Views on data spaces in industry

Sino-German Intelligent Manufacturing / Industrie 4.0 Standardisation Sub-Working Group

IMS 智能制造标准化 国家智能制造标准化总体组 National Intelligent Manufacturing Standardisation Administration Group

STANDARDIZATION COUNCIL INDUSTRIE 4.0

GPQI Dialogues for Innovation and Trade

Published by:
Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Registered offices
Bonn and Eschborn, Germany

Global Project Quality Infrastructure

Tayuan Diplomatic Office Building
No.14, Liangmahe Nanlu, chaoyang District
100600 Beijing, PR China
E info@gpqi.org I www.gpqi.org

Design:
Oliver Hick-Schulz

Photo credits:
yucelyilmaz/iStock

On behalf of
German Federal Ministry for Economic Affairs and Climate Action (BMWK)
Berlin, Germany 2023
Beijing, China 2023

Text:
Standardization Council Industrie 4.0
German Commission for Electrical, Electronic & Information
Technologies of DIN and VDE
60596 Frankfurt am Main

National Intelligent Manufacturing
Standardisation Administration Group
China Electronics Standardization Institute,
No.1 Andingmen East Street,
Dongcheng District, Beijing, 100007, China

Authors/Experts
Prof. Dr Ulrich Löwen — Siemens AG, Technology, Erlangen
Yuhang Cheng — CESI, Beijing
Dr Sha Wei — CAICT, Beijing
Fan Gao — CAICT, Beijing
Weiran Guan — CAICT, Beijing
Zhiman Chen — CRRC ZIC, Zhuzhou
Detlef Tenhagen — HARTING Stiftung & Co. KG, Espelkamp
Dr Sicco Lehmann-Brauns — Siemens AG, München
Fabian Hartner — Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen

The German Federal Ministry for Economic Affairs and Climate Action (BMWK) has commissioned the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH to implement the Global Project Quality Infrastructure (GPQI).

Implemented by:

Federal Ministry
for Economic Affairs
and Climate Action

Deutsche Gesellschaft
für Internationale
Zusammenarbeit (GIZ) GmbH

This publication is a result of the Sub-Working Group Industrie 4.0/Smart Manufacturing of the Sino-German Standardisation Cooperation Commission in cooperation with

中华人民共和国工业和信息化部

中国国家标准化管理委员会
Standardization Administration of the P.R.C.

Contributors

NATIONAL INTELLIGENT MANUFACTURING STANDARDISATION ADMINISTRATION GROUP

The National Intelligent Manufacturing Standardisation Administration Group (IMSG) was established to promote and accelerate the progress of intelligent manufacturing in China under the leadership of the Standardisation Administration of China (SAC) and Ministry of Industry and Information Technology (MIIT). It is responsible for carrying out practical work on intelligent manufacturing standardisation, including participation in international standard-setting on intelligent manufacturing as well as organising exchange and cooperation on international standards.

STANDARDIZATION COUNCIL INDUSTRIE 4.0

The Standardization Council Industrie 4.0 (SCI 4.0) was founded at the Hannover Messe 2016 as a German standardisation hub by Bitkom, DIN, DKE, VDMA and ZVEI. The initiative aims to initiate standards for digital production and to coordinate these standards nationally and internationally. SCI 4.0 orchestrates implementation of the standardisation strategy of the German Platform Industrie 4.0, which includes coordination with standardisation organisations (SDOs) and international partners as well as interlocking with pilot projects. The aim of this coordinated approach is to ensure that standards exploiting the potential of Industrie 4.0 are developed in a coordinated manner. SCI 4.0 is supported by DKE and the German Federal Ministry for Economic Affairs and Climate Action (BMWK).

GLOBAL PROJECT QUALITY INFRASTRUCTURE

The German Federal Ministry for Economic Affairs and Climate Action (BMWK) established the Global Project Quality Infrastructure (GPQI) to promote the development of well-functioning and internationally coherent quality infrastructures. GPQI supports political and technical dialogue and implements bilaterally agreed activities in collaboration with all relevant stakeholders. The project aims to reduce technical barriers to trade and enhance product safety through bilateral political and technical dialogue on quality infrastructure (QI) with some of Germany's key trading partners.

Contents

Introduction

Background

This report is the result of the cooperation line Use Cases and Applications in the context of the Sino-German Sub-Working Group Industrie 4.0/ Intelligent Manufacturing. The overall objective of the sub-working group is to discuss standardisation aspects in the focus of Germany and China, in order to intensify and broaden Sino-German cooperation by defining concrete issues and steps to be taken.

The objective of the cooperation line Use Cases and Applications is to analyse business strategies and customer needs in the manufacturing industries on the basis of specific customer projects. The findings will be compiled to create descriptions – so-called 'use cases' – based on proven best practices, e.g. the Industrial Internet Reference Architecture (IIRA), see [1]. These use cases facilitate a common understanding of markets, trends, drivers, concepts and solutions, which can then serve as a basis to articulate requirements for standardisation aspects.

Common understanding of use cases

The term use case has a diverse range of meanings. Its principal uses are as follows, see [2]:

- A use case in the sense of a business scenario, which describes business relationships within a value-added network in line with business model logic (e.g. a business canvas).

- A use case in the sense of a general understanding of how a technical system is viewed in the context of its application. This is used to describe the interaction of a technical system involving actors (such as other technical systems or humans). This is the understanding of the term use case in the cooperation line Use Cases and Applications.

- Use cases in the sense of specific projects.

The cooperation line Use Cases and Applications initially focused on (technical) use cases that describe how a technical system is viewed in the context of its application. It subsequently decided to deal also with use cases in the sense of business scenarios.

Objectives of this document

The topic of data spaces is complex and currently under discussion – with differing objectives – in various communities. This document adopts the following position on data spaces in order to lend structure to the topic:

- business viewpoint
- usage viewpoint
- functional viewpoint

This approach is in line with the guidance offered by the cooperation line Use Cases and Applications, see [2]. In this way, the various discussions are systematically integrated into a generally accepted framework to contribute to a better overall understanding.

Data spaces

General

Data spaces are discussed in the context of the digital economy, where an overarching goal is to generate additional value based on data, while maintaining and guaranteeing data sovereignty, data security and data integrity for all stakeholders involved. Data spaces offer participants common, trustworthy and controllable transaction spaces, through which data can be provided and jointly evaluated or managed. This creates a generally accepted framework for data exchange and data use. Implementation can be based on a centralised or decentralised infrastructure.

A data space defines technical, legal and business principles. There are already numerous examples in industry where companies create additional value based on data. A common challenge in all these examples is often to improve the economic scalability of these examples. Data spaces here create the potential that by offering common technical, legal, and business principles, specific added value in the individual examples can be provided uniformly based on a data space, and specific examples can thus be more easily transferred to other applications. In addition, data spaces generate the potential to open up new applications.

However, the specific requirements of data sovereignty, data security and data integrity depend very much on the application, so that in addition to the technical characteristics of a data space, one must always consider the associated cost-benefit ratio in each case. For this reason, in addition to offering appropriate features, a data space must provide an economic balance between market requirements and the regulatory framework.

Methodology

The cooperation line Use Cases and Applications has compiled and structured various activities and results in Germany and China on the subject of data spaces. Assessment of the cooperation line Use Cases and Applications shows that although many key activities were taken into consideration, the cooperation line Use Cases and Applications does not claim to be complete.

Structuring was based on the three viewpoints – business, usage and functional – proposed by the Industrial Internet Reference Architecture (IIRA), see [1] and Figure 1.

- Business viewpoint: the conceptual approach is fully aligned with the methodology proposed by the Digital Business Models working group of Plattform Industrie 4.0 to analyse practical examples, see [3], and examples for business scenarios in manufacturing industry, see [4].

- Usage viewpoint: the conceptual approach is fully aligned with the methodology proposed by IEC TC65 WG23 TF SM Use Cases, see [5], as well as with the use cases prepared by the cooperation line Use Cases and Applications, see [6] and [7].

- Functional viewpoint: the description of a functional view is usually strongly influenced by the specific application. In this respect, this document adopts a relatively broad understanding of the functional viewpoint, in comparison to the business and usage viewpoints. The key point in this document is to distinguish the functional viewpoint conceptually from the business and usage viewpoints.

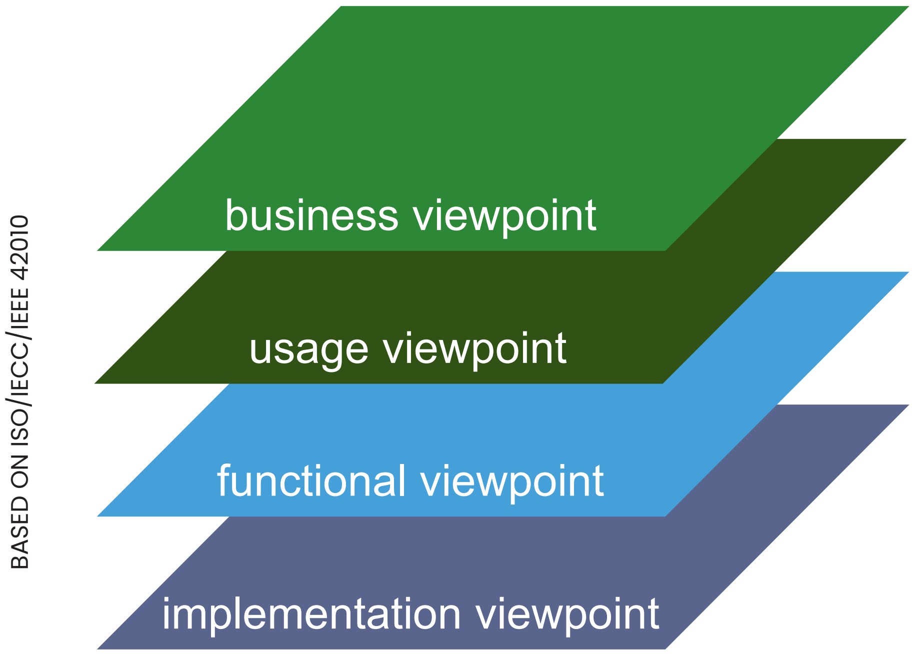

Figure 1: Industrial Internet Reference Architecture (IIRA), see [1]

Business view considerations

General

In the past, numerous business scenarios have been analysed and described internationally in various associations, initiatives and communities. The approach described in [3] was used as the basis for the methodology. Most of these business scenarios were triggered by the objective that additional added value was or should be generated based on data. In this respect, these examples contribute to the overarching objective of data rooms.

