

---

## 2 Datensouveränität durch Gaia-X – Grundlagen der Ergebnisse in der Förderlinie InGAIA-X

N. MAISCH, A. LECHLER, O. RIEDEL, F. MITSCHKE, T. SCHRAGE, J. EICHHOLZ, D. KARABULUT, M. RICHTER, P. PATOLLA, S. VOGT, C. FRIEDRICH, F. RUDOLPH, K. WENZEL, J. ZERBIN

### 2.1 Definitionen und grundlegende Konzepte

In diesem Unterkapitel werden die grundlegenden Konzepte und Begriffe von Gaia-X eingehend erläutert, um ein einheitliches Verständnis für die nachfolgende Kapitel zu schaffen. Zunächst wird aus dem zentralen Konzept von Gaia-X eingegangen, die *Interoperabilität*. Anschließend wird der *Datenraum* als zentraler Baustein für den sicheren und kontrollierten Datenaustausch in einem föderierten Umfeld vorgestellt. Darauf aufbauend wird auf die Bedeutung von Wertschöpfungsnetzwerken und Ökosystemen eingegangen, die die kooperative und dynamische Zusammenarbeit verschiedener Akteure zur Schaffung gemeinsamen Mehrwerts beschreiben. Die Interoperabilität der Systeme und Organisationen bildet dabei die Grundlage für eine effektive Zusammenarbeit in diesem Netzwerk. Ein weiterer Fokus liegt auf den Konzepten der Datensouveränität, Sicherheit und Verträge, die sicherstellen, dass der Datenaustausch den strengen Anforderungen an Datenschutz und Kontrolle entspricht. Schließlich wird Gaia-X selbst und das Konzept des "X" in Gaia-X detailliert erläutert, um die Vielseitigkeit und die unterschiedlichen Dimensionen dieser europäischen Initiative umfassend darzustellen.

#### 2.1.1 Interoperabilität

*Interoperabilität* bezeichnet im Allgemeinen die Fähigkeit der Zusammenarbeit unterschiedlicher Systeme, wobei sowohl technische Systeme als auch Organisationen, auf administrativer und personeller Ebene, berücksichtigt werden. Folglich ist Interoperabilität die Voraussetzung für eine reibungslose und effiziente Interaktion zwischen Akteuren, damit Daten übertragen und Informationen geteilt werden können [1].

Interoperabilität umfasst mehrere wesentliche Aspekte, die eine effektive Zusammenarbeit und Integration von Systemen und Organisationen ermöglichen. *Technische Interoperabilität* bezieht sich auf die Fähigkeit verschiedener Systeme, sich durch gemeinsame Standards und Schnittstellen zu verbinden und Daten auszutauschen [2, 3]. *Semantische Interoperabilität* sorgt dafür, dass die Bedeutung der Daten von allen beteiligten Systemen konsistent verstanden wird, oft durch Verwendung von standardisierten Formaten und Ontologien [4, 5]. *Organisatorische Interoperabilität* umfasst die Fähigkeit von Unternehmen, Leistungen sowohl untereinander als auch für ihre Kunden bereitzustellen. Zusammen ermöglichen diese Aspekte eine reibungslose und effiziente Integration und Kommunikation zwischen verschiedenen Systemen und Organisationen[6].

Im Zusammenhang mit *Datenräumen*, *Ökosystemen* und *Wertschöpfungsnetzwerken* ist Interoperabilität besonders wichtig, um eine nahtlose Integration und Zusammenarbeit zwischen den Teilnehmenden zu ermöglichen [7–9]. Auf diese Aspekte wird in den folgenden Abschnitte eingegangen.

## 2.1.2 Datenraum

*Datenräume* (engl. Data Spaces) sind zentrale Bestandteile moderner Datenökosysteme und spielen eine entscheidende Rolle bei der Verwaltung, dem Austausch und der Nutzung von Daten über organisatorische und sektorale Grenzen hinweg [10–12]. Im Unterschied zu zentralen Ansätzen zur Datenintegration, wie bei klassischen digitalen Plattformen, werden die Daten nicht physisch zusammengeführt, wodurch kein gemeinsames Datenbankschema erforderlich ist [11]. In diesen dezentral organisierten Umgebungen erfolgt der Zugriff auf Daten vielmehr durch einen kontrollierten Zugang zu den Daten über eine Plattform [12]. Vor diesem Hintergrund werden Datenräume auch als “eine föderierte, offene Infrastruktur für souveränen Datenaustausch, die auf gemeinsamen Vereinbarungen, Regeln und Standards beruht” [12] definiert.

Datenräume zeichnen sich vor allem durch *Dezentralität*, *Transparenz* und *Interoperabilität* aus [12–14]. Dabei verbleiben die Daten bei ihren Besitzern und können mithilfe standardisierter Modelle nahtlos ausgetauscht werden. Die Dateninhaber behalten die volle Kontrolle über ihre Daten, da diese durch entsprechende Sicherheitstechnologien geschützt sind und die Nutzung jederzeit nachvollziehbar bleibt [12, 13]. Um sowohl die Dezentralisierung von Datenräumen zu fördern als auch eine breite Vernetzung der Teilnehmenden zu gewährleisten, setzen die führenden Initiativen für Datenräume auf einen föderierten Ansatz [11, 12, 15]. Dies ermöglicht es zusammen mit der Einführung eines einheitlichen Katalogs politischer Richtlinien und funktionaler und technischer Spezifikationen, Datenräume interoperabel zu gestalten, sodass Austauschprozesse über die Grenzen eines einzelnen Datenraums hinaus stattfinden können [15]. Durch harmonisierte Prozesse, Formate und Semantiken, ermöglichen es föderierte Datenräume ihren Teilnehmenden mit geringem Aufwand auch an anderen Datenräumen teilzunehmen. Durch föderierte Datenräume wird ein reibungsloser und effizienter Informationsaustausch gewährleistet [12].

## 2.1.3 Wertschöpfungsnetzwerke und Ökosysteme

Ein *Wertschöpfungsnetzwerk* ist ein anpassungsfähiges System bestehend aus Unternehmen, Geschäftspartnern, Kunden und weiteren Beteiligten, die direkt oder indirekt an der Produktion und Lieferung von Waren oder Dienstleistungen zusammenarbeiten [16]. In diesem System wird ein Mehrwert durch enge Zusammenarbeit geschaffen, bei der die Akteure innerhalb des Netzwerkes ihre Ressourcen, Fähigkeiten und Fertigkeiten kombinieren [16, 17]. Die Interaktion und der Austausch zwischen verschiedenen Unternehmen steht dabei im Mittelpunkt [16]. Die Art dieses Austauschs führt dazu, dass die Ressourcen dorthin fließen, wo sie am wertvollsten sind und ermöglicht es den Unternehmen, ihre Ressourcen besser zu nutzen und sich auf die eigenen Kernkompetenzen zu fokussieren [16, 17]. Auf diese Weise kann sich das Netzwerk nicht nur kontinuierlich an Veränderungen in den Märkten oder der Technologie anpassen, sondern gleichzeitig auch langfristige und stabile Beziehungen fördern [16].

Der Begriff des *Ökosystems* (engl. Ecosystem) lässt sich im Kontext von Gaia-X aus zwei Blickwinkeln betrachten: *Business Ecosystem* und *Data Ecosystem*. Im Folgenden werden diese beiden Begriffe näher erläutert.

Ein *Business Ecosystem* beschreibt eine Gruppe von Unternehmen und Einzelpersonen, die miteinander in Beziehung stehen und sich gegenseitig ergänzen, während sie gemeinsam Waren, Technologien und Dienstleistungen für Kunden bereitstellen. Im Rahmen dieser gegenseitigen Ergänzung stehen vor allem *Komplementaritäten* und *Interdependenzen* zwischen den Teilnehmern eines Ökosystems im Fokus [11, 18–20]. Komplementaritäten beziehen sich auf die Art und Weise, wie sich verschiedene Akteure oder Komponenten innerhalb des Ökosystems ergänzen und zusammen einen gesteigerten Mehrwert erzeugen [19]. Interdependenzen beschreiben die gegenseitige Abhängigkeit der Teilnehmer im Ökosystem. Jede Einheit beeinflusst die anderen und wird von ihnen beeinflusst, was zu einer engen Verflechtung der Aktivitäten führt [19, 20]. Durch die Kombination von Komplementaritäten und

Interdependenzen entsteht ein dynamisches, anpassungsfähiges Ökosystem, das es den beteiligten Akteuren ermöglicht, auf Marktveränderungen zu reagieren und gemeinsam Innovationen voranzutreiben [19]. Diese synergetischen Beziehungen sind entscheidend für die Stabilität und das wirtschaftliche Wachstum des gesamten Ökosystems, da sie eine enge Zusammenarbeit und Koordination über die traditionellen Unternehmensgrenzen hinaus fördern [18].

