Fraunhofer

Incentives and Economics of Data Sharing

Fields of action of cross-company data exchange and status quo of the German economy

IEDS-Project partners

IW

technische universität dortmund

ZEW

Part of the Project

IEDS

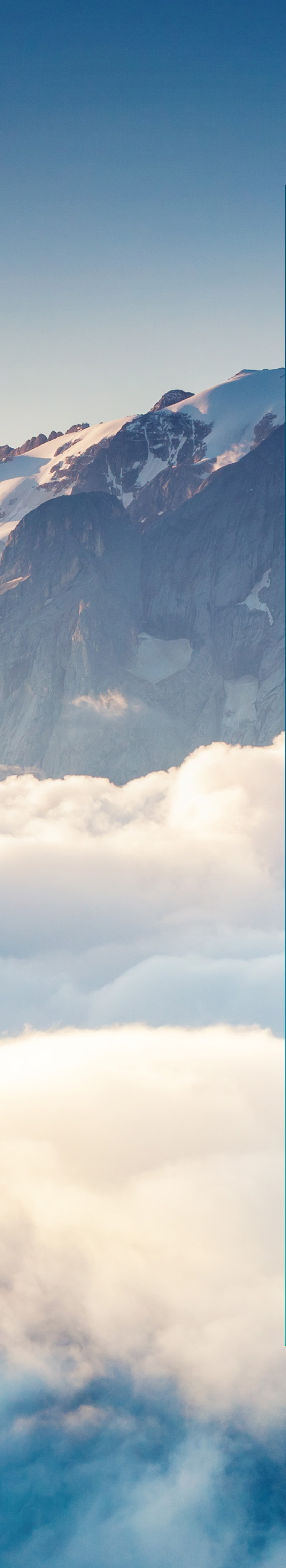


Table of Contents

Foreword 5

Zusammenfassung — Executive summary 6

1 Innovative data economy — Europe's opportunity through data rooms and data sharing 8
- 1.1 Using digitization and data sharing to harness innovation potentials 8
- 1.2 IDS and Gaia-X — European data infrastructure as the basis for data sharing . 9
- 1.3 Data sharing from a cross-company perspective 11
- 1.4 Data sharing from an internal company perspective 16

2 Status quo of the German data economy 20
- 2.1 Data economy readiness and data economy 20
- 2.2 Monitoring Gaia-X 24
- 2.3 Pretest survey experiment incentive mechanisms 28

3 Incentives and requirements to participate in data sharing — research overview 30
- 3.1 Research overview »data sharing« 30
- 3.2 Data strategies and efficient data management 33
- 3.3 Data-driven business models 39
- 3.4 Data evaluation 45
- 3.5 Data law 50

4 Outlook 54

5 IEDS research project overview and project partner presentation 56

6 Sources 58

Imprint 69

Foreword

Prof. Dr.-Ing. Boris Otto
Managing Director of Fraunhofer ISST and Chair of Industrial Information Management at TU Dortmund University

Prof. Dr. Michael Hüther
Director and Member of the Executive Board of the Institut der deutschen Wirtschaft (IW)

Prof. Dr. Irene Bertschek
Head of the ZEW Research Department »Digital Economy«

Prof. Dr.-Ing. Prof. e. h. Wilhelm Bauer
Executive Director Fraunhofer IAO

The digitization of our physical world is advancing continuously. The importance of data as a driver of the economy is evident in the innovation processes of new business models and digital product-service systems, which increasingly cannot be accomplished by one actor alone. Most recently, the Corona pandemic illustratively demonstrated that digitization and data play a central role in a functioning economy, science and society. Federated services are therefore moving into the focus of industries and sectors to take advantage of the opportunity of digital networking and combination of resources. The development of federated data spaces and data ecosystems are in focus, in both the data strategy of the Federal Republic of Germany and that of the European Commission. Initiatives such as Gaia-X and the International Data Spaces Association (IDSA) are working towards the vision of a distributed data infrastructure while maintaining data sovereignty. However, data space initiatives reveal that many companies are not yet empowered in their digital maturity to successfully and sustainably participate in the data economy. The Incentives and Economics of Data Sharing — IEDS project aims to address this. Scientific findings on enabling companies to participate in data sharing and intelligent incentive systems for data sharing can support the further development of the data economy and open up new potential for Germany and Europe.

Together with our partners and the Federal Ministry of Education and Research, we want to expand the possibilities in this project and provide blueprints that companies as well as science and society can use to increase their digital maturity and participate in data sharing.

Zusammenfassung – Executive summary

Zusammenfassung

Die enormen Innovationspotenziale, die Daten Unternehmen sowie der Wissenschaft bieten, werden in Deutschland überwiegend noch nicht genutzt. Häufig sind dabei die Potenziale sowie der Nutzen der jeweiligen Daten nicht bekannt. Zudem hegen vor allem Unternehmen große Bedenken hinsichtlich rechtlicher Fragestellungen und der Datensicherheit. Speziell im Hinblick auf die Entwicklung datengetriebener Geschäftsmodelle fehlt es Unternehmen an einer Betrachtung von Ökosystemen. Ebenso impliziert die Bewertung von Daten erhebliche Herausforderungen, sodass bestehende Verfahren zur Preisfindung von Gütern hier an ihre Grenzen stoßen. Gleichwohl zeigt das Engagement von mittlerweile über 300 Unternehmen innerhalb der Gaia-X-Initiative die zunehmende Bedeutung von Daten für die Wertschöpfung, des strategischen Umgangs mit Daten sowie der dafür notwendigen Infrastrukturen.

Das **Projekt IEDS – Incentives and Economics of Data Sharing** thematisiert die wirtschaftliche Bedeutung von Daten sowie die Möglichkeiten zu deren Austausch, Nutzung und Verwertung im unternehmerischen Kontext. Es zielt darauf ab, die Ausgestaltung von unternehmensübergreifendem Data Sharing voranzutreiben, Anreize für das Teilen von Daten abzuleiten sowie die Weiterentwicklung der Datenökonomie zu unterstützen. Das IEDS Projekt schafft hierzu ein Referenzdokument, in dem die Zusammenhänge von Datenstrategien, datengetriebenen Geschäftsmodellen, Datenbewertung und Datenrecht aufgezeigt werden. Unternehmen unterschiedlicher Größe sollen durch die erarbeiteten Ergebnisse dazu bewegt und befähigt werden, an der Datenwirtschaft und den damit verbundenen Ökosystemen teilzunehmen.

Executive summary

The enormous potential for innovation offered by data to companies and the scientific community is still largely unused in Germany. Often, this potential and the benefits of the respective data are not known. In addition, companies in particular have major concerns about legal issues and data security. Especially with regard to the development of data-driven business models, companies do not consider ecosystems. Likewise, the valuation of data poses considerable challenges, so that existing procedures for the pricing of goods reach their limits here. However, the involvement of more than 300 companies in the Gaia-X initiative demonstrates the increasing importance of data for value creation, the strategic use of data and the necessary infrastructures.

The **IEDS — Incentives and Economics of Data Sharing** project addresses the economic importance of data for companies as well as the possibilities for its exchange, use and exploitation in a business context. It aims to advance the design of cross-company data sharing, to derive incentives for data sharing and to support the further development of the data economy. For this purpose, the IEDS project creates a reference document in which the interrelationships of data strategies, data-driven business models, data valuation and data law are elucidated. Companies of different sizes are to be encouraged and enabled to participate in the data economy and the associated ecosystems by the results produced.

1 Innovative data economy — Europe's opportunity through data rooms and data sharing

New products and services can be generated through the use of data rooms and participation in data sharing (see Section 1.1). Initiatives such as IDSA and Gaia-X are already developing a standard or infrastructure for companies to participate in data sharing (see Section 1.2). In order to support companies in carrying out value-creating activities with the help of data sharing, the cross-company and internal action levels are then examined (see Sections 1.3 and 1.4).

1.1 Using digitization and data sharing for innovation potentials

Data form the basis on which the digital and technological transformation of society and the economy is taking place. It is the lifeline of economic development and provides the basis for many new products and services (Europäische Kommission 2020a). The increased availability of data and the development of new technologies enable a multitude of unique opportunities to analyze and use data in ways that create value (Wilberg et al. 2018). As data continues to be produced in previously unimaginable quantities, digitization promises additional shifts in the strategic landscape of companies and the evolution of existing business models.

Instead of a corporate strategy dictating what data should be collected and analyzed, in some cases a significant impact of the collected and analyzed data on corporate strategy (Mazzei and Noble 2017) or on societal structures occurs. The traditional closed innovation process has been transformed into a parallel and open process by the concept of co-creation (Guggenberger et al. 2020b). To this end, innovation activities must take place simultaneously, with information from different sources being processed and products and services already being conceived and designed at the same time. Linear innovation models in which tasks are processed sequentially are therefore unsuitable for this purpose (Wong et al. 2016). These new data-driven innovations are increasingly difficult to develop by a single organization and in traditional value chains. Instead, the increasingly interconnected world is leading to the combination, enrichment, and sharing of different data sources by different actors in cross-sectoral, socio-technical networks - known as data ecosystems (Gelhaar et al. 2021a; Oliveira and Lóscio 2018). Data ecosystems consist of complex networks of organizations and individuals that share and use data as a primary resource (see Section 1.3). Such ecosystems also provide an opportunity and basis to create, manage and sustain data-sharing initiatives (Oliveira and Lóscio 2018). The value of digitally transforming a society and industries is unmistakable. However, the value creation process must be thoughtful and deliberate (Mielli and Bulanda 2019).

The European Union estimates global growth in available data volume at 175 zettabytes for 2025 (Europäische Kommission 2020b).

The value of data for industry and society is also emphasized by the German government in its data strategy, which states that data form the basis of the digital society. With the strategy, the German government aims to increase the innovative and responsible provision and use of data, particularly in Germany and Europe. One of the strategy's goals is to make German and European data ecosystems attractive to more participants by expanding data infrastructures in an interoperable, energy- and resource-saving and decentralized manner. To this end, the cross-industry Gaia-X project (see Section 1.2) is to be driven forward in order to create open and transparent data ecosystems in which data and services can be made available, merged and shared in a trustworthy manner (Bundeskanzleramt 2021).

The European Commission has also focused on data ecosystems and data sharing in its data strategy. According to the European data strategy, "the EU [...] can become a model for a society that is able to make better decisions in the economy and in the public sector thanks to data." (European Commission 2020a, p. 1). One of the goals of the European data strategy is to create a single European data space in the sense of a single market for data. In this data ecosystem, data should be available securely and easily. Specifically, the intention is to drive forward the creation of EU-wide interoperable data spaces that will remove the legal and technical barriers that can accompany data sharing. In these data rooms, European regulations, in particular on privacy and data protection, as well as competition law, should be fully respected and the rules for access to and use of data should be fair, practical and clear (European Commission 2020b). The basis for the interoperable data spaces is to be the federated data infrastructure of Gaia-X and the standard for sovereign data exchange of the International Data Spaces Association, each of which is discussed below in Section 1.2.

The increasing amount of data generated by digitization and its exchange across companies offers a great deal of potential, but there are also a number of challenges and obstacles. Incentives for companies to engage in data sharing and to participate in ecosystems designed for this purpose can thus be derived from the literature, but they still have to be created with the help of political activities due to concerns. The IEDS project aims to identify these incentives for companies and also to derive incentive mechanisms and systems that could be used to encourage companies to engage in data sharing.

1.2 IDS and Gaia-X – European data infrastructure as the basis for data sharing

In October 2015, the Fraunhofer-Gesellschaft initiated the **International Data Spaces (IDS) research project**[1], funded by the German Federal Ministry of Education and Research (BMBF). The goal of the IDS initiative is to establish a global standard for trustworthy, secure, sovereign and interoperable data exchange. This endeavor is supported by the user association International Data Spaces Association e. V. (IDSA)[2]. In 2021, the association consists of more than 130 members who are jointly defining the IDS standard for data sovereignty. The members come from different domains and are testing the IDS architecture in a wide variety of areas.

1 https://internationaldataspaces.org/.

2 https://internationaldataspaces.org/we/the-association/.

In fall 2019, the **Gaia-X**[3] initiative was also launched, which aims to promote the development of a sustainable and innovative data economy in Europe. Gaia-X aims to create a unified data infrastructure based on European values related to data and cloud sovereignty (Gaia-X European Association for Data and Cloud AISBL 2021). Geographically, Gaia-X is not limited to Europe, but represents and strengthens European values for the data economy. The initiative is closely related to the European Data Strategy as well as the EU Recovery Plan[4]. Accordingly, Gaia-X supports innovative data applications and cross-industry innovations. By taking European values into account, Gaia-X can be seen as a step towards digital sovereignty and technological independence for Europe.

A key capability for the European and international economy in this context is data sovereignty. Data owners must be empowered to select desired data users, limit purposes of data use, and thus determine how their data is exploited. This requires appropriate information technology solutions to enable and exercise data sovereignty.

With the goal of promoting infrastructures for cross-company data exchange while ensuring data sovereignty, both initiatives each have central elements. A key component of the IDS initiative is connectors, which can be used to exchange or define both data and the terms of use for it (International Data Spaces Association 2019). The **Eclipse Dataspace Connector**[5] should be mentioned here in particular. This enables companies to query data, exchange data, and technically design and monitor the policies that apply to the parties involved in the respective scenarios. In this context, the Eclipse Dataspace Connector is extensible and suitable for connecting multiple cloud implementations as well as for use in IDS as well as Gaia-X environments. Connectors can thus be integrated into the overall Gaia-X concept, as shown in Figure 1.1, as a secure way to exchange data. The conformity of IDS components with Gaia-X will be explained in detail in the reference architecture model version 4 (RAM 4.0), which will be published in spring 2022.

In addition, the central elements of Gaia-X, the so-called **Federation Services**[6], also match other components from the IDSs. The Gaia-X Federation Services "Identity & Trust", "Federated Catalogue", "Sovereign Data Exchange" and "Compliance" (see Figure 1.1) represent the minimum technical requirements and services needed for the operation of federated Gaia-X ecosystems consisting of infrastructure and data. To do so, they will rely on existing standards and open

3 https://www.gaia-x.eu/what-is-gaia-x.
4 https://europa.eu/next-generation-eu/index_de.
5 https://projects.eclipse.org/proposals/eclipse-dataspace-connector.
6 https://www.gxfs.de/federation-services/.

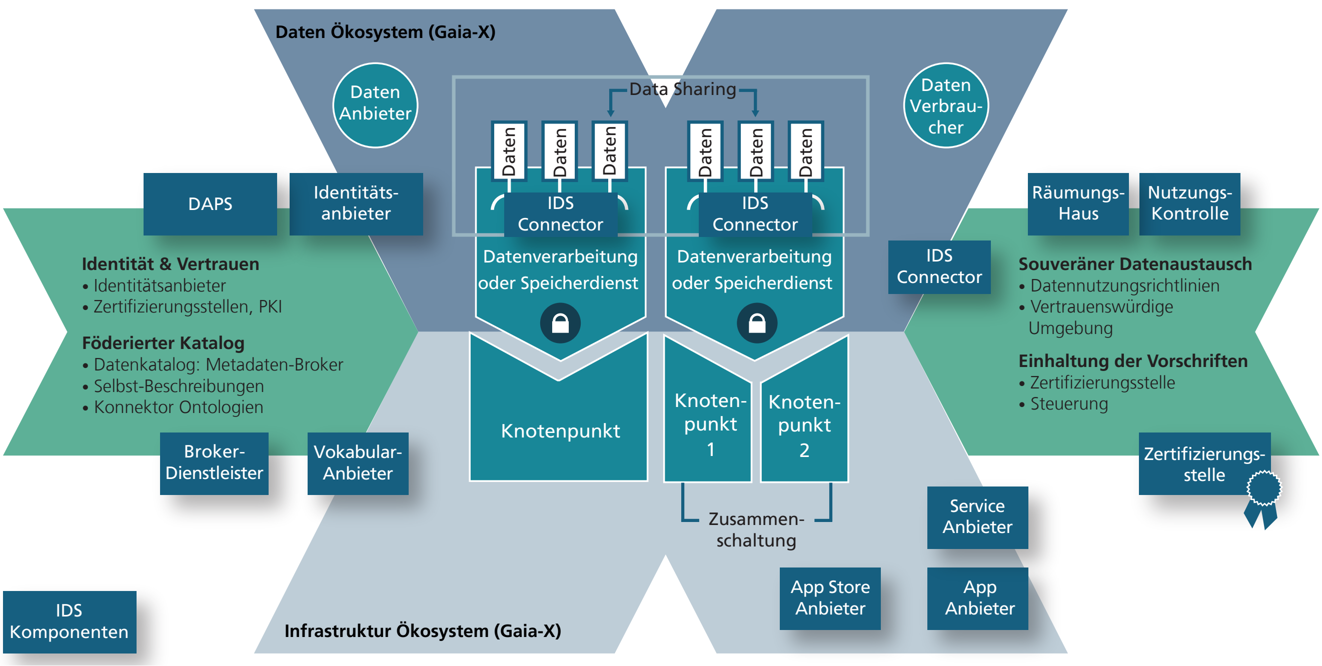


Figure 1.1: Integration and interaction of IDS and Gaia-X (own depiction based on (Otto et al. 2021)).

technologies, such as open source software. They will enable the trusted identification of participants and the exposure of existing data and service offerings. They will also provide solutions for the secure and transparent exchange of data as well as the technical verification of the compliance of players and their offered services.

Gaia-X and IDS thus show high potential for complementing each other to provide cloud and data sovereignty for end-to-end data value chains in federated ecosystems. Recognizing this exceeds the boundaries of the Gaia-X and IDS initiatives, they have joined forces with other participants as part of the Data Spaces Business Alliance (DSBA)[7] to pursue the building of data spaces and the promotion of the data economy as common goals (International Data Spaces Association 2021).

Through Federation Services, Gaia-X enables the merging of infrastructures where different actors can exchange data and its connected resources.

The Gaia-X architecture supports and enables data spaces for the development of advanced intelligent services. IDS provides an essential and attractive complement to Gaia-X to ensure secure and sovereign data exchange via connectors. Together, the IDS and Gaia-X initiatives thus contribute to breaking down barriers between companies and enabling complex value chains in which data can be exchanged and processed to enable innovative products and services.

Projects such as those of the IDS and Gaia-X initiatives create framework conditions for companies to promote their overarching data exchange. Section 1.3 also shows further conditions under which data can be generated and transferred.

1.3 Data sharing from a cross-company perspective

By taking a cross-enterprise perspective on data sharing, this Section aims to highlight factors and opportunities that generally apply or exist for companies sharing and exchanging data or seeking to do so in the future.

European initiatives, as presented in Section 1.2, aim to ensure the sovereign exchange of data, enable ecosystems in which companies and other actors can jointly develop innovative and data-based products and services. To comprehensively represent and explain such ecosystems, the IEDS project developed a representation that presents the different and interrelated ecosystem types as well as their technical components and interrelationships. The basis for this was the preliminary work from the DEMAND Use Case Report (Azkan et al. 2020b). The Building Blocks of Data Ecosystems presented in this paper have been further developed and specified in this regard based on literature such as Bansal and Kumar (2020); Oliveira and Lóscio (2018); Vargo et al. (2017), among others, as well as consideration of Gaia-X use cases. The aim of this is to develop or further sharpen companies' understanding of such ecosystems and also to identify points of reference where and how they could participate in the respective ecosystems. This results in the **Building Blocks of Ecosystems**, consisting of IoT (Internet of Things), Data and Service Ecosystems and their respective components. Figure 1.2 shows the different ecosystem types, in which the entire data value chain is considered, from data generation to data exchange in data ecosystems to the creation of innovative data-based products and services.

7 https://data-spaces-business-alliance.eu/.

Data Generators

Figure 1.2 shows the Building Block of data generators at the lowest level. They are the actual data sources in the form of technical components and systems. Today, data is increasingly generated at the edge of networks and by the devices located there (Asch et al. 2018). Such novel data sources here include sensors attached, for example, to physical objects such as machines, vehicles, or transportation infrastructure, which feed data into networks through them. As such, they become objects of the Internet of Things and thus essential components of IoT ecosystems. Other data sources include mobile devices such as smartphones and wearables, as well as social media.

IoT Ecosystem

The subsequent Building Block is that of the IoT ecosystems. The IoT layer connects the real world with the Internet. IoT ecosystems include large amounts of interconnected physical objects and form a system in which they are efficiently managed. In addition to the physical devices and the data they produce, they include the resources that enable and promote the networking of the real world with the virtual world, such as, among other things, the hardware and software solutions required for this purpose.

> In a business context, IoT ecosystems represent a community of interacting actors, such as companies or individuals, that use a common set of assets and resources to connect real physical objects to the virtual world of the Internet (Mazhelis et al. 2012).

Data Assets

Above this, the data assets are listed as the next Building Block in Figure 1.2. These were previously generated within the IoT ecosystems by the data generators. They represent the conditions of reality, provide virtual information about the states of physical objects such as machines, and form the basis for all further activities and the development of innovative products and services. Data assets can be, for example, machine operating data, position data or traffic data.

Actors and Roles

On top of the Building Block of data assets is the block of actors and roles. These can be companies, individuals or institutions. They can assume one or more roles and fulfill certain functions based on their respective competencies. All types of ecosystems include actors who interact with each other in their respective roles. These are also central components in data ecosystems. Generally, actors there provide data for others or use data sets themselves in order to carry out value-creating activities based on them.

Data Ecosystem

Data ecosystems are the next Building Block in this regard. A data ecosystem is a network in which the actors base their business relationships and interactions on data and exchange the data goods generated at the previous point across companies. In this context, data ecosystems encompass all stakeholders in the form of actors and roles, connecting them directly or indirectly within the network and existing value chains (Koskinen et al. 2019). In this way, they engage in co-creating value from which all actors involved can ultimately benefit. The actors' handling of data here takes place in a data value chain, which is shown in Figure 1.2 in the left margin. This can be divided into five stages in which the players involved can be active. However, this is not necessarily a linear process, as several loops may be run through the respective stages. The data value chain begins with the generation of data (1) before it is subsequently pre-processed and curated to increase data quality (2). To extract information from the data, it is then analyzed (3) and further combined with other information and experience from additional industries and domains (4). The resulting knowledge can now be applied to existing problems and used to make decisions in a business context (5).

> Data ecosystems and the data economy have a direct connection. Each data ecosystem represents a part of the data economy (Koskinen et al. 2019).

Data-driven Services

The Building Block based on the data ecosystems are the data-driven services. They result from the interaction of the various actors across multiple ecosystems, the data exchange carried out in the process, and the traversal of the data value chain, and represent innovative and data-driven services (Azkan et al. 2021). In data-driven services, the previously generated and further processed data assets are used as a central resource. These data-driven services aim to support decision-making in a strategic business context. The underlying value creation process for the creation of data-driven services can be divided into core and supporting processes that realize or support the development of data-driven services, such as the analysis of data or the provision of the required technical infrastructure. The cross-company exchange of data enables players to interact and their resources to be integrated within an ecosystem. Through this so-called value co-creation, data-driven services are created through which every actor benefits. This type of service is a central component of service ecosystems.

In service ecosystems, actors integrate their resources and are connected through mutual value creation and service exchange (Lusch and Vargo 2014).

Service Ecosystem

At the top level of the Building Blocks are the service ecosystems. As in the other ecosystem types, actors create mutual value in service ecosystems. In service ecosystems, this is achieved in particular by focusing on processes and outcomes to be provided or achieved through their interactions. To this end, they apply their respective competencies in the form of knowledge or skills and thus help each other to gain advantages. This is done in service ecosystems by sharing resources and exchanging the data-driven services previously developed on the basis of the generated data and through interaction within the framework of the Building Blocks presented as well as along the data value chain.

The Building Blocks of Ecosystems presented here depict the general framework in relation to data sharing in Figure 1.2. Subsequently, Section 1.4 also explains the requirements that companies must address internally if they want to share data beyond their borders.