In this context, special mention must be given to the study Data Sharing Models in the Electro and Digital Industry by ZVEI, see [9], in which various concepts of data sharing were clustered into five data sharing patterns.

The following subchapters describe selected examples of data spaces from a business perspective.

Example: Aspern Data Circle ECOM-X

Basic business idea

Thesis 1: Energy communities will increasingly establish themselves in the market in the years ahead.

Energy communities consist of members who negotiate and balance the energy needs and any excess energy among individual members. The basic benefit for members of an energy community is that they pay less for energy purchased within the energy community than on the external market and receive more money for energy supplied to the energy community than to the external market.

Since the energy market is regulated, energy communities are subject to various national regulations. It is assumed, however, that increasingly the boundary conditions for energy communities will be opened, so that different energy communities with different goals, offers and value propositions will establish themselves in the market, as well as probably in various legal setups.

These energy communities must solve the following challenges in particular:

- energy communities must market their specific offer and therefore need access to potential members;
- energy communities need access to their members' consumption, forecast and additional information as a basis for their offering to community members.

Thesis 2: In the context of energy communities, there is an opportunity for a data space operator to establish itself in the market.

Such a data space operator offers the following value propositions:

- It enables potential members of an energy community to publish their requirements regarding energy consumption and their offers regarding energy generation, while safeguarding their own interests. These interests may include usage policies for data with respect to published requirements and energy generation offers. For example, a forecast of factory utilisation could be derived from the forecast of energy consumption and the factory operator has a vested interest in sharing such data with others only under clearly defined contractual conditions. The provider mediated via the energy community could be a competitor who does not currently have sufficient capacity utilisation in its factory.

- It enables energy communities to access the requirements and offers of potential members, while protecting their interests with respect to usage policies for the data provided.

- It offers a legal framework and technical services that enable members of an energy community to share their specific energy consumption and forecasts within the energy community.

There are different ways in which such a data space operator could monetise the services offered, for example:

- Energy communities pay to access the published requirements for energy consumption and offers for energy generation. The publication of such requirements and offers by potential members of an energy community could be free of charge.

- The technical services used to process business transactions between energy producers and consumers within the energy community are monetised, for example based on subscription or usage. Whether there is a contractual relationship between the data space operator and individual members of the energy community or with the energy community alone depends on the business model of the energy community.

It is conceivable that both private households and companies can participate in an energy community. This mix creates further challenges, for example in terms of data protection and energy consumption or taxation.

Value network perspective

Figure 2 illustrates the current value network with no energy community and with two energy prosumers, i.e. entities that can act both as energy provider and energy consumer. Note that in some countries the operator of the electrical grid and the operator of the metering infrastructure are assumed by the same legal entity.

Figure 3 illustrates the future value network with an energy community. In addition, it shows an operator of an IT infrastructure, which provides the technical basis on which the operator of the data space executes its business model.

The operator of the data space is set up as a platform company. As an intermediary, this platform company connects the energy prosumer on the one hand with the energy communities on the other via a digital platform which facilitates their interaction. Thus, the operator of the data space creates a digital infrastructure and orchestrates an ecosystem for innovation, and – depending on its data sovereignty policy – strengthens the digital sovereignty of ecosystem participants.

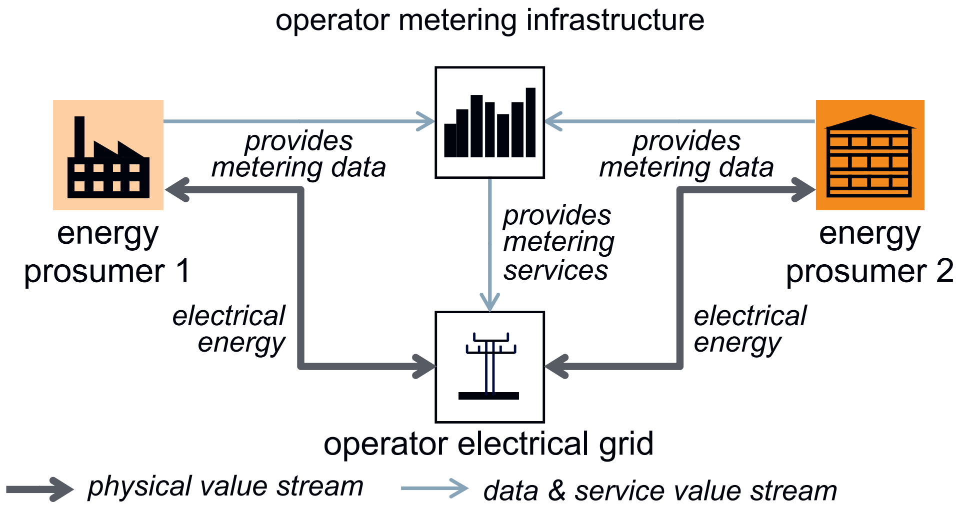

Figure 2: Current value network with no energy community

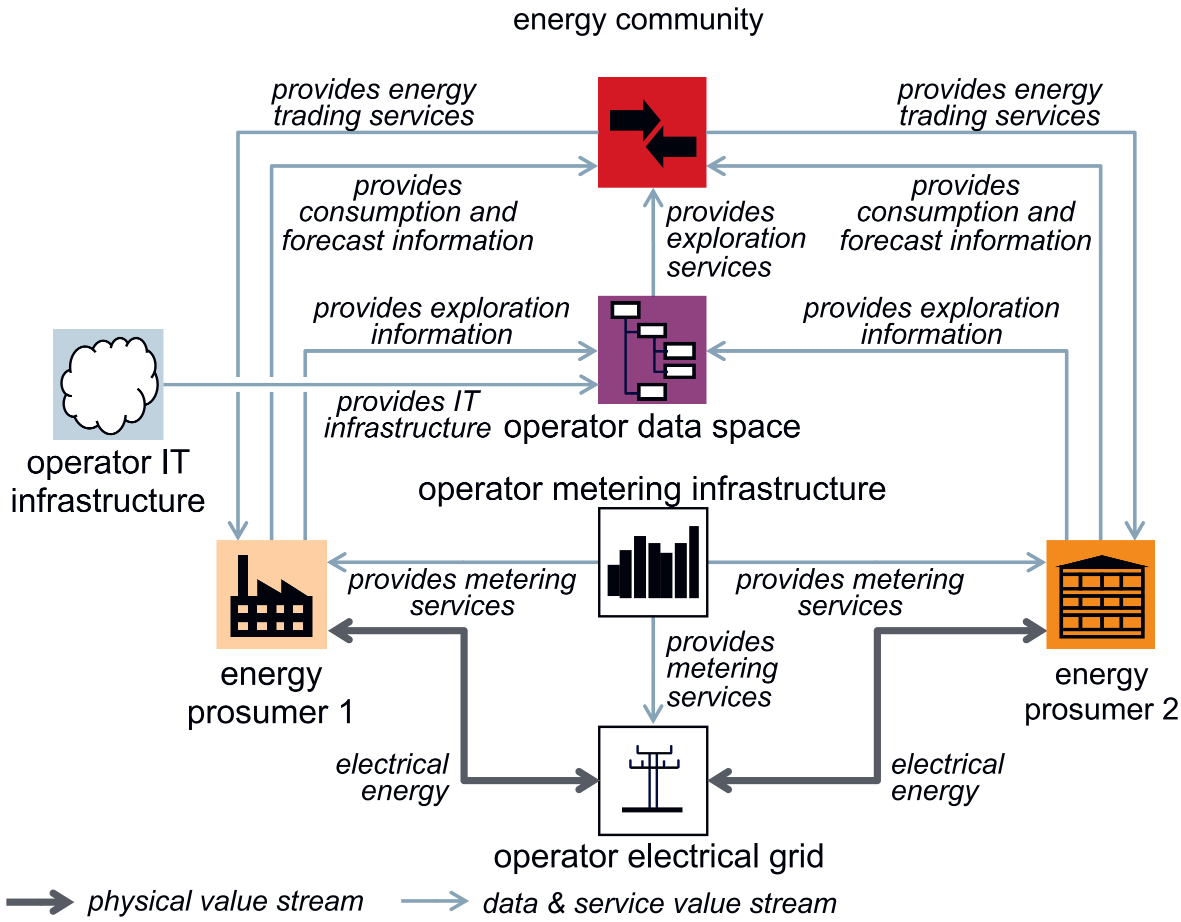

Figure 3: Future value network with an energy community

Example of an energy community

An energy community can be industrial and residential and benefit from the fact that excess energy from industry can be used by residents at weekends, while excess energy from residents can be used by industry during the week, see **Figure 4**.

In this example, an energy community of this type enables mutual benefits for industry and residents by using excess energy generation. This can happen thanks to increased sustainability resulting from locally generated green energy and cost savings by energy community trading.

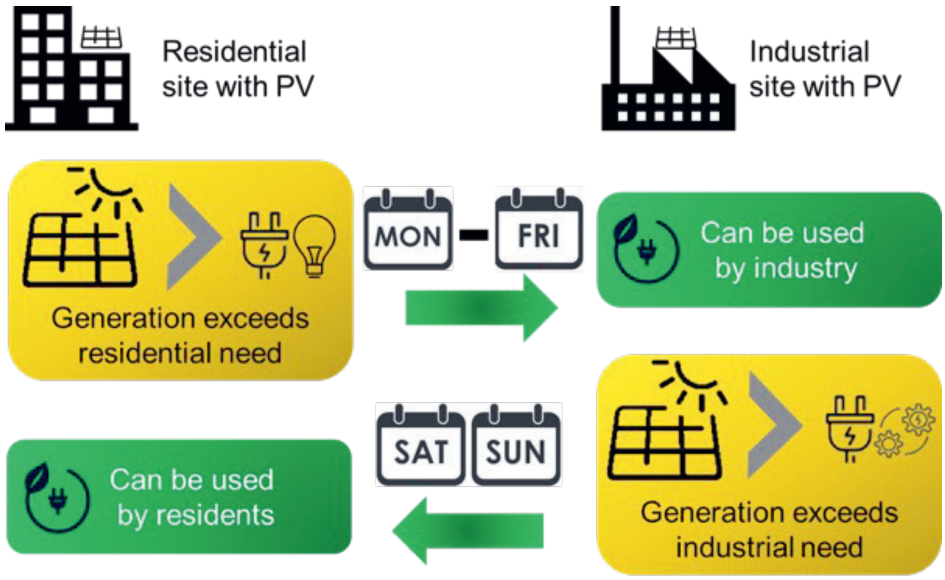

Figure 4: Example of an energy community with industrial and residential sites (Source: Siemens)

Example: Declaration of the product carbon footprint

Business requirements

In future, regulatory requirements will require a manufacturing company to declare a product carbon footprint (PCF) for its products. To calculate this PCF, a manufacturing company must determine the various contributions to the PCF, then add them up and make this specific value available to the customer:

- The company must determine its volume of direct emissions and any resources, such as energy, that are required by the company to manufacture the specific product.
- In addition, it must determine any indirect emissions arising from purchased components. Databases providing any such secondary data can also be taken into consideration. However, secondary data only reflect the facts imprecisely. For this reason, when suppliers make PCFs of supplied components available to their customers, primary data should be used.