Ein *Data Ecosystem* bezeichnet ein Netzwerk von miteinander verbundenen Akteuren, Systemen und Prozessen, dessen Fokus auf der gemeinsamen Nutzung, Erstellung, Verwaltung und dem Austausch von Daten liegt [11, 21]. Als miteinander verbundene Netzwerke unterstützen Data Ecosystems Unternehmen dabei, auf der Grundlage von Daten zusammenzuarbeiten und fundierte Entscheidungen zu treffen. Während Business Ecosystems ganze Branchen und die Beziehungen zwischen Unternehmen umfassen, konzentrieren sich Data Ecosystems eher auf die technologischen und informationstechnischen Aspekte, die den Geschäftsbetrieb unterstützen [21].

Eine Gemeinsamkeit zwischen beiden Ökosystemen ist der Faktor Interoperabilität. Sie spielt eine zentrale Rolle, da sie zwei oder mehreren Geschäftseinheiten die Möglichkeit bietet in einer vernetzten und heterogenen Umgebung Informationen oder Ressourcen auszutauschen und für die gemeinsame Wertschöpfung und Zusammenarbeit zu nutzen [7, 8]. Folglich steht analog zur ursprünglichen Begrifflichkeit aus der Biologie in beiden Formen der Ökosysteme die Schaffung eines harmonischen Ganzen im Vordergrund, das mehr ist als die Summe seiner Teile.

## 2.1.4 Datensouveränität

Datensouveränität ist ein Konzept, das sich auf die Kontrolle von Individuen, Organisationen und Staaten über ihre eigenen Daten bezieht. Es hat seine Wurzeln in der traditionellen Vorstellung von Souveränität, die die höchste Autorität eines Staates über sein Territorium und seine Bevölkerung beschreibt. In der digitalen Ära hat sich dieser Begriff jedoch auf den Umgang mit Daten ausgeweitet und ist zu einem entscheidenden Aspekt in den Diskussionen über digitale Rechte, Datenschutz und technologische Autonomie geworden. Dabei spielt Sicherheit im Kontext der Datensouveränität eine zentrale Rolle. Sicherheit umfasst hierbei den Schutz von Daten vor unbefugtem Zugriff, Verlust, Manipulation und Missbrauch. [22]

Der Gedanke der Datensouveränität gewinnt zunehmend an Bedeutung, da Daten zunehmend an Wert gewinnen. Die Kontrolle über Daten umfasst nicht nur deren Sammlung und Speicherung, sondern auch deren Verarbeitung, Verteilung und Nutzung. Die digitale Infrastruktur, auf der diese Prozesse basieren, ist häufig international vernetzt, was es schwierig macht, klare Grenzen der Datenhoheit zu ziehen. Diese Grenzüberlappungen bringen Herausforderungen mit sich, insbesondere wenn es um den Schutz der Daten vor Zugriffen durch ausländische Regierungen oder Unternehmen geht. Ein zentraler Punkt der Datensouveränität ist deshalb die Fähigkeit der Datenbesitzer, zu entscheiden, wo und wie ihre Daten gespeichert und verarbeitet werden. Diese Fähigkeit ist besonders wichtig, da in global verteilten Netzwerken oft unterschiedliche gesetzliche Rahmenbedingungen gelten. Beispielsweise können Daten, die in der EU gespeichert werden, unter die Datenschutz-Grundverordnung (DSGVO) fallen, die strengere Datenschutzanforderungen vorschreibt als Gesetze in anderen Teilen der Welt [23]. Dies gibt den Datenbesitzern in der EU eine gewisse Sicherheit, dass ihre Daten vor unbefugten Zugriffen geschützt sind. [22]

Trotz dieser Schutzmechanismen gibt es Herausforderungen, die die Datensouveränität beeinträchtigen können. Ein bedeutendes Problem ist die Abhängigkeit von Cloud-Diensten, die von großen internationalen Technologieunternehmen angeboten werden, insbesondere aus den USA [24]. Diese Unternehmen unterliegen den Gesetzen ihrer Heimatländer, die möglicherweise weniger strenge Datenschutzvorschriften haben oder sogar den Zugriff auf eigentlich geschützte Daten durch Behörden erlauben. Der US Cloud Act von 2018 beispielsweise ermöglicht es US-Behörden, auf Daten zuzugreifen, die von US-Unternehmen weltweit gespeichert werden, auch wenn diese Daten außerhalb der USA gespeichert sind [25]. Diese extraterritoriale Reichweite stellt ein Risiko für die Datensouveränität von

Unternehmen und Individuen in anderen Ländern dar, da sie der Kontrolle ihrer eigenen Daten beraubt werden könnten. Während Regierungen primär aus sicherheits- und ordnungspolitischen Gründen Daten sammeln, sind die Motive von Unternehmen kommerziell geprägt. [26]

Datensouveränität umfasst überdies auch die Fähigkeit, innovative und datengetriebene Geschäftsmodelle zu entwickeln, die auf einem fairen und sicheren Zugang zu Daten basieren [27]. In einer global vernetzten Wirtschaft sind Daten ein entscheidender Faktor für Innovation und Wettbewerb. Wer die Kontrolle über seine Daten hat, kann nicht nur bessere Entscheidungen treffen, sondern auch neue Märkte erschließen und sich Wettbewerbsvorteile sichern [27]. In Europa hat das Thema Datensouveränität dabei eine besondere politische und wirtschaftliche Bedeutung erlangt. Die Europäische Union betrachtet die digitale Souveränität als eine wesentliche Voraussetzung, um im globalen Wettbewerb bestehen zu können und gleichzeitig die Rechte ihrer Bürger zu schützen [24]. Initiativen wie Gaia-X sind Ausdruck dieser Bestrebungen.

Hindernisse für Datensouveränität können neben fehlender Regulierung auch in fehlender Interoperabilität zwischen verschiedenen IT-Systemen und Plattformen liegen. Interoperabilität bezeichnet dabei die Fähigkeit von verschiedenen technischen Systemen, Organisationen und Plattformen nahtlos und effizient im Rahmen eines Datenaustausch zusammenzuarbeiten, ohne dabei an Qualität oder Funktionalität der Daten zu verlieren. Viele Cloud-Dienste sind proprietär und erlauben es den Nutzern nicht, ihre Daten aufwandsarm von einem Anbieter zu einem anderen zu migrieren, ohne dabei die Kontrolle über ihre Daten zu verlieren. Diese „Lock-In“-Effekte machen es für Datenbesitzer schwierig, ihre Daten souverän zu verwalten, da sie an die Bedingungen des Dienstanbieters gebunden sind [28]. Verträge, die klare Klauseln zur Datenportabilität enthalten, ermöglichen es den Datenbesitzern, ihre Daten zu einem anderen Anbieter zu migrieren, ohne dabei die Kontrolle über die Daten zu verlieren. Dies stärkt die Datensouveränität, da es den Datenbesitzern die Flexibilität gibt, den Anbieter zu wechseln, wenn dies notwendig oder wünschenswert ist. Bisher sind solche Verträge allerdings nicht weit verbreitet. [29, 30]