BUILDING BLOCKS OF ECOSYSTEMS

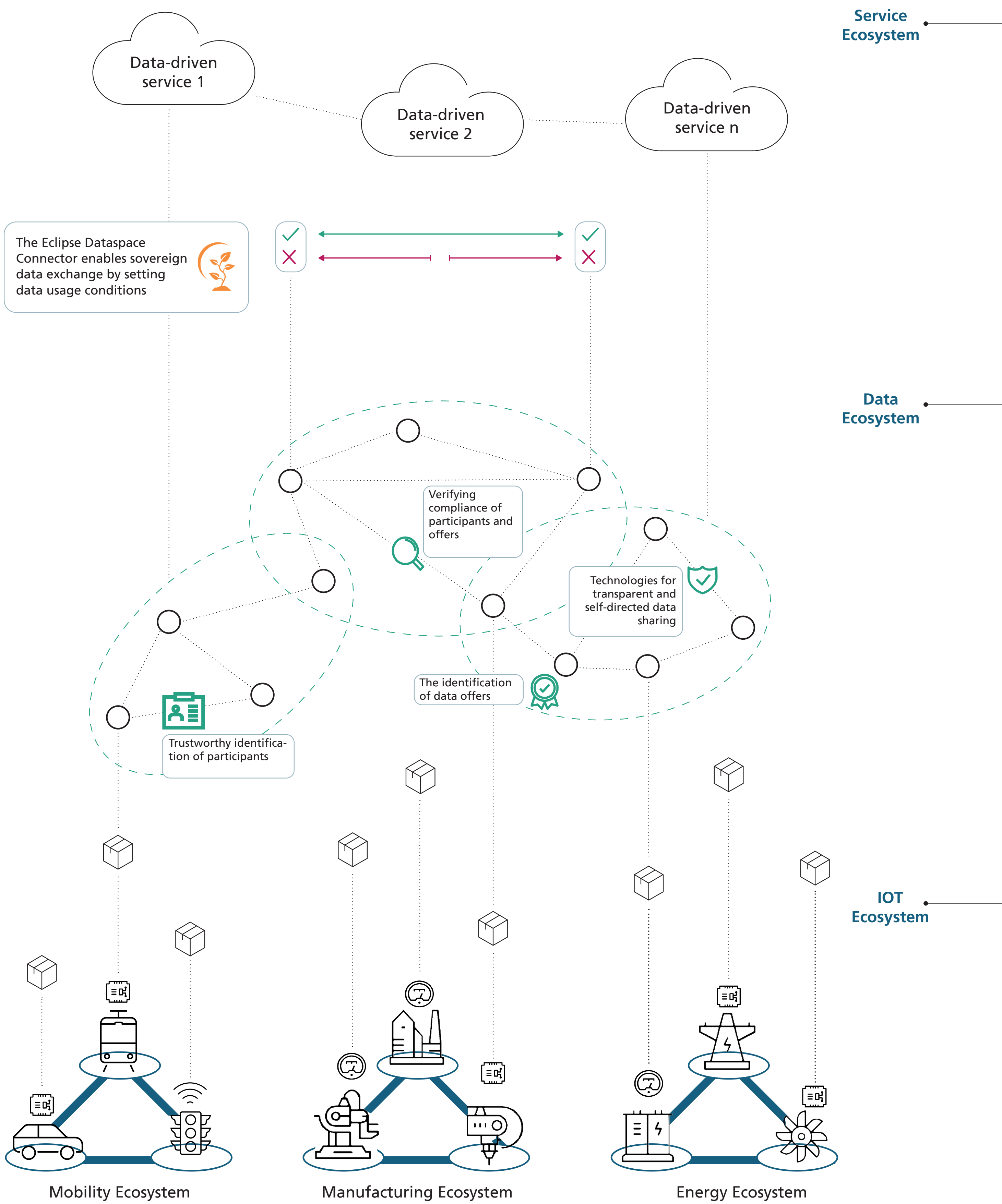


Figure 1.2: Building Blocks of Ecosystems from a company-external perspective on data sharing (own depiction based on (Azkan et al. 2020b))

SERVICE ECOSYSTEM

In service ecosystems, actors integrate their resources. Through joint agreements and the exchange of services, they create value together and mutually.

DATA-DRIVEN SERVICES

Data-driven services represent solutions and services based on generated and processed data. They are a combination of physical products and digital services and are able to meet needs individually and in line with requirements. They are both the result and the goal of the joint value creation of several players in data ecosystems.

DATA ECOSYSTEM

Within data ecosystems, data represent the central resources. The business relationships and interactions of the players are therefore also based on them.

ACTORS AND ROLES

Different actors, which can be both individuals and companies, act in ecosystems. They assume one or more roles in which they provide, mediate or consume data, for example.

DATA ASSETS

Generated data, which are subsequently used to create data-based services, represent data assets. These include vehicle or machine operating data.

IOT ECOSYSTEM

Internet of Things devices are the essential components of IoT ecosystems. These are physical objects equipped with sensors, through which data are generated and provided, consumed and further processed elsewhere.

DATA GENERATORS

They represent the sources of the data. More and more data is being generated at the edge of networks by the devices located there. These are, for example, IoT devices and a variety of sensors installed on physical objects.

1.4 Data sharing from an internal company perspective

In order for companies to be able to engage in value-creating activities through cross-company data exchange, it is necessary for companies wishing to participate in an ecosystem to consider various internal fields of action (here, the company-internal perspective). To this end, the IEDS project designed a three-layer framework model (see. Figure 1.3) consisting of the strategy, process and system levels. This model forms the basis for the development of incentives to encourage companies to participate in data sharing. The model supports the transformation of companies through the structured presentation of the essential fields of action. On the basis of this model, it is possible to work out internal company incentives that can be used to promote data sharing.

Strategy level

The strategy level represents the top level of this framework model and is elaborated by the company's management. Medium-term decisions with a time horizon of one to three years are defined at this level. The strategic handling of data within the company as well as the handling and exchange of data with other companies is determined. At the strategic level, a further distinction is made between data-driven business models and the data strategy. Business models are developed on the basis of strategic decisions that extend or complement the value proposition of the respective company. However, this first requires a data strategy in order to develop the data-driven business models on this basis and to adapt the processes in the company accordingly.

Data-driven business models (see Section 3.3):
Due to the increasing networking of companies, the basis for new types of business models based on data is being created. In the literature, these are described as entrepreneurial concepts of business activity that use data as a key resource for business activities and build activities on it (Azkan et al. 2020a; Hartmann et al. 2016). They offer new opportunities for creating new value streams in the form of business models, both in B2B and B2C environments.

Data strategy (see Section 3.2):
A company's data strategy describes a number of core aspects relating to defensive and offensive actions in dealing with data. Defensive actions are data security, integrity, quality, compliance and governance. In contrast, offensive actions involve improving one's competitive position and increasing profitability (DalleMule and Davenport 2017). When participating in ecosystems, the focus is particularly on the offensive components of the data strategy. Incentives of data sharing and the associated creation of data-driven business models are highlighted. Initially, the data strategy thus provides the basis for data-driven business models. Subsequently, there is a mutual influence, whereby the data strategy orchestrates the basic features of data-driven business models and developments of data-driven business models in relation to demand or changed demand influence the data strategy of the company.

Process level

The process level describes the current management and evaluation of the company's own data. This level is divided into data governance, data management, data evaluation and data analysis. In the project carried out, the focus is on the area of data management and data evaluation.

Data Governance:
Data governance provides the framework for data management in the company. In this context, roles and responsibilities are defined with regard to the management of data. This generates a basis for the processing, storage, maintenance and presentation of the data (Otto und Österle 2016).

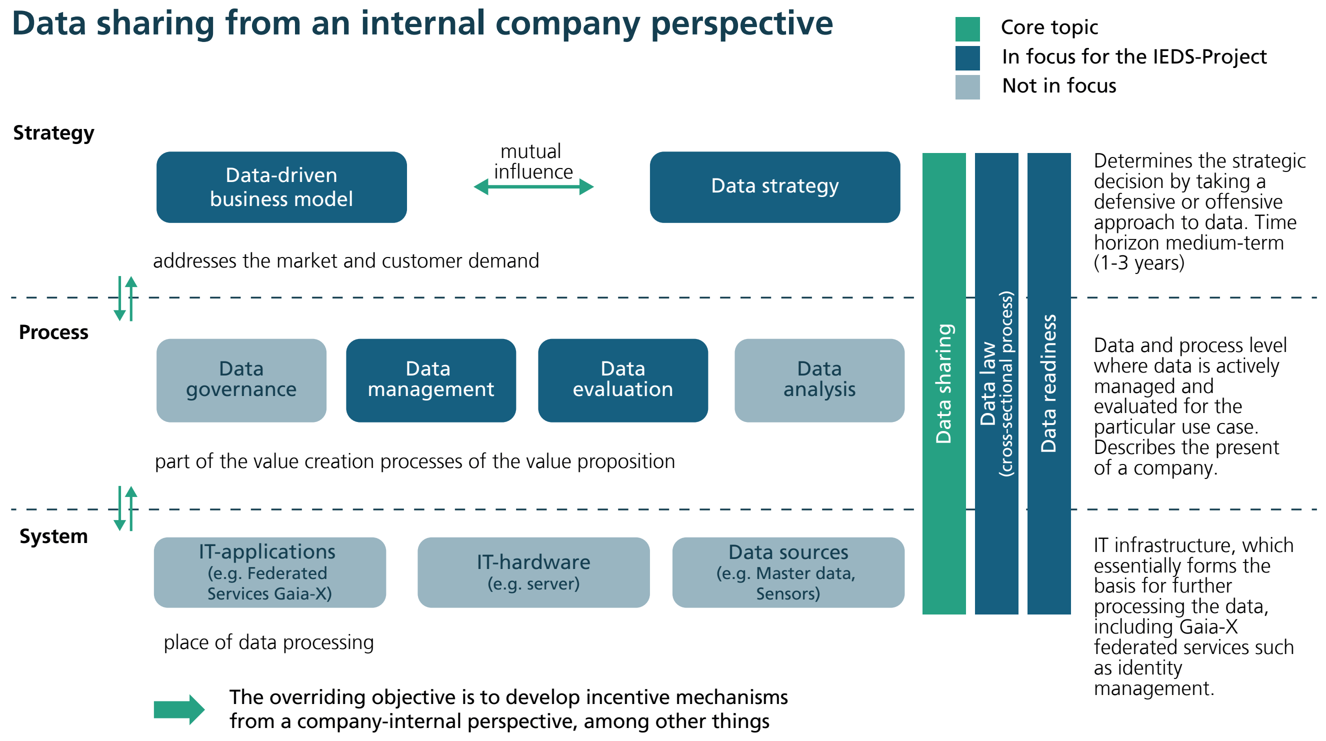


Figure 1.3: Data sharing from the company's internal perspective and project focus (own depiction).

Data management (see Section 3.2):
Data management is concerned with the development, execution, and monitoring of plans, policies, programs, and practices that provide, control, protect, and enhance the value of data and information assets throughout their lifecycle. In this context, the data management function is subject to the regulatory framework of data governance and thus to the lifecycle management of data, from planning and design to management as well as enhancement of data to data use and maintenance (Henderson und Earley 2017).

Data evaluation (see Section 3.4):
Data valuation represents another core aspect in the three-layer framework model with regard to incentives for data exchange. Here, data in the company is considered in terms of its quality, its processes, its performance, its costs, its benefits and its market values. In the context of data exchange with other companies, determining the market value is currently the most common method used to value data financially (Krotova et al. 2019).

Data analysis:
In data analysis, qualitative and quantitative data are used within the framework of various analysis options in order to gain insights for both the company's own processes and also for external processes. The various data analysis options include descriptive, predictive and prescriptive analyses which, depending on their complexity, can make statements about possible future events (Hupperz et al. 2021). For a company, this derives an incentive to invest in the analysis of its own data. Furthermore, the results can potentially be sold at a profit in addition to being used by the company itself.

System level

The lowest level of the framework model is the system level, which is formed by the application software, IT hardware and data sources that lie behind the processes. These lay the technical foundation for data exchange, but are not the focus of the IEDS project, but are nevertheless mentioned for the sake of full representation.

IT applications:
Application software, such as Gaia-X Federation Services Identity Management, forms the basis for the exchange of data within the Gaia-X ecosystem. Identity Management is a building block of the minimum technical requirements and services for the operation of the Gaia-X federated ecosystem (GAIA-X 2021).

IT hardware:
To be able to store, process and ultimately share data, companies first need IT hardware, for example, to enable data to be stored on-site or in a public cloud.

Data sources:
For data to be generated at all in the first step, data sources or generators and relevant data sets must first be identified. This concerns master data, metadata, but also, for example, sensors that generate data in the production or logistics process.

Cross-cutting issues

In addition, there are three cross-cutting topics in the framework model whose content must be considered across the board. These are data sharing, data law and data readiness.

Data sharing (see Section 3.1):
Data sharing encompasses all aspects of the model shown in Figure 1.3, from the data strategy that defines data sharing as a goal, to the data flow procedures, to the technical execution at the system level. This means that enterprise operations across all model levels must be aligned with the goal of data sharing in order for it to occur at all at the technical level and in alignment with the enterprise's data strategy. In data sharing, a distinction is made between internal company data exchange and data exchange with external companies.

Data law (see Section 3.5):
Data law must always be taken into account in the process of developing data-driven business models, the value creation process and data processing. Both in the definition of the data strategy and in the development of the data-driven business model, feasibility must be evaluated in the first step, taking into account regulations such as the GDPR in the area of data exchange. This also applies to the process and system level, where the handling of the relevant data must be carried out in accordance with data law (Pandit et al. 2018). Furthermore, upcoming legal requirements, such as the EU Data Act, must also be taken into account in the process (European Commission 2021).

Data Readiness (see Section 2.1):
Data readiness, or data economy readiness, describes a company's ability to process and use data. This ability is also called the maturity level of data management. The types of data that companies store play a role here, as do the forms of data management implemented and the purposes for which the company uses data.

The current status of data sharing by companies and their ability to participate in the data economy is discussed in more detail in the following Section 2.

2 Status quo of the German data economy

In the following, the status quo of data sharing and the data economy in Germany will be examined. To this end, we will first examine whether companies meet the requirements for sharing data and managing it jointly with other companies (see Section 2.1). Since the Gaia-X initiative is developing an infrastructure to enable companies to jointly manage data, Section 2.2 presents measures for analyzing the reach and acceptance of the initiative.

2.1 Data economy readiness and data economy

Companies must meet certain requirements in order to manage data jointly with other companies. On the one hand, they must be technically and organizationally capable of managing data efficiently. On the other hand, they must also be ready and willing to pass on their own data to external parties or to use data from other companies. Data economy readiness is a prerequisite for participating in Gaia-X (see Section 1.2).

In the Data Economy Readiness part of the project, a representative survey of 1,002 companies from industry and industry-related service providers (survey period September to November 2021) will be used to examine the extent to which companies in Germany are able to manage data efficiently. In addition to their own data economy readiness, the survey asks to what extent joint data management with other companies plays a role. In order to determine the number of companies for which the establishment of a sovereign and secure cloud infrastructure is relevant, the companies are also asked about their cloud usage behavior.

In particular, it is possible to identify the proportion of companies that meet the requirements for participation in the data economy and may also operate in-house data management, but have not shared data with other companies to date. Any barriers can be identified in this way.

Core results of the survey:
Building on and connecting to existing maturity surveys (Röhl et al. 2021; Demary et al. 2019), we will determine the extent to which companies are data economy ready or not, i.e., are able to manage their data efficiently. For this purpose, the response behavior of companies on the following three aspects of data management is relevant:

- Data storage
- Data management
- Data utilization

The data storage aspect asks which types of data companies store digitally. This includes, for example, product, process or personnel data that mostly relates to the company's own production or workforce, as well as supplier data or customer usage data that relates to corporate partners or actors outside the company.

The data management aspect examines how companies handle their data. For example, it considers whether internal company data is passed on via standardized and permanent interfaces, whether data is classified and quality checked, or whether companies regularly look for new data sources and ways of using data.

In terms of data use, it determines the purposes for which companies use data. The purposes queried include the (further) development of products, services or business models. In addition, data can be used for automation and control or offered for direct or indirect sale.

Based on the response behavior, the companies are divided into the groups 'data economy ready' or 'not data economy ready'. Figure 2.1 shows the results.

29 percent of all companies surveyed are data economy ready. The vast majority of 71 percent are not data economy ready, i.e., they cannot manage their data efficiently.

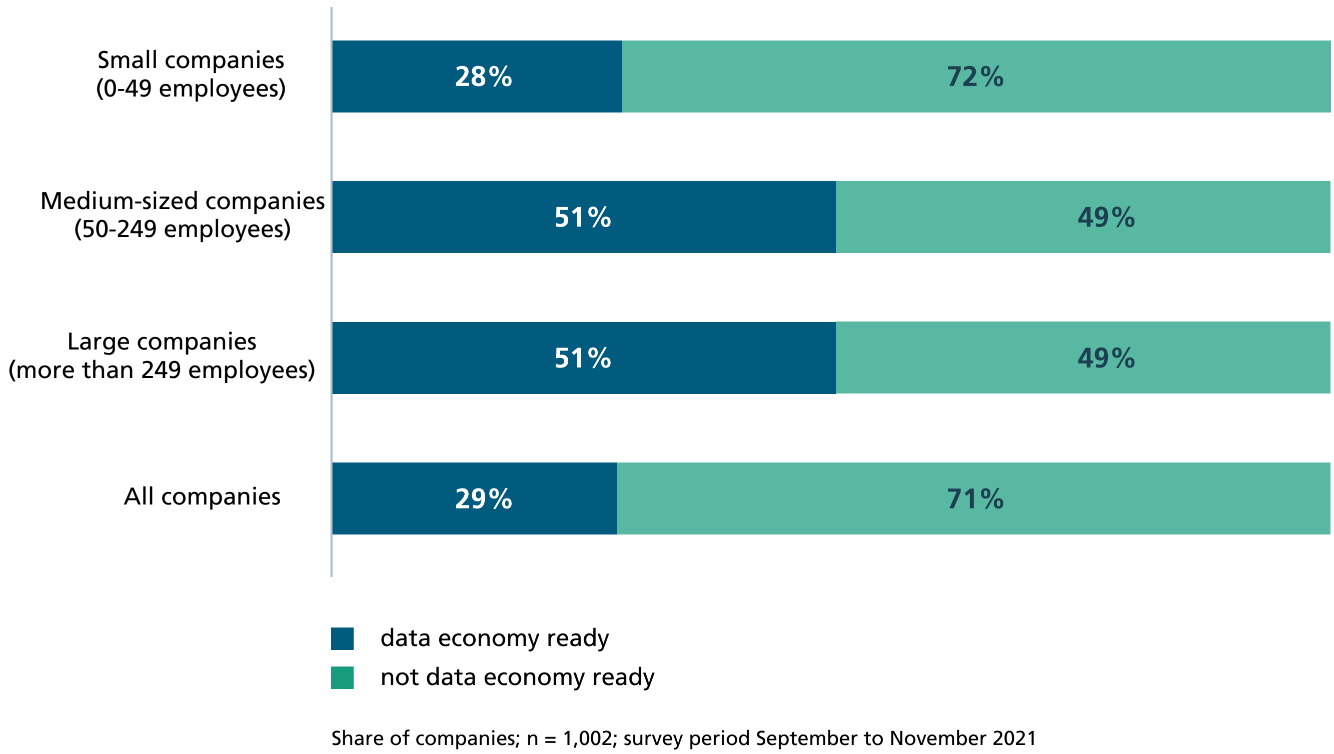


Figure 2.1: Data economy readiness (Institute of the German Economy)

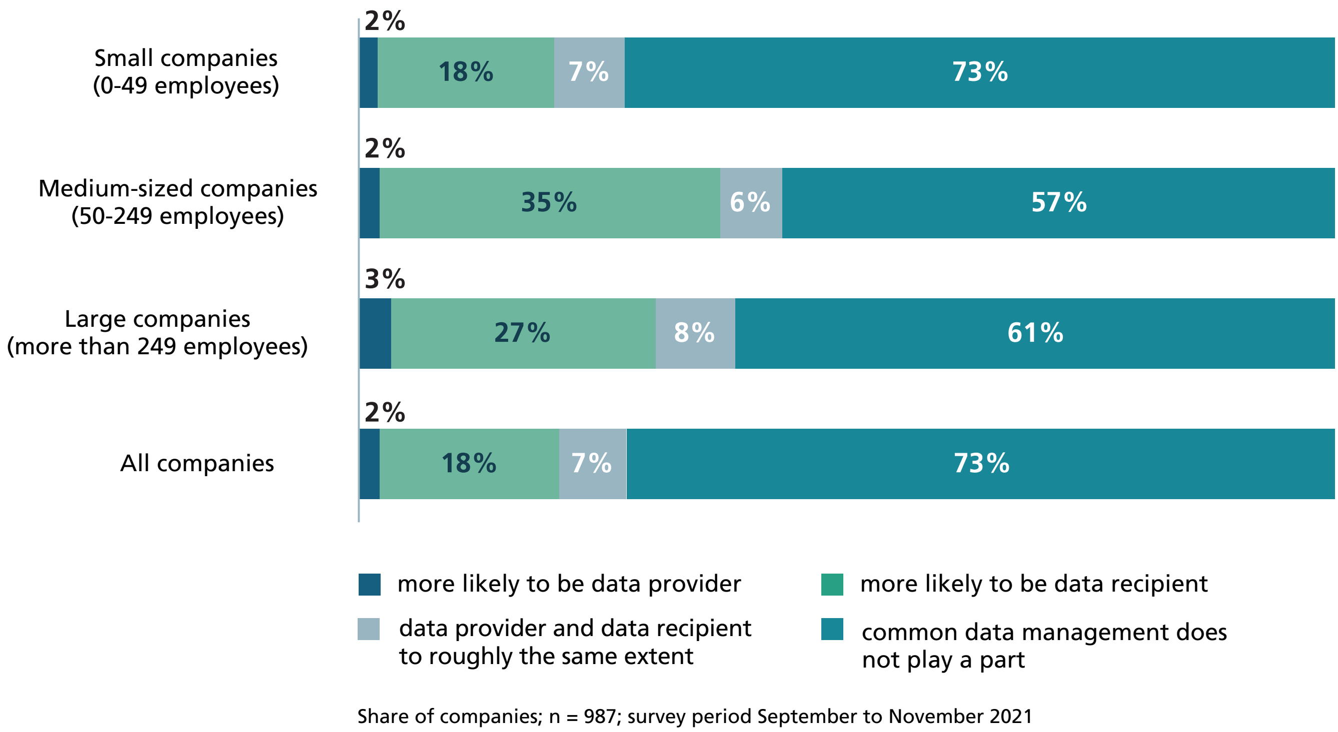


Figure 2.2: Data sharing (Institute of the German Economy)

Medium-sized and larger companies perform better than small ones. Around half of these companies with at least 50 employees are data economy ready. The figure for small companies with up to 49 employees is 28 percent.

In a second step, the companies are asked whether data sharing plays a role for them and whether they are more likely to be data providers or data recipients (see Figure 2.2).

For 73 percent of all companies, data sharing plays no role. 18 percent of the companies are more likely to be data recipients, two percent are more likely to be data providers and seven percent are data providers and data recipients to roughly the same extent.

As in the case of data economy readiness, the performance of small companies with fewer than 50 employees is close to that of Germany as a whole. This is due to the fact that the company size structure in Germany is characterized by a very large number of small companies (Destatis 2020).

At 43 percent, the proportion of companies for which shared data management plays a role is highest among medium-sized companies, ahead of large companies at 38 percent. This result is surprising in that the proportion of companies that are data economy ready is the same for both size classes. The higher relevance of shared data management among medium-sized companies compared with large companies is primarily due to the significantly higher proportion of data recipients at 35 percent (large companies: 27 percent). Medium-sized companies are thus often more open to the use of external data. One explanation could be that medium-sized companies are more dependent on external partners and are therefore more likely to rely on external data. On the other hand, large companies have slightly higher percentages of data providers and of "roughly equal numbers of data providers and data recipients".