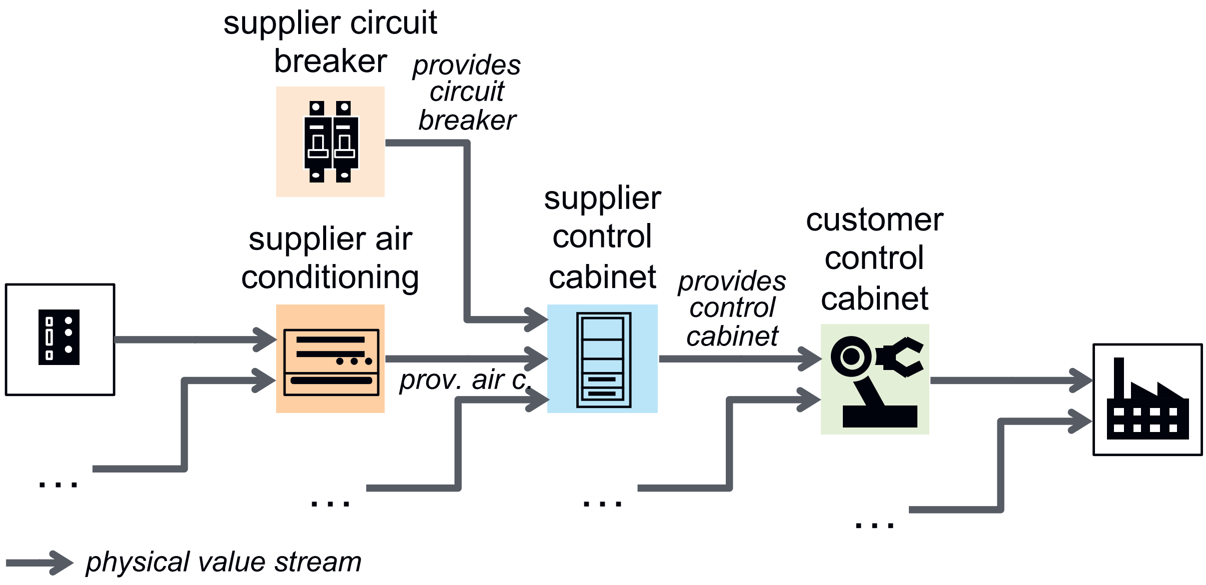

Figure 5: Illustration of a supply chain in manufacturing industries (Source: IEC 63278-4 CD)

Value network perspective

From a business perspective, the companies under consideration are organised in an overall supply chain, see **Figure 5** for illustration. The figure shows an example of a section from such a supply chain. Each of these companies declares the PCF for its products and shares this information with customers within the data space under consideration.

In addition, other business stakeholders must be taken into consideration:

- A software application provider offers software for managing PCFs. On the one hand, this software supports a manufacturing company in determining its direct emissions. This includes, among other things, how to allocate to specific products the overall emissions and resource consumption of a manufacturing company recorded, for example, in the company's Manufacturing Execution System. Furthermore, this software enables PCFs declared by the suppliers to be made available in a trustworthy manner. It also means the manufacturing company can make its own declared PCFs available to its customers in a trustworthy manner. In contrast to product properties such as weight, dimensions and performance, which a customer can measure and check independently, this cannot be done by the customer using the PCF. It is essential that the trust of a customer in a product supplier is taken into consideration when designing the data space.

- The additional business role of a trustworthiness services provider should also be considered, in order to address the core issue of trustworthiness. The core value proposition of a trustworthiness services provider is to manage information from various sources in such a way that it can be verifiably passed on to others in an unaltered manner.

- Finally, an independent certifier who certifies the processes of a manufacturing company and ensures that the PCF is calculated in accordance with the regulatory framework.

Figure 6 shows the value network for the business scenario 'declaration of product carbon footprint'.

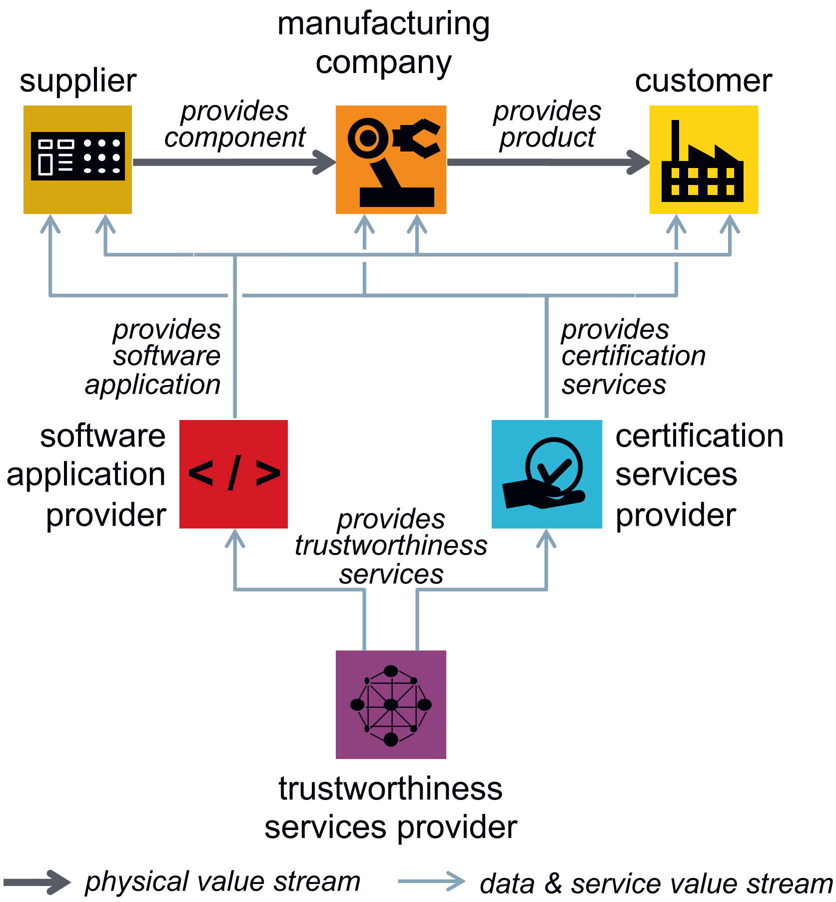

Figure 6: Value network for declaration of product carbon footprint

Example of data space for product carbon footprint declaration

Figure 7 shows how Siemens is working with partners to develop a data space for product carbon footprint.

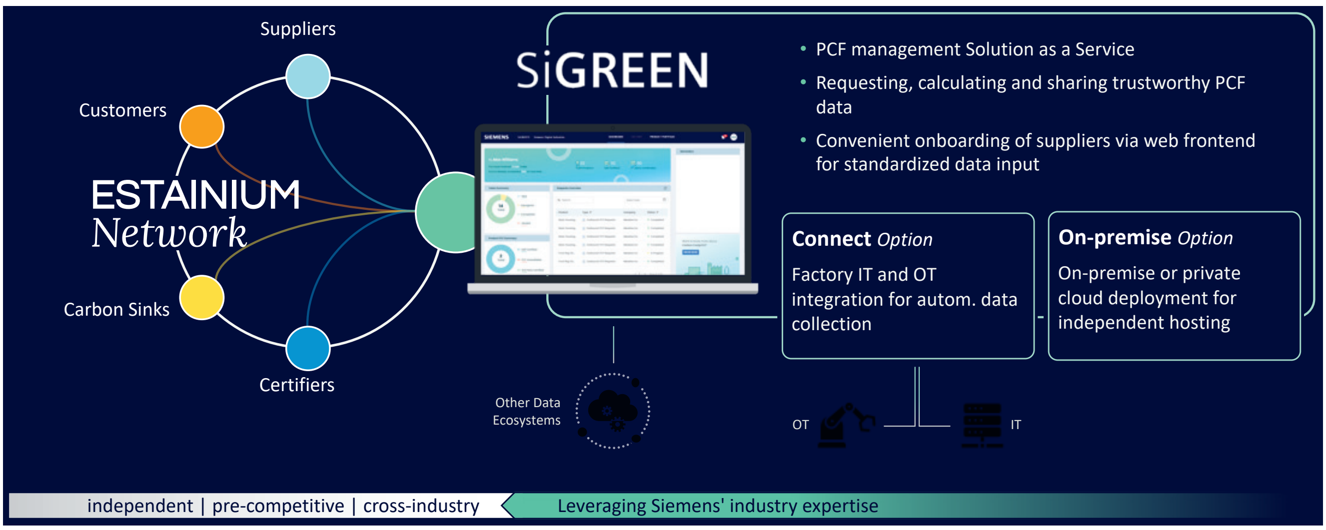

Figure 7: Data space for product carbon footprint declaration developed by Siemens (Source: Siemens)

To design the data space described, Siemens is working with partners to develop the following two portfolio elements:

- The software application SiGREEN, a PCF management solution offered as a service for requesting, calculating and sharing trustworthy PCF data. On the one hand, SiGREEN is offered as an on-premise option, in particular with the aim that customers can seamlessly integrate the software application into their existing IT infrastructure, such as Manufacturing Execution Systems and Product Lifecycle Management systems. This is typically a requirement of companies that offer a complex product portfolio. On the other hand, SiGREEN is also offered as a connect option. In this case, onboarding and operation of the software application are much easier. This is typically suitable for companies that offer a simple product portfolio, but it also serves as a low-effort entry offering to keep the hurdle of participation in the data space as low as possible with the option of successive expansion over time.