## 2.1.5 Gaia-X

Gaia-X ist ein europäisches Non-Profit-Projekt, das 2019 von Deutschland und Frankreich initiiert wurde, um eine offene, souveräne und vernetzte digitale Infrastruktur zu schaffen, die den europäischen Werten und Standards entspricht. Es entstand als Antwort auf die wachsende Abhängigkeit Europas von großen, meist nicht-europäischen Cloud-Anbietern, wie sie vor allem in den USA und China angesiedelt sind. Diese Abhängigkeit hat in Europa die Sorge verstärkt, dass die Kontrolle über Daten und digitale Infrastrukturen verloren geht. Mit Gaia-X will Europa seine technologische Autonomie zurückgewinnen und gleichzeitig die Innovationskraft und Wettbewerbsfähigkeit der europäischen Wirtschaft stärken. [31]

Gaia-X ist keine einzelne Instanz, sondern eine Föderation von Service- und Datenanbietern, die miteinander vernetzt sind und in Datenökosystemen interoperable Dienste anbieten. Das bedeutet, dass Gaia-X als eine Art übergeordnete Rahmenstruktur fungiert, die bestehende Cloud-Dienste und -Anbieter miteinander verbindet. Gaia-X beruht auf grundlegenden Prinzipien, die sicherstellen, dass die Nutzer jederzeit die Kontrolle über ihre Daten behalten. Datensouveränität spielt eine zentrale Rolle. Dies ist nicht nur theoretisch vorgesehen, sondern wird durch die Architektur von Gaia-X aktiv ermöglicht. Durch die Verwendung offener Schnittstellen und Standards fördert Gaia-X die Interoperabilität zwischen verschiedenen Cloud- und Datenanbietern. Transparenz ist ein weiterer Eckpfeiler der Plattform. Gaia-X setzt auf klare und offene Regeln, die allen Beteiligten gleiche Bedingungen bieten. Dies schafft Vertrauen und erleichtert die Zusammenarbeit zwischen den Akteuren. Zudem legt Gaia-X großen Wert auf höchste Sicherheitsstandards. Die Gaia-X-konformen Plattformen stellen sicher, dass alle gespeicherten und verarbeiteten Daten umfassend geschützt sind, um unbefugten Zugriff zu verhindern. Gleichzeitig verfolgt Gaia-X einen Ansatz der Dezentralisierung. Anstatt von einem zentralen Anbieter gesteuert zu werden, basieren die Plattformen auf einem Netzwerk von Teilnehmenden, die

diese in Kooperation betreiben. Diese Struktur fördert nicht nur die Resilienz des Systems, sondern unterstützt auch die Unabhängigkeit und Vielfalt innerhalb des Gaia-X-Ökosystems. [29]

Die Struktur von Gaia-X lässt sich in mehrere Ebenen unterteilen. Eine dieser Ebenen sind die sogenannten „Federated Services“. Diese Dienste bieten die grundlegenden Funktionen, die notwendig sind, um eine Interoperabilität zwischen den verschiedenen Cloud-Anbietern sicherzustellen. Dazu gehören unter anderem das Identitäts- und Zugangsmanagement, die Datenportabilität und -interoperabilität sowie die Einhaltung von Compliance- und Sicherheitsanforderungen. Eine weitere Ebene sind die so genannten „Datenräume“ oder Datenökosysteme. Diese sind spezielle Umgebungen, in denen Daten sicher und vertrauenswürdig ausgetauscht und verarbeitet werden können. Jedes Datenökosystem ist für einen bestimmten Sektor oder eine Branche ausgelegt, wie beispielsweise das Gesundheitswesen, die Industrie oder die öffentliche Verwaltung. Innerhalb dieser Datenökosysteme können Unternehmen und Organisationen sicher miteinander kooperieren und Daten austauschen, auch über verschiedenen Datenökosysteme hinweg. Gaia-X ermöglicht es auch, dass sich verschiedene Akteure zu sogenannten „Ecosystemen und Communities“ zusammenschließen, um gemeinsam Innovationen voranzutreiben. Diese Gemeinschaften spielen eine entscheidende Rolle bei der Entwicklung neuer Anwendungsfälle und der Definition von Standards. [29]

Entsprechend der ambitionierten Ziele steht Gaia-X vor einigen Herausforderungen. Dazu gehören die technische Komplexität, die Integration unterschiedlicher Standards und Technologien sowie die Notwendigkeit, eine breite Akzeptanz bei europäischen Unternehmen und Regierungen zu erzielen. Es wird entscheidend sein, wie gut es Gaia-X gelingt, ein starkes und dynamisches Ökosystem aufzubauen, das von einer breiten Basis von Teilnehmern unterstützt wird. Gaia-X ist kein fertiges Produkt, sondern ein ständig wachsendes Ökosystem, das sich an die Bedürfnisse der Teilnehmer anpasst und kontinuierlich weiterentwickelt.

## **Das „X“**

Das „X“ in „Gaia-X“ steht symbolisch für die Vernetzung und die Interoperabilität, die das Projekt schaffen möchte. Es signalisiert die Kreuzung oder das Zusammenspiel verschiedener Akteure, Dienste und Datenökosysteme innerhalb einer föderierten Struktur. Die Wahl des X ist bewusst gewählt, um die Vielseitigkeit und die Verbindungen darzustellen, die Gaia-X ermöglichen soll.

Zudem steht das „X“ auch für die Unbekannten oder Variablen, die in einer dynamischen und komplexen digitalen Infrastruktur eine Rolle spielen oder zukünftig spielen werden. Es symbolisiert die Offenheit und Flexibilität des Systems, das unterschiedliche Technologien, Anbieter und Standards miteinander verknüpfen kann, ohne auf eine starre Architektur festgelegt zu sein. Damit steht Gaia-X dem Aufkommen neuartiger Aspekte, Chancen, Anforderungen und Teilnehmer offen gegenüber. Hervorzuheben ist die Idee, dass das Projekt nicht nur eine Lösung ist, sondern ein Netzwerk von Lösungen, das sich ständig weiterentwickelt und anpasst, um den Anforderungen einer vielfältigen und dynamischen digitalen Welt gerecht zu werden.

## 2.2 Aufbau, Struktur & Technologien

In diesem Unterkapitel wird die grundlegende Struktur und Funktionsweise von *Datenräumen* sowie die Rollen beteiligter Akteure erläutert. Zudem werden zentrale Technologien – wie z. B. die *Gaia-X Federation Services* – vorgestellt, um ein grundlegendes Verständnis für die in den Projekten verwendeten Technologien zu schaffen.

### 2.2.1 Datenraum

Die Nutzung kollaborativer Datenräume ermöglicht neue datengetriebene, industrielle Anwendungsfälle. Datenräume sind dabei nicht zu verwechseln mit Clouds, auf die verschiedene Partner zugreifen können. Vielmehr werden durch das Konzept von Datenräumen gemeinsame Regeln für den gegenseitigen Datenzugriff geschaffen – unabhängig von der zugrundeliegenden Server- oder Cloud-Struktur. Durch die gemeinsam definierten Regeln entsteht eine Infrastruktur für den Datenaustausch, die als Basis für selbstsouveräne datengetriebene Services dient [32]. Datenräume bauen auf vier Prinzipien auf [14]:

1. *Datensovveränität*: Die Teilnehmer eines Datenraums treffen selbstbestimmte Entscheidungen hinsichtlich des Austauschs von Daten mit verschiedenen Partnern. Die Entscheidung über die Nutzung der Daten sowie die Gewährung des Zugriffs obliegen dem Teilnehmer selbst.
2. *Faire Wettbewerbsbedingungen*: Es existieren keine unüberwindbaren Einstiegshürden für neue Teilnehmer. Die Bildung von Datenmonopolen ist nicht zulässig.
3. *Dezentrale immaterielle Infrastruktur*: Analog zum Internet, das auf anerkannten Protokollen basiert, und zum Straßenverkehr, der durch eine Regelung der gültigen Fahrtrichtung gekennzeichnet ist, erfordert der Datenaustausch in Datenräumen gemeinsam anerkannte Regeln, um Offenheit, Interoperabilität oder Vertrauen zu gewährleisten.
4. *Zusammenspiel öffentlicher und privater Stakeholder*: Die gemeinsamen Regeln sollen durch öffentliche und private Stakeholder gemeinsam gestaltet und gepflegt werden.

