Oberfläche erleichtern sie den nahtlosen Datenaustausch und fördern die Zusammenarbeit.

4 Zusammenfassung

Der vorliegende technische Orientierungsbericht bietet eine eingehende Analyse der technischen Aspekte von Data Spaces, einschließlich ihrer Komponenten und gestaltenden Organisationen. Dabei liegt ein besonderer Fokus auf den technischen Anforderungen, die Data Spaces erfüllen müssen. Ein weiterer Schwerpunkt des Berichts liegt auf der detaillierten Darstellung der technischen Infrastruktur von Data Spaces, in Form von technologieunabhängigen Bausteinen bzw. Schlüsselkomponenten. Dazu gehören Protokolle, Konnektoren für Infrastruktur und Interoperabilität, Identitätsmanagement, Vertrauensmechanismen sowie Vertragssteuerung und Abrechnungsverfahren im Bereich Sicherheit und Trust. Zudem werden Aspekte der Datenbeschreibung und -angebote sowie die Funktionalitäten von Datenportalen und -marktplätzen zur Datenwertschöpfung behandelt. Zusammenfassend liefert der Bericht einen umfassenden Überblick über die technischen Grundlagen und Komponenten von Data Spaces.

Abbildungsverzeichnis

5 Verweise

Franklin, M., Halevy, A., & Maier, D. (2005). From databases to dataspaces: a new abstraction for information management. ACM Sigmod Record, 34(4), S. 27-33.

Gieß, A., Hupperz, M., Schoormann, T., & Möller, F. (2024). What Does it Take to Connect? Unveiling Characteristics of Data Space Connectors. Proceedings of the 57th Hawaii International Conference on System Sciences.

Groß, D. (2022). Vertragsdurchführung mit Smart Contracts – rechtliche Rahmenbedingungen und Herausforderungen. In M. B. Rohde, Datenwirtschaft und Datentechnologie (S. 145-159). Springer Vieweg. doi:https://doi.org/10.1007/978-3-662-65232-9_11

Hellmeier, M., & von Scherenberg, F. (2023). A Delimitation of Data Sovereignity from Digital and Technological Sovereignity. Proceedings of the 31st European Conference on Information Systems.

Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021). Data sovereignty: A review. Big Data & Society, 8(1). doi:https://doi.org/10.1177/2053951720982012

Olms, C., Nissen, C., Schier, A., Leveling, J., Rademacher, R., & Hompel, M. (2020). Architektur einer adaptiven Plattform für unternehmens-übergreifende datenbasierte Dienste mit dem International Da-ta Spaces.

Otto, B., Steinbuss, S., Teuscher, A., & Lohmann, S. (2019). IDS Reference Architecture Model (Version 3.0). International Data Spaces Association.

Pampus, J., Jahnke, B.-F., & Quensel, R. (2022). Evolving Data Space Technologies: Lessons Learned from an IDS Connector Reference Implementation. In T. Margaria-Steffen (Hrsg.), Lecture Notes in Computer Science: Vol. 13704. Leveraging applications of formal methods, verification and validation: 11th international symposium on leveraging applications of formal methods, ISoLA 2022 (S. 366–381). Rhodes, Greece: Springer.

Pettenpohl, H., Spiekermann, M., & Both, J. (2022). International Data Spaces in a Nutshell. In B. t. Otto, Designing Data Spaces (S. 29-40). Springer, Cham.

Pohlmann, N. (2019). Cyber-Sicherheit: Das Lehrbuch für Konzepte, Prinzipien, Mechanismen, Architekturen und Eigenschaften von Cyber-Sicherheitssystemen in der Digitalisierung. Springer Vieweg.

Schütte, J., Fridgen, P., Prinz, W., Rose, T., Urbach, N., Hoeren, T., & ... Nourak, A. (November 2017). BLOCKCHAIN UND SMART CONTRACTS: Technologien, Forschungsfragen und Anwendungen. Von https://www.iml.fraunhofer.de/content/dam/iml/en/documents/OE260/Fraunhofer-Positionspapier_Blockchain-und-Smart-Contracts.pdf abgerufen

Strüker, J., Urbach, N., Guggenberger, T., Lautenschlager, J., Ruhland, N., Schlatt, V., . . . Völter, F. (2021). Self-Sovereign Identity – Grundlagen, Anwendungen und Potenziale portabler digitaler Identitäten. Projektgruppe Wirtschaftsinformatik des Fraunhofer-Instituts für Angewandte Informationstechnik FIT, Bayreuth.

Zheng, Z., Shaoan, X., Dai, H.-N., & Chen, W. (2020). An overview on smart contracts: Challenges, advances and platforms. Future Generation Computer Systems, Volume 105, 475-491. doi:https://doi.org/10.1016/j.future.2019.12.019.