- Together with other companies, Siemens is involved in development of the ESTAINIUM network, which is organised cross-industry and pre-competitively. Among other things, the ESTAINIUM network lays the basis for trustworthiness services and addresses the challenge of standardising the PCF calculation and declaration.

Example: Model training for automatic quality inspection

Business requirements

Sometimes different manufacturing companies produce the same type of components for various assembling companies. Before providing the components to the assembling company, the manufacturing company examines the quality of the manufactured components to avoid sending defective components to the customer. Currently, most manufacturing companies hire employees to conduct selective inspections manually. However, not all components will be inspected – and this also creates increased cost. To improve this current practice, several manufacturing companies provide pictures of their components to an algorithm training platform to build automatic quality inspection models. The manufacturing companies use these quality inspection models to examine the manufactured components with the aim of avoiding the supply of defective components.

Value network perspective

From a business perspective, each manufacturing company provides pictures of the same type of components to the algorithm training platform. To train models for automatic inspection, the algorithm training platform requests pictures from more than one manufacturing company. The algorithm training platform also requires data from other databases and is obliged to store data in its own storage place, so that data usage can be governed by the data provider. When models are being trained, the algorithm training platform loads data from its own storage place, sets them into algorithms, then generates the data models accordingly. After model training, the algorithm training platform provides services for automatic inspection to all manufacturing companies.

Figure 8 illustrates model training for automatic quality inspection.

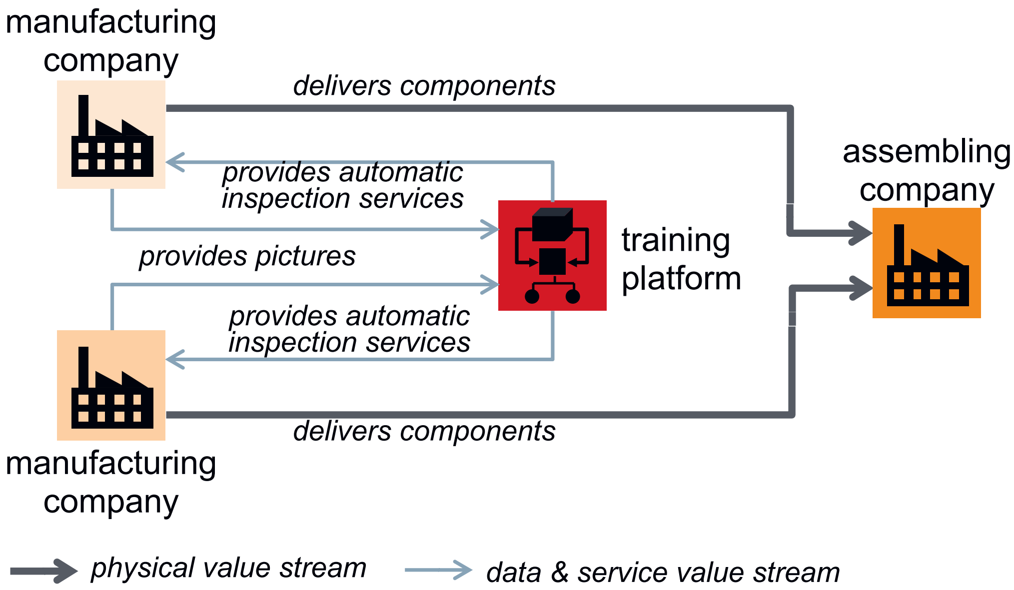

Figure 8: Illustration of automatic quality inspection services

In addition, other business stakeholders must be taken into consideration:

- A trusted data matrix client provider provides a 'trusted data matrix client' software application to ensure digital contracts are executed during data sharing. Based on consensus-made requirements, agreements and management tools, the trusted data matrix is a data and resource sharing digital infrastructure to facilitate trusted, secure, transparent and accountable the sharing, exchange, circulation and trading of data among different business stakeholders. The trusted data matrix is formed by multiple trusted data matrix clients and other third-party service providers. Using the trusted data matrix client from the same trusted data matrix client provider guarantees interoperability between data providers and data consumers.

- Service providers can offer many kinds of services, for example:

  - Data catalog providers offer a data catalog for data-providing entities to upload their metadata in a trusted data matrix. In this example, the training platform can search metadata from the data catalog of a trusted data matrix and use this metadata to decide which data to choose.

  - Brokerage service providers help data consumers and data providers to deal with contract issues. In this example, the training platform and data provider companies can negotiate a contract through a brokerage services provider. Having reached an agreement, the brokerage services provider issues digital contracts to the training platform and data-providing companies to be applied by the trusted data matrix.

  - Identity management service providers provide identities for all companies participating in a trusted data matrix.

  - Logging and recording service providers provide services to continually log actions. All activities such as login, uploading metadata, looking through metadata, searching for metadata, reading metadata, signing a digital contract and using data can be logged and saved.

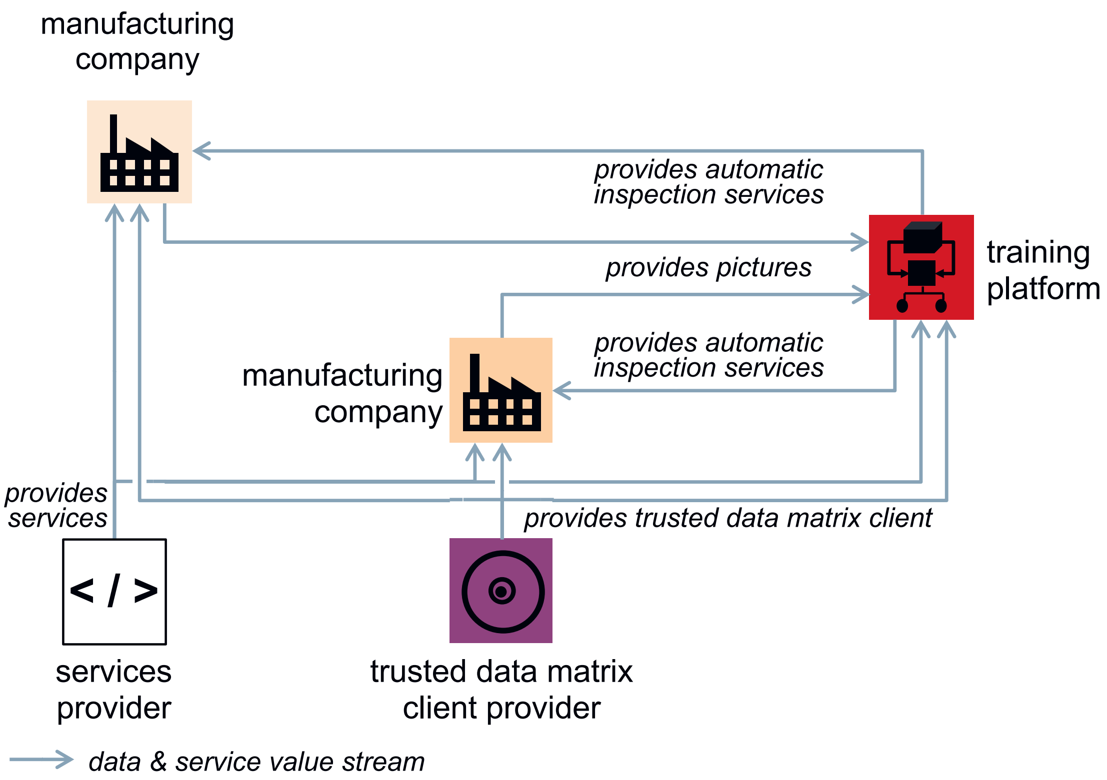

Figure 9: Trusted Data Matrix for algorithm training for automatic quality inspection with multiple data sources

Example: Data sharing of digital printing images in the textile industry

Business requirements

In home-used textiles, a textile company (data provider) sends a digital printing image (data asset) to a print factory (data consumer). The print factory (data consumer) will then use the digital printing image to print the textile product. The companies involved are generally small scale, have a low level of digitalisation and are labour intensive. With regard to data exchange characteristics, there is generally no data asset management in this industry. Once a company transmits image data to a print factory, there is always the possibility of a data leak. The traditional method for sharing an image is manual transmission by an employee of the textile company; this person supervises data usage. However, this not only increases costs, it also heightens the risk that data assets can be copied and transmitted second-hand to unauthorised data consumers. The trusted transmission and compliant use of images are still pain points for data providers.

In this use case:

- The textile company (data provider) sends the digital printing image (data asset) to the print factory (data consumer). The transmission of image data must be trustworthy and secured.

- The digital print factory will receive the digital printing image and store it as a file in local factory terminals (usually work PCs). The data storage phase must safeguard the integrity and confidentiality of the data asset.

- In the next step, a print factory employee uses professional digital printing software (e.g. Official Adobe Photoshop) to open the digital image file. After checking the image for correctness, the employee uses the digital printing software, which is linked to the factory's digital printing equipment. During data usage, there must be control of the scope of data circulation for digital printing images. There must

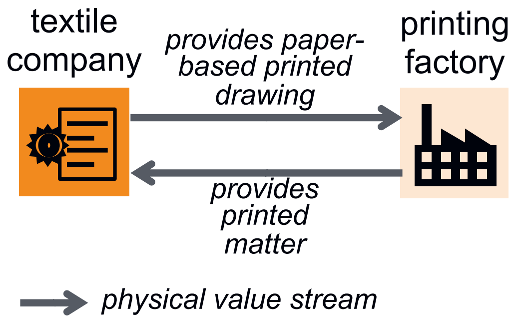

Figure 10: Illustration of supply chain in textile industry

be no transmission of copied files outside the print factory.

- After printing, the printed product is sent back to the textile company. In this final stage, the digital printing image and all copies must be destroyed. Any possibility of data leakage must be eliminated.