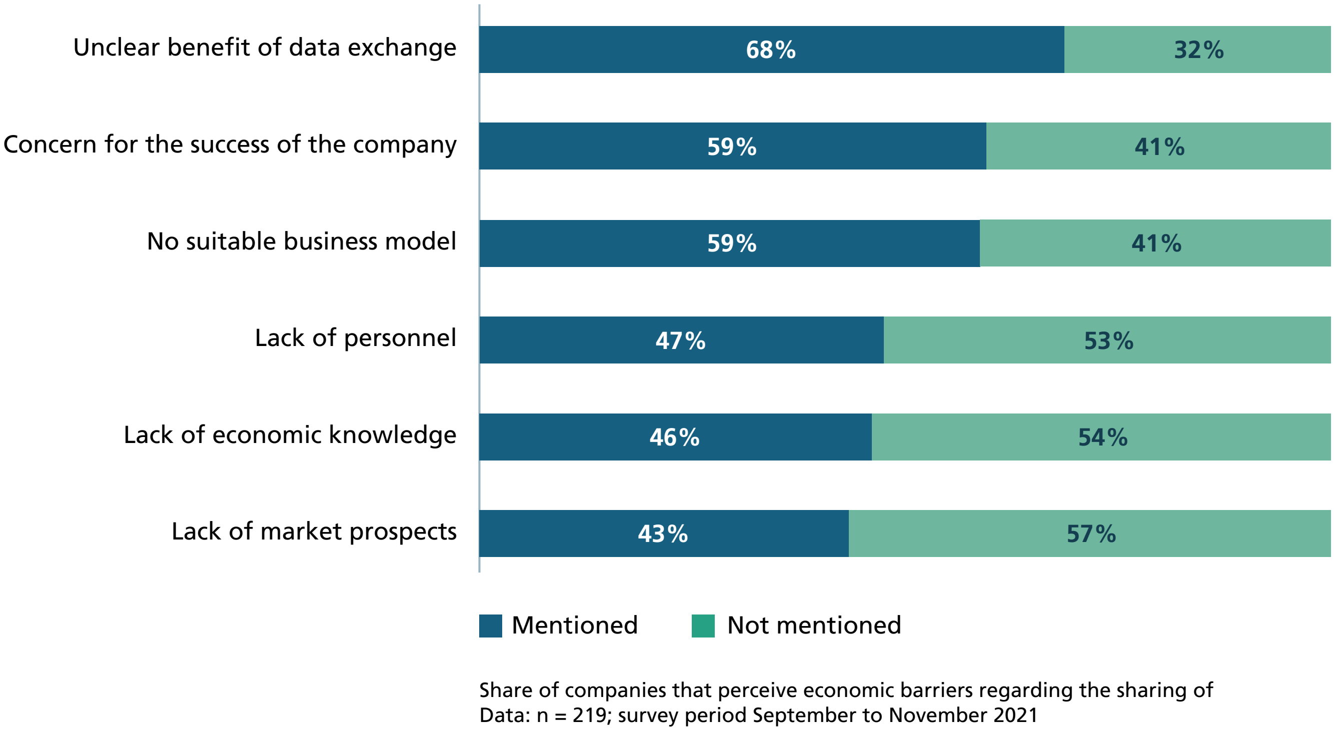


Figure 2.3: Economic barriers to data sharing (Institute of the German Economy)

Overall, the majority of all companies see themselves more as being recipients of data provided by third parties. Only two percent tend to be data providers. Sharing their own data plays a very minor role among the companies surveyed. This points to barriers to data sharing that are presumably more present in data sharing than in data use by third parties. Specifically, companies cite the following economic barriers related to data sharing (Figure 2.3).

68 percent of all companies that perceive economic barriers cite the unclear benefits of data sharing as an economic barrier. In each case, 59 percent of the companies cite concerns about their own company's success or the lack of a suitable business model. Other economic barriers include lack of personnel (47 percent), lack of economic knowledge (46 percent) and lack of market prospects (43 percent).

In the other work packages of the project, measures are being developed to help companies overcome these barriers. Legal barriers that also exist are described in more detail in Section 3.5.

In addition to promoting a secure and networked infrastructure for shared data management, the provision of trustworthy cloud services for enterprises makes up a key application area of Gaia-X. Cloud services can include, for example, mail, office and CRM software as well as virtual servers or on-demand computing power.

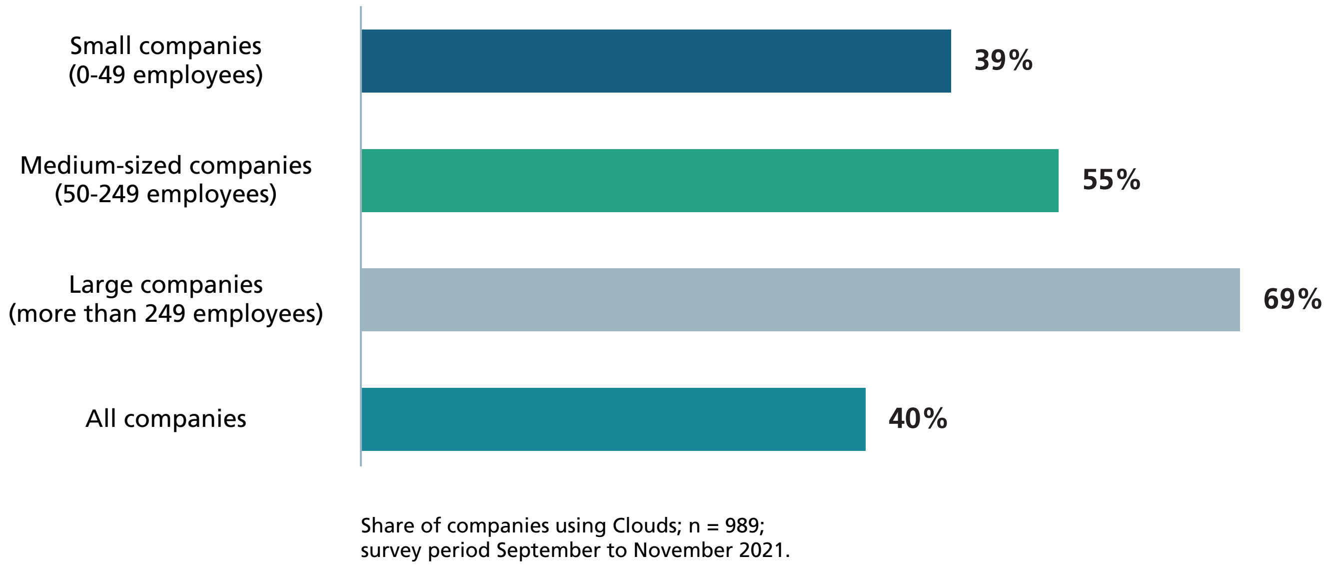


Figure 2.4: Cloud usage (Institute of the German Economy)

Figure 2.4 illustrates the status quo of cloud use by companies in Germany.

Across Germany as a whole, 40 percent of the companies surveyed use cloud services as users. This means that the share of cloud use is in line with the growth trend identified in other surveys. For 2020, for example, a Destatis survey determined a cloud usage share of 33 percent of companies with 10 or more employees, compared with 22 percent in 2018 (Destatis 2021).

The proportion of companies using cloud services increases with company size. 39 percent of small companies, 55 percent of medium-sized companies and 69 percent of large companies use cloud services. Cloud use has already become part of everyday business life for a large proportion of companies in Germany. However, there is still a lot of untapped potential, especially if companies currently only use rather low-threshold cloud services such as mail software and more advanced cloud services such as on-demand computing power do not (yet) play a role.

The points of contact between companies and cloud services are discussed again in the following Section and related to the awareness of Gaia-X.

2.2 Monitoring Gaia-X

Cloud computing offers numerous advantages for companies. For example, cloud solutions enable every employee to access the data and software they need from any location. In addition, storage and computing capacities can be easily added or removed in order to react cost-effectively to current requirements. Accordingly, the use of cloud computing has increased sharply in recent years (see Section 2.1). US and Asian companies in particular have benefited from this, as they account for the largest market shares in global and European sales

(Statista, 2021; KPMG, 2021). Against this backdrop, the call by Federal Minister of Economics Altmaier in 2019 for a European cloud infrastructure to ensure data sovereignty, among other things, can also be explained (BMWi, 2019). To this end, the Gaia-X initiative (see Section 1.2) was presented in the second half of 2019 (BMWi and BMBF 2019).

Core results:
Awareness and knowledge of Gaia-X among companies is still low. Based on the search queries for the terms "Gaia-X" and "cloud computing" on the Internet search engine Google (Figure 2.5), it is clear that, despite major fluctuations, searches for the topic of cloud computing are more or less constant. The period here covers from the second half of 2019 to the present day. It is clear that at each point in time there were fewer searches for Gaia-X than for cloud computing itself. Also, starting from a low level, there is no clear growth in search interest for Gaia-X. In this context, the results of a survey from September and October 2020 (Röhl et al. 2021), according to which only 6.5 percent of the surveyed companies from industry and industry-related services and ten percent of the surveyed companies classified as digital said they were aware of Gaia-X, are not surprising. In the survey for this project (see Section 2.1), more than nine percent of all companies surveyed stated that they were familiar with Gaia-X.

Against the background of the growing importance of cloud services in general as well as the intended benefits of Gaia-X, such as the secure sharing of data across different cloud providers, it can be assumed that awareness of Gaia-X will increase in the near future. In addition, Gaia-X is still in the development phase (Gaia-X European Association for Data and Cloud AISBL 2021), after which progress in terms of awareness and use can also be expected.

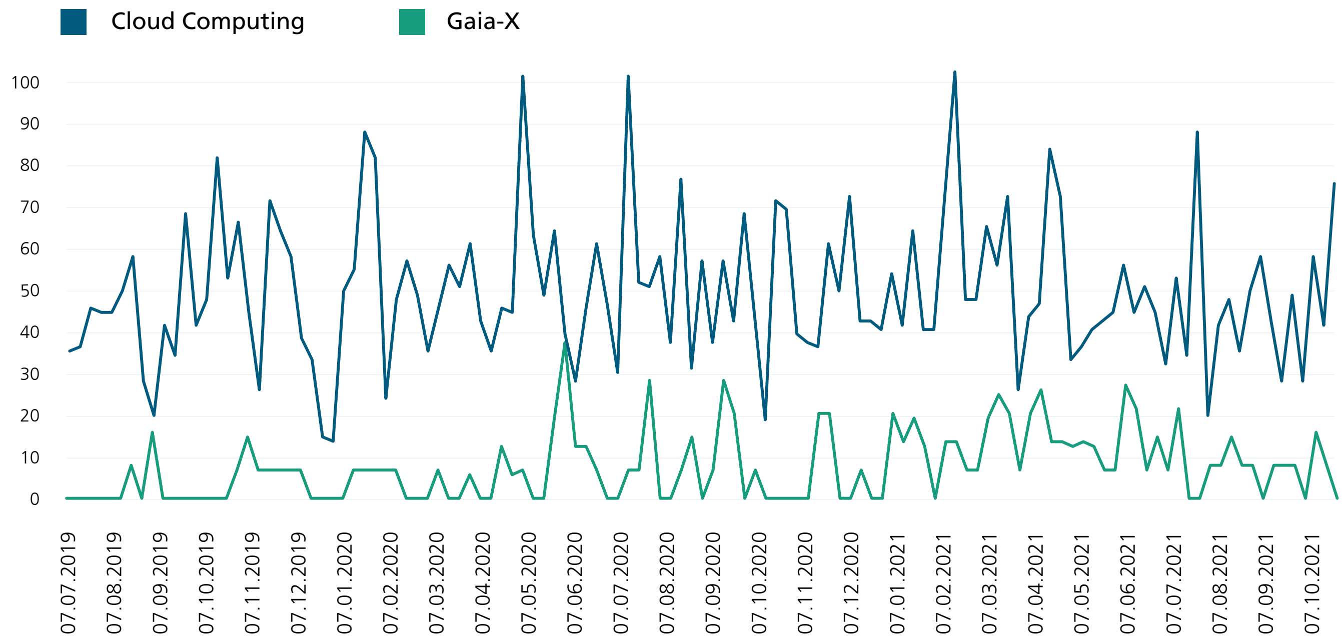


Figure 2.5 Indexed weekly searches for "Gaia-X" and "cloud computing" on Google Search from July 01, 2019 to October 24, 2021. As of October 28, 2021; searches from Germany relative to the maximum number of recorded searches in the period, which is given the value 100; date indicates the last day of the respective week.

In order to empirically track the future penetration of Gaia-X in business, the public and research, a Gaia-X monitoring system will be established. Using a dashboard, this will enable low-threshold, detailed and up-to-date tracking of awareness as well as the type and scope of discussion on Gaia-X. In addition, successes can be communicated and political recommendations for action can be made. In addition, successes can be communicated and political recommendations for action can be made. To ensure clarity, the dashboard will be limited to approximately 5 to 10 individual indicators that contain meaningful information about Gaia-X with regard to different dimensions or areas. These come from four categories that are critical to the success of Gaia-X (see Figure 2.6). These four categories are:

Public:
In the public category, indicators are used that allow statements to be made about the interest and receptiveness of society towards Gaia-X. The project can only be successful if the public and companies show interest in Gaia-X and, if necessary, perceive it as a quality feature. Possible indicators are the frequency that Gaia-X is mentioned in the news or in tweets on Twitter.

Research:
The role of Gaia-X in research projects and scientific publications can also provide information about the current importance of Gaia-X. Corresponding scientific findings represent possible added value for the use of Gaia-X and can thus drive its establishment.

Economy:
Businesses play an important role in both the demand and supply of data and applications via Gaia-X. Thus, the economy is also a central factor for the success of the project. Therefore, indicators from this area are also used. Possible indicators are the thematization of Gaia-X or more general data competencies in job advertisements or the use of corresponding terms on company websites.

Technology:
Technological development and the corresponding offering of services and software related to Gaia-X are also critical success factors. Therefore, activities relating to Gaia-X on the Github programmer platform are also included in monitoring.

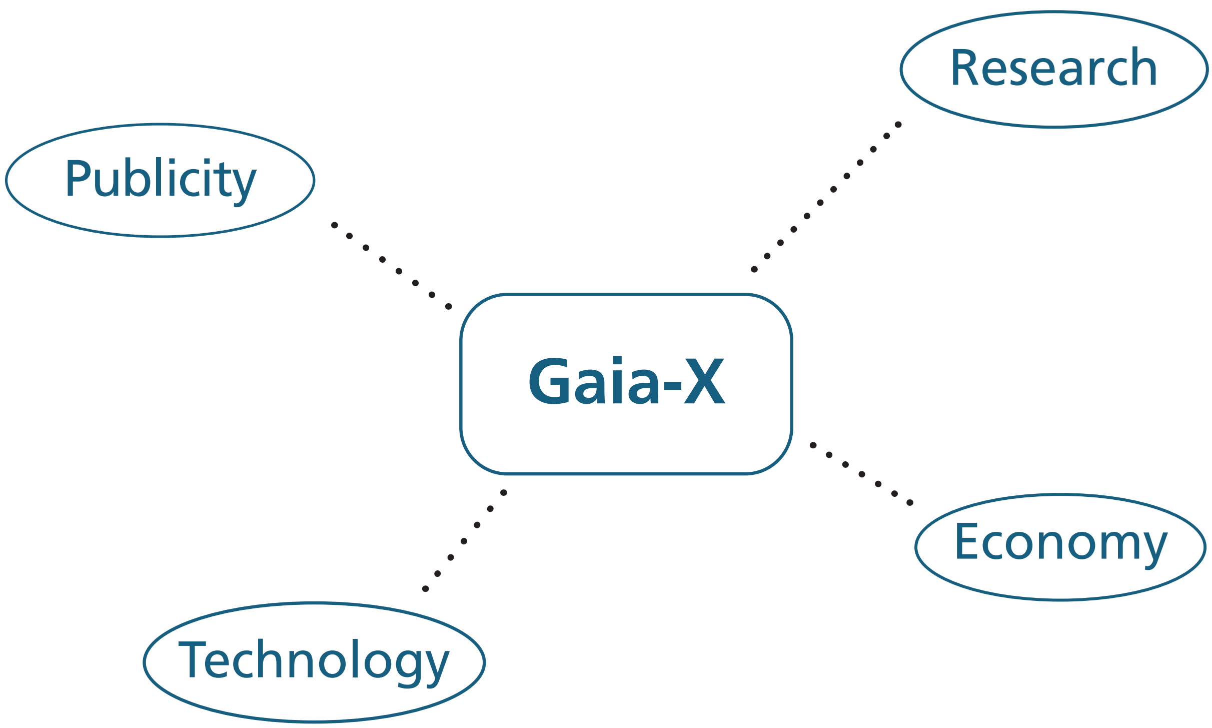


Figure 2.6 Dimensions of Gaia-X monitoring (own depiction

2.3 Pretest survey experiment incentive mechanisms

Another aspect of the project will investigate the response of German companies to various incentive mechanisms for sharing data. For this purpose, company surveys will be used. Specifically, survey experiments will be conducted to test the response to different incentive scenarios. As part of this methodology, companies will be randomly assigned to one of several groups, with each group receiving a different version of the questionnaire. The different versions of the questionnaire vary only in that they show different incentives to share data. Using statistical methods, it is possible to evaluate the extent to which decision-makers in German companies respond to different incentives for the questions that are identical in all versions.

In the course of the quarterly ZEW Business Survey Information Economy, a pretest was implemented to test this method of analysis. During the field phase (December 2021 - January 2022), about 750 companies from the information economy sector and about 450 companies from the manufacturing sector participated in the survey. The focus of the pretest was on the importance companies attach to collecting and analyzing data and what medium-term plans they are pursuing with regard to data use. A detailed investigation of whether the companies' assessment also depends on the information they receive before completing the questionnaire was to be performed. For this purpose, three different versions of the questionnaire were created, differing only in the introductory text at the beginning of the questionnaire. The companies were randomly assigned in advance to the control group or one of the two treatment groups and accordingly received one of the following three introductory texts at the beginning of the questionnaire:

Control group: "*With increasing digitization, more and more data is also being created. The following questions therefore relate to the use of data in your company.*"

Treatment group 1: "*With increasing digitization, more and more data is also being created. This opens up new opportunities for companies to collect and analyze data. However, as scientific studies show, the untapped potential in most companies is huge. The following questions therefore relate to the use of data in your company.*"

Treatment group 2: "With increasing digitization, more and more data is also being created. This opens up new opportunities for companies to collect and evaluate data in order to further develop processes, products or services. As scientific studies show, companies can increase their innovation activity, productivity and profits through data use. The following questions therefore relate to the use of data in your company."

While the introductory text for the control group only introduces the topic, the introductory texts for the two treatment groups contain additional information on the topic of data use. Treatment 1 emphasizes that although new opportunities for data use are now opening up for companies, at the same time the untapped potential in many companies is great. Treatment 2, on the other hand, focuses on the potential uses and positive effects of data use on corporate success. In order to measure the effect of these treatments, the questionnaire went on to record in what form the companies would like to use data in the future and to what extent they consider the respective use to be important for their company's success.

Core results:
The pretest described above first ensured that the randomization of companies to the three groups worked reliably and that the methodology used for allocation could be retained for follow-up experiments. In addition, the pretest showed that the sample size of around 1,200 participating companies in total was sufficient to identify statistically significant treatment effects in an experiment with three groups.

As an example of the content results of the pretest, Figure 2.7 presents a question block of the survey experiment. The graph indicates for which purposes the companies plan to use data in the next two years. The respective company shares are shown for both the control group and the two treatment groups. The majority of the companies surveyed in the information economy and in manufacturing plan to use data for the purposes covered in the questionnaire. It also shows that the approval ratings in both treatment groups tend to be higher than in the control group. In the information economy, for example, 82 percent of the companies surveyed in the control group plan to use data to control or improve processes in the next two years. In treatment group 1, on the other hand, the proportion is 89 percent and in treatment group 2 it is 86 percent. The values marked with an "*" in Figure 2.7 deviate from the value for the control group at a statistically significant level ($p<0.1$). All significant differences between the treatment groups and the control group are positive. Thus, the information contained in the introductory texts leads to a higher proportion of companies stating that they intend to use data for a specific purpose in the future.

In summary, the responses to the survey experiment offer a detailed view of how information can affect entrepreneurs' planned decisions. From such a focused view on the effects of information incentives, conclusions can also be drawn about the future effects that larger information campaigns or the dissemination of the results of the IEDS project could have. Furthermore, with the initial findings from the pretest, survey experiments can be designed to compare different incentives for sharing data. Sharing internally generated data across organizational boundaries can come with many caveats. Different treatments will be used to test which mechanisms can counteract these reservations, reduce hurdles and thus increase the willingness to share data.

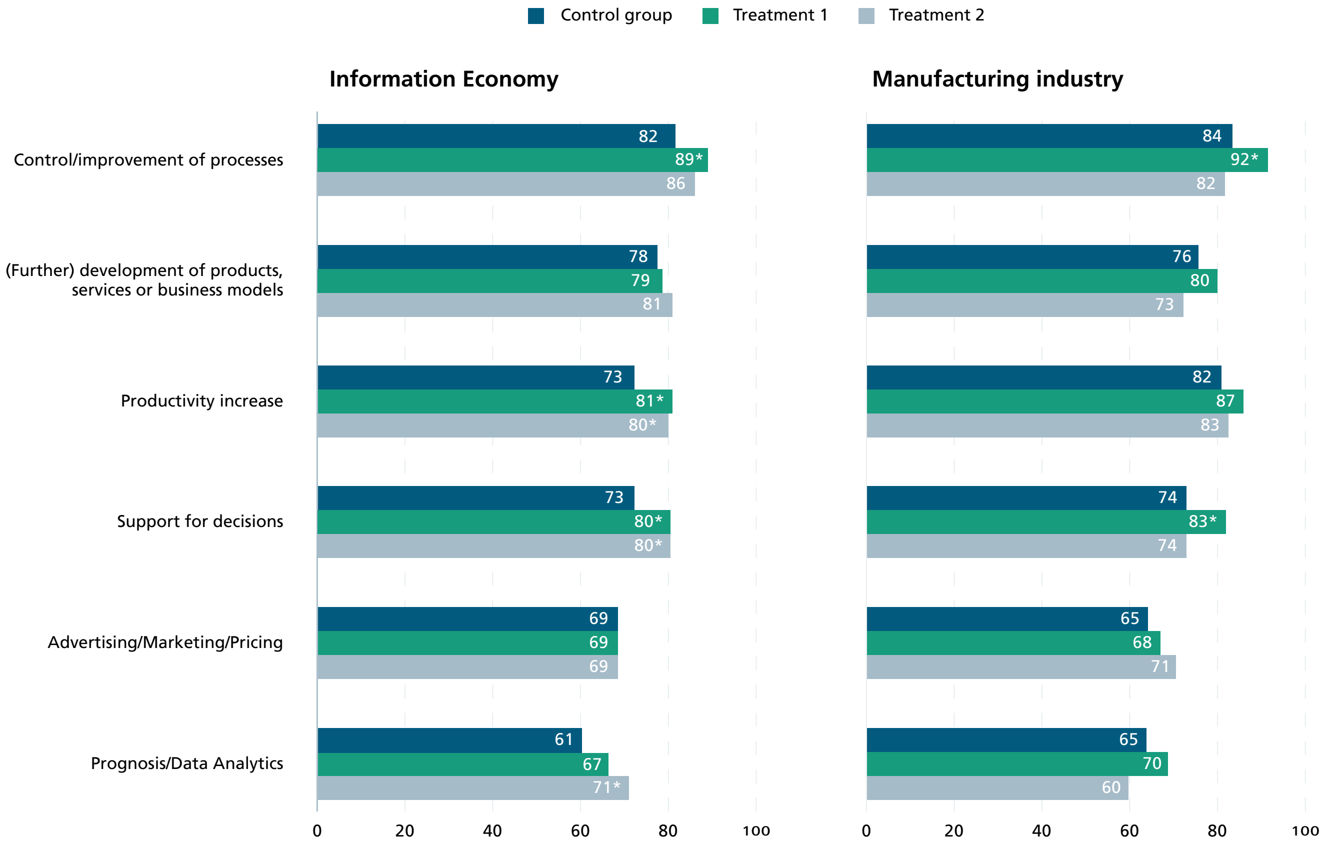


Figure 2.7: Purposes for using data, according to control and treatment groups. Reading aid: proportion of companies surveyed that answered yes to the question "Do you plan to use data for these purposes in the next two years?". In the control group, 82 percent of information economy respondents said they planned to use data to manage and improve processes in the next two years. In the group presented with Treatment 1, this figure was 89 percent. Values marked with an asterisk differ at least at the 10 percent level significantly from the control group.