Die Regeln der Datenräume werden entwickelt, um datengetriebene Anwendungsfälle gemäß dieser Werte in Ökosystemen zu ermöglichen. Um die Regeln und Richtlinien in der Praxis umsetzen zu können, bedarf es zunächst einer Abstraktion der Rollen der Teilnehmer [14] (Abbildung 2.1):

- *Data Owner*: Ein Data Owner besitzt die Rechte an den Daten. Als Eigentümer der Daten obliegt ihm die Entscheidung, ob die Daten privat gehalten oder mit Geschäftspartnern geteilt werden. Dabei ist es unerheblich, auf welchem Speichermedium die Daten gespeichert sind, ob on premise (vor Ort im eigenen Unternehmen), in einer Edge oder in einer Cloud.
- *Data Provider*: Ein Data Provider sammelt, speichert und verarbeitet Daten im Auftrag eines Data Owners. Er ermöglicht den sicheren Austausch und die Nutzung der Daten, während die Kontrolle über deren Verwendung beim Data Owner bleibt. In manchen Geschäftsmodellen nutzt der Data Provider anonymisierte Daten zur Weiterentwicklung eigener Services. Hostet ein Data Owner seine Daten selbst, entfällt der Data Provider.
- *Data Processor*: Ein Data Processor nutzt die Daten eines Data Owners, um neue Services zu entwickeln und diese auf dem Markt anzubieten. Dabei agiert der Data Processor als Schnittstelle zwischen den Daten und den Nutzern der daraus entstehenden Dienstleistungen.
- *Data Marketplace Operator*: Ein Data Marketplace Operator stellt die notwendige Infrastruktur und Governance für den Datenaustausch bereit und gewährleistet eine reibungslose und automatisierte Datennutzung, indem er die Komplexität von Back-Office-Vereinbarungen reduziert.

Die genannten Systeme erleichtern die Auffindbarkeit von Daten, verfolgen alle datenbezogenen Transaktionen auf transparente Weise und setzen die Einhaltung von Datennutzungsrichtlinien durch.

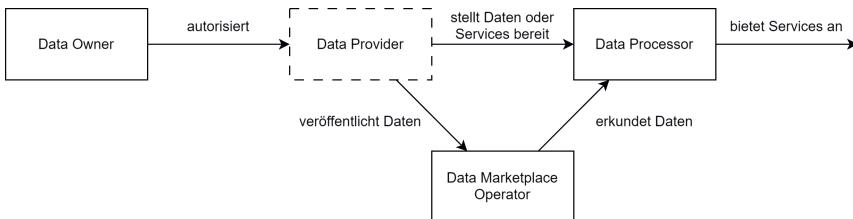


Abbildung 2.1: Rollen in einem Datenraum, eigene Abbildung nach [14]

Um das Zusammenspiel der Teilnehmer nach den oben genannten Regelungen zu organisieren, werden Protokolle und Regelungen benötigt, mit denen die Interaktionen implementiert werden können. Bezüglich der Umsetzung von Datenräumen gibt es verschiedene Ansätze, die von verschiedenen Partnern vorangetrieben werden. Die prominentesten Ansätze werden im Folgenden eingeführt.

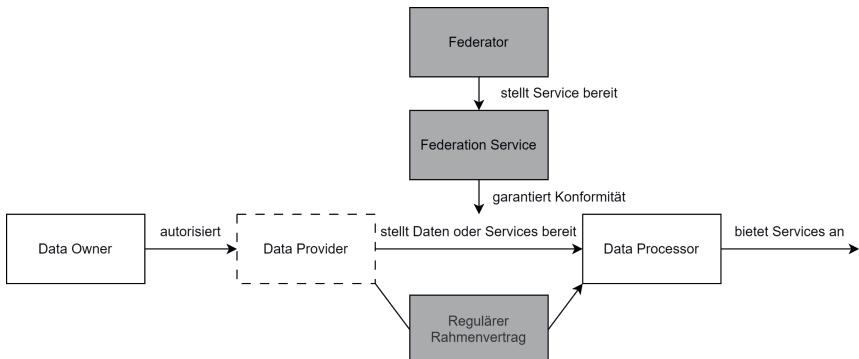
## 2.2.2 Gaia-X und Federation Services

Die Initiative Gaia-X hat sich zum Ziel gesetzt, die Prinzipien von Datenräumen und die immaterielle Infrastruktur in der Wirtschaft zu etablieren und deren Umsetzung voranzutreiben. Zu diesem Zweck erarbeiten Behörden, Unternehmen, Verbände und Forschungsinstitute Richtlinien, um die offenen Standards von Datenräumen zu implementieren [33]. Durch die *Gaia-X Federation Services* werden Software-Stacks entwickelt, die als Toolbox für die Umsetzung von Gaia-X und datengetriebenen Services dienen soll [34]. Hierbei stehen fünf Bereiche im Fokus [35]:

1. *Identität und Vertrauen*: Mit diesen Diensten können die Teilnehmer eines Datenraums sich gegenseitig identifizieren und autorisieren. Die eindeutige Identifikation eines Teilnehmers oder eines Services innerhalb eines Datenraums ist die Voraussetzung für die meisten Anwendungsfälle. Die Identifikation basiert auf Selbstbeschreibungen, welche den Vorgaben des W3C-Standards *Verifiable Credentials* entsprechen [34, 36]. Die Authentifizierung und Autorisierung erfolgt auf Basis einer aus Credentials selbst zusammengestellten *Verifiable Presentation*, die durch eine Reihe von Software-Stacks gemanagt werden kann. Die Bildung einer verteilten Vertrauensstruktur erfolgt durch die Etablierung allgemein anerkannter Vertrauensanker (je nach Anwendung z. B. Behörden, TÜV, Wirtschaftsprüfer, etc.) [34].
2. *Föderierter Katalog*: Der Föderierte Katalog dient innerhalb eines Datenraums als zentrales Repozitorium, in dem Teilnehmer die Informationen und Dienstangebote anderer Teilnehmer erkunden und finden können.
3. *Souveräner Datenaustausch*: Es werden Dienste entwickelt, mit denen Data Owner und Data Provider die Kontrolle über ihre Daten halten und Transparenz über die Nutzung der Daten erhalten. Zentrales Element ist die Möglichkeit, digitale Verträge über den Austausch von Daten abzuschließen und den Austausch zu dokumentieren [34].
4. *Compliance*: Dadurch wird die Überprüfung der Regelkonformität von gemeinsam genutzten Diensten ermöglicht, um sicherzustellen, dass Teilnehmer und Dienste die Gaia-X-Grundsätze einhalten.

5. *Portal und Integration:* Es werden Services geschaffen, um die Dienste von Datenräumen in zugänglichen Oberflächen bereitzustellen und zu orchestrieren.

Welche Federation Services in einem Use Case genutzt werden, obliegt der Absprachen der beteiligten Partner. Die Rollen innerhalb eines Datenraums werden bei Gaia-X um den *Federator* ergänzt, der die Federation Services bereitstellt (s. Abb. 2.2). Es können mehrere Federators pro Datenraum Services zur Verfügung stellen [34]. Die Federation Services werden neben der tatsächlichen Datenservices benötigt, um den Gaia-X-konformen Datenaustausch zu gewährleisten oder Services zu erkunden (analog zum Data Marketplace, vgl. Abb 2.1). Zusätzlich bestehen reguläre Verträge zur Festlegung der Rechte und Pflichten innerhalb des Datenraums.