Value network perspective

From a business perspective, the companies in question are organised in an overall supply chain, see **Figure 10** for illustration. This figure highlights the physical value stream for the example under consideration, which includes a textile company (data provider) and a digital printing factory (data consumer)

In addition, other business stakeholders must be taken into consideration:

- A trusted data matrix client provider provides a 'trusted data matrix client' software application to guarantee the conditions specified by a digital contract on data sharing. Based on consensus-made requirements, agreements and management tools, the trusted data matrix is a data and resource-sharing digital infrastructure to facilitate the trusted, secure, transparent and accountable sharing, exchange, circulation and trading of data among different business stakeholders. The trusted data matrix is formed by multiple trusted data matrix clients and other third party service providers. Using the trusted data matrix client from the same trusted data matrix client provider guarantees interoperability between data providers and data consumers.

- A brokerage services provider provides services to negotiate a digital contract specifying the conditions under which data provided by a data provider may be used by a data consumer.

- An identification and authentication services provider provides services to ensure that a system has credible user and equipment identities, for example a user enters a username, password or other identity factors to log into the system. The identification and authentication services provider can also provide services to identify data assets and verify identification in their circulation process, providing added value for secure data management.

- A logging and recording services provider offers services to collect and log metadata about data transmission, data storage, data usage and data destruction stages. This data logging serves as proof of legally compliant use for the print factory.

Figure 11 shows the value network for the business scenario 'Data sharing of digital printing images in textile industry'.

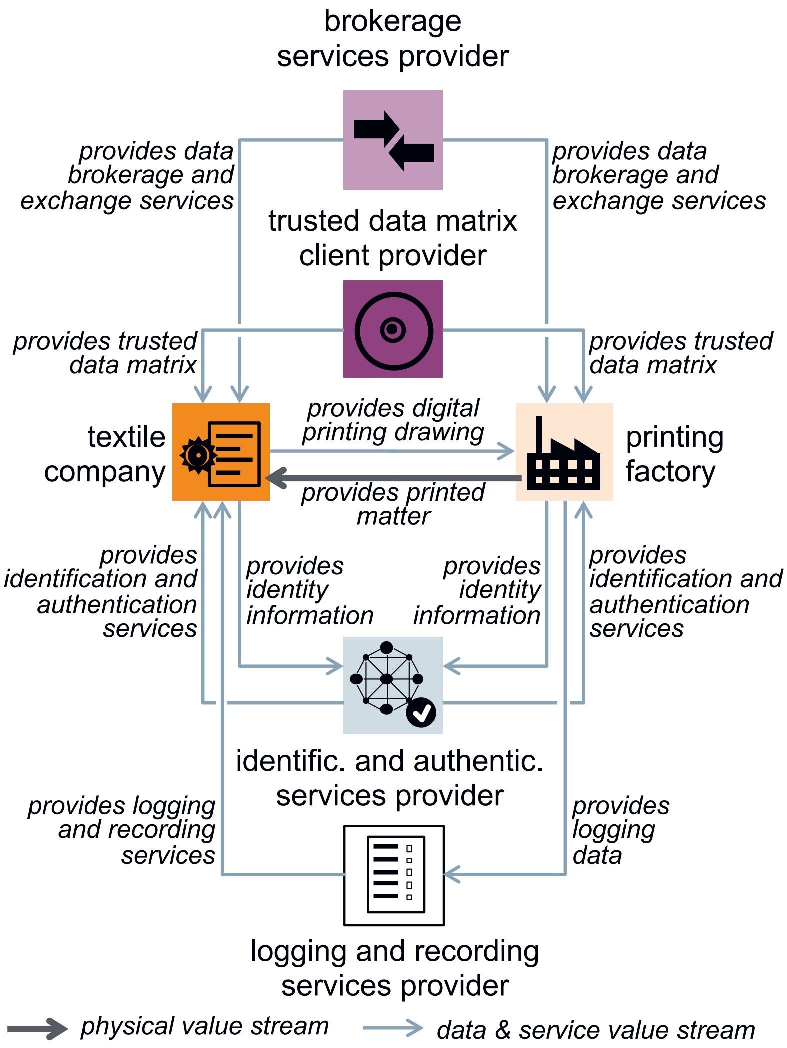

Figure 11: Value network for data sharing of digital printing images in textile industry

Usage view considerations

General

Use cases can be described at different levels of detail. The most comprehensive use case collection with regards to Industrie 4.0/Intelligent Manufacturing is the one prepared by the IEC TC65 WG23 task force on Smart Manufacturing use cases, see [5]. This collection is characterised by the fact that, on the one hand, it has international recognition as a technical report from the IEC and, on the other, that use cases are described based on a common terminology and a comparable level of detail.

Following public debate on the topic of data spaces in recent years, this use case collection [5] was analysed to establish whether further use cases should be included to better address the topic of data spaces. Several additional use cases have been identified and these are to be included as supplementary use cases in a planned revision of [5].

The following two subchapters outline two such use cases. These will be described in greater detail in the planned revision of [5]. Note that Data Brokerage describes a generalised usage view for the business scenario Aspern Data Circle ECOM-X, and Product Carbon Footprint Calculation describes a generalised usage view for the business scenario Declaration of the Product Carbon Footprint.

In the following two subchapters, two additional use cases will be added: 'Trusted model training with multiple data sources' and 'Trusted data sharing of digital printing drawings'. Both involve trusted data sharing between entities, utilising the trusted data matrix clients, logging services, identity services and metadata catalogue and brokerage services for data protection, monitoring and safeguarding of commercial interests and intellectual property.

Example: Data brokerage

Objective

A company acting as a data provider wants to provide own data to a company acting as data consumer in order to benefit in return, for example by optimising internal value-added processes or developing and providing additional services to its own customers.

A company acting as a data consumer wants to consume data from a company acting as data provider in order to benefit in return, for example by optimising internal value-added processes or developing and providing additional services to its own customers.

Overview

This use case considers two companies that want to exchange data with each other. The use case postulates another business role in the form of a data space operator, who enables and supports the data exchange between data provider and data consumer through appropriate services. The data space operator offers the following brokerage services:

- The data provider can use services provided by the data space operator to expose what data is made available to others and under what terms of use. The data provider will not expose the specific payload, only metadata describing the kind of data provided.

- The data consumer can use services provided by the data space operator to explore the metadata exposed by data providers, including the associated terms of use.

- If a data consumer has an interest in specific data, the data space operator forwards this request to the data provider(s) so that they can enter into bilateral negotiations. If agreement is reached, the data space operator can offer further services.

- The data space operator also supports the actual exchange of data between data provider and data consumer, in compliance with the underlying terms of use.

- A data space operator can offer further services, for example, to support a data provider in proactively marketing its data.

The actual exchange of data payload can take place conceptually in different ways:

- A common infrastructure is used for data exchange, which is operated either by the data space operator or in accordance with the specifications of the data space operator. The data provider and data consumer must be onboarded to this infrastructure in order to use it.

- Data provider and data consumer exchange the data bilaterally. For this purpose, both the data provider and data consumer operate a data connector, which they install in advance and certify in accordance with the specifications of the data space operator.

Business context

Figure 12 shows the business context for this use case. It should be noted that in the business scenario Aspern Data Circle ECOM-X, the energy community assumes the role of a data consuming company and the energy prosumers assume the role of a data providing company.

Technical perspective

Figure 13 (p. 21) shows the technical perspective of this use case.

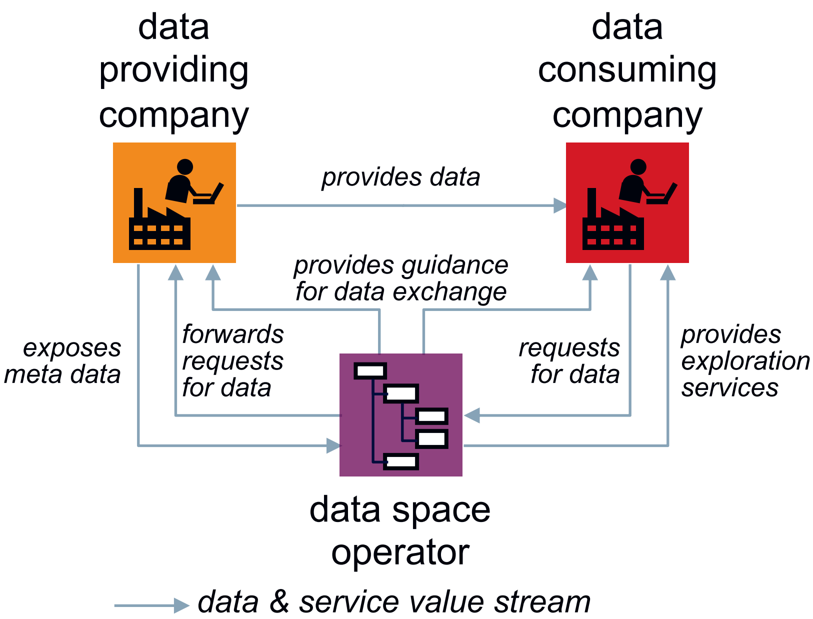

Figure 12: Business context of use case 'Data brokerage'

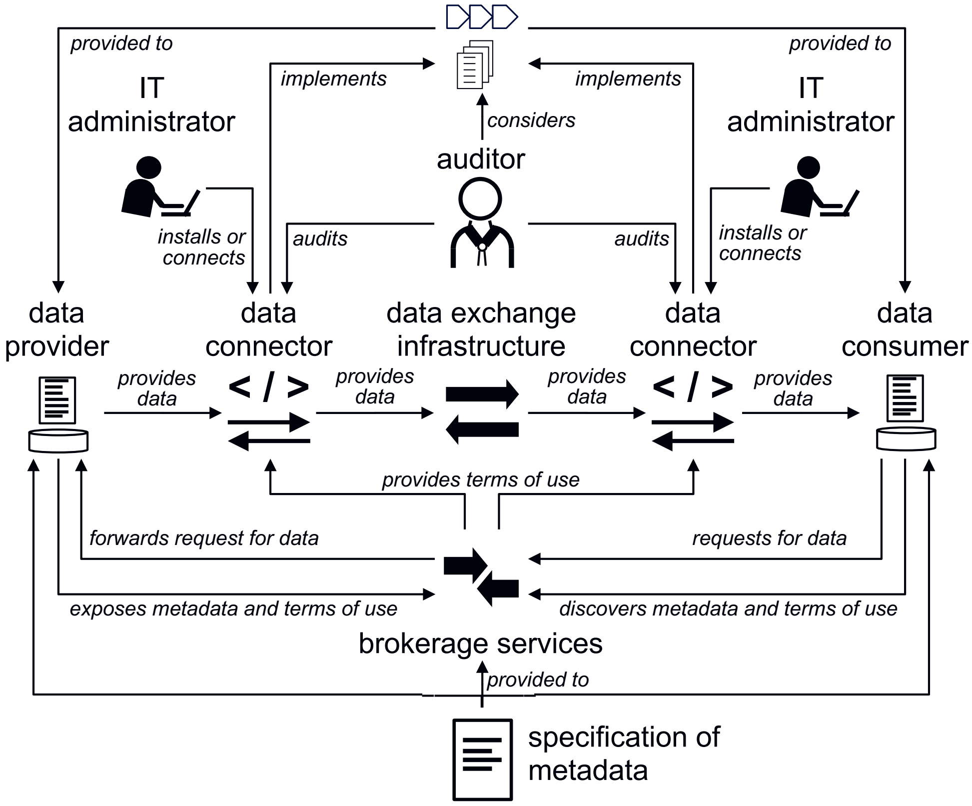

Figure 13: Technical perspective of use case 'Data brokerage'

