Federated Data Spaces für den Bahnsektor – Ein technischer Einblick

Durch Data Spaces entsteht eine digitale Infrastruktur für einen sicheren, EU-rechtskonformen Datenaustausch. In diesem Beitrag werden die technologischen Grundlagen föderierter Data Spaces, die zentralen Komponenten und das Trust Framework vorgestellt. Ein Anwendungsfall aus dem Bahnsektor zeigt zudem, wie domänenspezifische Labels die Datenproduktqualität und Interoperabilität fördern können.

Einleitung

Daten als wirtschaftliches Gut in Unternehmen, Datenprodukte, Services rund um Datenprodukte und die Notwendigkeit einer EU-rechtlich konformen und sicheren Datenökonomie haben Data Spaces branchenübergreifend zu einem vielversprechenden Ansatz werden lassen. So sind sie durch den Data Governance Act (Europäische Kommission 2022) und Data Act (Europäische Kommission 2023b) zu einem integralen Bestandteil der EU-Datenstrategie geworden [1]. Daraus sind auf EU-Ebene Strukturen erwachsen, die die Thematik durch Initiativen sowie Forschungs- und Industrieprojekte technologisch und wirtschaftlich resilient weiterentwickeln. Die innovative Grundlage bildet dabei die föderierte und dezentrale Infrastruktur: So wird eine Vernetzung heterogener Datenquellen ermöglicht, bei der Organisationen ihre Datensouveränität wahren und zugleich von einem gemeinsamen Ökosystem profitieren. Die Daten verbleiben auf den IT-Systemen der Eigentümer, während standardisierte Schnittstellen und Brokerservices einen kontrollierten und sicheren Zugriff ermöglichen. Dies basiert auf einem für den Data Space geltenden Trust Framework, das Regeln und Standards für einen sicheren und vertrauenswürdigen Austausch von Daten formalisiert. In der Folge werden Daten nicht nur isoliert in Organisationen verwaltet, sondern können auf Basis standardisierter Schnittstellen und transparenter Zugriffsregeln geteilt und somit kollaborativ für verschiedene Anwendungsfälle nutzbar gemacht werden.

Nachdem in zwei ETR-Artikeln dieses Jahr (vgl. ETR 3/2025 und ETR 6/2025) bereits auf Grundlegendes eingegangen wurde, widmet sich dieser Artikel der technologischen Perspektive auf die zentralen Komponenten der föderierten Data Space-Architektur. Ebenso wird das Trust Framework als Formalisierung geltender Richtlinien vorgestellt. Zuletzt wird am Beispiel von bahndomänenspezifischen Labels erörtert, wie die in Data Spaces gehandelten Datenprodukte zukünftig ein zusätzliches Zertifizierungsinstrument erhalten und so die Datenqualität und Interoperabilität erhöht werden können.