**Abbildung 2.2:** Rollen in einem Gaia-X Datenraum, eigene Abbildung nach [14] und [34]

### 2.2.3 Eclipse Dataspace Components (EDC)

Für den sicheren und datenschutzkonformen Austausch von Daten innerhalb von Datenräumen ist eine Lösung erforderlich, die sowohl Interoperabilität als auch die Steuerung von Zugriffsrechten und Richtlinien ermöglicht. Das Eclipse Dataspace Components (EDC) Framework [37] erfüllt diese Anforderungen und spielt eine zentrale Rolle im Aufbau vertrauenswürdiger Datenräume. Das Eclipse Dataspace Components (EDC) Framework ist ein Open-Source-Projekt, welches den Prinzipien von Gaia-X folgt und den Aufbau, den Zugriff sowie die Verbindung von vertrauenswürdigen Datenräumen ermöglicht. Das Framework wird von der Eclipse Foundation verwaltet. Seine zentrale Komponente ist der Eclipse Dataspace Components Connector (EDC-Connector), der für einen sicheren, selbstverwalteten Datenaustausch zwischen den Teilnehmern eines Datenraums sorgt und die Einhaltung von Datenschutzanforderungen, wie der EU-Datenschutz-Grundverordnung (DSGVO), unterstützt. Jeder Partner benötigt einen eigenen EDC-Connector, der entsprechende interoperable Schnittstellen bietet und den Austausch von Daten auf eine richtlinienbasierte Weise steuert. Durch die Durchsetzung von Zugangs- und Nutzungsrichtlinien behält jeder Datenanbieter die volle Kontrolle über seine Daten. Diese Richtlinien werden mithilfe der Open Digital Rights Language (ODRL) definiert und regeln in Form von Erlaubnissen, Verboten und Verpflichtungen, wer auf die Daten zugreifen darf, wie sie verwendet werden dürfen und unter welchen Bedingungen sie gelöscht werden müssen. Möglich sind u. a. zeitliche, geografische und zweckgebundene Begrenzungen. Beispielsweise kann die Verwendung auf den Zweck der Qualitätskontrolle begrenzt werden, bei der die Daten nur innerhalb eines bestimmten Zeitraums nach vorheriger Anonymisierung, analysiert werden dürfen und innerhalb der EU verbleiben müssen.

Ein wesentliches Merkmal des EDC-Connectors ist die Unterteilung in die Ebenen Control Plane und Data Plane. Die Control Plane steuert den Datenaustausch, indem sie Richtlinien durchsetzt und den Zugriff auf Daten autorisiert. Sie koordiniert die Kommunikation zwischen den beteiligten Parteien, um sicherzustellen, dass alle Datenbewegungen den festgelegten Nutzungsregeln folgen. Für jeden EDC-Connector gibt es genau eine Control Plane. Von der Data Plane, die für den eigentlichen Datentransfer verantwortlich ist, können hingegen mehrere Instanzen parallel existieren, die auf unterschiedliche Protokolle, Datenarten oder Transferanforderungen zugeschnitten sind. Dies ermöglicht es, verschiedene Transferwege getrennt zu behandeln, während die Sicherheit und Interoperabilität gewahrt bleiben.

Für den Transfer müssen die Daten zunächst über den EDC-Connector des Datenanbieters katalogisiert und mit entsprechenden Zugriffs- und Nutzungsrichtlinien in einem Vertrag verknüpft werden. Die Daten verbleiben hierbei an ihren Quellen, welche u. a. lokale Dateisysteme, Cloud-Speicher oder über APIs zugängliche Datenquellen umfassen können. Der Katalog des Anbieters kann anschließend vom EDC-Connector eines vertrauenswürdigen Datenkonsumenten im Datenraum abgerufen werden. Wenn dieser einen gewünschten Datenbestand beziehen möchte, muss er dem Anbieter ein entsprechendes Vertragsangebot zur Verhandlung zusenden. Dieses muss die vom Anbieter geforderten Richtlinien erfüllen. Kommt es durch die Zustimmung des Datenanbieters zu einer Einigung kann der Datentransfer auf zwei Wegen erfolgen. Entweder sendet der Anbieter die Daten sofort oder er vergibt einen Token mit dem die Daten zu einem späteren Zeitpunkt transferiert werden können. Beim Transfer dient der EDC-Connector für beide Seiten als Proxy und verbirgt alle Datenquellen / -referenzen und Ziel-Speicherorte.

Das EDC-Framework umfasst außerdem zahlreiche Funktionserweiterungen des EDC-Connectors, z. B. für Persistenz, unterschiedliche Authentifizierungsverfahren oder für die Konnektivität zu bestimmten Cloud-Diensten. Eigene Erweiterungen können durch den Open-Source Ansatz ebenfalls integriert werden.

## 2.2.4 Sovereign Cloud Stack

Der *Sovereign Cloud Stack* (SCS) ist ein Open-Source-Projekt vom Bundesministerium für Wirtschaft und Klimaschutz (BMWK), welches im Rahmen von Gaia-X gefördert wird [38]. Das Kernziel des SCS besteht darin, eine Cloud-Infrastruktur bereitzustellen, die den spezifischen Anforderungen von europäischen Unternehmen und öffentlichen Institutionen gerecht wird. Der SCS dient dabei als Entwicklungsplattform für die Gaia-X Federation Services [39–41]. Das Projekt trägt zu den Bemühungen der europäischen Union bei, die technologische Souveränität unter Beachtung der europäischen Datenschutzrichtlinien zu stärken und eine unabhängige Alternative zu den globalen Cloud-Anbietern zu schaffen [38, 42].

Der SCS basiert auf offenen Standards und Open-Source-Technologien, was Transparenz, Sicherheit und Interoperabilität zwischen verschiedenen Cloud-Anbietern und -Diensten gewährleistet [38, 42]. Diese Interoperabilität wird durch zertifizierbare Standards erreicht, die sowohl eine organisationsübergreifende Föderation als auch eine einfachere Migration von Diensten und Daten aus bestehenden Umgebungen ermöglicht [38, 40, 43]. Damit können Unternehmen ihre Aufgaben und/oder Anforderungen an eine Cloud auf die Clouds mehrerer Betreiber verteilen, was eine hohes Maß an Flexibilität ermöglicht [40]. Darüber hinaus vereinfacht SCS den Aufbau und Betrieb moderner Cloud-Infrastrukturen. Aufgrund seiner modularen Struktur ermöglicht SCS zudem eine schrittweise Umstellung sowie den stetigen Einsatz einzelner Komponenten innerhalb von bereits existenten IT-Infrastrukturen [38]. Ein Beispiel für die erfolgreiche Implementierung von SCS ist die *PlusCloudOpen* von PlusServer.

## 2.3 Weitere Technologien

In diesem Unterkapitel werden ergänzende Technologien beschrieben, die in den Projekten erprobt und angewendet wurden, aber nicht zu den zentralen Technologien von Gaia-X zählen. Ein Beispiel hierfür ist die die *Open Platform Communications Unified Architecture* (OPC UA), die zur Unterstützung der Interoperabilität und Kommunikation in den Systemen beiträgt.

### 2.3.1 Asset Administration Shell (AAS)

Die AAS ist ein Metamodell zur Standardisierung von Digitalen Zwillingen, welches von der *Industrial Digital Twin Association e.V.* (IDTA) veröffentlicht wird. Zweck des Modells ist es, Informationen einheitlich zu strukturieren und diese über definierte Schnittstellen zu kommunizieren.

Es wird eine hierarchische Gliederung angewendet, die eine klare Trennung der verschiedenen Ebenen von Informationen ermöglicht und die Zuordnung von Daten zu spezifischen Aspekten eines Assets erleichtert. In der hierarchischen Gliederung wird zunächst definiert, welches Asset verwaltet werden soll. Anschließend erfolgt die Erstellung von spezifischen Submodellen, die die verschiedenen Aspekte und Eigenschaften des Assets in logische Gruppen einteilen. Beispiele für solche logischen Gruppen können sein: *Technical Data*, *Operational Data* oder auch *Documentation*. Hierbei unterstützt die IDTA mit bereits vordefinierten *Submodel Templates*. Diese können – und sollten – für den Aufbau der eigenen AAS verwendet werden, um eine standardisierte Zuordnung zwischen inhaltlich gleichen Submodellen zu ermöglichen. Den kleinsten Informationsbausteine bilden die sogenannten *Submodel Elements*, welche die Eigenschaften eines Assets innerhalb eines Submodells abbilden. Für die Erstellung der Submodelle lassen sich auch existierende Standardisierungsformate, wie sie etwa vom *ECLASS e.V.* bereitgestellt werden, importieren.

Schlussendlich wird das Metamodell als eine strukturierte Textdatei im JSON oder XML Format abgelegt oder als Bestandteil des AAS-spezifischen Containerformats AASX gespeichert. Durch den Austausch dieser Dateien können Informationen über digitale Zwillinge geteilt werden. Für die Bereitstellung von Asset Administration Shells als digitalen Service wird ein entsprechender AAS Server benötigt. Folgende Implementierungen wurden in den Projekten erprobt und verwendet:

**IDTA AAS<sup>1</sup>** wird von der *Industrial Digital Twin Association e.V.* entwickelt und bereitgestellt. Sie setzt den von der IDTA definierten Standard um.

