Fraunhofer

DATA SOVEREIGNTY

THE INDUSTRIAL DATA SPACE

AUTHOR

Prof. Dr.-Ing. Boris Otto
Fraunhofer Institute for Software
and Systems Engineering ISST
in Dortmund

PUBLISHER

Fraunhofer-Gesellschaft zur Förderung
der angewandten Forschung e.V.
Hansastr. 27 c
80686 München
Germany

Internet:
https://www.fraunhofer.de/en.html
E-Mail: info@zv.fraunhofer.de

COORDINATION

Communications department:
Janis Eitner
Design: Fraunhofer-Gesellschaft e.V.
Typesetting and page layout:
www.ansichtssache.de

PHOTO ACKNOWLEDGMENTS

Page 9: Robert Bosch GmbH
Page 13: iStock
All other photos:
© Fraunhofer-Gesellschaft e.V.

SPONSORED BY THE
Federal Ministry
of Education
and Research

DLR Project Management Agency

INDUSTRIAL
DATA SPACE

Grant ID 01IS15054

TABLE OF CONTENTS

PRESERVING DATA SOVEREIGNTY IN THE DATA ECONOMY

The digital transformation currently underway in all areas of the economy and society has changed the way data is viewed. Data is not considered just a by-product of business activities anymore, but a strategic resource that constitutes the basis for developing and offering novel digital products and services, and for establishing novel, digital business models.

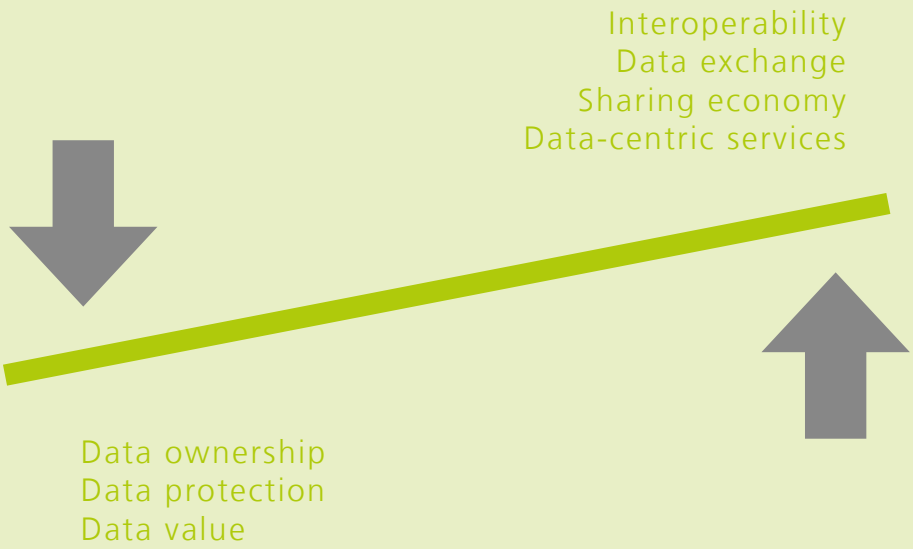


Figure 01: Conflicting goals in the data economy

1.1 Data in the Smart Service Welt

Novel digital products and services often emerge in business ecosystems, which companies enter to jointly fulfill the needs of customers better than they can do on their own. In such ecosystems, which emerge and dissolve much faster than traditional value creating networks, the participating companies have a clear focus on end-to-end customer processes in order to jointly develop innovative products and services [1]. Examples of business ecosystems are numerous and can be found across all industries; many of them have been analyzed and documented by the Smart Service Welt working group [2].

1.2 Data sovereignty as a key capability

From these two developments – data turning into a strategic resource, and companies increasingly collaborating with each other in business ecosystems – results a fundamental conflict of goals as a main characteristic of the digital economy: one the one hand, companies increasingly need to exchange data in business ecosystems; on the other hand, they feel the need to protect their data more than ever before, since the importance of data has grown so much. This conflict of goals is even more intensified when a company engages in business ecosystems, and the value of the data contributes to the overall success.

Can this conflict of goals be resolved? The answer is yes – if companies manage to preserve their data sovereignty.

Data sovereignty can be defined as a natural person's or legal entity's capability of being entirely self-determined with regard to its data.

Data sovereignty is about finding a balance between the need for protecting one's data and the need for sharing one's data with others (see Fig. 01). It can be considered a key capability for companies to develop in order to be successful in the data economy.

To find that balance, it is important to take a close view at the data itself, as not all data requires the same level of protection, and as the value contribution of data varies, depending on what class or category the data can be subsumed under.