3 Incentives and requirements to participate in data sharing – research overview

The central aspect of the IEDS project is to identify incentives for and requirements of organizations to participate in data ecosystems and data sharing. The analysis of these research aspects takes place in focus topics that take a holistic view of data sharing and identify possible potentials, challenges and incentives. The elaboration of the results in the IEDS project will be carried out in the form of collaborative research by analyzing the incentives and requirements for data sharing in a business context from both technical and economic perspectives. The analysis considers the focus topics of data strategy, data management, data-driven business models, data valuation and data law (see Figure 1.3). In addition, to ensure a holistic view of data sharing and its incentives and requirements, a research overview of data sharing was developed (see Section 3.1), which provides a catalog of requirements and a research map. In this Section, the first results of the IEDS research are presented.

3.1 Research overview "data sharing"

Cross-company data exchange opens up new strategic opportunities for companies to use data as a resource for the further development of their own processes and products (Richter und Slowinski 2019). According to current literature, there is not yet a uniform and differentiated definition of the term 'data exchange'. However, the following definition within the framework of the German government's data strategy describes the central aspects:

> In data sharing, »[...] different actors [...] share data with third parties or use them jointly on the basis of commercial or non-commercial agreements, or on the basis of mandatory legal requirements« (Bundeskanzleramt 2021, p. 110).

The value of the data economy in Europe at the end of 2019, estimated by the European Commission at around 324.86 billion euros, shows why the exchange of data is associated with such great potential and is also of particular economic interest to companies. In the next few years, growth is forecast to a value of an estimated 829 billion euros within the EU-27 states by 2025 (Mildebrath 2021). Nevertheless, companies are often unaware of the intrinsic value of their own data resources and the benefits of participating in cross-enterprise exchange (Parvinen et al. 2020; Azkan et al. 2020c).

Legal standards such as the General Data Protection Regulation and the planned Data Governance Act define the framework for ensuring the exchange and protection of data in accordance with European standards. (European Union 2016; European Commission 2020c). Thus, the existence of an appropriate legal framework forms one of numerous requirements for cross-company data exchange. Based on this requirement, the European initiative Gaia-X, launched in 2019, is addressing goals such as the creation of a secure data infrastructure (see Section 1.2).

Catalog of requirements for cross-company data exchange

A catalog of requirements for cross-company data exchange serves to list the necessary conditions that should be met in order to motivate the various actors to exchange data and also to support them during the process. Initial requirements were derived from literature, which were adapted and supplemented in an iterative process through workshops and Gaia-X use cases.

Examples of general requirements for cross-enterprise data exchange are trust and transparency, which have been discussed particularly prominently in the literature (Dahlberg and Nokkala 2019). Still, a lack of trust between organizations is a major obstacle to data sharing, which trust-based structures as well as identification processes and the use of encryption technologies attempts to counter (Gelhaar and Otto 2020; Dahlberg and Nokkala 2019). Equally significant, but at the same time difficult to realize, is the requirement for transparency. This poses major challenges for data exchange, as transparency must be enabled by the individual actors themselves through insights into their data. On the other hand, the data must nevertheless always remain protected.

Subsequently, the literature-based result was further differentiated in several iterations and thus completed in stages. This included conducting two workshops and aligning the requirements with the conceptual use cases of the Gaia-X project. The two workshops were conducted with differently composed groups of experts, all working in different work packages of the IEDS project. Through the workshops, the requirements of the literature were complemented, further specified with additional relevant sub-aspects, and assigned to higher-level categories. Selected requirements for cross-company data exchange are shown in Figure 3.1. It can be seen that the requirements can be categorized according to their thematic focus. In the course of the individual iterations, it became clear that a distinction must be made between general requirements (1) and requirements relating to the legal framework (2) and technical implementation (3). In addition to the workshops that were held, a targeted link was created by comparing the literature-based requirements with a large number of Gaia-X use cases. In the process, additional requirements emerged as

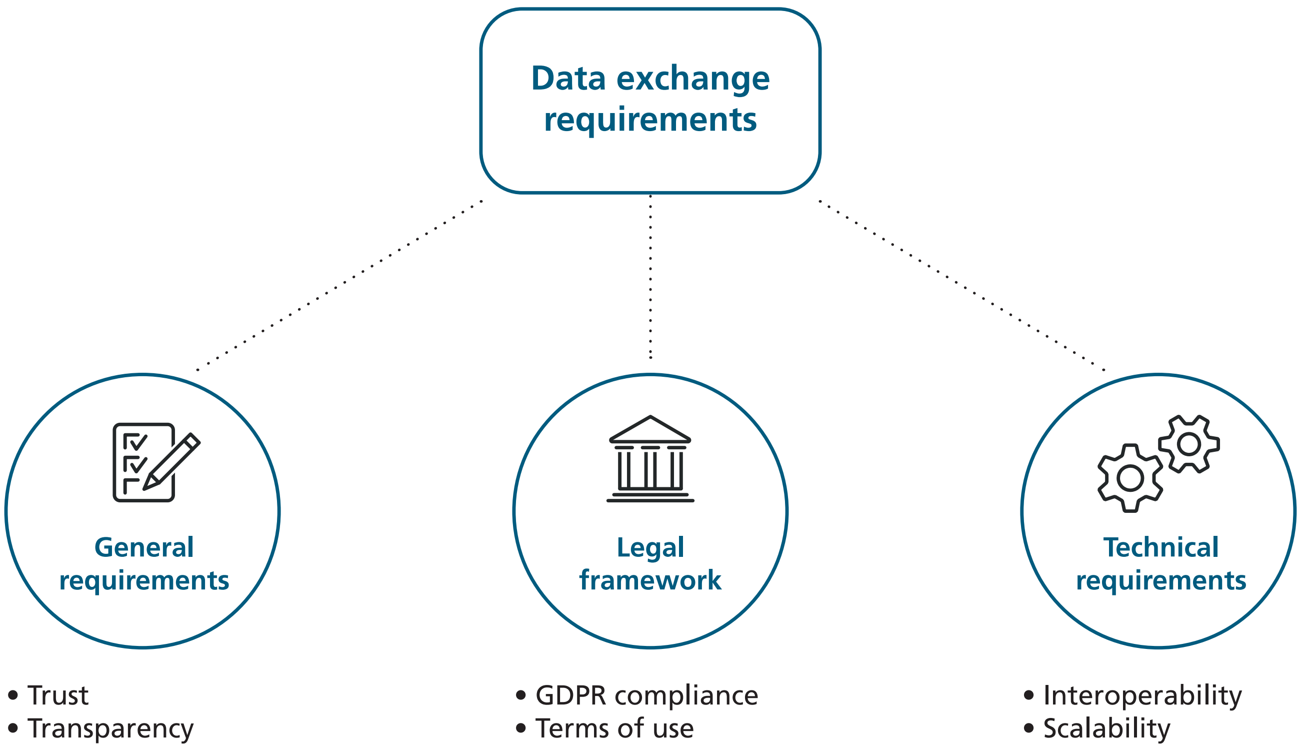


Figure 3.1 Exemplary requirements for cross-company data exchange (own depiction based on the results of the evaluation carried out)

significant for data exchange, such as scalability, interoperability or conformity with European legislation such as the General Data Protection Regulation (GDPR). Furthermore, the interoperability of data and services is of great importance, as this has a huge impact on data exchange as well as on the usability of the data. In general, it is clear that the requirements catalog must contain some requirements independent of the application context, while isolated requirements must be defined depending on the specific framework conditions of the use case.

Research map for cross-company data exchange

To clarify the relevant topics in the field of data economy as well as the focus topics of the IEDS project, a research map was designed that brings together approaches from the literature with the work packages of the project and identifies possible topics for Module B (see Figure 3.2). This conceptualizes the 16 innovation drivers and future relevant research needs within the four core dimensions of cross-enterprise data exchange. The research map provides a structured insight into the crucial topics and incentives of the data economy in the future.

Data strategy and **data management** enable companies to manage their data securely and efficiently. Data governance mechanisms are a central aspect of this objective (see Section 1.4). The focus of data governance is on the quality and security of data resources. Both within and across companies, strategic frameworks and structures must be designed to help shape data exchange and data management in a responsible manner and to build competencies in handling data (van den Broek and van Veenstra 2015; Lis and Otto 2020). From a strategic perspective, companies must develop tools that enable them to process data successfully. A sustainable data strategy ensures the development of necessary competencies in the long term, such as the design of digital twins or adapted tools for data analysis (Cirullies and Schwede 2021). Increasingly, privacy- and security-preserving technologies are also coming into focus

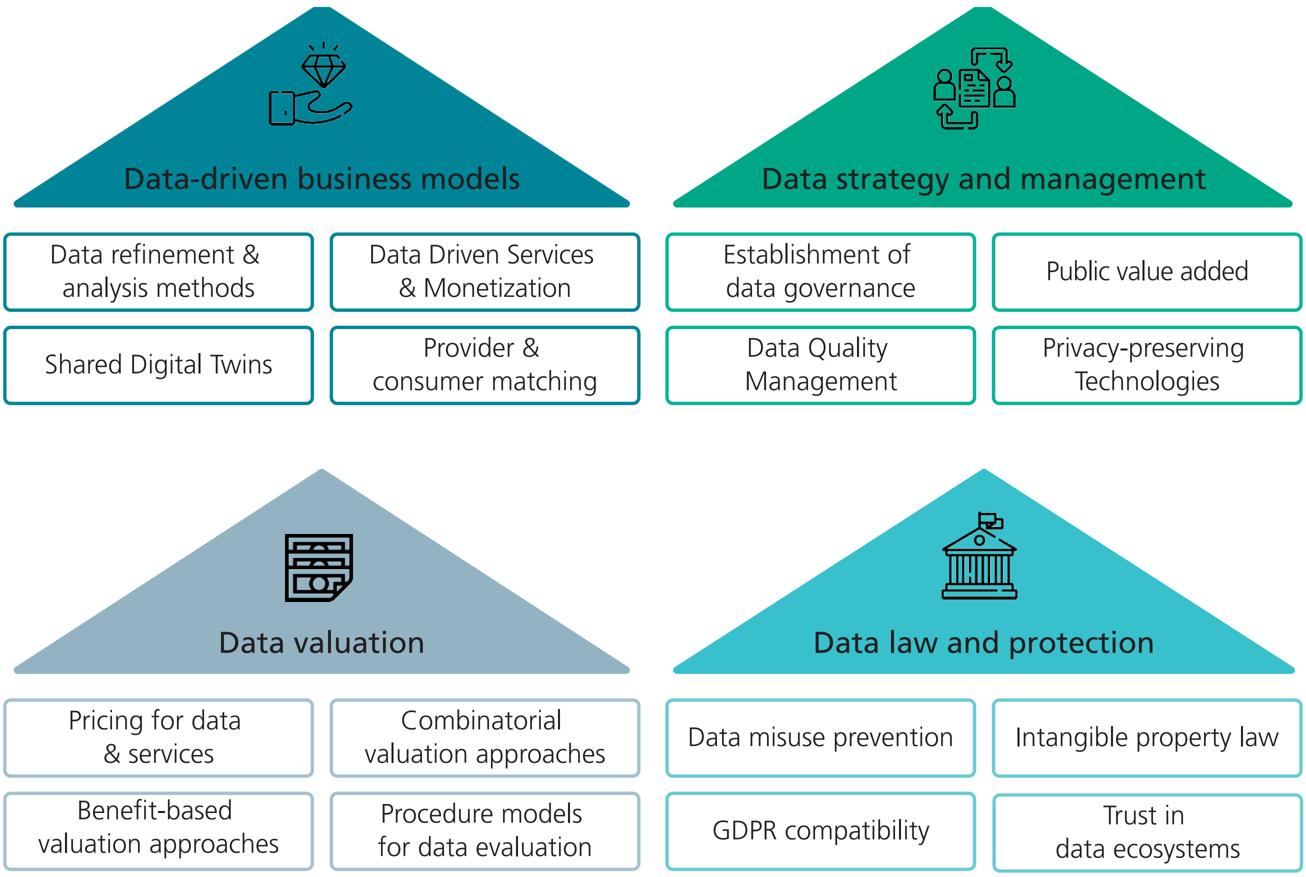


Figure 3.2 The four core dimensions of the research map (own depiction)

in cross-company data exchange. In particular, cryptographic methods provide a high level of security and make intermediary data brokers obsolete (Agahari et al. 2021; Dahlberg and Nokkala 2019). The Covid-19 crisis manifested the relevance of cross-enterprise data sharing through the need to aggregate health data. In public administration in particular, data sharing can contribute to collaborative value creation (Susha et al. 2019).

In the topic area of **data-driven business models**, the focus is on data-based services and their economic significance (see Section 1.4). Increasing computing and storage capacities support the development of data-driven business models (Parvinen et al. 2020). In the future, mechanisms and data refinement techniques, such as analytics and information services, will need to be developed to increase the utility of raw data (Fruhwirth et al. 2020). The economic incentives for sharing data and providing the services as part of business models represent an important element of cross-enterprise data sharing. Data-driven business models can be viewed from different perspectives and accordingly offer different implications for different actors in data ecosystems. Based on this, archetypal patterns for data marketplaces and their participants can be developed (Gelhaar et al. 2021b). In particular, there is a need to explore tools for adequately merging data demand and supply through metadata and matching algorithms (Susha et al. 2017). As data sharing initiatives emerge, multi-layered reference models and architectures are emerging.

Data valuation focuses on determining the value as well as the economic impact and economic significance of data exchange. This includes, for example, the determination of the combinatorial value of data and the exchange of data for mutual benefit. Upcoming research can focus on the pricing of datasets, for example, considering utility-based or combinatorial valuation approaches (Badewitz et al. 2020). In addition, procedural models can be developed to simplify the process of data valuation.

In the area of **data law**, the focus is on compliance rules, contractual and liability concepts of data management, intangible property rights, and data security, each with a subsequent legal-economic evaluation.

The aforementioned and explained core dimensions of the research map also represent focus topics of the IEDS project (see Figure 1.3). Their initial results, starting with the topic area of data strategy and efficient data management, are presented below.

3.2 Data strategies and efficient data management

Viewing data as a strategic resource enables completely new forms of value creation and optimization potentials, so that companies often align their considered data volume not according to the strategy, but the strategy according to the potentially usable data. In this context, a data strategy provides a reference on methods, tools, services, architectures, and usage patterns for managing and using data (Gurevich and Dey 2018). A data strategy can be defined as a plan that requires setting goals, identifying data sources, and using analytics to ask the right questions through strategic thinking in collaboration with technological expertise to create value for internal and external stakeholders (Gür et al. 2021). In this context, data management has the corporate function of planning, controlling, and providing data (Mosley et al. 2010). It aims at using data efficiently, as it includes several functions for formulating a data strategy, defining management processes, measures, and standards, assigning roles and responsibilities (Otto 2011), and managing applications and systems (Pentek et al. 2017). In the following, we identify ways to efficiently perform data management for data sharing and to align an organization's strategic and operational direction accordingly.

Another aspect of data management, which is particularly important for the reuse of data, is finding the required data in internal and external sources. Data is typically managed in directories or offered via data marketplaces and made available via a search function (Spiekermann 2019). Due to their characteristics, data assets pose special requirements for data search. The following is the approach taken in the IEDS project to improve data search using current technologies to better match data supply and demand.

Agile data management for data sharing: DataOps

To achieve long-term, sustainable success and competitive advantage, companies need to reach a higher level of maturity in managing and using their data (Altamony et al. 2012) so that they can participate in data ecosystems and data sharing. However, companies face a variety of difficulties. These challenges occur not only at the technology level, when it comes to outdated technologies and an extensive software landscape (Figure 3.3), but also at the organizational level. Often, data is stuck in data silos within the organization, or the team of Data Scientists and Data Engineers is unable to quickly leverage new data sources (Hurley 2018). Organizations today need data teams that are agile to respond to stakeholder data needs as quickly as possible (Sparapani 2019). At the same time, data must be trustworthy so that decision makers can use it without worry. Studies show that companies need to overcome a number of obstacles to maximize their strategic value from data (Hurley 2018; Nexla Inc. 2018).

DataOps, short for Data Operations, is an enterprise-wide data management practice that manages the flow of data from source to value with the goal of accelerating the process of creating value from data (Nexla 2018). DataOps connects data creators with data consumers — both humans and machines — to accelerate collaboration and digital innovation, making it particularly effective for the massive amounts of high-value data required for AI processing (Sparapani 2019). As analyses show (Gür 2021), the underlying concept and terminology of DataOps is very new. To date, there is no universally accepted definition of DataOps. According to (Mainali et al. 2021), the term DataOps was first used in (Lenny 2014), where the importance of performing data analysis tasks quickly with easy collaboration and assured quality results in various Big Data and cloud computing environments is discussed (Mainali et al. 2021).

According to a study conducted by Experian ltd, 89% of organizations face challenges with data due to long delays in gaining insights, lack of trust in data, or lack of ability to use the data (Experian ltd. 2019).

A widely used definition of DataOps is offered by Ereth (2018), who defines DataOps as "a set of practices, processes, and technologies that combine an integrated and process-oriented view of data with automation and methods from agile software engineering to improve quality, speed, and collaboration and foster a culture of continuous improvement" (Ereth 2018, p. 5). As part of the IEDS research project, extensive research of data management practices using DataOps is taking place to determine how organizations can embrace these techniques to successfully participate in data ecosystems and data sharing. Inspirations and borrowed principles for DataOps data management practices come from Lean Management, DevOps, short for Development Operations, Agile methodologies, and Total Quality Management (TQM). An extensive and in-depth analysis of DataOps and its core building blocks is presented by the IEDS project in our publication "DataOps for Data Sharing" (Gür 2021).

While DataOps is closely linked to operational efficiency, quality, and agility, studies show how a mature DataOps approach in the corporate culture can create critical business advantages for long-term and sustainable competitive advantage. DataOps practices leverage automation and standardization to deliver significant impact on data curation services, metadata management, data governance (Madera and Aguilera 2020), and other business functions. Organizations are leveraging their data to drive efficiencies and advanced capabilities such as AI-driven and data-intensive applications like IoT, advanced R&D efforts, and complex financial analytics (Sparapani 2019). Studies show multiple benefits that can be observed through successful implementation of DataOps practices in organizations. Some examples of significant DataOps benefits are highlighted in the 451 Research study (Aslett 2019). In this study, the authors conduct a survey with 150 representatives of organizations with more than 1000 employees. The most commonly cited benefit resulting from implementing DataOps practices is facilitating security and compliance as cross-functional concerns related to data management. Given the core tenets of DataOps, an obvious benefit was increased business agility and faster time to market.

However, for an organization to be successful with DataOps, potential issues and challenges must also be considered, especially if DataOps practices are to replace outdated methods. (According to (Mainali et al. 2021), these include changing corporate culture, low-risk innovation, the cost of DataOps, transitioning from expert teams to cross-functional teams, managing multiple environments, knowledge sharing, tools and technology diversity, and security and quality).

DataOps supports highly productive teams with automation technologies to achieve efficiency gains in both project outcomes and delivery times. To reap the benefits, continuous adaptation and evolution of internal corporate culture is required. As more and more business units require data to gain contextual insights (Madera and Aguilera 2020), and participation in data ecosystems is almost essential in today's economy, implementing DataOps practices in companies is advisable.

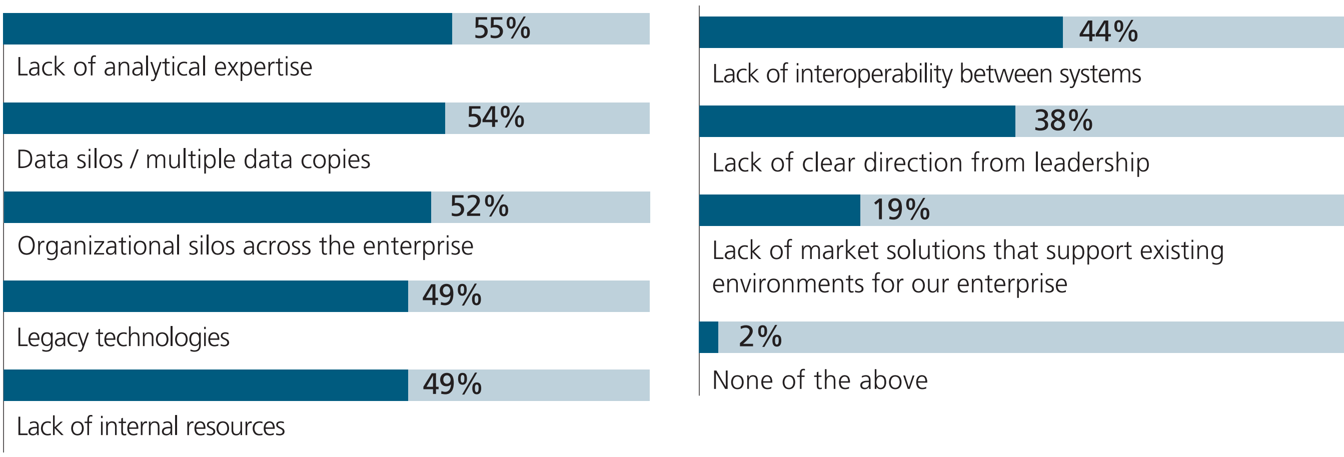


Figure 3.3 Barriers to maximizing the strategic value of data. Source: (Harvard Business Review Analytic Services 2019)

Agent-based and AI-assisted matching techniques to improve data search.

The goal of data management in general and of DataOps (see Section 3.2) in particular is to make data usable within a company. Once the value of data (see Section 3.4) has been recognized and it is systematically collected and managed, the next step is to exploit the data and make it usable within the framework of cross-company data exchange and trade. The prerequisite for a successful exchange of data between different parties is to bring together data supply and data demand. This central mediation task is usually performed by data marketplaces, which have a catalog of available data assets and provide information about the existing data supply via a search function for potential data users.

For physical products, successful search procedures have already existed for many years that are adapted to the specific characteristics of the search object. Major online marketplaces, such as Amazon, trade products that have metadata descriptions with fixed attributes, usually dependent on the product type. In addition to keyword searches, these marketplaces offer the ability to further narrow search results via filters by allowing users to specify allowed values or ranges of values for individual product attributes (Wei et al. 2013). Complementing the methods described above, user behavior is also increasingly being recorded and evaluated using AI methods to create user profiles and incorporate user preferences into search results and suggestion systems (Ai et al. 2017).

Data assets offered on data marketplaces have special characteristics compared to physical products that also affect searches. In media data, for example, advanced AI techniques can be used to classify image content and understand and extract spoken text. With this information, media data can be found by its content. Temporal information, such as the date of creation of a data asset or the time range of a sensor measurement, plays a major role in finding many data assets such as financial and business data, address data, Internet-of-Things and sensor data, or research data. Many data goods, such as map and geospatial data, sensor data, and also survey data, also have a strong reference to location. In addition, the interpretation of data assets such as sensor data, medical data, or research data usually requires information about the type of origin, the (measurement) methods used, and the processing steps that have already been taken (Koutroumpis et al. 2017).

These particularities of data as an intangible tradable good therefore pose special challenges for marketplaces and especially for data search. These include enabling detailed temporal and spatial narrowing of search results (Kacprzak et al. 2018), evaluating the reusability of data assets and assessing their fitness for purpose in advance (Kacprzak et al. 2018), complex analysis and linking of information from multiple datasets (Koesten et al. 2017), and deep extraction of metadata and analysis of datasets to deliver particularly precise search results (Chapman et al. 2020).