**BaSyx AAS<sup>2</sup>** ist ein vom *Fraunhofer IESE* entwickeltes AAS Projekt, welches sich nah am IDTA Standard orientiert.

**FAAAST AAS<sup>3</sup>** ist ebenfalls ein Projekt zur Bereitstellung von AAS Metamodellen und wird vom *Fraunhofer IOSB* entwickelt.

### 2.3.2 OPC UA

*Open Platform Communications Unified Architecture* (OPC UA) ist ein industrieller Kommunikationsstandard, der in der Automatisierungstechnik und Fertigungsinustrie weit verbreitet ist. Entwickelt wurde er von der OPC Foundation und stellt eine moderne, plattformunabhängige Weiterentwicklung der ursprünglichen OPC-Standards dar. Sein Ziel ist es, eine sichere, zuverlässige und herstellerunabhängige Kommunikation zwischen Maschinen, Anlagen, Geräten und IT-Systemen zu ermöglichen. Im Vergleich zu den vorherigen OPC-Standards, die auf Microsoft-Technologien wie COM/DCOM basierten, ist OPC UA plattformunabhängig und verwendet moderne Kommunikationsprotokolle wie TCP/IP oder HTTP. Es unterstützt verschiedene Implementierungen, von kleinen eingebetteten Systemen bis hin zu komplexen IT-Umgebungen, und kann auf verschiedenen Betriebssystemen wie

<sup>1</sup><https://github.com/eclipse-aaspe/server>

<sup>2</sup><https://github.com/eclipse-basyx/basyx-java-server-sdk>

<sup>3</sup><https://github.com/FraunhoferIOSB/FAAAST-Service>

Windows, Linux oder sogar in Cloud-Umgebungen ausgeführt werden. Ein wesentlicher Vorteil von OPC UA ist seine *Service-orientierte Architektur* (SOA). Das bedeutet, dass Systeme über klar definierte Dienste miteinander kommunizieren, die beispielsweise Daten lesen und schreiben, Ereignisse überwachen oder Methoden ausführen können. Dies ermöglicht eine flexible und skalierbare Struktur, die sowohl für einfache Datenübertragungen als auch für komplexe Steuerungsprozesse geeignet ist. Sicherheitsaspekte spielen eine zentrale Rolle bei OPC UA. Der Standard bietet End-to-End-Verschlüsselung, Authentifizierung, Autorisierung und Datenschutzmechanismen, die es ermöglichen, Daten sicher über verschiedene Netzwerke und Systemgrenzen hinweg zu übertragen. Dies ist besonders in der Industrie 4.0 und dem *Industrial Internet of Things* (IIoT) wichtig, wo Maschinen, Sensoren und IT-Systeme eng miteinander vernetzt sind. Ein weiterer Vorteil von OPC UA ist die Unterstützung einer semantischen Datenmodellierung. Durch die Möglichkeit, komplexe Datenstrukturen zu beschreiben und zu übertragen, können Maschinen und Systeme nicht nur einfache Werte, sondern auch deren Bedeutung und Kontext austauschen. Dies ist entscheidend für die Interoperabilität in heterogenen Umgebungen.

Insgesamt stellt OPC UA einen wichtigen Baustein für die Realisierung von intelligenten Fertigungssystemen dar, da es nicht nur die technische Kommunikation zwischen Geräten sicherstellt, sondern auch die Grundlage für die Integration von IT-Systemen und die Umsetzung datengetriebener Prozesse bietet. Dies macht OPC UA zu einem Schlüsselfaktor der digitalen Transformation und der Weiterentwicklung hin zu intelligenten Fabriken im Kontext von Industrie 4.0.

**OPC UA Companion Specifications** sind Erweiterungen des OPC UA-Standards, die spezifische Anforderungen und Datenmodelle für bestimmte Branchen, Technologien oder Anwendungen definieren. Während OPC UA als universeller Kommunikationsstandard fungiert, stellen die Companion Specifications sicher, dass branchenspezifische Informationen und Funktionen nahtlos und standardisiert integriert werden können. Sie dienen als verbindendes Element, um die Interoperabilität zwischen unterschiedlichen Systemen und Geräten herstellerübergreifend zu gewährleisten.

Jede Companion Specification wird in enger Zusammenarbeit zwischen der OPC Foundation und verschiedenen Industrieverbänden, Normungsorganisationen oder Konsortien entwickelt. Der Zweck dieser Erweiterungen besteht darin, branchenspezifische Datenmodelle zu definieren, die sicherstellen, dass Maschinen, Anlagen und IT-Systeme einer bestimmten Branche nicht nur über OPC UA kommunizieren, sondern auch die Bedeutung und Struktur der übertragenen Daten einheitlich interpretieren können. Beispiele für Companion Specifications gibt es in den Bereichen der Robotik, Automatisierung, der Pharmaindustrie oder des Energiewesens.

Ein wesentlicher Aspekt der Companion Specifications ist die semantische Interoperabilität. Dies bedeutet, dass die Bedeutung der ausgetauschten Daten durch standardisierte Modelle und Beschreibungen klar definiert ist, was es Systemen ermöglicht, nicht nur Daten zu verstehen, sondern auch deren Kontext zu erkennen. Dies ist besonders wichtig in heterogenen Systemlandschaften, in denen unterschiedliche Technologien und Maschinen miteinander agieren müssen.

Ein typisches Beispiel für eine Companion Specification ist die Zusammenarbeit zwischen dem Verband Deutscher Maschinen- und Anlagenbau (VDMA) und der OPC Foundation, bei der spezifische Datenmodelle für die Maschinenbauindustrie entwickelt wurden. Diese Companion Specifications ermöglichen es den Maschinenherstellern, ihre Maschinen und Anlagen in Industrie 4.0-Umgebungen effizienter und standardisiert zu integrieren.

Durch die Verwendung von OPC UA Companion Specifications wird die Interoperabilität auf einem hohen Niveau sichergestellt, was für den Erfolg von Industrie 4.0 und dem IIoT von zentraler Bedeutung ist. Sie bieten einen strukturierten Rahmen, um branchenspezifische Anforderungen in eine globale, standardisierte Kommunikationsarchitektur zu überführen und tragen so zur Effizienzsteigerung, Flexibilität und Skalierbarkeit industrieller Systeme bei.

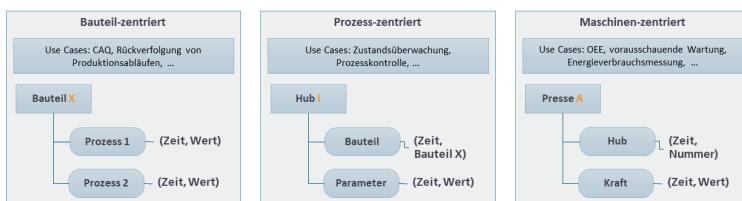
### 2.3.3 LinkedFactory

Die Kombination von Prozess-, Produkt- und Ressourcendaten ist unerlässlich, um die Einflussfaktoren auf Produktqualität und Produktionsleistung besser zu verstehen. Unternehmen stehen dabei jedoch vor vielfältigen Herausforderungen. Zum einen werden Bezugsobjekte wie Produkte, Prozesse und Ressourcen in unterschiedlichen Strukturen (z. B. Stücklisten, CAx-Daten) und Datentypen (z. B. numerische Werte, Strings, Datensätze) aus verschiedenen Quellen wie Steuerungssystemen und Sensoren erfasst. Zum anderen erzeugen die Systeme zwar große Mengen an Informationen, diese sind jedoch häufig weder universell verständlich aufbereitet noch leicht zugänglich und auswertbar.

Semantische Modelle können die Organisation der Daten unterstützen, indem sie Struktur- und Herkunftsinformationen bereitstellen. Gleichzeitig sollte eine flexible Datenarchitektur die Integration von beobachteten Objekten und deren Attributten zu jedem Zeitpunkt ermöglichen.