Ein Blick auf die föderierte Data Space-Infrastruktur

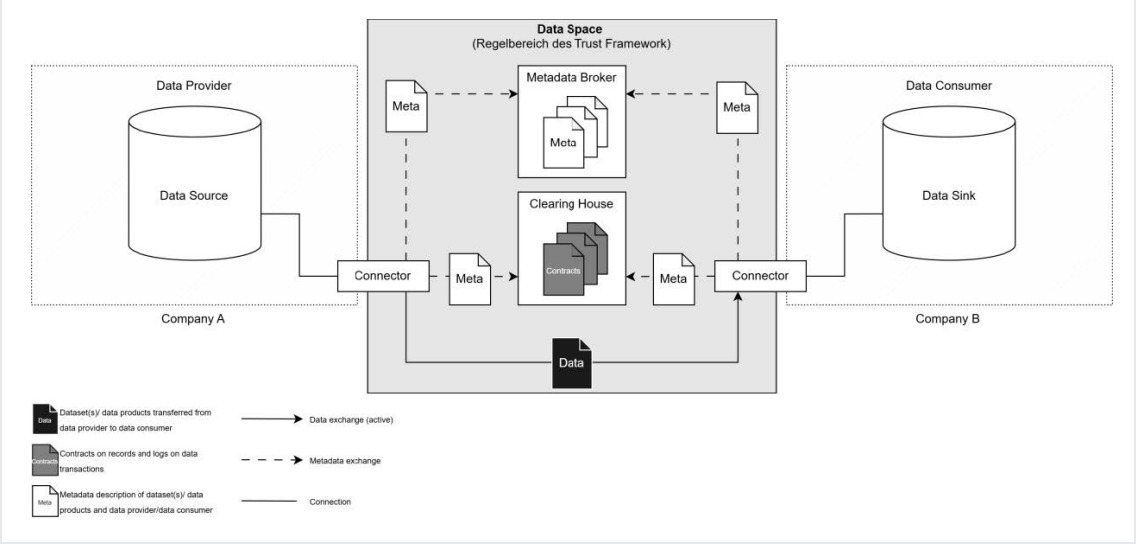
Die föderierte Infrastruktur von Data Spaces bezeichnet den dezentralen Ansatz zur Schaffung von Datenökosystemen, bei dem Daten nicht an einem zentralen Ort zusammengeführt werden, sondern in ihren ursprünglichen Quellen verbleiben. Der Zugriff, die Nutzung und der Austausch von Daten wird durch standardisierte Protokolle, Metadaten und Sicherheitsmechanismen ermöglicht. Data Spaces verfolgen das Prinzip der Dezentralisierung, sodass lediglich Metadaten als Beschreibung der angebotenen Daten(produkte) an Datenbroker(dienste) weitergegeben werden, die diese Angebote katalogisieren und Matchmaking-Dienste für Data Provider und Data Consumer anbieten [2, 3]. Diese Architektur bildet die infrastrukturelle Grundlage, die es teilnehmenden Parteien ermöglicht, Daten in einer sicheren, transparenten und vertrauenswürdigen Weise auszutauschen. Dabei steht die Datensouveränität im Vordergrund, die durch die semantische und technische Formalisierung von Datennutzungsbedingungen implementiert und sichergestellt wird [4]. Auf diese Weise werden Daten erst transferiert, wenn ein Konsens über die Datennutzung und -vergütung erreicht wurde [3].

Nicolás Ortiz, M. Sc.
Wissenschaftlicher Mitarbeiter, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Institut für Verkehrssystemtechnik
nicolas.ortizkammrath@dlr.de

Dipl.-Geolnf. Christian Linder
Wissenschaftlicher Mitarbeiter, Gruppenleiter, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Institut für Verkehrssystemtechnik
christian.linder@dlr.de

Angela Uschok, M. Sc.
Wissenschaftliche Mitarbeiterin, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Institut für Verkehrssystemtechnik
angela.uschok@dlr.de

Dipl.-Ing. Michael Kürschner
Wissenschaftlicher Mitarbeiter, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Institut für Verkehrssystemtechnik
michael.kuerschner@dlr.de



1: Vereinfachte Darstellung der Data Space Architekturkomponenten, Daten- und Metadatenflüsse

Quelle: eigene, vereinfachte Darstellung nach IDSA RAM 4.0 [9]

Föderierte Systeme geben Teilnehmern eine erhöhte Autonomie und Selbstbestimmung [5], sodass die Kontrolle über die eigenen Daten besteht und die Bedingungen für deren Nutzung festgelegt werden können. Dies erfolgt in Abgrenzung zu Datenintegrationsmodellen mit zentralen Instanzen, die als Intermediäre Daten unautorisiert an Dritte veräußern könnten [6]. Föderationen können in dem Kontext als heterogene, lose gekoppelte Zusammenschlüsse autonomer Akteure definiert werden, die durch die Bereitstellung von Infrastrukturdiensten und der Implementierung von Interoperabilitätsmechanismen eine virtuelle Nutzungsebene in einem Data Space bilden und teilen [7, 8].