2

VIEWING DATA AS AN ASSET

Companies that consider data a strategic resource need to view and treat data as an asset. The concept of viewing data as an asset is not new; first theoretical reflections on that idea date back to the 1980s [3, 4]. The discussion in those days focused primarily on the intra-organizational use of data, however. Today, the idea of viewing data as an asset is discussed with a strong focus on the use of data in business ecosystems – leading to the conflict of goals described above.

2.1 Properties and views of data

It is indisputable that data has a value and that data management produces costs. Today, data is traded in the market, it has a price, and many companies monitor the costs incurred for data management. However, data as an intangible asset differs from tangible assets with regard to a number of properties, among them [5]:

- **Wear and tear:** Unlike tangible assets, data is not subject to wear and tear as a result of being used.
- **Scarcity:** Unlike tangible assets, data is not a rare good. The value of data increases as it is being used (and, in many cases, as the number of users increases).
- **Integration:** The value of data increases when data is integrated with other data (as can be seen from the emergence of »data lakes«, in which as much data as possible is centrally stored to conduct big-data analysis [6]).

At the same time, data is often viewed in analogy to tangible goods. Among the views taken for categorizing data are:

- **Market view:** Data can be distinguished into private data (e.g. product master data) and public data (e.g. geo information or address information). A mixed form can be found if data is treated like a »club good« (i.e. if it is accessible by a certain group of users only, e.g. track-and-trace data in a supply chain).
- **Production view:** Data can be viewed as the raw material of information products. As a consequence, different stages of production and »data supply chains« can be found.
- **Demand view:** Data can be interpreted as substitute data (e.g. weather information from different data sources of similar quality) or complementary data (e.g. location and traffic information to calculate the time of arrival of goods).

As data goods contribute to the development of innovative products and services (depending on what category the data can be assigned to), the need for protection of data is not the same across all categories. Public data, for example, which can be accessed by any company, requires a lower level of protection than private data or club data.

Because of these differences and distinctions made with regard to data, a generally accepted understanding of the value of data has not been established so far. Nevertheless, there is a growing need to determine the value of data, given the rapid developments taking place in the Smart Service Welt [7].

2.2 Determining the financial value of data

Similar to assigning data to different categories reflecting its nature, different approaches can be used to determine the value of data [8]:

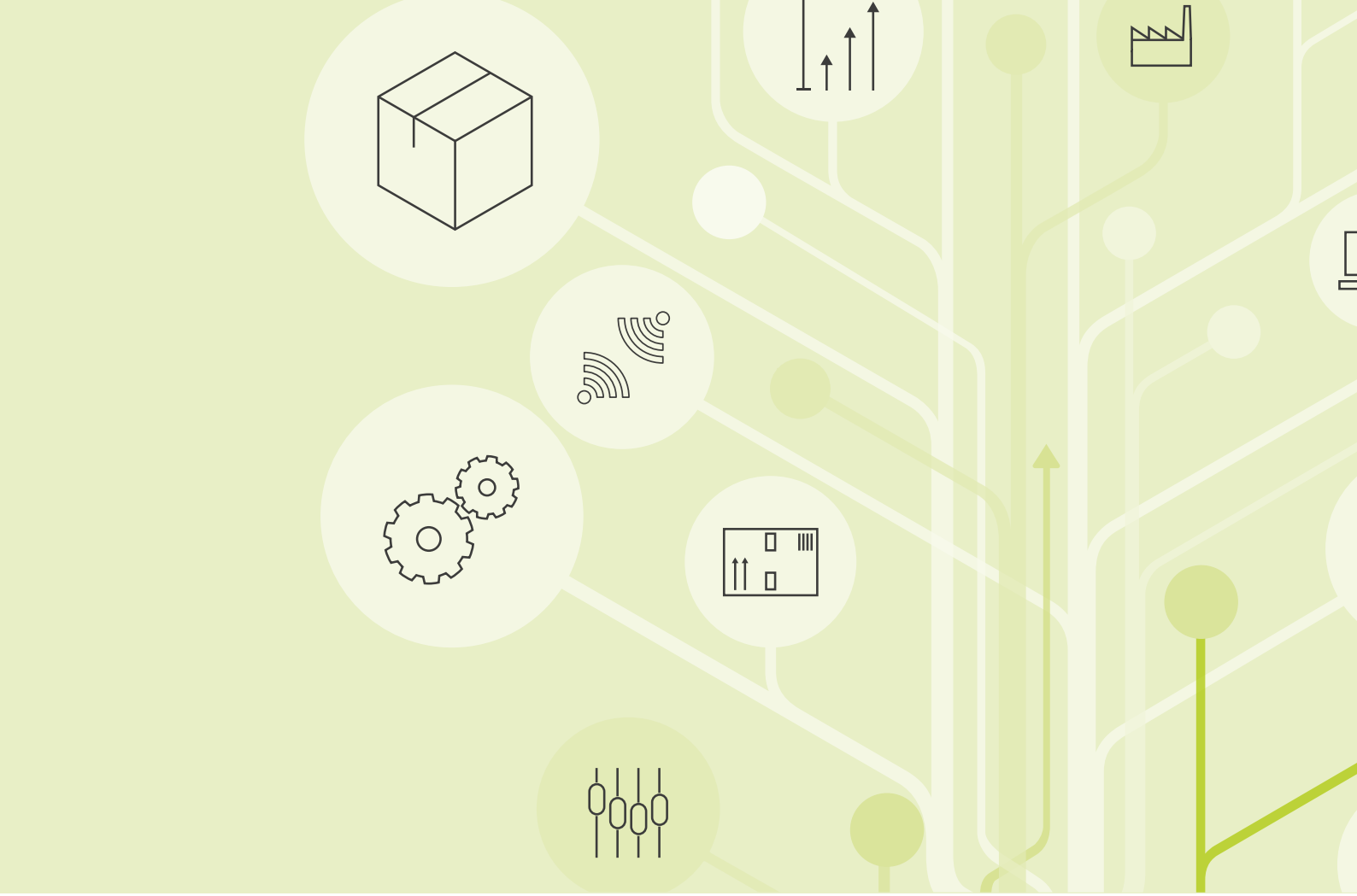
- **Cost approach:** A lot of data has no market value (simply because there is no market for it). Such data typically occurs and is used at an early stage of the data supply chain (e.g. master data on semi-finished products manufactured by the company itself). The value of such data can be determined by means of the cost incurred for collecting it, entering it into systems and databases, and maintaining it over time.
- **Income approach:** In many business processes, data has a »value-in-use«, which is significantly determined by the quality the data has [9]. In the case of customer master data, for example, this means that e.g. sales and distribution controlling is all the more effective the more data is consistent, complete, and up-to-date.
- **Market approach:** If there is a market for data, the value of that data can be determined by its market price. Regarding business partner data, for example, a number of commercial data providers have already established themselves in the market [10].

For German companies operating in highly competitive international markets it is critical not to be stuck with the costs of data management while letting others benefit from the monetarization of data.

Companies undergoing digital transformation therefore must be able to distinguish between the different approaches of data valuation, to measure and control the value of data, and to understand the »mechanics« of digital business models.

3 PERSPECTIVES OF DATA SOVEREIGNTY

Data sovereignty can be considered a key capability for companies to develop. Companies need to preserve the sovereignty over their data in order to stay competitive in times of digital transformation.



3.1 Legal perspective

Today, in many companies »data owners« are responsible and accountable for data (similar to »product owners« and »process owners« being responsible/accountable for certain products and processes, respectively) [11]. From a legal point of view, it is important to clearly define the term »data ownership«, as this concept determines how data is to be used in business ecosystems [13].

What »data ownership« means is currently being controversially discussed also on a European level. For instance, the European Commission has organized a number of round tables in order to get advice by legal experts, which was received critically by various special interest groups [14].

Although there are still obstacles to overcome in the legal implementation of the term »data ownership«, there can be no doubt that generally applicable recommendations for action are required to ensure legally binding control over data [12].

3.2 Economic perspective

While it is good to know that data has a value, that value is of no use if it cannot be compared to other values and if its measurement cannot be repeated. Without standardized methods for data valuation, the data economy lacks the necessary rules with regard to determining the value of data.

While methods for determining the value of data exist and are well applicable, accounting for data is still not reality, however. This is mainly due to the fact that these methods have not found their way into the everyday practice of accountants and auditors yet [7]. The following three-step procedure could help overcome this situation:

- First, comparable methods and tools for data valuation need to be developed.
- Second, these methods and tools need to be evaluated by the community of accountants and auditors, and standardized by the respective bodies and boards.
- Third, on the basis of this evaluation and standardization, the principles of accounting for data need to be developed and established.

The overall vision behind these reflections is to include data on the balance sheet as an asset, alongside with tangible assets (e.g. fixed assets and current assets).

3.3 Information technology perspective

While both the legal perspective and the economic perspective pose a mandatory condition for companies to succeed in preserving their data sovereignty, the implementation of these concepts on the level of information technology can be considered a sufficient condition for achieving that goal.