Recent years have shown that numerous data marketplaces, such as Azure Data Marketplace, have failed to establish themselves and have had to cease operations (Spiekermann 2019). Reasons include data providers' concerns about data security, protection of intellectual property, unclear value of data assets, but also that no suitable data sets could be found (Röhl et al. 2021).

One of the goals within the IEDS project is to address some of the aforementioned challenges by using innovative techniques from the fields of artificial intelligence and agent-based simulation, and to increase their acceptance via the improvement of data search on data marketplaces in order to ultimately create incentives for cross-company data exchange from them.

Agent-based simulation and test environment: The concept

The approach pursued in the IEDS project to improve the matching of data supply and demand is composed, on the one hand, of the improved search procedures and, on the other hand, of a simulation and test environment that makes it possible to measure the performance of the search procedure and to systematically compare different approaches.

In order to be able to map different aspects of supply and demand matching from the perspective of the data providers and the data users, the planned simulation environment should meet the following requirements:

- The typical search functionality of a marketplace with keyword and faceted search is to serve as a comparative reference for all experiments. This search functionality must be extendable with regard to the processing and indexing of new data and the technical handling of search queries in order to be able to extend the existing search in a targeted manner (using AI-based processes) and to specialize it on data assets.
- The special characteristics of data goods are taken into account in the scenarios examined. For this purpose, it may be necessary to store certain behavioral patterns for data providers and data users and to enable the interaction of multiple actors.
- For the evaluation of the experiments, a framework is created that logs all relevant data of a simulation and provides tools for the analysis of the results.

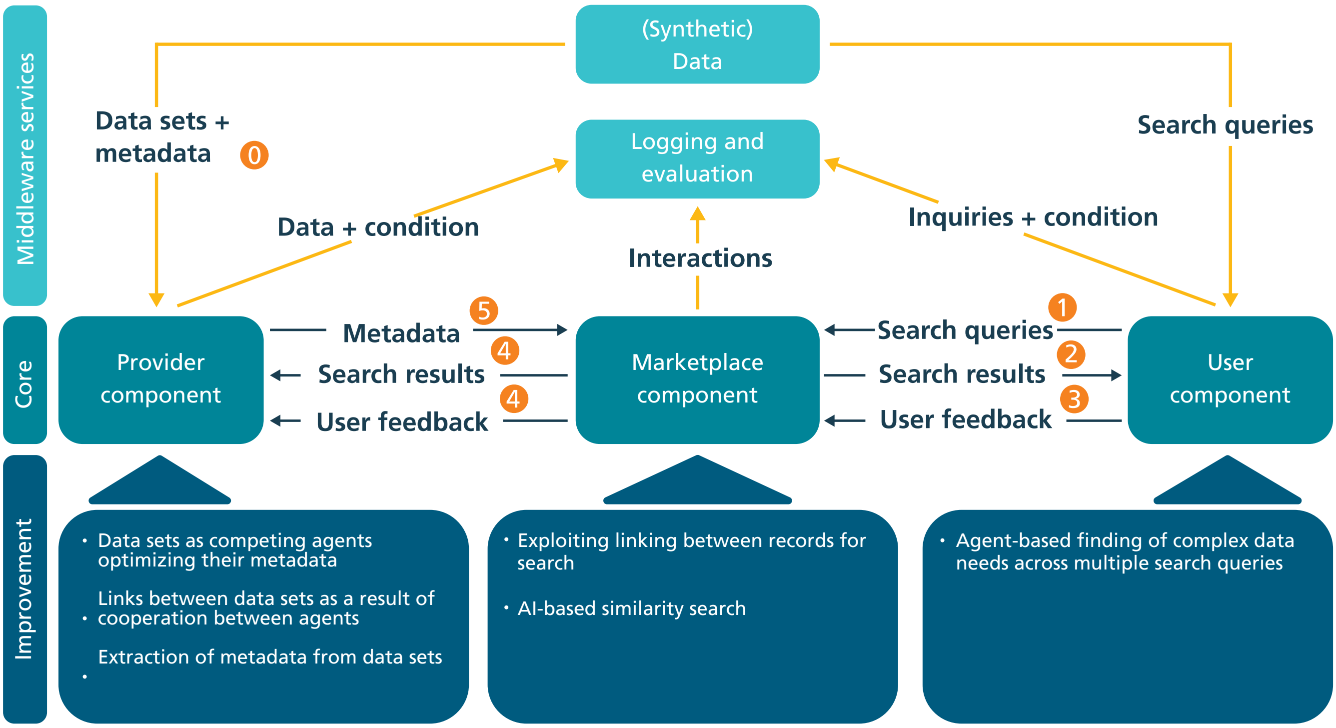


Figure 3.4: Simulation environment and possible approaches to improve data search (own depiction).

These requirements were translated into the concept of the simulation and test environment shown in Figure 3.4. The central simulation core is formed by the marketplace component, which contains the search functionality, and the data provider and data user components. If the approach to be tested requires complex behavior of data providers and data users and interactions between them, the provider and user components are implemented on the basis of an agent-based simulation environment. Here, an agent is understood as an autonomous entity that has its own state (e.g., data asset), interacts with other agents (e.g., data providers or data users) and its environment (e.g., the marketplace), and makes autonomous decisions (Jackson et al. 2017). The marketplace component has an interaction model that allows data providers and data users to communicate with the marketplace. The technical basis of the marketplace component is formed by an existing and extensible search system such as Apache SOLR[1]. This simulation core is supplemented by supporting services, which are responsible, for example, for generating synthetic data and search queries as well as for logging and evaluating the experiments.

The challenges described previously are addressed with the agent and AI-based approach to improving data search described below (Figure 3.4, blue area).

The simulation cycles through several phases to gradually improve the data descriptions (the individual steps are shown as numbered circles in Figure 3.4). At the beginning, the data provider agents publish their initial data descriptions to the marketplace (0). Then, data user agents submit representative search queries to the marketplace (1). The data user agent evaluates the search results (2) and selects matching data sets (3). This information is fed back to the data provider agents (4), who then modify their description and update it on the marketplace (5). The data marketplace records all interactions and uses them to improve search.

In addition to the described approach, other improvement approaches are also being investigated, such as the implementation of an AI-based similarity search as the basis for a suggestion system, the AI-assisted assignment of intended uses to data assets, or the search of complex data needs across multiple search queries.

1 https://solr.apache.org/.

The targeted use and exploitation of data addressed in this section and the interaction of the supply and demand sides are also discussed in the following Section 3.3 in the context of data-driven business models in the context of ecosystems.

3.3 Data-driven business models

The fact that data is accorded great importance in this day and age is certainly not a fact that comes as a great surprise. Large corporations such as Meta[2] and LinkedIn[3], whose business models are based entirely on the use of data, are no longer an exception. Such data-driven business models are based on the key resource of data and use it to generate added value for themselves and their customers (Schüritz et al. 2017; Kühne and Böhmann 2018; Guggenberger et al. 2020a). Since companies in today's world often no longer act alone, it is a natural reaction for so-called ecosystems to form. These ecosystems offer companies the space to act with each other in a simple way and to market their services or products. The advantage to such an association is that all participants are motivated by their own interests to make the best possible contribution to the ecosystem (Oliveira and Lóscio 2018; Oliveira et al. 2018). Now, to understand how a company builds its business model based on data in such a data ecosystem, it is important to understand which components are of high importance and which opportunities are available for companies to participate in the ecosystem.

Data-driven business models in the environment of data ecosystems

In order to understand a data-driven business model, it is important to first look at its environment. Since a data-driven business model often uses data in various forms as a key resource, it is natural to first take a look at data ecosystems. According to Oliveira and Lóscio (2018), a data ecosystem is defined as a network of diverse actors that leverage their synergies to share resources, usually data or similar products, with each other (Oliveira and Lóscio 2018). When considering individual actors in an ecosystem, it can be noted that each one must have a business model in order to define its strategy for value generation. Consequently, a data-driven business model in the environment of an ecosystem can be classified as an actor in a network (ecosystem) that uses the core resource of data (see Figure 3.5). The respective company can produce this data itself, receive it from other companies and consume it, and/or pass it on to other players

> Data-driven business models can be described as business models that focus on the use and further processing of data to create the value proposition. To add value to these value propositions, data analysis processes are used in particular to generate new insights and knowledge.

A data ecosystem consists of four main components: Actors, Roles, Relationships and Resources (Oliveira and Lóscio 2018; Oliveira et al. 2018). In order for a company to participate in an ecosystem, it must first decide which role it wants to take in an ecosystem, as this fundamentally determines the further direction of the business model. In addition, it must determine how the relationships with the other players are to be defined and which resources are to be used. Within these areas, further characteristics that are important for a business model, such as the range of services, are defined. The combination of a large number of different companies and the cross-company exchange of data ultimately results in an ecosystem in which value-creating activities are driven based on data (see Figure 3.5).

2 https://about.facebook.com/de/meta/.

3 https://about.linkedin.com/de-de?trk=homepage-basic_directory_aboutUrl&lr=1.

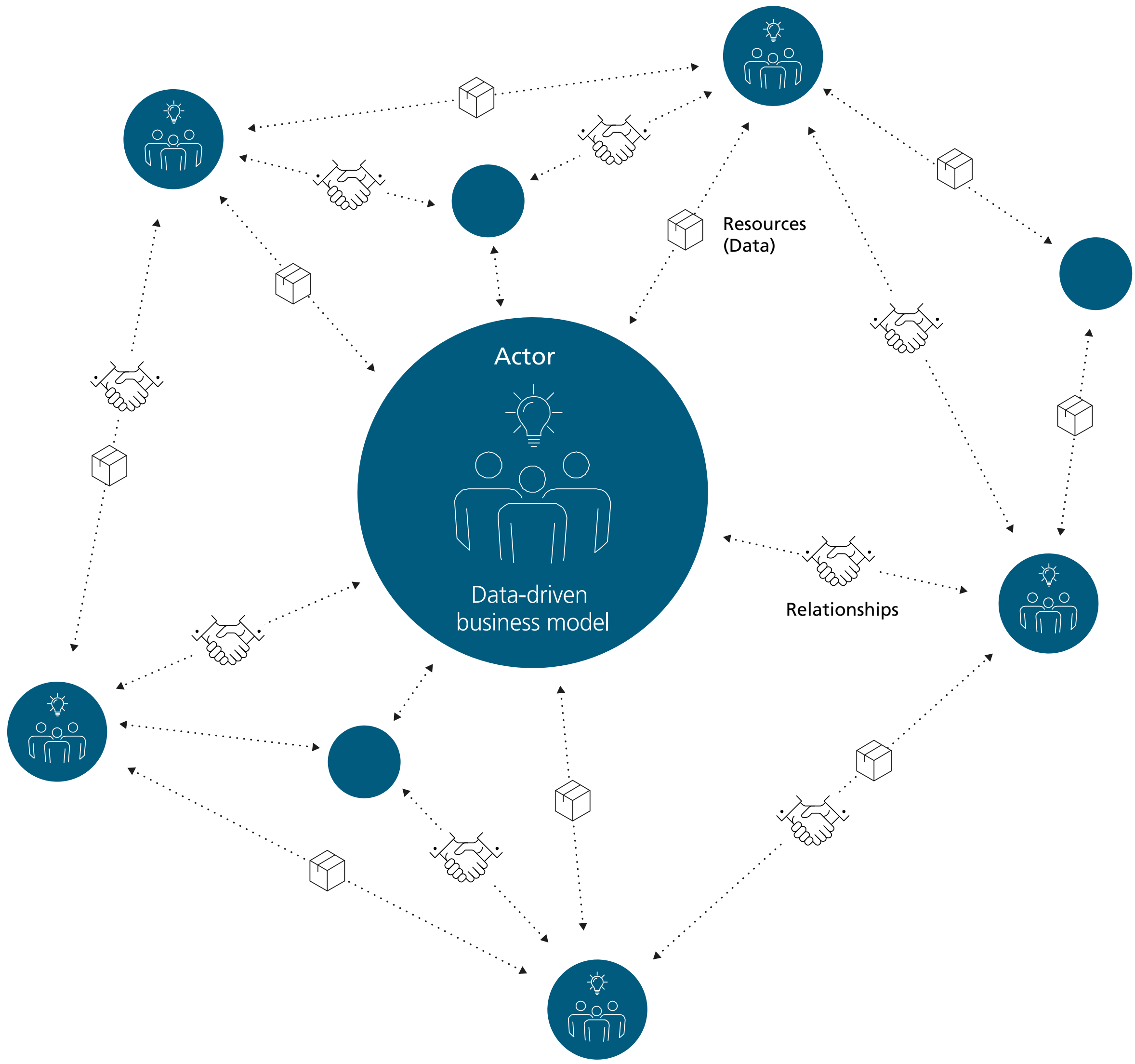


Figure 3.5 Actors and their data-driven business models in the ecosystem context (own depiction).

As mentioned earlier, the choice of roles for companies is an important decision at the outset. Initially, there are three possible roles that an actor can assume: Consumer, Intermediary, or Producer (Oliveira and Lóscio 2018). However, many other manifestations of these roles can be found in practice, often defining more specific tasks for themselves. In general, it can be stated that there is at least one actor that consumes the data or services provided to it that another actor has produced. However, since the exchange between two actors often needs to be regulated, there is an intermediary between the two actors who provides the necessary infrastructure and ensures compliance with the applicable rules and regulations.

Roles in ecosystems

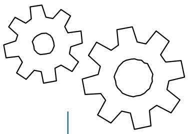
In order to be able to more precisely define the different roles in an ecosystem, 64 use cases of the Gaia-X Hub Germany[4] (as of August 2020) were examined. These use cases were selected because they come from different industries and provide a comprehensive insight into current data-driven business model architectures and ecosystems based on the Gaia-X infrastructure (see Section 1.2).

It was worked out which roles exist in the respective use cases and which actor assumes which specific role in the ecosystem. Based on this, role descriptions were prepared in relation to the four business model dimensions of value offering, value generation, value provision and value capture. The value proposition describes the overarching value or benefit that the other actors are to receive from the services or products offered, while the value generation describes how the creation of these values takes place (Osterwalder and Pigneur 2002). Value provision defines the channels through which the value proposition is made available to other actors, and value capture defines how revenues can be generated from it (Azkan et al. 2020a). In total, eight different roles could be identified based on the research conducted, which, based on the four business model dimensions, are summarized in Figure 3.6.

As stated in this chapter, data is valuable to companies and drives them to base their business models on it. Determining the value of data is therefore considered in more detail below in Section 3.4.

4 https://www.bmwi.de/Redaktion/DE/Dossier/gaia-x.html.

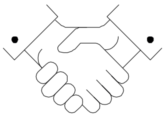
Ecosystem Roles



Service Provider

provides services to players in the ecosystem, usually building on the data provided by the data provider.

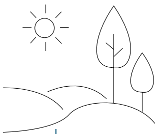
- The value proposition is characterized by the generation of knowledge, recommendations for action and the development of applications.
- The acquisition of knowledge and recommendations for action is carried out in particular through the management of data and the analysis of these data.
- To offer its services to the other players, it uses APIs, cloud platforms or makes them available as a direct download.
- For its services, it charges a fee in the form of one-time payments, subscription agreements, pay-per-use agreements or performance-based contracting.



Data Trustee

mediates between the data provider and the service provider, ensures the availability of the data.

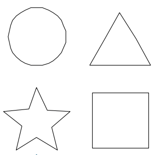
- The value proposition is characterized by the guarantee of secure and legally compliant data exchange.
- Implementation through the management of access rights, the storage, encryption and transmission of the data
- Cloud solutions are used for the transfer of the data, to which the users of the data trust service have access.
- In return, the data trustees usually receive fixed amounts for their services or are paid according to the pay-per-use principle.



Data-Infrastructure Provider

provides the IT infrastructure on which the cloud platform providers operate and exch-ange sensitive data.

- Guarantees the highest possible security, data sovereignty, and decentralized and interoperable data management with media-break-free provision of data.
- Provides an infrastructure based on standards and policy rules on which compliance with regulatory requirements for security standards is realized.
- It is provided as a cloud solution.
- In the case of GAIA-X, no monetary consideration is required.



App-Store Provider

provides an exchange platform where developers can offer their tools and third parties can obtain them.

- The value proposition is characterized by developers being able to market their applications and third parties being able to find and source them.
- This value proposition is enabled by the creation of an exchange platform that typically includes search and filter functions.
- The app store is provided as a cloud platform.
- In return for its intermediary function, the provider usually receives a transaction fee for transactions between developers and customers.

Figure 3.6: Roles in ecosystems (own depiction)

Data Provider

provides data or metadata to the ecosystem and shares it with other actors.

- The value proposition is characterized by the provision of data that can be reused by the other players in the ecosystem.
- Data is collected using technical tools, either as a main product (e.g. weather data) or as a by-product (e.g. data for monitoring the company's own facilities).
- Data is automatically transmitted to the service provider (e.g., via smart wearables) or made available through APIs, cloud platforms or download portals.
- Added value is generated by selling the data directly, through process optimization, cost savings or quality improvements based on data analysis.

Ecosystem Orchestrator

created the ecosystem by bringing stakeholders together and has access to information about the entire ecosystem.

- Ensures that all stakeholders involved in the ecosystem have the opportunity to participate in a value-creating manner and the highest possible shared value is created.
- Identifies the different roles in the ecosystem, creates connections between them, motivates them to work together, and takes action to curate the ecosystem if necessary.
- Either enters into direct contact with the actors or modifies the platform used by the ecosystem.
- Since the actor who assumes the role of Ecosystem Orchestrator usually also assumes the role of Cloud Platform Provider, the same revenue models apply for it.

Cloud Plattform Provider

provides a platform through which stakeholders participate in the ecosystem and interact with each other in a value-added manner.

- The value proposition includes bringing players together on the platform, but can also include storing, analyzing, and displaying data.
- The value proposition is realized by providing the necessary exchange platform, applications and tools.
- The services are provided on the basis of a cloud platform.
- The cloud platform provider may charge a usage fee, although intermediary fees are not uncommon either.

Data-Marketplace Operator

provides a platform through which data providers can offer their data and third parties can consume it.

- The value proposition is characterized by bringing data providers and users together and allowing data to move freely and securely.
- It provides the trusted data sharing infrastructure to offer and find data and its meta-information.
- The data marketplace can be accessed by stakeholders and other participants via APIs or a cloud platform.
- It usually receives a transaction fee for successful transactions, but a pure usage fee for the use of the data marketplace is also possible.

Excursus: The impact of Big Data on business performance

The digitization of companies is a key factor in boosting corporate performance. For example, the use of information and communication technologies (ICT) not only offers opportunities to produce and market already familiar products and services more efficiently, but also to create new products and services and even entirely new business models.

At the macroeconomic level, the influence of ICT on productivity has been discussed at least since the Solow Paradox was postulated in 1987. This states that "the computer age can be found everywhere except in productivity statistics" (Solow, 1987). Oliner & Sichel (2000) was one of the first studies to show an economically significant contribution of ICT to productivity at the macro level. The positive influence of digitization on productivity development was confirmed in later studies by Jorgenson & Stiroh (2000) and Byrne et al. (2013) as well as Oliner & Sichel (2000) for the USA and Inklaar et al. (2005) for the USA and four EU countries.

At company level, too, numerous studies show a positive correlation between the degree of digitization of a company and its performance. An overview can be found in Draca et al. (2007, 2018), Cardona et al. 2007, Biagi (2013), and Schweikl & Obermaier (2020). The first robust empirical evidence on this at the company level is shown by Brynjolfsson & Hitt (1995), although small number studies and case studies have already named positive effects in earlier years, see Brynjolfsson & Yang (1996).

However, it is not only internal digitization that is relevant for a company's success. Factors external to the company, such as the availability of broadband connections, also have an influence on company performance. An overview of the contribution of broadband connections to corporate success is provided by Bertschek et al. (2015). Gal et al. (2019) also show that companies benefit not only from in-house digitization projects, but also from a digital environment.

The overall positive effects of ICT manifest themselves in different ways. For example, companies with high ICT use are more likely to increase the number of their employees and less likely to exit the market (van Reenen et al., 2010). Cappelli's (2010) analysis of specific ICTs shows, among other things, that the use of corporate intranets and data warehousing procedures has a significant positive impact on turnover per employee. Cardona et al. (2013) further find that the relationship between ICT and the productivity of firms increases over time. More recent studies have also concluded that investments in ICT lead to higher gross value added (Dhyne et al., 2018).

Big Data analysis / data-driven decision making

The increasing digitization of companies is leading to a large increase in data volumes. The data here comes from various sources: Both internal company data from sensors, machines or enterprise resource planning software, as well as data from (potential) customers through social media or website clicks are available in ever greater quantities. The analysis of this data (Big Data analysis) provides a lot of additional information for decision making and strategy development that was previously based on explicitly collected data (Constantiou & Kallinikos, 2015). As concrete impact mechanisms through which Big Data analysis can have a positive influence on business performance, Engels & Goecke (2019) discuss cost savings, risk minimization, revenue increases, and changed business models. The European Commission is therefore "[...] convinced that the use of data can enable EU businesses and the public sector to make better decisions" (European Commission, 2020a).

Empirical literature also associates Big Data analysis with benefits at the firm level. Niebel et al. (2019) demonstrate a positive association of Big Data analysis with product innovation in Germany. Brynjolfsson et al. (2011) show that companies that make data-driven decisions have higher market value, higher return on equity, and are more productive. Brynjolfsson & McElheran (2019) also find positive productivity effects of data-driven decision making. Based on a panel of publicly traded U.S. companies, Müller et al. (2018) find a direct relationship between Big Data analytics and productivity. Wu et al. (2020) also find a complementary relationship of Big Data analysis and process innovations in their productivity effects.

3.4 Data evaluation

Data takes on value through its use for specific purposes, for example in the context of data-driven business models. They thus become economic goods. However, data differ from other goods in several respects. Against this background, the combinatorial valuation of data as well as the valuation of data as intangible assets will be examined within the framework of the IEDS project.

According to a study by Ocean Tomo (2020), up to 90% of the market value of companies in the S&P 500 consists of intangible assets. For the S&P Europe 350, this value is somewhat lower at 75%. It can thus be stated that the value of companies today consists to a large extent of non-physical assets.

Combinatorial data evaluation

An essential characteristic of data sets as an economic good is that their value results from complex substitution and complementary relationships of the data points they contain. This must be determined combinatorially using costly procedures.

The need for combinatorial data valuation is based on the observation that, when optimally combined, a few data points can achieve the same or even greater benefit than many data points. Figure 3.7 illustrates this with the example of quality control in paint shops. Two paint shops use photo data for quality control of painted surfaces. Here, photos of defective and defect-free paint spots are used to calibrate a machine learning algorithm to automatically detect defective paint spots. The higher the defect detection rate of the algorithm, the better. In principle, it can be expected that the marginal utility of a larger photo or data set will decrease. This means, for example, that doubling the size of the data set will not lead to a doubling, but to a smaller improvement in the error detection rate. Accordingly, merging the photo sets from the two paint shops should lead to an improvement, but not a doubling. In contrast, an optimal selection of photos from the photo set of both paint shops may well achieve higher defect detection rates, so that higher defect detection rates can be achieved with photo sets of the same or even smaller size. This results from the fact that data points can have substitutive or complementary relationships to each other.

These complementarity and substitution relationships between data points can be explained using the paint shop example. Assume that there are three regularities in the data. On the one hand, coatings with the color red would generally be more susceptible to defects. Secondly, coatings with the color blue would be less susceptible to defects. Furthermore, red and blue coatings are less frequent than other coatings with other colors. If one then starts with a small random selection of photos when compiling a photo set for the calibration of the algorithm, photos of paint finishes with the colors red and blue are likely to occur too rarely for the algorithm to fully capture this relationship. The higher or lower susceptibility to defects associated with these colors would thus only be detectable to a limited extent. If we now add images of red paint spots, the algorithm should learn the relationship between susceptibility to defects and the color red. However, the further addition of images of red paint spots should not result in any major improvements in the defect detection rate, since the images of red paint spots are likely to be substitutive for one another. The algorithm would not be able to learn anything new. It has already captured that red paint jobs are more prone to defects. In contrast, adding more shots of blue paint spots should further increase the defect detection rate, since these shots should act complementary to the previous set of photos. The algorithm could still learn in this case, namely that blue paint spots are less prone to defects. In real-world use cases and data sets, such substitution and complementarity relationships are usually much more complex and — as data sets get larger and larger — highly dimensional.