Die zentralen Komponenten eines föderierten Data Spaces sind in Abbildung 1 veranschaulicht: Der Connector ist eine Softwarekomponente der föderierten Infrastruktur und dient den Teilnehmern als Gateway zum Data Space, um nach gemeinsamen Regeln auf Daten zuzugreifen [2]. Um verfügbare Daten zu katalogisieren und auffindbar zu machen, ohne die eigentlichen Daten zentralisieren zu müssen, sammelt der Metadata Broker semantisch beschriebene Informationen über Datenressourcen und Dienste. Er verwaltet die Metadaten über verfügbare Datenangebote und ermöglicht es den Teilnehmern so, relevante Daten zu entdecken und deren Eigenschaften zu verstehen. Es ist somit ein zentraler Dienst und fungiert als Vermittlungskomponente. Diese registriert zudem Schnittstellen, Nutzungsbedingungen und weitere Informationen zum Datenangebot und ermöglicht auf der anderen Seite semantische Suchen nach Daten. Datenangebot- und -nachfrage finden auf diese Weise zusammen und können Verträge über Datentransaktionen abschließen, sofern die Datenzugangs- und Datennutzungsbedingungen von beiden Parteien erfüllt werden. Die Überprüfung dessen erfolgt durch das Clearing House, welches Datenvermittlungsprozesse wie Verträge und Datentransaktionsprotokollierungen koordiniert. Das Clearing House ist auch eine wichtige Infrastrukturkomponente zur wirtschaftlichen Abwicklung und Vertrauensbildung im datenbasierten Ökosystem. So werden Datenflüsse geloggt, wodurch die Transparenz und Nachvollziehbarkeit der Vorgänge steigt und die Teilnehmeridentitäten eindeutig mit Aktivitäten im Data Space in Verbindung gebracht werden können. Alle o.g. Komponenten der föderierten Data Spaces implementieren standardisierte Protokolle und Sicherheitsmechanismen, wofür bereits Referenzimplementierungen existieren. [9, 10, 11]

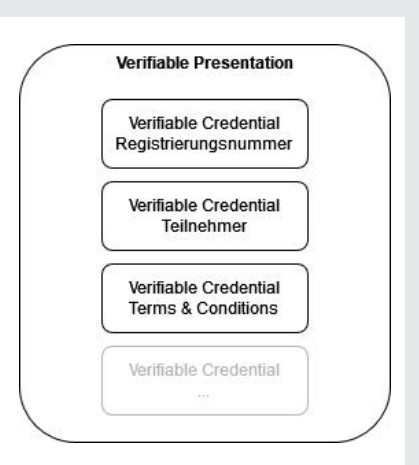
Somit ergeben sich nach Abb. 1 von Datenquelle bis -senke folgende beispielhafte Daten- bzw. Metadatentransfers: Der Metadata Broker katalogisiert die verfügbaren Metadatenbeschreibungen der Datenangebote und stellt diese für entsprechend Berechtigte im Data Space bereit. Bei erfolgreichem Matchmaking zwischen Data Provider und Data Consumer wird anschließend vom Clearing House ein Vertrag aufgesetzt, welcher die Datennutzungsbedingungen bindend festschreibt. In einem letzten Schritt erfolgt dann schließlich erst der eigentliche Datentransfer zwischen den Connectoren von Data Provider (und der angebundenen Datenquelle) sowie Data Consumer (und der angebundenen Datensenke).

Die föderierte Data Space-Architektur erfordert ein Datenökosystemregelwerk für den Regelraum zwischen Data Provider und Data Consumer (siehe Abb. 1). Da das Ziel der Vertrauensbildung fokussiert wird, fußen Data Spaces auf gemeinsam vereinbarten Prinzipien zwischen den teilnehmenden Parteien [12]. Diese werden in einem Trust Framework formalisiert, durch einen Trust Anchor als Vertrauensinstanz verifiziert und schließlich durch ein Clearing House im laufenden Betrieb kontrolliert, sodass eine mehrdimensionale Durchsetzung des Datenökosystemregelwerks, also des Trust Frameworks, sichergestellt ist.

Ein Blick auf das Trust Framework

Während die föderierte IT-Infrastruktur eines Data Space die Grundlage für Interoperabilität, Datenverfügbarkeit und skalierbare Vernetzung bildet, ist das Trust Framework essenziell für die Etablierung von Vertrauen zwischen den teilnehmenden Akteuren. Erst im Zusammenspiel beider entsteht ein belastbares und vertrauenswürdiges Datenökosystem. So versteht man unter einem Trust Framework ein regelbasiertes Rahmenwerk, das technische, organisatorische und rechtliche Mechanismen umfasst, um sichere, souveräne und regelkonforme Datenaustausche innerhalb des Data Space zu gewährleisten.