Interactions of roles

The mediation and possible resulting data exchange between data provider and data consumer is based on the following workflow:

- A data provider exposes metadata and corresponding terms of use using the brokerage services. When describing the metadata, the data provider follows the specification of metadata; when describing the terms of use, the data provider follows the usage policies.
- The brokerages services provide the metadata and terms of use exposed by a data provider to third party data consumers for the purpose of discovery. Due to its business interests, the data space operator typically designs the specification of metadata and usage policies.
- A data consumer discovers the metadata and terms of use using the brokerage services with regard to the purpose of its interests. The data consumer takes the underlying specification of metadata and usage policies as a basis.
- If the discovery is successful, the data consumer sends a request to the brokerage services, so that the brokerage services can forward this request to the associated data provider.
- The data consumer and mediated data provider enter into a business negotiation with the aim of reaching agreement on the technical and business conditions. This applies to the specific data payload that the data provider makes available to the data consumer.
- If agreement is reached, the brokerage services configure the connectors of the data provider and data consumer to ensure compliance of the specific data exchange with the required usage policies.
- The data provider then transmits the specifically agreed data payload to the data exchange infrastructure using the appropriately configured connector. The data exchange infrastructure provides suitable services, so that

the specific data payload can be transferred from the data provider's connector to the data consumer's connector. The data consumer receives the specifically agreed data payload via the appropriately configured connector, to then use it for its own purposes.

In order to use the brokerage services and execute a data exchange, the data provider and data consumer must use or install a connector. Usage with respect to installation depends on the underlying concept of the exchange of data payload:

- A common infrastructure is used for data exchange: In this case, the data provider and data consumer are each provided with a connector offering capabilities to provide or receive the specifically agreed data payload. The connection with the connector is established by an IT administrator, whereby the data provider or data consumer can use the services of a third party service provider for this purpose. The connectors are operated either by the data space operator or in accordance with the specifications of the data space operator. From a technical point of view, an audit by an independent auditor is not necessary.

- The data provider and data consumer exchange the data bilaterally: In this case, the data provider and data consumer install their own connector as a basis for a peer-to-peer bilateral data exchange. This installation is carried out by an IT administrator, whereby the data provider or data consumer can use the services of a third party service provider for this purpose. The installation is then assessed by an independent auditor to ensure that the connectors are used in compliance with the usage policies required for a specific data exchange.

Expected change and impact

Today, this use case is mostly implemented individually for a data provider and data consumer. In addition, there are very few overarching and comprehensive brokerage services available which handle the offering of data, especially for industries. For this reason, potential business cases between data provider and data consumer are often not economically viable.

The use case postulates that in the future a new company will establish itself on the market, providing on the one hand a standardised implementation for data exchange between data provider and data consumer, and on the other hand offering comprehensive brokerage services for data providers and consumers, especially for industries. Data providers and data consumers can therefore focus on the value-added aspects of using data and can build on a technical basis provided by another company. This means lower investment costs for the data provider and data consumer, making the business case between data provider and data consumer more likely to pay off, since the other company's investment can be allocated to numerous applications.

Example: Product carbon footprint calculation

Objective

A manufacturing company wants to declare the product carbon footprint of its products.

Overview

The product carbon footprint (PCF) is a measure of the total amount of carbon dioxide emissions caused either directly or indirectly by an activity. The PCF of a product considers all phases of a product's lifecycle, up to and including delivery to the customer. Since there are different ways to calculate PCF, a manufacturing company not only requests a supplier's actual declared PCF, but also the calculation method used.

The information required to calculate direct emissions is usually available to a manufacturing company, for example in corresponding systems for enterprise management and production execution. However, this information is often aggregated, so a breakdown is requested for specifically manufactured products. With regard to indirect emissions, the challenge is that manufacturing companies must be able to rely on PCFs declared by their suppliers.

In the specific use case, both the supplier and manufacturing company use trustworthiness management services to validate their PCFs. The manufacturing company and/or product customer use trustworthiness management services to validate PCFs declared by the supplier and/or manufacturing company with regard to trustworthiness.

The certifier uses trustworthiness management services to document certification; the customer uses trustworthiness management services to ensure that the manufacturing company has calculated the PCF as declared.

Business context

The business context for this use case is shown in **Figure 6**.

Technical perspective

Figure 14 shows the technical perspective for this use case.

Interactions of roles

The PCF coordinator defines the guidelines which the manufacturing company uses as a basis for calculating PCFs and which it complies with both technically and organisationally.

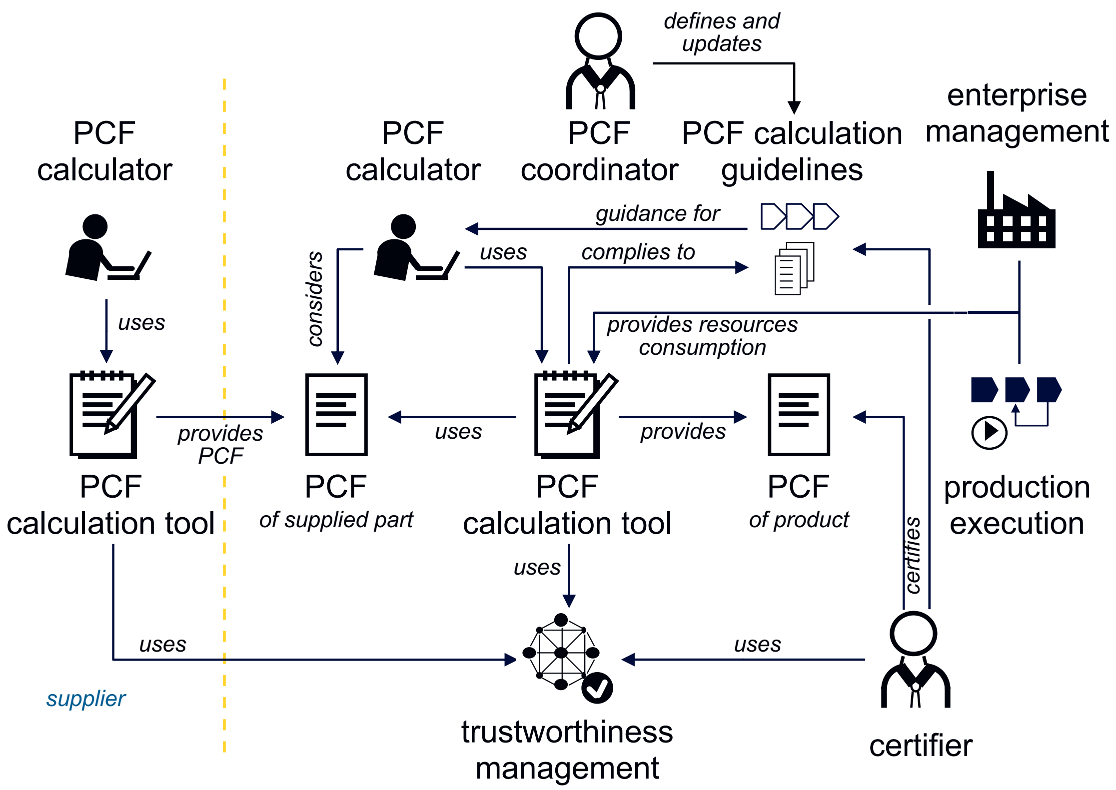

Figure 14: Technical perspective of use case 'Product carbon footprint calculation'

The PCF calculator performs the following activities as part of product design:

- The PCF calculator, in the role of a product developer, decides how the PCF of a product is calculated in principle. This is part of the product design when developing a product series. On the one hand, this concerns which method(s) are used for the PCF calculation. In addition, as part of product design, the process determines which product components are manufactured in-house and therefore a source of direct emissions and energy consumption within the manufacturing company, and which product components are purchased from a supplier and thus affect the supply chain.

- With regard to the calculation of direct emissions, the PCF calculator determines how typically aggregated information about direct emissions and energy consumption can be broken down to specific direct emissions and energy consumption caused by a specific product. This can result in a request from the PCF calculator to adjust enterprise management and production execution in order that these systems can deliver the information required to calculate direct emissions.

- The PCF calculator complies with the PCF calculation guidelines of the manufacturing company, which can be documented accordingly.

Based on the PCF calculation guidelines and the technical requirements of product design, the manufacturing company installs and integrates a PCF calculation tool into its overall IT landscape.

The PCF calculator performs the following activities as part of production execution, for which there are two alternatives:

- The PCF calculator first initiates the PCF calculation, then the PCF calculation for the individual products takes place automatically without any manual intervention by the PCF calculator. This requires that the PCF calculation tool is suitably integrated into the overall IT landscape and that all information required for the PCF calculation is made available by suitable IT systems.

- The PCF calculator intervenes manually in the PCF calculation using the PCF calculation tool. It is even conceivable that the PCF calculator carries out the specific PCF calculation outside of the PCF calculation tool and enters the result of the calculation manually into the PCF calculation tool. In this latter case, there is no need for the PCF calculation tool to be integrated into the overall IT landscape of the manufacturing company.

In both cases, the PCF calculation tool provides a PCF value, which the manufacturing company then makes available to the customer together with the product. In addition, the PCF calculation tool uses trustworthiness management to create a certificate stating that the specific PCF value has been calculated and made available by the manufacturing company in accordance with its PCF calculation guidelines.