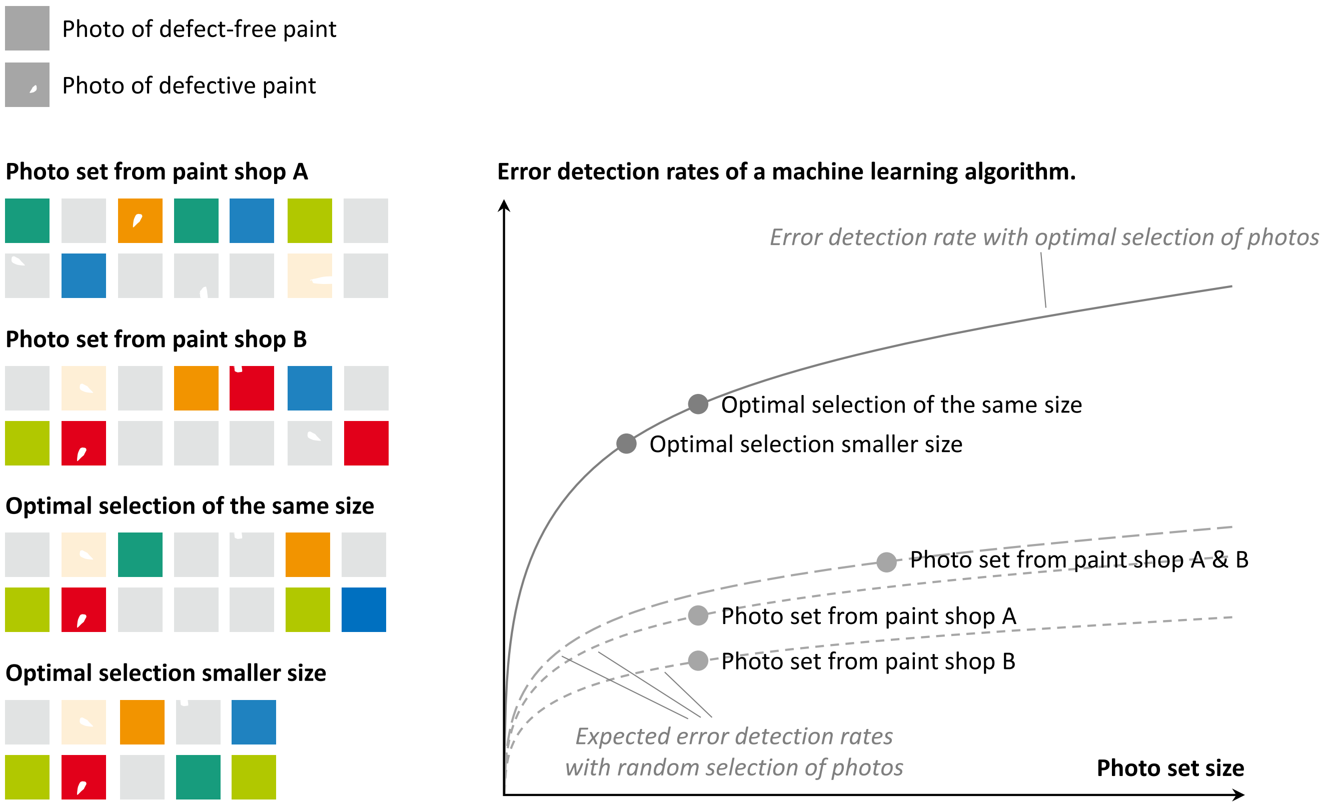


Figure 3.7 Example of combinatorial data evaluation (own depiction)

However, selecting optimal data bundles for maximum benefit is very computationally expensive and usually involves high costs. Essentially, every conceivable combination of data points must be tested for utility for a given size of data bundle or to meet a given budget limit. For example, using data to calibrate machine learning algorithms would require a completely new calibration of the algorithm for each conceivable data bundle. Due to the large number of conceivable data bundles and the high computational effort for the benefit determination, the combinatorial evaluation of data thus becomes a cost factor of its own.

Recently, however, some technical innovations have been achieved in academic research that make it possible to consider combinatorial effects when evaluating data. For example, in the field of machine learning, game-theoretic approaches are used in which the average utility or value contribution of individual data points is estimated in combination with other data points (Ghorbani and Zou 2019; Kwon et al. 2021). Furthermore, attempts are being made to determine the value contribution of individual data points or data bundles of this data set already during the initial calibration of machine learning algorithms to a data set, for example, using reinforced learning methods (Yoon et al. 2020). Approaches for efficient sampling from very large datasets for the purpose of statistical estimation are also relevant (Lee and Ng 2020). Modern combinatorial optimization techniques, which have recently included machine learning methods (Bengio et al. 2018), also offer opportunities for efficient combinatorial evaluation of data.

These new capabilities for combinatorial data valuation can make markets for data more efficient. Efficient markets for data could in turn result in strong incentives to collect, share, and trade data. For example: two data providers have different data. If these data are substitutive for data demanders and generate similar benefits, these data providers are competitors in this market segment. This creates price pressure and lower prices for the buyers of the data. If the data are complementary for the data demanders and generate greater benefits together than alone, these data providers are potential cooperation partners in this market segment. Together they can achieve higher prices than alone and enable new applications. However, the basic requirement of such increases in market efficiency is that the combinatorial value of data can be clearly determined. Thus, solving this technical problem has a large potential economic impact.

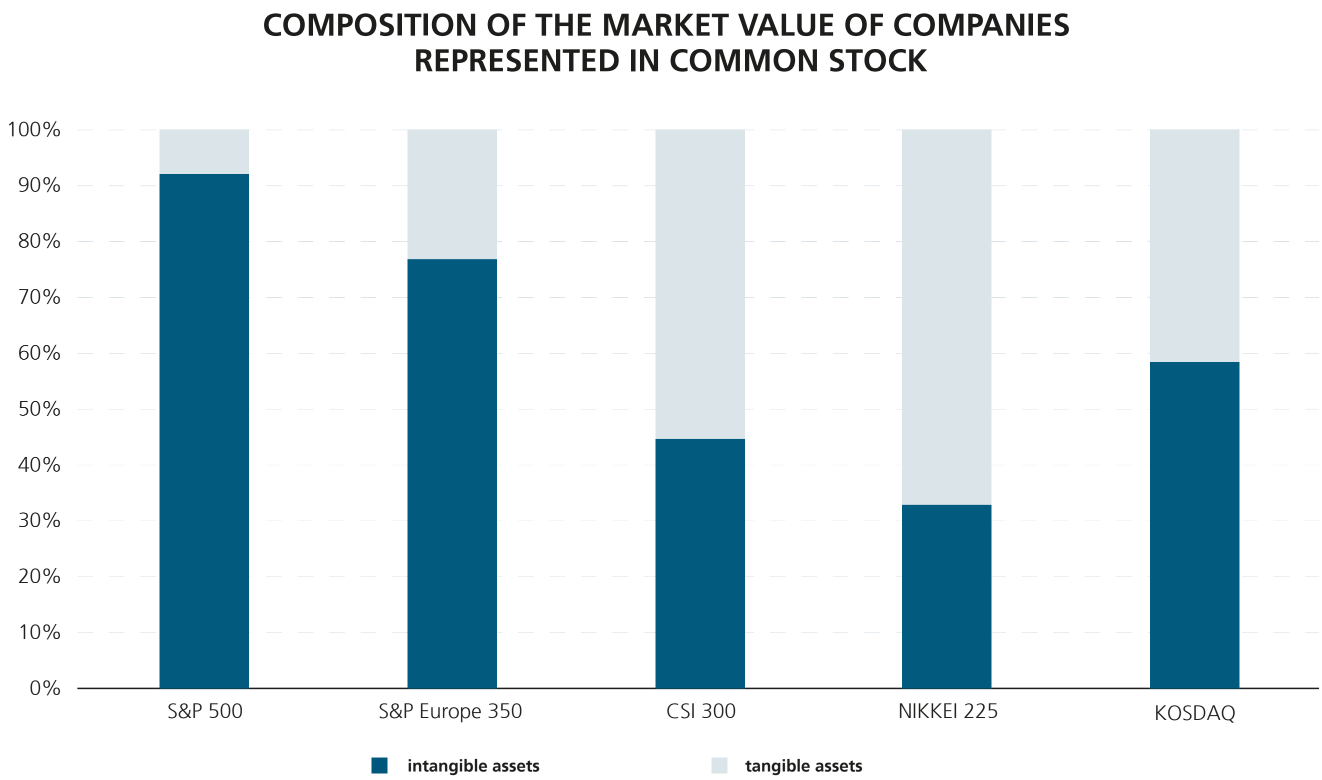


Figure 3.8 Distribution of intangible and tangible assets in popular stock indices (own depiction based on (Ocean Tomo 2020))

Valuation of data as intangible assets

Data as a central component of value creation could be treated as intangible assets in accounting. By capitalizing intangible assets in the balance sheet, it would be conceivable for companies to include the associated profit expectations in their balance sheets. This could have a positive impact on the market value of companies. As a result, there are incentives to trade intangible assets, as an increase in revenue can be achieved through the sale (Singapore Digital 2019). Furthermore, by measuring the impact of data on processes, landmarks for reducing manufacturing costs can be highlighted and data could be made available for public or collaborative use (e.g., within supply chains) (ibid.). However, for data to be capitalized on the balance sheet, it needs to be valued, should it be able to be defined as an intangible asset.

Intangible assets (computerized information, intellectual property and economic competencies) already account for a large share of the market value of companies. According to (Ocean Tomo 2020), 90% of the market capitalization of companies included in the S&P 500 stock index is composed of intangible assets (see Figure 3.8) . According to International Accounting Standards (IAS) 38.8, an intangible asset is an "identifiable non-monetary asset without physical substance." According to IAS 38.12, other key characteristics include the possibility of distinguishing the asset from other assets (identifiability), which would be provided by the possibility of individual sale or transferability. In addition, according to IAS 38.13, intangible assets would have to be controllable by the respective company in order to restrict their use by third parties. Intangible assets that could be recognized in the balance sheet would be characterized by a clear link to the value creation of a company (Kristandl and Bontis 2007). Furthermore, databases and software (computerized information) could be recognized as intangible assets.

In addition to the requirements for intangible assets, the requirements for their recognition are also regulated within the framework of International Accounting Standards and the German Commercial Code. According to IAS 38.21, an intangible asset would only be recognized in the balance sheet if, on the one hand, it is probable that the expected benefits will flow to the company in the future and if, on the other hand, it is possible to reliably determine the acquisition and production costs of the intangible asset. In accordance with Section 253 (1) of the German Commercial Code, assets would also be recognized at no more than their acquisition or production cost, less depreciation. In the case of internally generated intangible assets, the expenses incurred in developing the intangible asset should be taken as the basis for this (Section 255 (2a) German Commercial Code). This would mean that under both the IAS and the German Commercial Code, cost-oriented methods would be required to recognize data in the balance sheet as immaterial assets.

The transferability of these considerations on intangible assets to data is only possible to a very limited extent. This is due to the fact that, from a legal perspective, individual data are neither capable of ownership or possession (see Section 3.5). If an actor has data at their disposal, for example by owning the associated storage medium, this actor can also determine the further use of the data. The decisive factor here is the de facto control of the data, without an exclusive right being required for this (Scheufen 2020a). With regard to the control of intangible assets as required by the IAS, rights to use the data could be specified and transfered, however these can only be regulated individually by means of strict legal and contractual access rights. An exclusive property right that corresponds to the protection of an intangible property right (e.g. patent or copyright) is, however, only possible to a limited extent in the context of data and is at most applicable to e.g. a database, but not for the data itself (i.e. only at a structural level, and not syntactic or semantic levels) (see Section 3.5).

For commercial and tax purposes, the contribution of intangible assets to value creation is crucial, as previously mentioned. Considering data as computerized information, information characteristics such as accessibility, usability, timeliness, context, accuracy, relevance, and trust level (Tang et al. 2008) are likely to be relevant in assessing its contribution to value creation. Table 3.1 presents major types and sources of data, or computerized information, as found in organizations. These types of information could be valued to account for them.

Table 3.1: Main data types and data sources in companies (own depiction based on (Dakova et al. 2018))

Type	Definition	Examples
Intellectual capital	Human capital and intellectual property	Competences, processes, image
Knowledge	Value in use/evaluation of knowledge as action	Product management, customer service, partner management
Supply Chain Data	Supply chain mapping and utilization	Inventories, capacity planning in production
Business process information	Information as business transformation	Information Lifecycle Management
Decision Support Systems	Value contribution to decisions	Bayesian network

As with all assets, the value contribution or value of data could be measured very effectively via an active market for these assets (Oppenheim et al. 2003). For intangible assets, however, active markets are often lacking. In addition to such a market price-oriented method, cost-oriented and use-oriented methods can also be used to estimate the value added (Laney 2018). Cost-oriented methods determine the financial value based on the production costs or acquisition costs of goods. Cost-oriented methods are characterized by low complexity and are based on physical value chains. Use-oriented methods assess the financial benefit over the entire period of use. This entails a relatively high degree of complexity, but offers the possibility of taking future value creation into account (Zechmann 2018).

A fine-grained cost-oriented valuation of data creates the basis for determining a selling price, so that this data generates income in the event of disposal. The production costs determined from this can also be reduced in a targeted manner, as their composition is now known. In addition, data can be managed efficiently along a value chain through valuation, in that it can optimize existing processes in the event of data exchange.

The valuation of data also touches on legal issues, especially with regard to accounting. The area of data law, which deals with this in more detail, is presented in more detail in the next section.

3.5 Data law

Management in general and the sharing of data in particular ultimately always take place in the context of a (territorial) regulatory framework. This immediately raises the question of the legally compliant use and exploitation of data. Especially with regard to innovative (partly AI-supported) data-driven business models, however, the digital age raises questions that never arose before the analogous context of emergence of the civil law regulatory framework (Fries and Scheufen 2019; Scheufen 2020b). Legal concerns and data security issues also pose a significant hurdle to the internalization of the (economic) potential of data and the (incentive-driven) willingness of companies to share data (Demary et al. 2019; Krotova et al. 2020; Röhl et al. 2021).

The data law work package is divided into two components:

- **Legal analysis:** The legal analysis outlines and analyzes the status quo of the legal regulatory framework for data sharing. The individual legal areas for data management in general and data sharing (in the Gaia-X context) in particular are not only described, but are also discussed against the background of ideal-typical examples of use. The legal analysis is supplemented by a subsequent legal-economic evaluation, which identifies existing legal gaps and derives from them the need for legal policy action.

- **Requirements analysis:** The requirements analysis analyzes in detail the various barriers to data sharing from an economic, legal, technical and organizational perspective. The significance of the legal perspective as well as individual areas of law for the willingness to share data are to provide orientation and guidelines for a legal policy agenda of an incentive system for data sharing on the basis of a company survey.

In a joint and practice-oriented view, both modules show the status quo and the relevance of different (civil) legal standards as well as solution approaches for a transaction cost-reducing implementation.

Status quo of the legal regulatory framework

Data law is a cross-cutting issue that is relevant in all areas of data use and extends along the entire value chain. The legally compliant use of data must be viewed against the background of the distinction between personal and non-personal data (see Figure 3.9). If data is related to a person — even if the data can only be used to refer to persons in a broader sense (so-called personal data) — the scope of the General Data Protection Regulation (GDPR) generally applies (Fries and Scheufen 2019). This is associated with a tightly meshed catalog of obligations for the generation, collection, storage, analysis, and utilization of personal data, which hardly allows for an economic use of the data beyond the immediate contractual purpose — and thus also the sharing of this data (Fries and Scheufen 2019). In particular, the right to be forgotten (Art. 17 GDPR) and thus the need for potential deletion of individual data pose serious problems for data sharing in corporate practice (see Figure 3.9) — also because data sovereignty, i.e., de facto dominion over the data, of data once shared appears to be technically almost impossible to implement (Farke et al. 2019).

While personal data is thus subject to a narrow regulatory framework, there are hardly any specific regulatory standards for non-personal data outside of antitrust barriers or the protection of trade secrets. Data cannot be owned or possessed within the meaning of Sections 903 and 854 of the German Civil Code and are only protected by copyright law in exceptional cases. The decisive factor for the right of use, however, is merely the factual control over the data — a genuine right of exclusion is not required. Consequently, the data owner also decides on the use and, if necessary, the transfer of individual rights of use (Scheufen 2020a). Thus, in addition to narrow legal options, access rights (e.g. for sharing data) can be granted in particular by contractual agreement.

Against this background, the management of data in general and the sharing of non-personal data in particular can be implemented in a legally compliant manner by means of individual contracts or general terms and conditions. However, in the case of a general terms and conditions agreement for

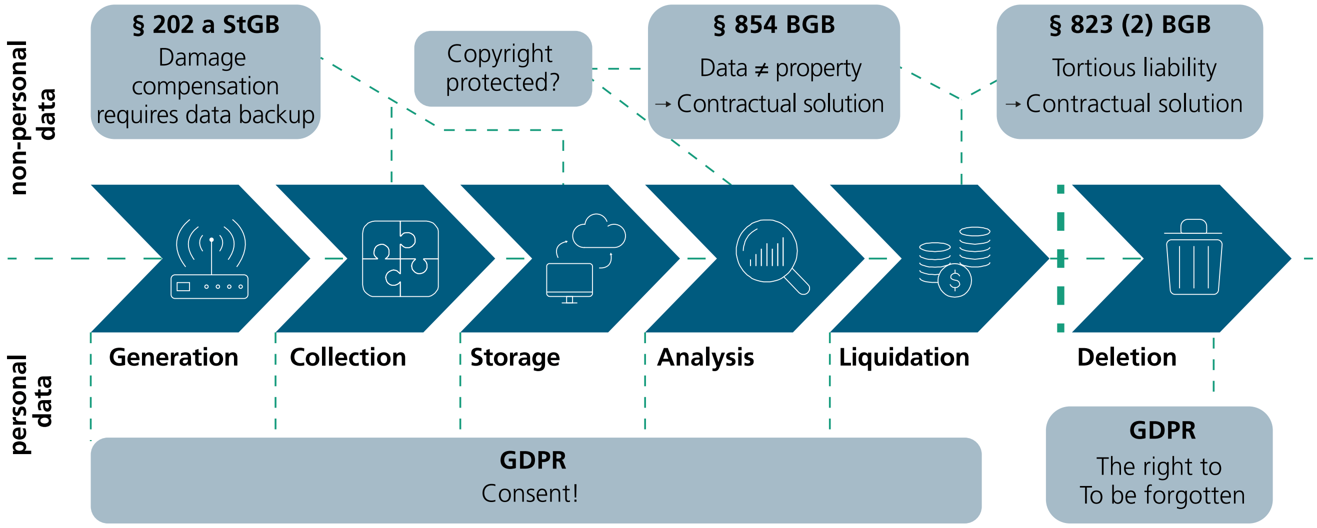


Figure 3.9: Law as a cross-cutting issue along the value chain (own depiction based on Fries and Scheufen 2019)

data (cross)-licensing, a possible review of the general terms and conditions by the court must be taken into account, which could result in legal uncertainty under certain circumstances. In addition to a restriction of the rights of use in terms of content, time or territory, such a data licensing agreement also permits (individual) regulations on liability issues (Rosenkranz and Scheufen 2021). Since claims for damages are linked to data protection under criminal law (Section 202a of the German Criminal Code), any responsibilities for such data protection can also be contractually fixed (Fries und Scheufen 2019).

Analysis of legal-economic barriers, consequences and needs

In addition to a general overview of the (civil) legal regulatory framework for managing and sharing data, evidence-based knowledge of the key economic, technical, organizational and legal barriers is crucial in order to derive concrete needs and recommendations for action for corporate practice. In this context, studies show that in Germany, legal issues represent significant hurdles for companies (Röhl et al. 2021). In addition to data law regulation, international studies have also identified data-related and technological barriers (difficulties in using and exchanging data, lack of tools and knowledge for Big Data analyses), economic barriers (expected high costs for necessary investments with unclear returns), and a general aversion to data-related technologies and data use, which may also have a cultural background (Dremel 2017; Moktadir et al. 2019; Mosig et al. 2021; Sun et al. 2016; Malaka und Brown 2015).

Based on the survey presented in Section 2.1, the economic, legal, technical and organizational barriers to data sharing are examined in detail. Based on the question "In which of these areas do you see the greatest obstacles with regard to data sharing?", it becomes clear that legal obstacles in particular are opposed to a greater willingness to share data. Around 68 percent of the companies surveyed see legal barriers, while organizational (25.8 percent), technical (22.1 percent) and economic (21.9 percent) barriers to data sharing are seen far behind (see Figure 3.10).