Die LinkedFactory (LF)-Architektur<sup>4</sup> stellt solch eine flexible Architektur zur Verfügung und definiert dazu ein einfaches Datenformat und Protokoll zur Erfassung, Speicherung und Abfrage von Produktionsdaten. Dadurch wird ein generisches Datenmodell realisiert, welches es ermöglicht Zeitreihendaten und die zugehörigen Metadaten durch Nutzung der Linked-Data-Prinzipien zu speichern und zu verknüpfen. LF verwendet ein an JSON-LD angelehntes leichtgewichtiges Linked-Data-Format, das eine graphenbasierte Struktur zur Darstellung von Objekten und ihren Attributen ermöglicht. Weiterhin werden HTTP-basierte APIs zur Aufnahme, zum Abruf und zur Abfrage der Daten definiert [44]. Darüber hinaus ermöglicht ein SPARQL-Endpunkt die gemeinsame Abfrage und Aggregation des Wissensgraphs und der Zeitreihendaten. Letztere werden durch das Bezugsobjekt, die erfasste Eigenschaft, einen Kontext (z. B. Organisationseinheit, Simulationslauf) und einem Zeitstempel qualifiziert. Durch die Anwendung der Linked-Data-Prinzipien ist es möglich, Verknüpfungen zwischen den Werten eines Elements und anderen Elementen herzustellen (z. B. Teile, die von einer Maschine zu einem bestimmten Zeitpunkt produziert werden, und das Teil selbst mit den zugehörigen Prozessdaten). Dieser weborientierte Ansatz erlaubt die Kombination von semantischen Beschreibungen für Produktionssysteme (gespeichert als RDF) mit großen Mengen von Produktionsdaten (gespeichert in einer Zeitreihendatenbank) [45].

Innerhalb dieses flexiblen Ansatzes ist es darüber hinaus möglich, verschiedene Arten von semantischen Beschreibungen zu integrieren, unabhängig davon, ob sie auf Ontologien basieren, einfache RDF-Beschreibungen verwenden oder aus anderen Frameworks, wie dem Semantic Aspect Meta Model (SAMM), abgeleitet sind. Dies ermöglicht die Darstellung von Daten für mehrere Anwendungsfälle, z. B. in bauteil-, prozess- und maschinenzentrierten Sichten. In der Abbildung 2.3 ist dies exemplarisch am Beispiel eines Produktionsprozesses mit einer Presse dargestellt.



**Abbildung 2.3:** Verschiedene Sichten auf Produktionsdaten, eigene Abbildung

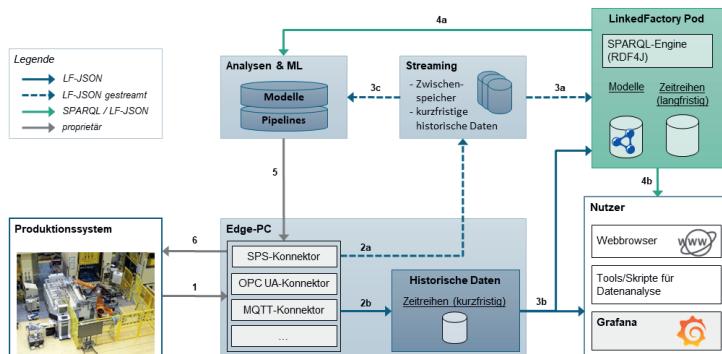
Abbildung 2.4 zeigt eine mögliche Referenzarchitektur. Die Daten werden in der Regel über konfigurierbare Konnektoren erfasst, die Industriestandards wie OPC UA, S7, MQTT und Modbus un-

<sup>4</sup><https://github.com/linkedfactory/specification>

terstützen. Diese Daten werden in das LF-JSON-Format umgewandelt und entweder per Streaming (2a) weitergeleitet oder über die LF-HTTP-API in einem lokalen oder entfernten historischen Daten-dienst (2b) gespeichert. Die LF-HTTP-API ermöglicht zudem das Batch-Laden mehrerer Datensätze über POST-Anfragen, bei denen LF-JSON- oder CSV-Formate als Nutzdaten genutzt werden können.

Die LF-Architektur unterstützt verschiedene Speicherlösungen zur Verwaltung von Zeitreihen-daten, darunter SQL, NoSQL und Parquet. Für eigenständige Implementierungen kann die Key-Value-Datenbank wie LevelDB in Kombination mit einem Parquet-basierten Archiv verwendet werden. Je nach Anforderung lässt sich Streaming entweder mit klassischem Message Queueing, wie etwa MQTT, oder mit Kafka realisieren.

Eine Open-Source-Referenzimplementierung ist als LinkedFactory Pod<sup>5</sup> verfügbar, die die Aufnah-me von (Streaming-)Daten (3a) sowie SPARQL-basierte Abfragen auf verteilten Datenservices über die LF-HTTP-API (3b) unterstützt. Durch die Verwendung der Abfragesprache SPARQL wird eine integrierte Abfrage sowohl der Zeitreihen- als auch der Metadaten ermöglicht. Abhängig von den zeitlichen Anforderungen können Online-Daten per Streaming (3c) oder durch eine Kombination aus der LF-HTTP-API und SPARQL-Abfragen (4a) bereitgestellt werden. Letztere Option erlaubt benutzerde-finierte Analysen und Visualisierungen, beispielsweise mit Grafana und einer LF-Datenquelle (4b). Der Regelkreis wird geschlossen, indem Parameter über einen Konnektor (5) an das Produktionssystem zurückgesendet werden (6).



**Abbildung 2.4:** Architekturentwurf für den Aufbau von Regelkreisen für Produktionsprozesse, eigene Abbildung

<sup>5</sup><https://github.com/linkedfactory/linkedfactory-pod>

## Literaturverzeichnis

- [1] Uwe Bracht, Dieter Geckler, and Sigrid Wenzel. *Digitale Fabrik: Methoden und Praxisbeispiele*. VDI-Buch. Springer Vieweg, Berlin and Heidelberg, 2., aktualisierte und erweiterte auflage edition, 2018. ISBN 3662557827. doi: 10.1007/978-3-662-55783-9.
- [2] Michael Schneider, Benjamin Hippchen, Sebastian Abeck, Michael Jacoby, and Reinhard Herzog. Enabling iot platform interoperability using a systematic development approach by example. In *2018 Global Internet of Things Summit (GIoTS)*, pages 1–6, 2018. doi: 10.1109/GIOTS.2018.8534549.
- [3] Mark Kasunic and William Anderson. Measuring systems interoperability: Challenges and opportunities. Technical report, Software Engineering Institute, Carnegie Mellon University, 2004.
- [4] Sandra Heiler. Semantic interoperability. *ACM Computing Surveys*, 27(2):271–273, 1995. ISSN 0360-0300. doi: 10.1145/210376.210392.
- [5] Farhad Ameri, Dusan Sormaz, Foivos Psaromatis, and Dimitris Kiritsis. Industrial ontologies for interoperability in agile and resilient manufacturing. *International Journal of Production Research*, 60(2):420–441, 2022. ISSN 0020-7543. doi: 10.1080/00207543.2021.1987553.
- [6] François B. Vernadat. Technical, semantic and organizational issues of enterprise interoperability and networking. *Annual Reviews in Control*, 34(1):139–144, 2010. ISSN 13675788. doi: 10.1016/j.arcontrol.2010.02.009.
- [7] Gabriel S. S. Leal, Wided Guédria, and Hervé Panetto. Enterprise interoperability assessment: a requirements engineering approach. *International Journal of Computer Integrated Manufacturing*, 33(3):265–286, 2020. ISSN 0951-192X. doi: 10.1080/0951192X.2020.1736636.
- [8] Daniel Hodapp and André Hanelt. Interoperability in the era of digital innovation: An information systems research agenda. *Journal of Information Technology*, 37(4):407–427, 2022. ISSN 0268-3962. doi: 10.1177/02683962211064304.
- [9] Damianos Sakas and Konstadinos Kutsikos. An adaptable decision making model for sustainable enterprise interoperability. *Procedia - Social and Behavioral Sciences*, 148:611–618, 2014. ISSN 18770428. doi: 10.1016/j.sbspro.2014.07.087.
- [10] Gaia-X. Gaia-x - architecture document. Technical report, Gaia-X, 2022.
- [11] Boris Otto. The evolution of data spaces. In Boris Otto, Michael ten Hompel, and Stefan Wrobel, editors, *Designing Data Spaces*, pages 3–15. Springer International Publishing, Cham, 2022. ISBN 978-3-030-93974-8. doi: 10.1007/978-3-030-93975-5{\textunderscore}1.
- [12] Abel Reiberg, Crispin Niebel, and Peter Kraemer. What is a data space? White Paper 1/2022, Gaia-X Hub Germany, 2022. Gaia-X Hub Germany White Paper.
- [13] Boris Otto, Jan Jürjens, Jochen Schon, Sören Auer, Nadja Menz, and Sven Wenzel. Industrial data space - digitale souveränität über daten. Technical report, Fraunhofer-Gesellschaft, 2016.
- [14] Lars Nagel and Douwe Lycklama. Design principles for data spaces, 2021. URL <https://doi.org/10.5281/zenodo.5244997>.