Trustworthiness management provides a user with the following services – in this use case, the user is either the PCF calculation tool or the auditor:

- A user can issue a certificate.

- A user can validate an issued certificate and therefore can be certain that the certificate comes from the certificate issuer. In this use case, the customer is interested in validating the PCF value declared by the manufacturing company. On the other hand, the manufacturing company is also interested in having a validation of PCF values for the supplied parts declared by the suppliers.

The certifier performs the following activities:

- The certifier analyses the manufacturing company's PCF calculation guidelines and thus certifies that the manufacturing company is acting in accordance with the regulatory requirements.
- The certifier approves the specific PCF calculation of the manufacturing company.
- Using trustworthiness management, the certifier creates a certificate stating that the manufacturing company calculates and makes its PCF values available in compliance with the specifications.

Expected change and impact

This is an emerging topic that is also heavily influenced by current social and political discussions.

In this respect, the focus of this use case is initially on designing a high-level solution framework to create the basis for a common technical understanding. In addition, however, specific technical implementations are necessary to set out in concrete terms the resulting challenges from the technical, organisational and economic perspective.

Example: Trusted model training with multiple data sources

Objective

Manufacturing companies want to make data available to receive models trained with this data in return for being able to improve their quality inspection, for example, but without this data being accessible to other companies.

The algorithm training platform needs sufficient data from different companies to be able to train their models.

Overview

Most manufacturing companies cannot use their own data to train models to carry out their own automatic inspections, for example. For this reason, they would prefer to use an algorithm training platform which can train models for such purposes. Manufacturing companies send their data to the algorithm training platform. Once model training is complete, the manufacturing companies can use the trained models for automatic quality inspection.

In this case, different manufacturing companies request trained models. So, each manufacturing company shares data with the algorithm training platform in order to have a model trained for their use. These manufacturing companies have an interest in their data being protected against data sharing with other manufacturing companies, while the algorithm training platform has an interest in receiving data from as many manufacturing companies as possible, since this enables them to generate better models.

Every company therefore operates a trusted data matrix client. Together these form an overall trusted data matrix, within which the manufacturing companies provide data to enable the algorithm training platform to train their models using this database. Additionally, all participants use logging services to store all activities and results regarding data exchange and usage. Furthermore, all participants use identity services for identification purposes. Finally, the manufacturing companies and algorithm training platform use metadata catalogue and brokerage services. Manufacturing companies can expose metadata, including the terms of use, so that the algorithm training platform can discover the metadata and terms of use and can request specific data from a manufacturing company. The data can then be exchanged based on a digital contract in accordance with the terms of use. The digital contract is provided by the metadata catalogue and brokerage services and enforced by the trusted data matrix clients.

Business context

The business context for this use case is shown in **Figure 9**.

Technical perspective

Figure 15 shows the technical perspective for this use case.

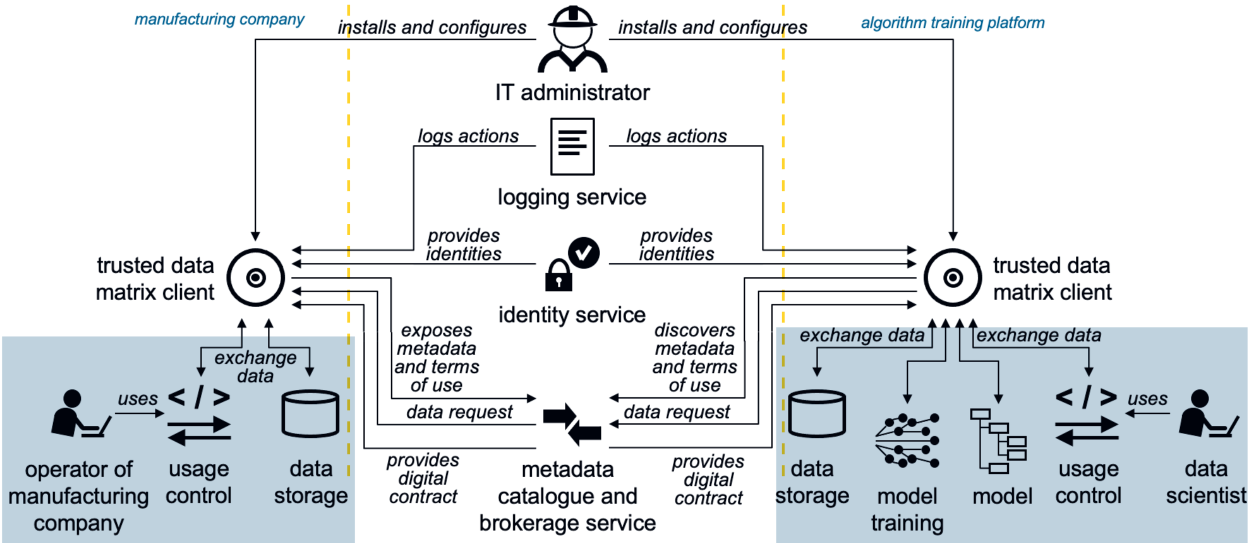

Figure 15: Technical perspective of use case 'Trusted model training with multiple data sources'

Interactions of roles

Any company wanting to participate in a trustworthy data exchange as a data-providing or data-consuming company uses a trusted data matrix client installed and configured by an independent IT administrator. This ensures that all companies participating in the overall trusted data matrix follow the same policies regarding data sharing. However, this also requires that all software applications in companies participating in the overall trusted data matrix use the data provided via the trusted data matrix exclusively in accordance with the digital contract concluded with the data-providing company.

- A digital contract represents a binding digital agreement crafted between a data provider and data consumer before any data sharing occurs. It offers the ability to set sharing conditions and configure control policies. A digital contract can set restrictions on various factors during the use of shared data, including time, location, subjects, actions and objects. The control policies for sharing can be pre-established by the data provider, or collaboratively decided upon by both the provider and consumer. Being machine-readable, the digital contract controls, monitors and enforces the scope and method of data object usage, addressing security, trust and data content protection issues throughout the data lifecycle.

- Depending on the digital contract in question, this can have very far-reaching consequences in terms of data exposure, data accountability, data access, data integrity and data usage. To implement associated requirements, it is recommended that an evaluation is carried out of the technical and organisational measures that can be used. Examples of technical measures are encryption and decryption capabilities to protect data during transmission and storage. An example of an organisational measure is the defining of employee roles with specific restrictions on accessing and changing data.

In this use case various services are considered:

Logging services: During data sharing activities, all request and data sharing actions caused or provided by the manufacturing companies and the algorithm training platform – for example, request, result, time, user identity – will be recorded and stored by these logging services. All such logging information that is recorded and stored can subsequently be provided for audit purposes.

Identity services: These services include identity authentication for data scientists from the algorithm training platform as a data-consuming company, and for operators of the manufacturing companies as data-providing companies. In this way, user identities can be trusted and reduce the risk of participation by a malicious party. In addition, these services provide authentication for the trusted data matrix clients.

Metadata catalogue and brokerage services: As data providers, manufacturing companies can expose and upload metadata, including terms of use, when sharing data with a catalogue. As a data consumer, the algorithm training platform can search for and discover metadata in the catalogue. The algorithm training platform uses the catalogue metadata to send a data request to a manufacturing company through the metadata brokerage service. Following this, the algorithm training platform and manufacturing company can negotiate data usage limitation based on the metadata brokerage service. If a mutual agreement is reached, the algorithm training platform and manufacturing company sign a digital contract provided by the metadata brokerage service. The metadata brokerage service distributes the signed digital contract to the trusted data matrix clients of the algorithm training platform and manufacturing company in order that the trusted data matrix clients enforce compliance with the digital contract.

In this specific use case, the algorithm training platform stores the data provided in an internal data storage system, trains the specific models with the help of model training algorithms, which in turn access the internal data storage, see areas with grey background in **Figure 15**.

Expected change and impact

Technologies for cross-company data sharing are currently – and very controversially – under discussion. The trusted data matrix is one possible concept through which companies are motivated to share data across companies, with the value proposition that after sharing their data they still have a certain amount of control over the shared data. It is not possible at the present time to assess the extent to which the various concepts for cross-company data sharing will prevail in the market.

Example: Trusted data sharing of digital printing images

Objective

A digital printing image provider wants to be supplied with printed matter provided by a print factory without the digital printing image being leaked.

The print factory wants to be a trustworthy partner of the digital printing provider.

Overview

This use case considers one company sharing data with another. As the data provider, a digital printing image provider sends a digital printing image to a print factory, the data consumer. The print factory uses the digital printing image to produce a product using its own equipment and delivers the printed matter to the digital printing provider.

This use case addresses the problem that the owner of images often cannot control the scope of data circulation of its images. The use case prevents the leakage of digital printing images and the risk of infringement of commercial interests and intellectual property rights caused by data sharing and exchange. It provides support for companies to increase their competitiveness and protect their intellectual property rights for digital printing images. Furthermore, it also provides experience in sharing and exchanging high-value data.

Both the data provider and data consumer therefore operate a trusted data matrix client, together forming an overall trusted data matrix in which the data provider provides data in a way that can be used by the data consumer. This includes an IT infrastructure for data transmission, storage and computing. Data control, trusted storage and the necessary environment abilities are executed in each trusted data matrix client.

The trusted data matrix is complemented by the following additional services:

- The data provider and data consumer can use metadata catalogue and brokerage services. The data provider can expose metadata, including the terms of use, so that a data consumer can discover the metadata and terms of use and can request specific data from a data provider. The data consumer can then negotiate a digital contract with a data provider. This contract includes data usage policies, which clarify the data usage subject, object, action, time factor, environment factor and further entities. For example, an employee (subject) of a printing company may only load (action) the digital printing image (object) into a printing resource in the printing company's PC (environment factor) within 14 days (time factor). Based on such a digital contract the data can then be exchanged.
- The data provider and data consumer can use identity services to prove that their identity is trusted, which is the basic entry condition for a trusted data matrix.
- Logging services can be used to collect and store all activities and results regarding data exchange and usage. Logging includes the data usage subject, object, action, time, environment factor and further entities. The logged data can be used as evidence in clearing and auditing.

Business context

The business context for this use case is shown in **Figure 11**.

Technical perspective

Figure 16 shows the technical perspective of this use case.