Das Trust Framework der europäischen Initiative Gaia-X (vgl. ETR 6/2025) definiert die grundlegenden Prinzipien und An-



2: A Gaia-X Verifiable Presentation
Quelle: eigene Darstellung nach [14, 18]

forderungen für den Austausch und die Nutzung von Daten in einem föderierten Datenökosystem. Es umfasst Regeln für Sicherheit, Compliance, Identitätsmanagement, Transparenz und Interoperabilität der beteiligten Akteure zur Wahrung von Vertrauen und EU-rechtlicher Sicherheit innerhalb des Data Space. Hierfür existieren verschiedene Gaia-X-Spezifikationen: Die Anforderungen zum Aufbau eines Gaia-X-Ökosystems liefert das Architecture Document [13], die eindeutige Identifizierung, Authentifizierung und Autorisierung der Teilnehmenden beschreibt das Identity Document [14] und der sichere Datenaustausch ist im Data Exchange Document [15] spezifiziert. Das Compliance Document [16] definiert Sicherheits-, Datenschutz- und Interoperabilitätsstandards und legt fest, wie diese Standards geprüft werden. Außerdem beschreibt es die Compliance-Level durch eine Labelvergabe. Diese Labels implizieren eine Bewertung der Ziele hinsichtlich Transparenz, Autonomie, Datenschutz, Sicherheit, Interoperabilität, Portabilität, Nachhaltigkeit und europäische Kontrolle. Es folgt nun die Skizzierung der digitalen Identifizierung einer „juristische[n] oder natürliche[n] Person, die am Gaia-X-Ökosystem teilnimmt" [17]. Das dient der Verdeutlichung der zu durchlaufenden digitalen Sicherheitsprozesse. Gaia-X folgt hier den Empfehlungen der W3C zu Verifiable Credentials [14, 18], die als digitale Nachweise fungieren. Das Ziel dafür ist eine Art digitales Selbstportrait, eine s. g. Verifiable Presentation (siehe Abb. 2), in der die Verifiable Credentials der gesetzlichen Registrierungsnummer, der Bedingungen des Ökosystems und des Teilnehmers enthalten sind.

In den Gaia-X Credentials sind Teilnehmerinformationen enthalten, die zunächst als Behauptung angesehen werden. Um diese Behauptungen zu prüfen, muss für die Zukunft sichergestellt sein, dass diese unverändert und unverfälscht vorliegen. Dazu werden private und öffentliche digitale Schlüssel verwendet, die der Teilnehmer erstellen muss. Die eigene Identität hält der Teilnehmer in einem Decentralized Identifier (DID) fest, wodurch der Bezug zum öffentlichen Schlüssel hinterlegt ist. Der Teilnehmer muss zudem seine gesetzliche Registrierungsnummer vom Clearing House prüfen lassen, welches nach positiver Prüfung mit der Registrierungsnummer ein Verifiable Credential erstellt, das durch den Trust Anchor Service als Vertrauensinstanz manipulationssicher signiert wird. Dies ist die erste Datei, die in der Verifiable Presentation hinterlegt wird. Die Datei mit den Informationen zum Teilnehmer (Verifiable Credential Teilnehmer) und die der Ökosystembedingungen (Verifiable Credential Terms & Conditions) in Abbildung 2 müssen ebenfalls als Credentials in das digitale Selbstportrait einfließen. Ein Clearing House prüft das digitale Selbstportrait und jede Signatur der enthaltenen Verifiable Credentials sowie, ob die bereitgestellten Informationen konsistent sind. Geprüft wird u. a. auf die vorausgesetzten Standards und ob der Compliance-Level erreicht ist. Nach positiver Prüfung wird die Compliance vom Trust Anchor ausgestellt. Das digitale Selbstportrait und die ausgestellte Compliance zusammen weisen die Teilnehmer im Ökosystem aus. [18, 19]

Die vier Gaia-X Compliance Level definieren abgestufte Anforderungen an Teilnehmer und Dienste im Ökosystem, wobei jeweils steigende Vertrauens-, Prüf- und Nachweisniveaus durch Selbstdeklaration, technische Nachweise oder externe Auditierung erreicht und in Form von verifizierbaren Nachweisen (s. o. Verifiable Credentials) dokumentiert werden.