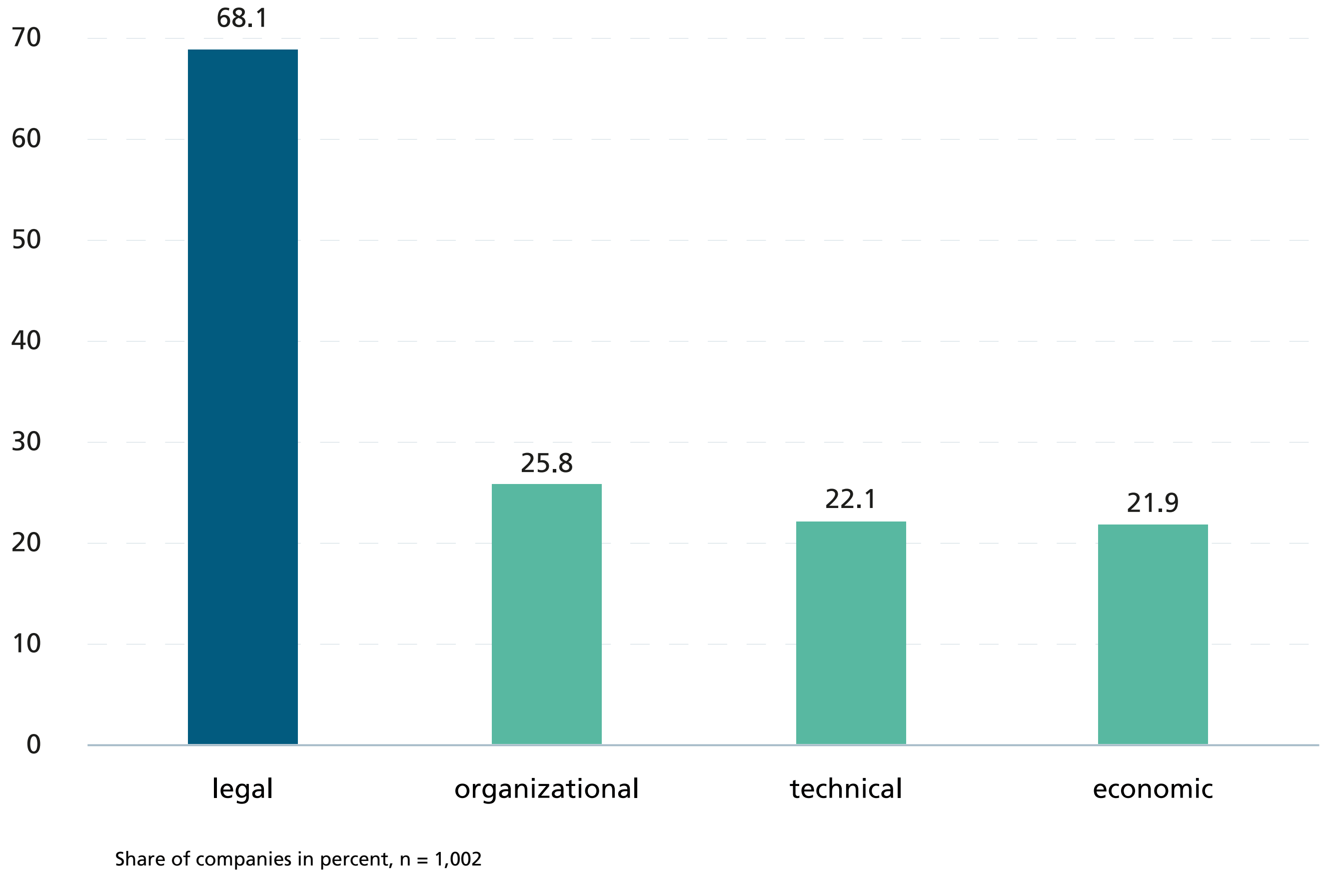
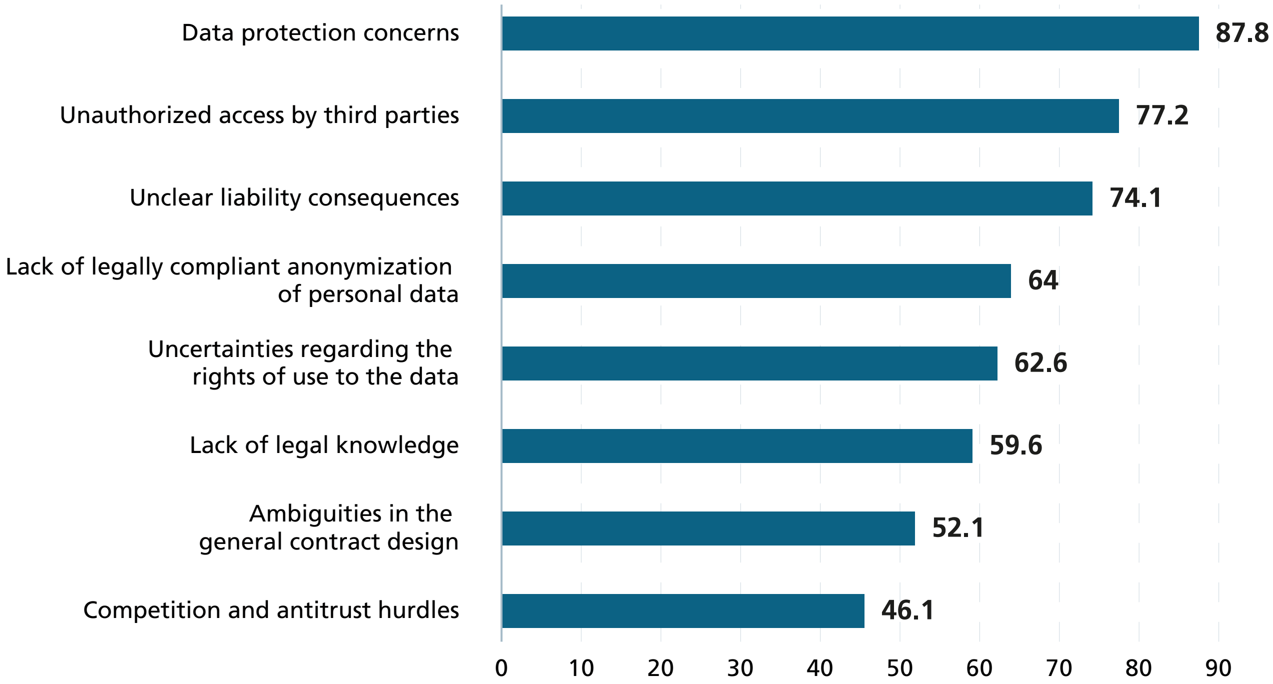


Figure 3.10: Types of inhibitors to data sharing (Institute of the German Economy, IW Consult)

The particular importance of legal obstacles to data sharing motivates further analysis of the various legal areas and thus the different focal points for the needs of companies. Figure 3.11 illustrates that for just under 88 percent of the companies surveyed, data protection concerns in particular limit their willingness to share data. As already discussed, the rights of use for personal data are very tightly regulated, so that further possible uses — especially the sharing of data — are hardly possible. Against the backdrop of the survey results, the companies seem to be aware of the tight regulation, so that a lack of knowledge of how to use this data in a way that complies with data protection regulations presumably leads to this assessment. Further analysis shows that there are significant differences between small (< 50 employees) and medium-sized (between 50 and 249 employees) or large companies (more than 249 employees). One reason for these differences could be that small companies are less likely to have their own legal department and therefore the legal hurdles are estimated to be even higher.



Share of companies that mentioned the corresponding items in response to the question "To what extent are the following legal aspects a hindrance for your company to share your data today or in the future?", in percent, n = 723.

Figure 3.11: The importance of different legal inhibitions to data sharing (Institute of the German Economy, IW Consult)

In addition to data privacy concerns, 77 percent of companies cite unauthorized access by third parties as a legal obstacle, followed by unclear liability consequences (74.1 percent), lack of legally secure anonymization of personal data (64 percent), lack of clarity regarding rights to use data (62.6 percent), lack of legal knowledge (59.6 percent), and lack of clarity regarding general contract terms (52.1 percent). For a minority (46.1 percent), hurdles under competition and antitrust law pose a problem (see Figure 3.11).

Consequently, there are not only legal uncertainties with regard to personal data, but also liability issues and contractual questions relating to the rights of use — also with regard to protecting these rights of use against unauthorized access by third parties — to non-personal data. Against this background, further work should be directed at possible dispositive norms (e.g., a data contract law with model contracts that specifically address the sharing of data) in order not only to describe data contracts but also to make them applicable to business practi-ce. The following section provides a further outlook on further work within the project.

4 Outlook

The findings and action areas presented in this white paper are the first steps in a holistic and comprehensive project aimed at identifying incentives for companies to participate in data sharing and data spaces, as well as lowering the associated hurdles and challenges. In addition, the scientific findings highlight opportunities to actively participate in the data economy in a way that adds value. In the further course of the project, incentives for, potentials of, as well as requirements for data sharing will be explored through the following strategic activities:

- **Data Economy Readiness:** The study results presented in the white paper regarding the status quo of data management in Germany ((see Section 2) will be evaluated annually. In the further course of the project, the data management capability of companies, their willingness to share data and their cloud usage will be analyzed in more detail. This will be done on the basis of the survey results, which will include further detailed questions. A comparison of different industries will also be made. The extent to which successful and unsuccessful companies differ in terms of their maturity of data management as well as their willingness to share data will also be examined. In the coming years, the survey will be repeated. In this way, the development of the maturity of data management, the role of data sharing and cloud usage in German companies can be regularly tracked and analyzed. Furthermore, the presented approaches to monitoring Gaia-X will be put into practice and the collected data will be analyzed and presented. After the selection of indicators for the dashboard has been largely completed, the indicators will be defined in more detail in the next step. This includes, for example, the frequency of updating and the specific figures and delimitations used. Subsequently, the automated procedures for collecting the indicators are programmed and tested in detail. This allows the data to be collected and secured in high quality on a sufficient scale. In a next step, this enables the data to be processed and analyzed. At the same time, the dashboard is designed, programmed and implemented in order to optimally present the data and the insights gained. Furthermore, survey experiments regarding the incentives for sharing data will take place in order to track developments and test new incentive constellations.

- **Data strategy and management** (see Section 3.2): In the project, procedures and opportunities for lean and agile data management were demonstrated in order to build up data pipelines flexibly and efficiently and thus make data available to the respective consumers (Gür 2021). In addition, it will be explored how such data management influences a company's data strategy and furthermore how an approach to develop a data strategy for data sharing can look like. In addition, mechanisms and systems will be explored to bring together AI-based data providers and consumers. These systems will be further evaluated and prototyped in the further course of the project.

- **Data-driven business models** (see Section 3.3): Based on extensive research, various business model patterns and roles in data ecosystems and data sharing were identified. addition, tools for designing these data-driven business models will be developed. To this end, visual inquiry tools that enable intuitive, collaborative, and visually supported work on business models will be developed. The tools will be developed scientifically using the methodology of design-oriented research and will be based on preliminary work from the literature as well as on the competencies of experts and users.

In addition, the IEDS project uses microeconometric estimation methods to investigate the causal influence of Big Data Analysis on company performance. Since the empirical literature has not yet come to a clear conclusion on this relationship and previous studies have mainly shown correlations, the analysis of the causal effect is intended to provide a decision-making basis for companies to evaluate the benefits of Big Data Analysis.

The results can thus contribute to strategy development as well as to the adaptation and expansion of business models.

- **Data evaluation** (see Section 3.4): An economic model will be developed as part of the project to weigh up the costs and benefits of the combinatorial evaluation of data. It should be possible to calibrate the model with experimental data in order to represent the modeled relationships as realistically as possible. Among other things, it will compare the benefits generated by data bundles with the effort required to determine optimal data bundles. This enables data users, for example, to estimate under which conditions the determination of optimal data bundles is worthwhile, which methods they should use for this purpose, and which willingness to pay they should have for certain data bundles. For data providers, this allows them to anticipate such relationships and take this into account, for example, when collecting data. An essential prerequisite for the calibration of such a model is information on the current technical possibilities for combinatorial data valuation. The status quo of these possibilities will be surveyed in an online competition. With regard to the valuation of data as intangible assets, in the further course of the project it will be worked out how data as intangible assets could be valued in a cost-oriented manner. This should help companies to understand data as an intangible asset and as a tradable good. Based on this, it could be taken into account in accounting. The focus of further work is less on the details of accounting and more on an evaluation of the value contribution of data.

- **Data law** (see Section 3.5): The project will analyze the status quo of the legal regulatory framework and conduct a follow-up legal-economic evaluation. In addition, in the sense of a needs analysis, it will be shown which legal — but also economic, technical and organizational — obstacles have so far stood in the way of data sharing. Deeper insights into the types and significance of various obstacles can reveal needs for action and provide impetus for concrete recommendations for action for political decision-makers. Further future work of the Data Law Work Package is directed at enabling companies to manage and share data in a legally compliant manner. This is to be fostered by the following two key elements: a best practice catalog, which, in addition to a classifying ontology of data law, also outlines examples against the background of use cases in order to provide orientation and navigation for legally compliant use of data (1), and a contract generator, i.e., an interactive toolbox for easy generation and individual composition of contracts that can be integrated into existing data transfer architectures, such as Gaia-X (2).

The aforementioned strategic activities are intended to demonstrate incentives for companies to participate in data sharing and data spaces, and to lower the associated hurdles and challenges. To cover the technological, economic and legal aspects, the collaboration of five institutions results (see Section 5). The project partners involved are presented below.

5 IEDS research project overview and project partner presentation

The IEDS research project highlights functionalities of the data economy and also presents incentives for sharing and exchanging data in order to participate in it.

The multifaceted and interdisciplinary topics of the data economy make it clear that the coverage of technological, economic and legal aspects is equally necessary. Based on this, the following institutions emerge as participating institutions, combining the necessary expertise in a spectrum of competence:

Fraunhofer ISST

The **Fraunhofer Institute for Software and Systems Engineering (ISST)** in Dortmund, as the applicant, assumes the role of the lead institution and coordinates and controls all activities in the course of the project. Fraunhofer ISST has been researching the value of and confident handling of data for over 25 years. The competencies of the Data Science Department lie in the consulting, conception and implementation of data strategies, the development of solutions for data management, the construction of data architectures, the evaluation of data assets as well as in the field of data analysis and artificial intelligence. Through the applied research of Fraunhofer ISST, the latest scientific findings are developed in cooperation with industrial companies and transferred into practice.

Fraunhofer IAO

The **Fraunhofer Institute for Industrial Engineering (IAO)** develops strategies, business models and solutions for digital transformation together with companies, institutions and public sector organizations. The Digital Business Services research team supports organizations in the digital transformation of business models, service offerings and business processes. A methodical and model-based combination of strategic and technical aspects (digital business models, smart service offerings, data and service ecosystems, etc.), technical aspects (IT architectures, Internet of Things, etc.) for the conception and implementation of smart services and artificial intelligence applications is the focus of the activities.

The **German Economic Institute (IW)** is a private, non-commercial research institute. It is supported by employers' associations, trade associations and companies. On a scientific basis, it develops analyses and statements on all issues of economic and social policy, the education and training system, and the labor market. Characteristic of the IW's work is the close combination of scientific analysis based on sound theoretical knowledge as well as empirical research and target group-oriented public relations work. The IW has experience in the economic research of relevant topics such as new data-driven business models and platforms, challenges in the digital transformation for companies, development of digital maturity models as well as data economics including the context-dependent economic analysis of the law.

Digital technologies are changing the way we work and are having a profound impact on business and society. Long-established methods and processes are being modernized and revolutionized by digitization in the shortest of timeframes.

technische universität
dortmund

The **Chair of Industrial Information Management (IIM)** of the Faculty of Mechanical Engineering at the Technical University of Dortmund researches innovative concepts, processes, architectures and solutions for business and logistics networks. The work is characterized by an interdisciplinary approach to the research subject at the interface of engineering, business administration and computer science. A special focus of the chair is on basic research in the areas of data management and data-driven business models. Due to its connection to the Technical University, the chair offers numerous opportunities for knowledge transfer into university education as well as the promotion of young scientists. A graduate support network already exists at the chair as a support and further education measure with regard to doctoral studies at TU Dortmund University as well as at Fraunhofer ISST. In addition, there is participation in the Graduate School of Logistics, in which doctoral students are guided methodically and in terms of content towards a doctorate.

ZEW

The **ZEW - Leibniz Centre for European Economic Research** in Mannheim is a non-profit economic research institute in the legal form of a limited liability company and a member of the Leibniz Association. The overarching research guiding principle at ZEW is the analysis and design of functioning markets and institutions in Europe. ZEW is open to interdisciplinary cooperation and perspectives. The Digital Economy research area investigates how digitization influences economic processes. It analyzes the effects of digitization on production, innovation and the world of work as well as the functioning of digital markets and platforms. Methodologically, the research area follows an empirical-quantitative approach. Data from its own company surveys and from Internet platforms as well as macroeconomic databases are evaluated using statistical and econometric methods. With this profile, the department is a central point of contact in Germany for the economic analysis of digitization.

6 Sources

Agahari, Wirawan; Dolci, Riccardo; Reuver, Mark de (2021): Business Model Implications of privacy-preserving technologies in data marketplaces: The case of multi-party computation. In: Twenty-Ninth European Conference on Information Systems (ECIS 2021), p. 1–16.

Ai, Qingyao; Zhang, Yongfeng; Bi, Keping; Chen, Xu; Croft, W. Bruce (2017): Learning a Hierarchical Embedding Model for Personalized Product Search. In: Noriko Kando (Hg.): Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '17: The 40th International ACM SIGIR conference on research and development in Information Retrieval. Shinjuku Tokyo Japan, 07.08.2017. ACM Special Interest Group on Information Retrieval. New York: ACM (ACM Digital Library), p. 645–654, last retrieved on 04.11.2021.

Altamony, Hamzah; Alshurideh, Muhammad Turki; Masa'deh; Ra'ed; Obeidat, Bader (2012): Information Systems for Competitive Advantage: Implementation of an Organisational Strategic Management Process. In: Innovation and Sustainable Competitive Advantage: From Regional Development to World Economies.

Asch, M.; Moore, T.; Badia, R.; Beck, M.; Beckman, P.; Bidot, T. et al. (2018): Big data and extreme-scale computing: Pathways to Convergence-Toward a shaping strategy for a future software and data ecosystem for scientific inquiry. In: The International Journal of High Performance Computing Applications 32 (4), S. 435–479. Available online at https://journals.sagepub.com/doi/full/10.1177/1094342018778123, last retrieved on 30.11.2021.

Aslett, Matt (2019): DataOps Lays the Foundation for Agility, Security and Transformational Change. Hg. v. 451 Research.

Azkan, Can; Iggena, Lennart; Gür, Inan; Möller, Frederik; Otto, Boris (2020a): A Taxonomy for Data-Driven Services in Manufacturing Industries. In: Proceedings of the 24th Pacific Asia Conference on Information Systems (PACIS). Dubai, United Arab Emirates.

Azkan, Can; Iggena, Lennart; Meisel, Lukas; Spiekermann, Markus; Korte, Tobias; Otto, Boris et al. (2020b): DEMAND – Perspektiven der Datenwirtschaft (Use Case Report). Available online at https://www.demand-projekt.de/.

Azkan, Can; Iggena, Lennart; Möller, Frederik; Otto, Boris (2021): Towards Design Principles for Data-Driven Services in Industrial Environments. In: Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS). Hawaii, USA.

Azkan, Can; Möller, Frederik; Meisel, Lukas; Otto, Boris (2020c): SERVICE DOMINANT LOGIC PERSPECTIVE ON DATA ECOSYSTEMS – A CASE STUDY BASED MORPHOLOGY. In: Proceedings of the 28th European Conference on Information Systems (ECIS).

Badewitz, Wolfgang; Kloker, Simon; Weinhardt, Christof (2020): The Data Provision Game: Researching Revenue Sharing in Collaborative Data Networks, p. 191–200. DOI: 10.1109/CBI49978.2020.00028.

Bansal, Sharu; Kumar, Dilip (2020): IoT Ecosystem: A Survey on Devices, Gateways, Operating Systems, Middleware and Communication. In: International Journal of Wireless Information Networks 27 (3), p. 340–364. Available online at https://link.springer.com/article/10.1007/s10776-020-00483-7, last retrieved on 23.12.2021.

Bengio, Yoshua; Lodi, Andrea; Prouvost, Antoine (2018): Machine Learning for Combinatorial Optimization: a Methodological Tour d'Horizon. Available online at http://arxiv.org/pdf/1811.06128v2.

Bertschek, I., Briglauer, W., Hüschelrath, K., Kauf, B., & Niebel, T. (2015). The Economic Impacts of Broadband Internet: A Survey. Review of Network Economics, 14(4), 201–227.

Biagi, F. (2013). ICT and Productivity: A Review of the Literature. Institute for Prospective Technological Studies Digital Economy Working Paper No. 2013/09.

BMWi: Altmaier: „Wir brauchen eine eigene europäische Dateninfrastruktur!“. Available online at https://www.de.digital/DIGITAL/Redaktion/DE/Meldungen/2019/2019-10-28-altmaier-wir-brauchen-eine-eigene-europaeische-dateninfrastruktur.html, last retrieved on 28.10.2021.

BMWi; BMBF (2019): Das Projekt GAIA-X, Eine vernetzte Dateninfrastruktur als Wiege eines vitalen, europäischen Ökosystems. Berlin.

Borowiecki, Martin; Pareliussen, Jon; Glocker, Daniela; Kim, Eun Jung; Polder, Michael; Rud, Iryna (2021): The impact of digitalisation on productivity: Firm-level evidence from the Netherlands. In: OECD Economics Department Working Papers. DOI: 10.1787/e800ee1d-en.

Brynjolfsson, Erik; Hitt, Lorin M.; Kim, Heekyung Hellen (2011): Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? In: SSRN Journal. DOI: 10.2139/ssrn.1819486.

Brynjolfsson, Erik; McElheran, Kristina (2019): Data in action: data-driven decision making and predictive analytics in US manufacturing. In: Rotman School of Management Working Paper (3422397).

Bundeskanzleramt (2021): Datenstrategie der Bundesregierung. Available online at https://www.bundesregierung.de/breg-de/suche/datenstrategie-der-bundesregierung-1845632, last retrieved on 02.12.2021.

Byrne, D., Oliner, S., & Sichel, D. (2013). Is the Information Technology Revolution Over? International Productivity Monitor, 25, 20–36.

Chapman, Adriane; Simperl, Elena; Koesten, Laura; Konstantinidis, George; Ibáñez, Luis-Daniel; Kacprzak, Emilia; Groth, Paul (2020): Dataset search: a survey. In: The VLDB Journal 29 (1), p. 251–272. DOI: 10.1007/s00778-019-00564-x.

Cirullies, Jan; Schwede, Christian (2021): On-demand Shared Digital Twins – An Information Architectural Model to Create Transparency in Collaborative Supply Networks. In: Tung Bui (Hg.): Proceedings of the 54th Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences: Hawaii International Conference on System Sciences (Proceedings of the Annual Hawaii International Conference on System Sciences).

Constantiou, Ioanna.; Kallinikos, Jannis (2015): New Games, New Rules: Big Data and the Changing Context of Strategy. In: Journal of Information Technology 30 (1), S. 44–57. DOI: 10.1057/jit.2014.17.

Dahlberg, Tomi; Nokkala, Tiina (2019): Willingness to Share Supply Chain Data in an Ecosystem Governed Platform – An Interview Study. In: BLED 2019 Proceedings, p. 619–638. DOI: 10.18690/978-961-286-280-0.33.

Dakova, J.; Antunes, P.; Chiu, Y. (2018): A Pluralistic Approach to Information Valuation. In: 22nd Pacific Asia Conference on Information Systems.

DalleMule, Leandro; Davenport, Thomas H. (2017): What's Your Data Strategy? In: Harvard Business Review.

Demary, Vera; Fritsch, Manuel; Goecke, Henry; Krotova, Alevtina; Azkan, Can; Krote, Tobias et al. (2019): Readiness Data Economy. Bereitschaft der deutschen Unternehmen für die Teilhabe an der Datenwirtschaft.

Destasis (2020): Unternehmensregister. Rechtliche Einheiten und abhängig Beschäftigte nach Beschäftigtengrößenklassen und Wirtschaftsabschnitten. Available online at https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Unternehmen/Unternehmensregister/Tabellen/unternehmen-beschaeftigtengroessenklassen-wz08.html, last retrieved on 25.11.2021.

Destasis (2021): Jedes dritte deutsche Unternehmen nutzte 2020 Cloud Computing. Available online at https://www.destatis.de/DE/Presse/Pressemitteilungen/2021/05/PD21_241_52911.html, last retrieved on 25.11.2021.

Dhyne, Emmanuel; Konings, Jozef; van den Bosch, Jeroen; Vanormelingen, Stijn (2018): IT and productivity: A firm level analysis. In: NBB Working Paper No. 346.

Dremel, Christian (2017): Barriers to the Adoption of Big Data Analytics in the Auto-motive Sector. In: Twenty-third Americas Conference on Information Systemss, p. 1–10.

Draca, M., Martin, R., & Sanchis-Guarner, R. (2018). The evolving role of ICT in the economy. Report by LSE Consulting for Huawei.

Draca, M., Sadun, R., & van Reenen, J. (2007). Productivity and ICT: A Review of the Evidence. In R. Mansell (Ed.), The Oxford Handbook of Information and Communication Technologies (pp. 100–147). Oxford University Press.

Engels, B., & Goecke, H. (2019). Big Data in Wirtschaft und Wissenschaft: Eine Bestandsaufnahme. IW-Analysen: Vol. 130. Institut der Deutschen Wirtschaft Köln Medien GmbH.

Ereth, Julian (2018): DataOps – Towards a Definition. In: Lernen, Wissen, Daten, Analysen.

European Commission (2020a): MITTEILUNG DER KOMMISSION AN DAS EUROPÄISCHE PARLAMENT, DEN RAT, DEN EUROPÄISCHEN WIRTSCHAFTS- UND SOZIALAUSSCHUSS UND DEN AUSSCHUSS DER REGIONEN: Eine europäische Datenstrategie. Available online at https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52020DC0066&from=DE, zuletzt geprüft am 02.12.2021.

European Commission (2020b): THE EUROPEAN DATA STRATEGY. https://ec.europa.eu/commission/presscorner/detail/en/fs_20_283

European Commission (Hg.) (2020c): Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates über europäische Daten-Governance. Available online at https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52020PC0767&from=EN, last updated on 25.11.2020, last retrieved on 21.08.2021.

European Commission (2021): Data Act & amended rules on the legal protection of databases. Available online at https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13045-Data-Act-&-amended-rules-on-the-legal-protection-of-databases_en.

European Union (2016): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, p. 47–194. DOI: 10.9785/9783504385781-007.