- [15] Hubert Tardieu. Role of gaia-x in the european data space ecosystem. In Boris Otto, Michael ten Hompel, and Stefan Wrobel, editors, *Designing Data Spaces*, pages 41–59. Springer International Publishing, Cham, 2022. ISBN 978-3-030-93974-8. doi: 10.1007/978-3-030-93975-5{\textunderscore}4.
- [16] Martin Johanson and Torkel Wedin. Value creation in industrial networks. In *16th Annual IMP Conference*, 2000.
- [17] Maila Herrala and Pekka Pakkala. Value-creating networks – a conceptual model and analysis, 2009.
- [18] Shaker A. Zahra and Satish Nambisan. Entrepreneurship and strategic thinking in business ecosystems. *Business Horizons*, 55(3):219–229, 2012. ISSN 00076813. doi: 10.1016/j.bushor.2011.12.004.
- [19] Rahul Kapoor. Ecosystems: broadening the locus of value creation. *Journal of Organization Design*, 7(1), 2018. doi: 10.1186/s41469-018-0035-4.
- [20] James F. Moore. Business ecosystems and the view from the firm. *The Antitrust Bulletin*, 51(1):31–75, 2006. ISSN 0003-603X. doi: 10.1177/0003603X0605100103.
- [21] Svetlana Maltseva and Peter Golubtsov. Business ecosystem & data ecosystem: Introduction to international workshop on big data for business ecosystems. In *International Conference on Information Systems 2022 Special Interest Group on Big Data Proceedings*. Association for Information Systems, 2022. URL <https://aisel.aisnet.org/sigbd2022/1>.
- [22] Patrik Hummel, Matthias Braun, Max Tretter, and Peter Dabrock. Data sovereignty: A review. *Big Data & Society*, 8(1):2053951720982012, 2021.
- [23] Alexander Roßnagel. Datenschutz-grundverordnung. *Datenschutzaufsicht nach der EU-Datenschutz-Grundverordnung: Neue Aufgaben und Befugnisse der Aufsichtsbehörden*, pages 21–39, 2017.
- [24] Huw Roberts, Josh Cowls, Federico Casolari, Jessica Morley, Mariarosaria Taddeo, and Luciano Floridi. Safeguarding european values with digital sovereignty: an analysis of statements and policies. *Internet Policy Review, Forthcoming*, 2021.
- [25] Jean Galbraith. Congress enacts the clarifying lawful overseas use of data (cloud) act, reshaping us law governing cross-border access to data. *The American Journal of International Law*, 112(3):487–493, 2018.
- [26] Edoardo Celeste. Digital sovereignty in the eu: Challenges and future perspectives. In Federico Fabbrini, Edoardo Celeste, and John Quinn, editors, *Data protection beyond borders*. Hart, Oxford and London and New York and New Delhi and Sydney, 2022. ISBN 978-1-50994-066-0. doi: 10.5040/9781509940691.ch-013.
- [27] Christian Jung and Jörg Dörr. Data usage control, 2022.
- [28] Markus Eurich and Michael Burtscher. The business-to-consumer lock-in effect. *Cambridge Service Alliance*, 14, 2014.
- [29] Gaia-X Hub. Gaia-x explained, 2024. URL <https://gaia-x-hub.de/en/gaia-x-explained/>.

- [30] Justice Opara-Martins, Reza Sahandi, and Feng Tian. Critical analysis of vendor lock-in and its impact on cloud computing migration: a business perspective. *Journal of Cloud Computing*, 5 (1), 2016. doi: 10.1186/s13677-016-0054-z.
- [31] Alice Pannier. *The changing landscape of European cloud computing: Gaia-X, the French national strategy, and EU plans*. Briefings de l'Ifri. Ifri, Paris, 22 July 2021. ISBN 9791037303882. URL [https://www.ifri.org/sites/default/files/atoms/files/pannier\\_european\\_cloud\\_computing\\_2021.pdf](https://www.ifri.org/sites/default/files/atoms/files/pannier_european_cloud_computing_2021.pdf).
- [32] Boris Otto. Gaia-x and ids: Position paper, version 1.0, 2021. URL <https://doi.org/10.5281/zenodo.5675897>.
- [33] Francesco Bonfiglio. Vision & strategy, 2021. URL <https://gaia-x.eu/wp-content/uploads/2021/12/Vision-Strategy.pdf>.
- [34] Gaia-X European Association for Data and Cloud AISBL. Gaia-x – architecture document: 22.10 release, 2022. URL <https://docs.gaia-x.eu/technical-committee/architecture-document/22.10/>.
- [35] eco – Verband der Internetwirtschaft e.V. Strategien zum aufbau von gaia-x ökosystemen mithilfe der gaia-x föderationsdienste: Im dialog mit den gaia-x förderprojekten, 2022. URL <https://www.gxfs.eu/de/downloads/>.
- [36] World Wide Web Consortium. Verifiable credentials data model v1.1, 03 March 2022. URL <https://www.w3.org/TR/vc-data-model>.
- [37] Eclipse Foundation. Eclipse dataspace components connector, 2024. URL <https://github.com/eclipse-edc>.
- [38] Manuela Urban and Kurt Garloff. Sovereign cloud stack: Technische souveränität durch offene, föderierbare und modulare cloud-plattformen vollständig auf basis von open-source-software. *Datenschutz und Datensicherheit-DuD*, 46:616–621, 2022. URL <https://link.springer.com/article/10.1007/s11623-022-1669-5>.
- [39] Bundesministerium für Wirtschaft und Klimaschutz. Fortschrittsbericht 2022 – industrie 4.0 gestalten. resilient, nachhaltig, wettbewerbsstark. Technical report, Bundesministerium für Wirtschaft und Klimaschutz, 2022.
- [40] OSB Alliance. Sovereign cloud stack veröffentlicht release 6, 2024. URL <https://osb-alliance.de/pressemitteilungen/sovereign-cloud-stack-veroeffentlicht-release-6>.
- [41] Sovereign Cloud Stack. Über sovereign cloud stack, 2024. URL <https://scs.community/de/about/>.
- [42] OSB Alliance. Osb alliance entwickelt sovereign cloud stack (scs) für das gaia-x projekt, 2021. URL <https://osb-alliance.de/sovereign-cloud-stack-scs>.
- [43] SCS-Community. Frequently asked questions, 2024. URL <https://docs.scs.community/docs/faq>.
- [44] Marko Friedemann, Ken Wenzel, and Adrian Singer. Linked data architecture for assistance and traceability in smart manufacturing. *MATEC Web of Conferences*, 304:04006, 01 2019. doi: 10.1051/matecconf/201930404006.

- [45] Jossy Milagros Grützmann, Franziska Rudolph, Martin Boesler, and Ken Wenzel. Flexible data architecture for enabling ai applications in production environments. In Martin Krockert and Torsten Munkelt, editors, *Workshop Proceedings AI in Production - 1, KI 2024 - 47th German Conference on Artificial Intelligence*, volume 1 of *AI in Production*, Würzburg, Germany, 2024. Hochschule für Technik, Wirtschaft und Kultur Leipzig. doi: 10.33968/2024.77.