Erweiterung des Trust Frameworks um industriespezifische Labels (Domain Extensions)

Auf Basis dieser Vertrauenswürdigkeitslevel der Data Space-Teilnehmer lassen sich in konkreten Anwendungsdomänen spezielle domänenspezifische oder geographiespezifische Labels im Ökosystem etablieren. Diese kommende Erweiterung des Trust Frameworks wird aktuell von der Gaia-X AISBL vorangetrieben und zielt darauf ab, anwendungs- und industriebezogene Labels zu erzeugen und einem Datenangebot hinzuzufügen. Ein Beispiel aus dem Bahnsektor soll dies erläutern.

Ein Anwendungsfall im Bahnbetrieb ist, dass ein Eisenbahnverkehrsunternehmen (EVU) mit dem Betrieb seiner Züge stetig Strecken befährt und dabei ein Sensorsetup mit sich führt, das Continous Track Monitoring (CTM) unterstützt [20]. Dabei sammelt der Zug Daten über den Streckenzustand, um eine Gleislagebewertung gemäß der dafür gültigen Norm EN13848 durchzuführen [21]. Diese Daten liegen anschließend beim EVU, könnten jedoch für einen Netzbetreiber potenziell interessant sein, um dessen Instandhaltung zu unterstützen. In diesem Anwendungsfall bieten Data Spaces nun folgende Möglichkeit: Das EVU kann dem Netzbetreiber die gesammelten Daten als Datenprodukt anbieten.

Dazu fügt das EVU dem Datenangebot zusätzliche industriespezifische Labels hinzu, wie z. B., dass der Datensatz der Industrienorm EN13848 entspricht. Ein Infrastrukturbetreiber würde entscheiden können, ob die Daten zur Unterstützung der Instandhaltung eingekauft werden sollen. Die Entscheidungsfindung wird dadurch unterstützt, dass die Daten einer vertrauenswürdigen Quelle entstammen (s. o. generelle Labels), aber auch, dass Semantik und Syntax bekannten Industriestandards entsprechen, was durch Domain Extensions der Bahndomäne erreicht werden kann.

Auf diese Weise werden Situationen erzeugt, in denen für das EVU ein sekundärer Revenue Stream etabliert werden kann, der zusätzlich zum operativen Streckenbetrieb die Wirtschaftlichkeit erhöht. Weiterhin kann der Infrastrukturbetreiber entscheiden, ob der Preis zum Erwerb der Daten wirtschaftlich ist und gegenüber dem erwarteten Nutzen in einem positiven Verhältnis steht. Die Antizipation dieses Nutzens fällt indes leichter, wenn bereits im Datenangebot klar erkennbar ist, dass dieses vertrauenswürdig und standardkonform ist. Ein weiterer Aspekt der domänenspezifischen Labels ist die Möglichkeit, diese für die Suche innerhalb des Data Spaces zu nutzen. Infrastrukturbetreiber könnten gezielt nach Datenangeboten suchen, die das benötigte Label aufweisen. Die Realisierung dieser Domain Extension für den Bahnbereich bedarf allerdings einer Instanz, die die Konformität des angebotenen Datensatzes mit dem angegebenen Label überprüft und ausweist. Hierzu eignen sich

neutrale Stellen wie akademische Partner, Verbände oder andere vertrauenswürdige Dritte.

Fazit

Data Spaces bieten durch die föderierte Infrastruktur Datensouveränität und Autonomie für teilnehmende Parteien. Die vorgestellten technologischen Grundlagen, zentralen Komponenten und das Trust Framework bilden dabei die Basis für vertrauenswürdige Kooperationen. Der Anwendungsfall aus dem Bahnsektor verdeutlicht, wie domänenspezifische Labels die Qualität von Datenprodukten steigern und die Interoperabilität innerhalb des Sektors fördern können.