Experian ltd. (2019): 2019 Global data management research – Taking control in the digital age.

Fabritz, Nadine (2015): ICT as an Enabler of Innovation. Evidence from German Micro-data. Munich: ifo Institute - Leibniz Institute for Economic Research at the University of Munich (ifo Working Paper, 195). Available online at https://www.econstor.eu/handle/10419/108762.

Farke, Florian; Rensinghoff, Jan; Dürmuth, Markus; Gostomzyk, Tobias (2019): Recht auf Vergessenwerden. Chancen und Grenzen der technischen Umsetzung. In: Datenschutz und Datensicherheit 43 (11), p. 681–685.

Fries, Martin; Scheufen, Marc (2019): Märkte für Maschinendaten: Eine rechtliche und rechtsökonomische Standortbestimmung. In: MMR 22 (11), p. 721–726.

Fruhwirth, Michael; Rachinger, Michael; Prlja, Emina (2020): Discovering Business Models of Data Marketplaces. In: Tung Bui (Hg.): Proceedings of the 53rd Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences: Hawaii International Conference on System Sciences (Proceedings of the Annual Hawaii International Conference on System Sciences).

GAIA-X (2021): GAIA-X Federation Services. Available online at https://www.gxfs.de/.

Gaia-X European Association for Data and Cloud AISBL (2021): Gaia-X Architecture Document. 21.09 Release. Brüssel. Available online at https://www.gaia-x.eu/sites/default/files/2021-10/Gaia-X_Architecture_Document_2109.pdf, last retrieved on 29.11.2021.

Gal, P., Nicoletti, G., Renault, T., Sorbe, S., & Timiliotis, C. Digitalisation and productivity: In search of the holy grail – Firm-level empirical evidence from EU countries. OECD Economics Department Working Papers, 2019. https://doi.org/10.1787/18151973

Gelhaar, J., Groß, T., Otto, B., 2021. A Taxonomy for Data Ecosystems, in: Proceedings of the 54th Hawaii International Conference on System Sciences, Hawaii, USA

Gelhaar, Joshua; Gürpinar, Tan; Henke, Michael; Otto, Boris (2021b): Towards a taxonomy of incentive mechanisms for data sharing in data ecosystems. In: PACIS 2021 Proceedings, p. 1–14.

Gelhaar, Joshua; Otto, Boris (2020): Challenges in the Emergence of Data Ecosystems. In: PACIS 2020 Proceedings, p. 1–14.

Ghorbani, A.; Zou, J. (2019): Data Shapley: Equitable Valuation of Data for Machine Learning. Available online at http://arxiv.org/pdf/1904.02868v2.

Guggenberger, Tobias; Möller, Frederik; Boualouch, Karim; Otto, Boris (2020a): Towards a Unifying Understanding of Digital Business Models. In: Twenty-Third Pacific Asia Conference on Information Systems, p. 1–14.

Guggenberger, Tobias; Möller, Frederik; Haarhaus, Tim; Gür, Inan; Otto, B. (2020b): ECOSYSTEM TYPES IN INFORMATION SYSTEMS. In: Twenty-Eighth European Conference on Information Systems.

Gür, Inan (2021): DataOps for Data Sharing – Challenges and Requirements for interorganizational Data Sharing. Hg. v. Fraunhofer ISST (ISST-Report).

Gür, Inan; Spiekermann, Markus; Arbter, Michael; Otto, Boris (2021): Data Strategy Development A Taxonomy for Data Strategy Tools and Methodologies in the Economy. In: Internationale Konferenz Wirtschaftsinformatik.

Gurevich, Aleksey; Dey, Srijani (2018): Defining a data strategy. Hg. v. DXC.Technology.

Hartmann, Philipp Max; Zaki, Mohamed; Feldmann, Niels; Neely, Andy (2016): Capturing value from big data – a taxonomy of data-driven business models used by start-up firms. In: International Journal of Operations and Production Management 36 (10), p. 1382–1406.

HARVARD BUSINESS REVIEW ANALYTIC SERVICES, 2019. Critical success factors to achieve a better enterprise data strategy in a multi-cloud environment. Harvard Business Review Analytic Services.

Henderson, Deborah; Earley, Susan (Hg.) (2017): DAMA-DMBOK. Data management body of knowledge. Data Administration Management Association. Second edition. Basking Ridge, New Jersey: Technics Publications.

Hupperz, Marius Johannes; Gür, Inan; Möller, Frederik; and Otto, Boris (2021): What is a Data-Driven Organization? In: AMCIS 2021 Proceedings. 6, Bd. 6.

Hurley, John (2018): WHY YOUR DATA STRATEGY IS YOUR B2B GROWTH STRATEGY. In: Harvard Bus. Rev.

Inklaar, R., O'Mahony, M., & Timmer, M. P. (2005). ICT and Europe's Productivity Performance: Industry-Level Growth Account Comparisons with the United States. Review of Income and Wealth, 51(4), 505–536.

International Data Spaces Association (2019): Reference Architecture Model. Version 3.0 April 2019. Berlin. Available online at https://internationaldataspaces.org/download/16630/, last retrieved on 29.11.2021.

International Data Spaces Association (2021): BDVA, FIWARE, GAIA-X and IDSA Launch an Alliance to Accelerate Business Transformation in the Data Economy. Brüssel, Berlin. Available online at https://internationaldataspaces.org/download/28798/, last retrieved on 30.11.2021.

Jackson, Joshua Conrad; Rand, David; Lewis, Kevin; Norton, Michael I.; Gray, Kurt (2017): Agent-Based Modeling. In: Social Psychological and Personality Science 8 (4), p. 387–395. DOI: 10.1177/1948550617691100.

Jorgenson, D., & Stiroh, K. (2000). Raising the Speed Limit: U.S. Economic Growth in the Information Age. Brookings Papers on Economic Activity, 31(1), 125–236.

Kacprzak, Emilia; Koesten, Laura; Ibanez, Luis-Daniel; Blount, Tom; Tennison, Jeni; Simperl, Elena (2018): Characterising Dataset Search – An Analysis of Search Logs and Data Requests. In: SSRN Journal. DOI: 10.2139/ssrn.3287149.

Koesten, Laura M.; Kacprzak, Emilia; Tennison, Jenifer F. A.; Simperl, Elena (2017): The Trials and Tribulations of Working with Structured Data. In: Gloria Mark (Hg.): Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. CHI '17: CHI Conference on Human Factors in Computing Systems. Denver Colorado USA, 06 05 2017 11 05 2017. Association for Computing Machinery-Digital Library; ACM Special Interest Group on Computer-Human Interaction. New York, NY: ACM (ACM Digital Library), p. 1277–1289, last retrieved on 21.10.2021.

Koskinen, Jani; Knaapi-Junnila, Sari; Rantanen, Minna M. (2019): What if we Had Fair, People-Centred Data Economy Ecosystems? In: Institute of Electrical and Electronics Engineers (Hg.): 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (Smart-World/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), p. 329–334. Available online at https://ieeexplore.ieee.org/document/9060350, last retrieved on https://ieeexplore.ieee.org/document/9060350.

Koutroumpis, Pantelis; Leiponen, Aija; Thomas, Llewellyn D. W. (2017): The (Unfulfilled) Potential of Data Marketplaces. In: ETLA Working Papers (53). Available online at http://pub.etla.fi/ETLA-Working-Papers-53.pdf, last retrieved on 07.07.2021.

Kristandl, Gerhard; Bontis, Nick (2007): Constructing a definition for intangibles using the resource based view of the firm. In: Management Decision 45 (9), p. 1510–1524. DOI: 10.1108/00251740710828744.

Krotova, Alevtina; Mertens, Armin; Scheufen, Marc (2020): Open data and data sharing – An Economic Analysis. In: IW-Policy Paper, Bd. 21. Köln.

Krotova, Alevtina; Rusche, Christian; Spiekermann, Markus (2019): Die ökonomische Bewertung von Daten. Verfahren, Beispiele und Anwendungen = The economic evaluation of data : procedures, examples and applications. Köln: Institut der deutschen Wirtschaft Köln Medien GmbH (IW-Analysen, 129).

Kühne, Babett; Böhmann, Tilo (2018): Requirements for Representing Data-Driven Business Models – Towards Extending the Business Model Canvas. In: Proceedings 24th Americas Conference on Information Systems. New Orleans: USA.

Kwon, Yongchan; Rivas, Manuel A.; Zou, James (2021): Efficient Computation and Analysis of Distributional Shapley Values. In:. International Conference on Artificial Intelligence and Statistics: PMLR, p. 793–801. Available online at http://proceedings.mlr.press/v130/kwon21a/kwon21a.pdf.

Laney, Douglas B. (2018): Infonomics. How to monetize, manage, and measure information as an asset for competitive advantage. New York, New York: Bibliomotion Inc. Available online at https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5023889.

Lee, Sokbae; Ng, Serena (2020): An Econometric Perspective on Algorithmic Subsampling. In: Annu. Rev. Econ. 12 (1), p. 45–80. DOI: 10.1146/annurev-economics-022720-114138.

Lenny, L. (2014): 3 reasons why DataOps is essential for big data success — IBM Big Data and Analytics Hub. IBM. Available online at https://www.ibmbigdatahub.com/blog/3-reasons-why-dataops-essential-big-data-success, last updated on 12.08.2021.

Lusch, Robert F.; Vargo, Stephen L. (2014): Service-Dominant Logic. Premises, Perspectives, Possibilities. Cambridge: Cambridge Unitversity Press.

Madera, Karen; Aguilera, Diana Paredes (2020): Deliver business-ready data fast with DataOps. An introduction to the IBM DataOps methodology and practice. Hg. v. IBM.

Mainali, Kiran; Ehrlinger, Lisa; Matskin, Mihhail; Himmelbauer, Johannes (2021): Discovering DataOps: A Comprehensive Review of Definitions, Use Cases, and Tools. In: DATA ANALYTICS 2021 : The Tenth International Conference on Data Analytics.

Malaka, Iman; Brown, Irwin (2015): Challenges to the Organisational Adoption of Big Data Analytics: A Case Study in the South African Telecommunications Industry. In: Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists.

Mazhelis, Oleksiy; Luoma, Eetu; Warma, Henna (2012): Defining an Internet-of-Things Ecosystem. In: Sergey Andreev, S. Balandin und Yevgeni Koucheryavy (Hg.): Internet of Things, Smart Spaces, and Next Generation Networking. 12th International Conference, NEW2AN 2012, and 5th Conference, ruSMART 2012, St. Petersburg, Russia, August 27-29, 2012. Proceedings. Berlin, Heidelberg: Springer-Verlag, p. 1–14. Available online at https://link.springer.com/chapter/10.1007/978-3-642-32686-8_1, last retrieved on 30.11.2021.

Mazzei, Matthew J.; Noble, David (2017): Big data dreams: A framework for corporate strategy. In: Business Horizons 60 (3), p. 405–414. DOI: 10.1016/j.bushor.2017.01.010.

Mielli, Fabio; Bulanda, Nicole (2019): Digital Transformation: Why Projects Fail, Potential Best Practices and Successful Initiatives. In: 2019 IEEE-IAS/PCA Cement Industry Conference (IAS/PCA). 2019 IEEE-IAS/PCA Cement Industry Conference (IAS/PCA). St Louis, MO, USA: IEEE, p. 1–6.

Mildebrath, Hendrik (2021): Data Governance Act. Briefing – EU Legislation in Progress. Available online at https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/690674/EPRS_BRI(2021)690674_EN.pdf, last updated on 06.2021, last retrieved on 21.08.2021.

Moktadir, Md. Abdul; Ali, Syed Mithun; Paul, Sanjoy Kumar; Shukla, Nagesh (2019): Barriers to Big Data Analytics in Manufacturing Supply Chains: A Case Study from Bangladesh. In: Computers & Industrial Engineering 128, p. 1063–1075.

Mosig, Tim; Lehmann, Claudia; Neyer, Anne-Katrin (2021): Data-Driven Business Model Innovation. About Barriers and New Perspectives. In: International Journal of Innovation and Technology Managerment (02), 2040017-1-20400174-16.

Mosley, Mark; Brackett, Michael; Earley, Susan (Hg.) (2010): The DAMA guide to the data management body of knowledge. (DAMA-DMBOK guide). First edition. Brad-ley Beach, NJ: Technics Publications LLC.

Müller, Oliver; Fay, Maria; vom Brocke, Jan (2018): The effect of big data and analytics on firm performance. In: Journal of management information systems.

Nexla (2018): The definitive Data Operations Report. Nexla Inc.

Nexla Inc. (2018): The Definitive Data Operations Report.

Niebel, Thomas; Rasel, Fabienne; Viete, Steffen (2019): BIG data – BIG gains? Understanding the link between big data analytics and innovation. In: Economics of Innovation and New Technology 28 (3), p. 296–316. DOI: 10.1080/10438599.2018.1493075.

Ocean Tomo (2020): Intagible Asset Market Value Study.

Oliner, S., & Sichel, D. (2000). The Resurgence of Growth in the Late 1990s: Is Information Technology the Story? Journal of Economic Perspectives, 14(4), 3–22.

Oliner, S., & Sichel, D. (2002). Information Technology and Productivity: Where are we now and Where are we Going? Economic Review, Q3, 15–44.

Oliveira, Marcelo Iury S.; Lóscio, Bernadette Farias (2018): What is a data ecosystem? In: Marijn Janssen, Soon Ae Chun und Vishanth Weerakkody (Hg.): Proceedings of the 19th Annual International Conference on Digital Government Research Governance in the Data Age – dgo '18. the 19th Annual International Conference. Delft, The Netherlands, 30.05.2018 - 01.06.2018. New York, New York, USA: ACM Press, p. 1–9.

Oliveira, Marcelo Iury S.; Oliveira, Lairson Emanuel R. A.; Batista, Marlos G. Ribeiro; Lóscio, Bernadette Farias (2018): Towards a meta-model for data ecosystems. In: dg.o '18: dg.o 2018: Proceedings of the 19th Annual International Conference on Digital Government Research, May 30-June 1, 2018, Delft, Netherlands Anneke Zuiderwijk and Charles C. Hinnant (Eds.). ACM, New York, NY, USA, 10 pages., p. 1–10. DOI: 10.1145/3209281.3209333.

Oppenheim, Charles; Stenson, Joan; Wilson, Richard M. S. (2003): Studies on Information as an Asset I: Definitions. In: Journal of Information Science 29 (3), p. 159–166. DOI: 10.1177/01655515030293003.

Osterwalder, Alexander; Pigneur, Yves (2002): An e-Business Model Ontology for Modeling e-Business. In: Proceedings of the 15th Bled Electronic Commerce Conference. Bled, Slovenia.

Otto, B. (2011): A MORPHOLOGY OF THE ORGANISATION OF DATA GOVERNANCE. In: ECIS 2011.

Otto, Boris; Österle, Hubert (2016): Corporate Data Quality. Voraussetzung erfolgreicher Geschäftsmodelle. Berlin: Springer Gabler.

Otto, Boris; Rubina, Alina; Eitel, Andreas; Teuscher, Andreas; Schleimer, Anna Maria; Lange, Christoph et al. (2021): Gaia-X and IDS. Hg. v. International Data Spaces Association. Dortmund. Available online at https://internationaldataspaces.org/download/19016/, last retrieved on 13.12.2021.

Pandit, Harshvardhan J.; O'Sullivan, Declan; Lewis, Dave (2018): GDPR Data Interoperability Model.

Parvinen, Petri; Pöyry, Essi; Gustafsson, Robin; Laitila, Miikka; Rossi, Matti (2020): Advancing Data Monetization and the Creation of Data-based Business Models. In: Communications of the Association for Information Systems 47, p. 25–49. DOI: 10.17705/1CAIS.04702.

Pentek, Tobias; Legner, Christine; Otto, Boris (2017): Towards a reference model for data management in the digital economy. In: Proceedings of the 12th International Conference on Design Science Research in Information Systems and Technology.

Richter, Heiko; Slowinski, Peter R. (2019): The Data Sharing Economy: On the Emergence of New Intermediaries. In: IIC – International Review of Intellectual Property and Competition Law 50, p. 4–29. DOI: 10.1007/s40319-018-00777-7.

Röhl, Klaus-Heiner / Bolwin, Lennart / Hüttl, Paula, 2021, Datenwirtschaft in Deutschland. Wo stehen die Unternehmen in der Datennutzung und was sind ihre größten Hemmnisse?, Gutachten im Auftrag des Bundesverbands der Deutschen Industrie e.V. (BDI), Köln Available online at https://www.iwkoeln.de/studien/klaus-heiner-roehl-lennart-bolwin-wo-stehen-die-unternehmen-in-der-datennutzung-und-was-sind-ihre-groessten-hemmnisse.html.

Rosenkranz, Frank; Scheufen, Marc (2021): Rechtliche Rahmenbedingungen der Lizenzierung von nicht-personenbezogenen Daten. IGEDI Kolloquium. Ruhr-Universität Bochum. Bochum, 06.12.2021.

Scheufen, Marc (2020a): Angewandte Mikroökonomie und Wirtschaftspolitik: Mit einer Einführung in die ökonomische Analyse des Rechts. 2. Aufl. Wiesbaden: Springer.

Scheufen, Marc (2020b): Urheberrechtliche Fragen der KI. In: Andreas Leupold, Andreas Wiebe und Silke Glossner (Hg.): IT-Recht. Recht, Wirtschaft und Technik der digitalen Tranformation. 4. Aufl. München: C.H.Beck, p. 1033–1039.

Schüritz, Ronny; Seebacher, Stefan; Dorner, Rebecca (2017): Capturing Value from Data: Revenue Models for Data-Driven Services. In: Proceedings of the 50th Hawaii International Conference on System Sciences. Hawaii, p. 5348–5357.

Schweikl, S., & Obermaier, R. (2020). Lessons from three decades of IT productivity research: towards a better understanding of IT-induced productivity effects. Management Review Quarterly. 70(4), 461-507.

Singapore Digital (2019): Guide to Data Valuation for Data Sharing.

Solow, R. M. (1987). We'd Better Watch Out. New York Times Book Review, 36.

Sparapani, Jason (2019): DataOps and the future of data management. Hg. v. MIT Technology Review Insights.

Spiekermann, Markus (2019): Data Marketplaces: Trends and Monetisation of Data Goods. In: Intereconomics 54 (4), p. 208–216. DOI: 10.1007/s10272-019-0826-z.

Spiezia, Vincenzo (2011): Are ICT Users More Innovative? In: OECD Journal: Economic Studies 2011 (1), p. 1–21. DOI: 10.1787/eco_studies-2011-5kg2d2hkn6vg.

Statista: Marktanteile der führenden Unternehmen am Umsatz im Bereich Cloud Computing weltweit von Juli 2018 bis Juni 2019. Available online at https://de.statista.com/statistik/daten/studie/150979/umfrage/marktanteile-der-fuehrenden-unternehmen-im-bereich-cloud-computing/, last retrieved on 28.11.2021.

Sun, Shiwei; Cegielski, Casey G.; Jia, Lin; Halla, Dianne J. (2016): Understanding the Factors Affecting the Organizational Adoption of Big Data. In: Journal of Computer Information Systems, p. 193–203.

Susha, Iryna; Grönlund, Åke; van Tulder, Rob (2019): Data driven social partnerships: Exploring an emergent trend in search of research challenges and questions. In: Government Information Quarterly 36 (1), p. 112–128. DOI: 10.1016/j.giq.2018.11.002.

Susha, Iryna; Janssen, Marijn; Verhulst, Stefaan (2017): Data Collaboratives as a New Frontier of Cross-Sector Partnerships in the Age of Open Data: Taxonomy Development. In: Proceedings of the 50th Hawaii International Conference on System Sciences. Hawaii, p. 2691–2700.

Syverson, Chad (2011): What Determines Productivity? In: Journal of Economic Literature 49 (2), p. 326–365. DOI: 10.1257/jel.49.2.326.

Tang, Llewellyn C.M.; Zhao, Yuyang; Austin, Simon; Darlington, Mansur; Culley, Steve (2008): A characteristic based information evaluation model. In: Katsumi Tanaka (Hg.): Proceedings of the 2nd ACM workshop on Information credibility on the web. Proceeding of the 2nd ACM workshop. Napa Valley, California, USA, 10/30/2008 – 10/30/2008. Association for Computing Machinery; ACM Special Interest Group on Hypertext, Hypermedia, and Web; ACM Special Interest Group on Information Retrieval. New York, NY: ACM (ACM Conferences), p. 89.

van Reenen, J., Bloom, N., Draca, M., Kretschmer, T., & Sadun, R. (2010). The Economic Impact of ICT. Centre for Economic Performance, London School of Economics.

Vargo, Stephen L.; Akaka, Melissa Achpru; Vaughan, Claudia M. (2017): Conceptualizing Value: A Service-ecosystem View. In: Journal of Creating Value 3 (2), p. 1–8. Available online at https://journals.sagepub.com/doi/10.1177/2394964317732861, last retrieved on 23.12.2021.

Wei, Bifan; Liu, Jun; Zheng, Qinghua; Zhang, Wei; Fu, Xiaoyu; Feng, Boqin (2013): A survey of faceted search. In: Journal of Web Engineering Volume 12 (1-2), p. 41–64. DOI: 10.5555/2481562.2481564.

Wilberg, Julian; Kalla, Tobias; Fetscher, Manuel; Rimböck, Franz; Hollauer, Christoph; Omer, Mayada (2018): "Development of a Use Phase Data Strategy for Connected Products: A Case Study in Industry". Available online at https://ieeexplore.ieee.org/document/8481828.

Wong, Tracey Y.T.; Peko, Gabrielle; Sundaram, David; Piramuthu, Selwyn (2016): Mobile environments and innovation co-creation processes & ecosystems. In: Information & Management 53 (3), p. 336–344. DOI: 10.1016/j.im.2015.09.005.

Wu, Lynn; Hitt, Lorin M.; Lou, Bowen (2020): Data analytics, innovation, and firm productivity. In: Management Science, 66(5), 2017-2039.

Yoon, Jinsung; Arik, Sercann; Pfister, Thomas (2020): Data Valuation using Reinforcement Learning. In: Proceedings of the 37th International Conference on Machine Learning 119, p. 10842–11085.

Zechmann, Andreas (2018): Nutzungsbasierte Datenbewertung. Entwicklung und Anwendung eines Konzepts zur finanziellen Bewertung von Datenvermögenswerten auf Basis des AHP. Berlin: epubli.

Imprint

1st edition, March 2022

Editor
Fraunhofer Institut für Software und Systemtechnik ISST
Emil-Figge-Straße 91
44227 Dortmund

Authors
Fraunhofer-Institut für Software- und Systemtechnik ISST
Dr.-Ing. Can Azkan
Inan Gür
Marius Hupperz
Joshua Gelhaar
Anna Gieß
Tobias Groß

Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO
Sandra Frings
Dr.-Ing. Holger Kett
Damian Kutzias
Oliver Strauß

Institut der deutschen Wirtschaft
Jan Büchel
Dr. Vera Demary
Barbara Engels
Dr. Henry Goecke
Dr. Armin Mertens
Dr. Klaus-Heiner Röhl
Dr. Christian Rusche
Dr. Marc Scheufen
Bjarne Schröder

Technische Universität Dortmund
Valentin Dahms
Ilka Jussen
Dr.-Ing. Frederik Möller
Julia Schweihoff

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung
Raphaela Andres
Dr. Daniel Erdsiek
Dr. Thomas Niebel
Dr. Dominik Rehse
Vincent Rost
Johannes Walter
Sebastian Valet

Typesetting and layout
Elisa Kadelka

The original version of the publication is available at
www.ieds-projekt.de

This research and development project is funded and supervised by the German Federal Ministry of Education and Research (BMBF). The responsibility for the content of this publication lies with the authors.

SPONSORED BY THE

Federal Ministry
of Education
and Research

Contact

Dr.-Ing. Can Azkan
Tel. +49 231 976770
pmo@ieds-projekt.de

Fraunhofer-Institut für Software- und Systemtechnik ISST
Emil-Figge-Str. 91
44227 Dortmund
www.isst.fraunhofer.de