Without appropriate information technology making self-determination and control over data electronically feasible, data sovereignty remains just a vision.

Owners of data need to be enabled to preserve the sovereignty over their data, even if they share this data with their partners in business ecosystems.

The need to consider different levels of protection to be applied for different categories of data is relevant also when it comes to designing and implementing the necessary information technology applications.

Inferior data (i.e. data that has a low value) poses fewer requirements on data sovereignty than data of high value. This will result in a coexistence of different approaches of information technology architectures for data sovereignty.

The spectrum may range from data lakes, in which data from a multitude of owners is centrally collected by third-parties without special data protection mechanisms in place, to decentral architectures, in which only data processing routines are exchanged, but not data itself.

The Industrial Data Space is a decentral architecture for exchanging data between trustworthy participants.

4

THE INDUSTRIAL DATA SPACE

4.1 Goals

The Industrial Data Space initiative aims at data sovereignty in business ecosystems. Participants in the Industrial Data Space can exchange data between each other under the premise that self-determination and control over their data is always guaranteed.

The information technology architecture providing the basis of the Industrial Data Space is demonstrated in use case projects in terms of applicability and usability.

The Industrial Data Space initiative is organized in two branches: an association named »Industrial Data Space e.V.« and a research project. The research project is led by Fraunhofer and funded by the German Federal Ministry of Education and Research. Both the project and the association are pre-competitive and not for profit.

The initiative promotes a broad dissemination of the Industrial Data Space. It welcomes commercial paths of exploitation, which are open to each individual participant.

Trusted Platform Module (TPM)	No TPM		TPM 1.2		TPM 2.0
Certification	No certificate	Self-signed certificate	CA-based certificate	CA-based certificate from external CA (cross-certified)	CA-based certificate based on IDS-PKI
Container Management Layer (CML)	No CML		CML & Core Container Attestation		CML & Core Container & Container Attestation
Isolation	Various				
Software Assurance Level	Unknown stack			IDS-certified stack	

Fig. 02: Security profiles in the Industrial Data Space

4.2 Design principles

From the goal of making data sovereignty feasible on an information technology level, a number of requirements can be derived, which can be translated into design principles regarding the architecture of the Industrial Data Space. Among these design principle are

- distributed storage of data,
- certification of participants,
- secure data exchange,
- governance on the basis of commonly agreed rules,
- open development process,
- networking effects for data providers and data consumers,
- economies of scale through standardization.

These principles guide the design of a reference architecture model for the data economy.

The Industrial Data Space initiative acknowledges the coexistence of different types of architectures, ranging from data lakes for central data storage to architectures for distributed data storage. This approach takes into account that different types and categories of data imply different levels of data protection.

4.3 Architecture

The Industrial Data Space is designed as a network of data endpoints. Each data endpoint is a certified implementation of a central software component of the Industrial Data Space: the Industrial Data Space Connector.

Each organization operating an Industrial Data Space Connector is a participant in the Industrial Data Space. The Connector works as a container into which different Data Apps can be loaded.

Two types of Data Apps can be distinguished: while »Basic Data Apps« provide basic services (e.g. functionality for searching or publishing data, or for transforming data from one format into another), »Smart Data Apps« provide special services (e.g. functionality for sending a warning if the temperature of certain goods exceeds a certain threshold value during transportation). Data Apps can be combined to Data App Packages (e.g. the »Broker Package«).

Regarding the Industrial Data Space Connector, different versions can be used to implement different levels of security (see Fig. 02), depending on the requirements of data in terms of data protection.

4.4 Participation and cooperation

The Industrial Data Space initiative is open to companies and organizations for participation and cooperation. While the research project is concerned with the design and prototype implementation of the reference architecture model, the association takes up user requirements, promotes standardization efforts, and organizes the transfer of knowledge. The non-profit association has an international focus and already counts over forty members from eight countries.

The Industrial Data Space initiative has already established relations with cognate initiatives, such OPC Foundation or Plattform Industrie 4.0.

Both the Industrial Data Space research project and the association seek collaboration and exchange of ideas with other research projects and with standardization bodies. In designing the reference architecture model, the Industrial Data Space research project aims at using and building upon existing technologies (e.g. Docker, for system virtualization) and results from other research projects (e.g. Theseus) to the extent possible.