Literatur

[1] European Commission, "Common European Data Spaces". [Online]. Verfügbar: https://digitalstrategy.ec.europa.eu/en/policies/data-spaces. [Zugriff: 20.06.2025].
[2] B. Otto, A. Burmann, „Europäische Dateninfrastrukturen", Informatik Spektrum, Vol. 44, No. 4, pp. 283 - 291, 2021.
[3] N. Jahnke, I. Jussen-Lengersdorf, T. Schoormann, F. Möller, „Designing Federated Data Marketplaces in Industrial Production: Findings from a Prototypical Implementation", 19th International Conference on Wirtschaftsinformatik, Würzburg, Germany, 2024.
[4] F. Möller, I. Jussen-Lengersdorf, V. Springer, A. Gieß, J. Schweihoff, J. Gelhaar, T. Guggenberger, und B. Otto, „Industrial data ecosystems and data spaces," Electronic Markets, Vol. 34, pp. 1 - 17, 2024.
[5] D. Heimbigner, D. McLeod, „A federated architecture for information management", ACM Trans. Inf. Syst., Vol. 3, No. 3, pp. 253 - 278, July 1985.
[6] R. Abraham, J. Schneider, J. vom Brocke, „A taxonomy of data governance decision domains in data marketplaces", Electronic Markets, Vol. 33, No. 1, 2023.
[7] N. Duan, „Design Principles of a Federated Service-oriented Architecture Model for Net-centric Data Sharing", The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, Vol. 6, No. 4, pp. 165 - 176, 2009.
[8] C. A. Lee, R. B. Bohn, M. Michel, „The NIST Cloud Federation Reference Architecture", NIST, Gaithersburg, 2020.
[9] International Data Spaces Association, „IDS Reference Architecture Model (RAM) 4.0," GitHub repository, 2024. [Online]. Verfügbar: https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0. [Zugriff: 20.06.2025].
[10] G. Giussani, S. Steinbuss, N. Gras und T. Prasse, „Data Connector Report, No. 16 (September 2024)", International Data Spaces Association, Tech. Rep., Sept. 25, 2024. [Online]. Verfügbar: https://internationaldataspaces.org/data-connector-report/ [Zugriff: 20.06.2025].
[11] A. Gieß, T. Schoormann, F. Möller, I. Gür, „Discovering data spaces: A classification of design options", Computers in Industry, Vol. 164, 2025.
[12] L. Nagel, D. Lycklama, U. Ahle, H. Bastiaansen, K. Bengtsson, M. Caballero, S. Castellvi, F. Ette, M. Faraldi, J. Gelhaar, A. Graziani, A. Grguric, S. Gusmeroli, K. Helmholt, J. Hierro, D. Hoppenbrouwer, T. Hülsmann, S. Krco und J. Valiño, „Design Principles for Data Spaces – Position Paper", 2021.
[13] Gaia-X European Association for Data and Cloud, "Gaia-X Architecture Document", Version 25.05, 2025.
[14] Gaia-X European Association for Data and Cloud, "Gaia-X Identity Document", Version 24.07, 2024.
[15] Gaia-X European Association for Data and Cloud, „Gaia-X Exchange Document", Version 23.11, 2025.
[16] Gaia-X European Association for Data and Cloud, „Gaia-X Compliance Document", Version 25.03, 2025.
[17] Gaia-X Hub Germany, „Vertrauen ist kein Zufall – Wie vier Dokumente das Fundament digitaler Souveränität bilden," 2025. [Online]. Verfügbar: https://gaia-x-hub.de/gx-essentials/vertrauen-ist-kein-zufall-wie-vier-dokumente-das-fundament-digitaler-souveraenitaet-bilden. [Zugriff: 20.06.2025].
[18] W3C, „Verifiable Credentials Data Model 2.0", W3C Recommendation, 19 Apr. 2023. [Online]. Verfügbar: https://www.w3.org/TR/vc-data-model-2.0/. [Zugriff: 20.06.2025].
[19] Gaia-X, „Gaia-X Framework", 2025. [Online]. Verfügbar: https://docs.gaia-x.eu/#/framework. [Zugriff: 20.06.2025].
[20] K. U. Wolter, L. Huang, F. Löffler, J. Heland, G. Baumann, „Continuous Track Monitoring – neue Wege der Instandhaltung", EI - Der Eisenbahningenieur, 12/2024.
[21] Bahnanwendungen – Oberbau – Gleislagequalität – Teil 1: Beschreibung der Gleisgeometrie; Deutsche Fassung EN 13848-1:2019.

Summary

Federated Data Spaces for the railway sector - a technical insight

Data spaces provide a digital infrastructure for secure data exchange in compliance with EU law. This article presents the technological foundations of federated data spaces, the central components and the trust framework. A use case from the rail sector also shows how domain-specific labels can promote data product quality and interoperability.