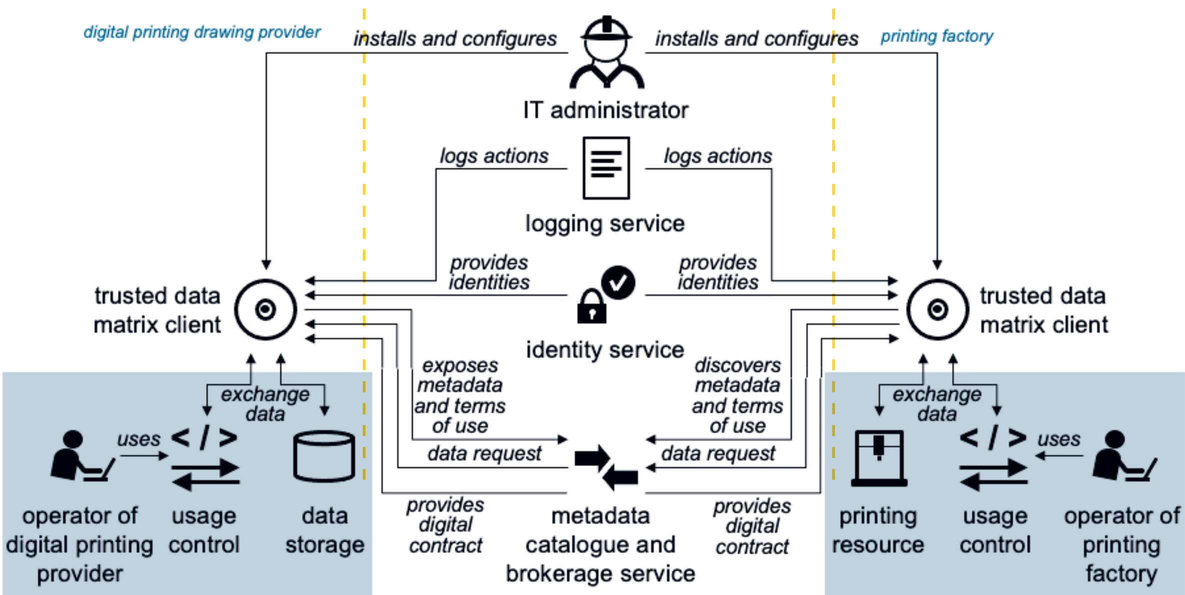

Figure 16: Technical perspective of use case 'Trusted data sharing of digital printing drawings'

Interactions of roles

See interaction of roles for the use case 'Trusted model training with multiple data sources'.

In this specific use case, the printing resource uses data provided by the trusted data matrix client and is operated by the operator of the print factory (see areas with grey background in **Figure 16**).

Expected change and impact

See expected change and impact of the use case 'Trusted model training with multiple data sources'.

Functional view considerations

At a functional level, many discussions are currently in progress. A comprehensive and consolidated overview is given in the position paper Design Principles for Data Spaces, see [10]. **Figure 17** shows one of the central images from [5] and illustrates the key building blocks of data spaces.

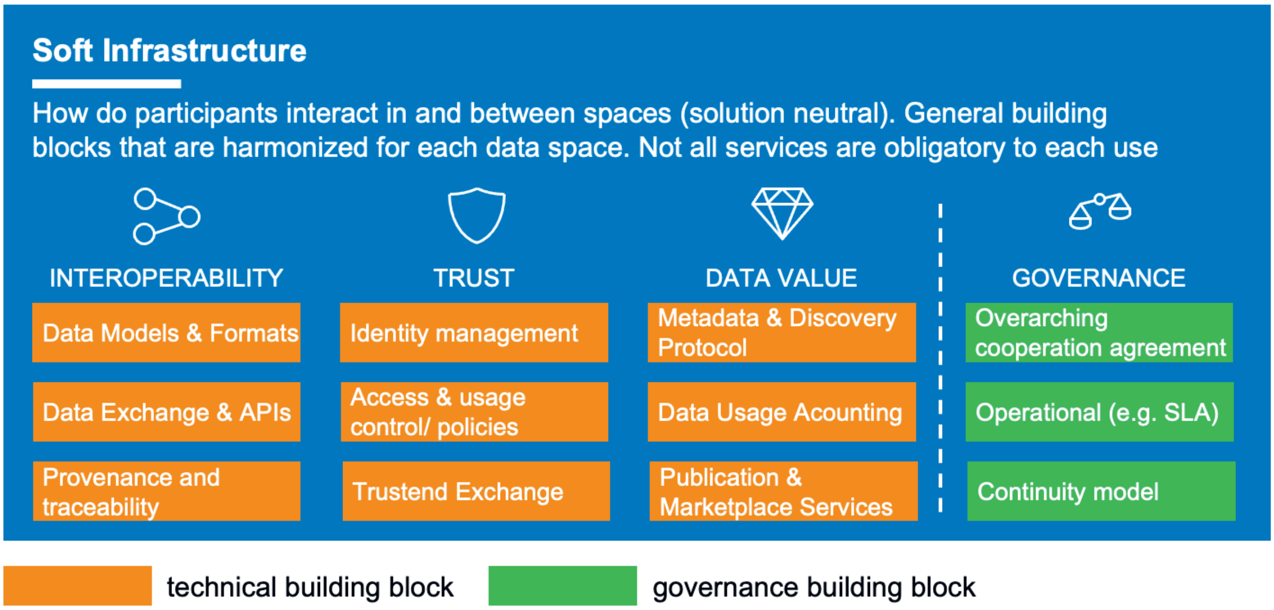

Figure 17: Data spaces building blocks, see [9]

Another approach is the trusted data matrix system developed by the Trusted Data Matrix Consortium, see **Figure 18.**

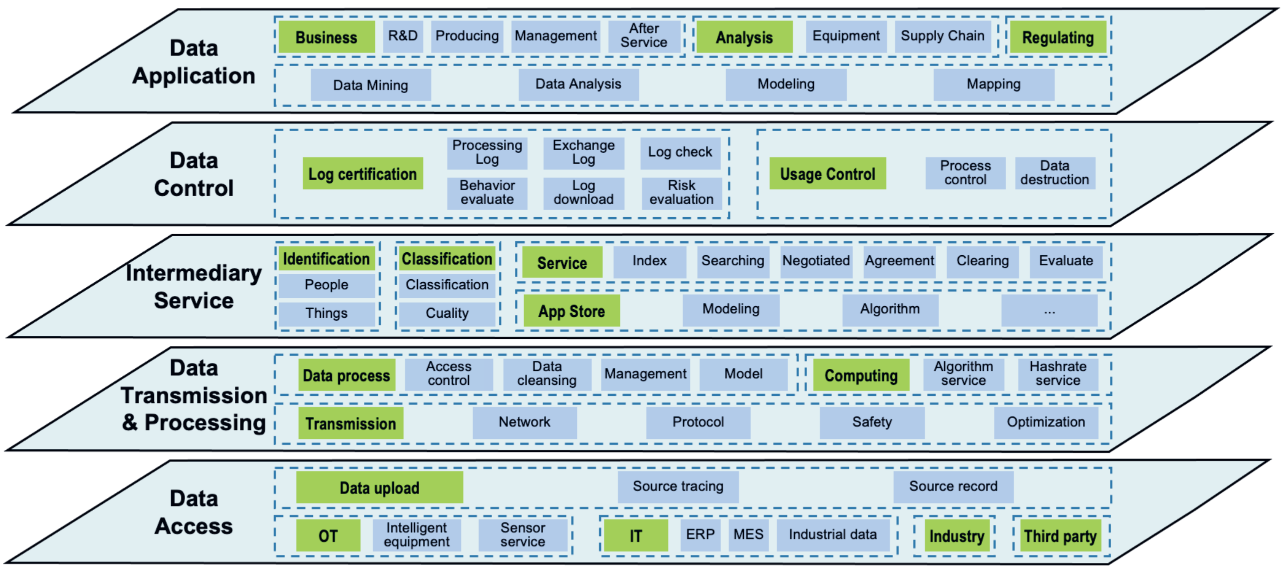

Figure 18: Functional view of Trusted Data Matrix Consortium, see [10]

The Trusted Data Matrix Consortium features a five-layered functional architecture designed to manage and utilise data effectively and securely:

- The first layer, the data access layer, acts as the data source, including industrial data from intelligent equipment and perception equipment from OT layers, and ERP and MES from the IT layer. This layer also records and traces data sources.

- The second layer is the transmission and processing layer, responsible for data transmission, processing and calculation. It provides core services such as network protocol and security, network performance optimisation, data access control, and data cleaning, storage, management and processing.

- The intermediary service layer, the third layer, is mainly comprised of third-party services, including participant and access authentication, data catalogue management and resource retrieval, alongside data contract management, transaction clearing, post-use evaluation and application store services.

- The fourth layer, the data control layer, provides log collection, certificate storage, behaviour evaluation, risk evaluation and audit report generation. It also controls data access and usage during the entire data lifecycle, including data use process control and post-use data destruction processes.

- Finally, the data application layer encompasses business operation, application innovation-related functions for companies and regulatory applications for governments. It provides services through data mining, data analysis, data modelling and knowledge mapping.

A detailed consideration of the functional view of data spaces is not within the scope of the cooperation line Use Cases and Applications.

Conclusion

The example of data spaces has shown that the guidance provided in the cooperation line Use Cases and Applications, see [2], is appropriate for preparing a complex topic in a systematic way and thus contributes to a better overall understanding.

We recommend that further examples are prepared in order to continue the discussion on data spaces and establish a clear idea of the viewpoint adopted in these examples.

In addition, there should also be discussion of the specific technical features that contribute to individual business goals and applications, with a view to anchoring technical implementations in a business and usage viewpoint.

References

[1] The Industrial Internet Reference Architecture Technical Report,
[2] Guidance on Use Cases and Applications,
[3] Digital business models for Industrie 4.0,
[4] Examples for Business Scenarios in Manufacturing Industry,
[5] IEC/TR 63283-2 ED1: Industrial-process measurement, control and automation – Smart Manufacturing – Part 2: Use cases
[6] Usage Viewpoint of Application Scenario Value-Based Service,
[7] Use Case Equipment Lifecycle Management,
[8] Data Sharing Models in the Electro and Digital Industry,
[9] Design Principles for Data Spaces,
[10] Trusted Industrial Data Matrix,