For more information please visit

Fraunhofer Research project:
https://www.fraunhofer.de/en/research/lighthouse-projects-fraunhofer-initiatives/industrial-data-space.html
Industrial Data Space e.V.:
http://www.industrialdataspace.org/en/

4.5 Industrial Data Space e.V.

Founded January 26th, 2016, Industrial Data Space e.V. is a user association headquartered in Berlin. Its main activities are the bundling of user requirements, promoting standardization, organizing the transfer of knowledge, and communicating the activities and results of the Industrial Data Space research project.

The following organizations are founding members of Industrial Data Space e.V.:

- Allianz SE
- Atos IT Solutions and Services GmbH
- Bayer HealthCare AG
- Boehringer Ingelheim Pharma GmbH & Co.KG
- Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.
- KOMSA Kommunikation Sachsen AG
- LANCOM Systems GmbH
- PricewaterhouseCoopers AG
- REWE Systems GmbHRobert Bosch GmbH
- Salzgitter AG
- Schaeffler AG
- Setlog GmbH
- SICK AG
- thyssenkrupp AG
- TÜV Nord AG
- Volkswagen AG
- ZVEI – Zentralverband Elektrotechnik- und Elektronikindustrie e.V.

4.6 Collaboration with Plattform Industrie 4.0

The activities of the Industrial Data Space initiative are closely aligned with Plattform Industrie 4.0. While Plattform Industrie 4.0 deals with all aspects of digitization, the Industrial Data Space initiative has a special focus on the data layer of the reference architecture. At the same time, the Industrial Data Space initiative follows a cross-industry approach, whereas Plattform Industrie 4.0 has a clear focus on manufacturing enterprises.

The collaboration between the Industrial Data Space initiative and Plattform Industrie 4.0 takes place mainly within the scope of two Plattform Industrie 4.0 working groups, namely

- Working Group 1: Reference architectures, standards, and norms
- Working Group 3: Security of networked systems

5

LIST OF REFERENCES

[1] MOORE, James F.: Business ecosystems and the view from the firm. In: Antitrust Bulletin 51 (2006), Nr. 1, S. 31–75
[2] ARBEITSKREIS SMART SERVICE WELT: Smart Service Welt: Umsetzungsempfehlungen für das Zukunftsprojekt »Internetbasierte Dienste für die Wirtschaft«. Berlin, Heidelberg, 2014
[3] HORNE, Nigel W.: Information as an asset – The board agenda. In: Computer Audit Update 1995 (1995), Nr. 9, S. 5–11
[4] OPPENHEIM, Charles; STENSON, Joan; WILSON, Richard M. S.: Studies on Information as an Asset I: Definitions. In: Journal of Information Science 29 (2003), No. 3, S. 159–166
[5] MOODY, Daniel; WALSH, Peter: Measuring The Value Of Information: An Asset Valuation Approach. In: Proceedings of the 7th European Conference on Information Systems (ECIS 1999).
[6] O'LEARY, Daniel E.: Embedding AI and Crowdsourcing in the Big Data Lake. In: IEEE Intelligent Systems 29 (2014), Nr. 5, S. 70–73
[7] ATKINSON, Keith; MCGAUGHEY, Ronald: Accounting for Data: A Shortcoming in Accounting for Intangible Assets. In: Academy of Accounting and Financial Studies Journal 2 (10), S. 85–95
[8] OTTO, Boris: Quality and Value of the Data Resource in Large Enterprises. In: Information Systems Management 32 (2015), Nr. 3, S. 234–251
[9] ZECHMANN, Andreas; MÖLLER, Klaus: Finanzielle Bewertung von Daten als Vermögenswerte: Methode und Anwendung eines nutzenorientierten Ansatzes. In: Controlling: Zeitschrift für erfolgsorientierte Unternehmenssteuerung 28 (2016), Nr. 10, S. 558–566
[10] OTTO, Boris; AIER, Stephan: Business Models in the Data Economy: A Case Study from the Business Partner Data Domain. In: Proceedings of the 11th International Conference on Wirtschaftsinformatik (WI 2013), S. 475–489
[11] OTTO, Boris; ÖSTERLE, Hubert: Corporate Data Quality: Voraussetzung erfolgreicher Geschäftsmodelle. 1. Auflage. Berlin, Heidelberg: Springer Gabler, 2016
[12] DENNIS JLUSSI: Industrie 4.0 und Dateneigentum. 2016
[13] BAUER, Robert; OTTO, Boris: Dateneigentum und Datenbesitz im Kontext des »Industrial Data Space«. – Waldkirch, Dortmund. 2016
[14] DIGITALEUROPE: Statement on Data Ownership, Access and Re-Use of Data. Brussels, 2016