HOW TO SHARE DATA?
DATA SHARING PLATFORMS FOR ORGANIZATIONS

ECONOMIC AND LEGAL FUNDAMENTALS, CURRENT PRACTICAL PROJECTS, FIRST RECOMMENDATIONS FOR ACTION

Study commissioned by the Federal Ministry for Economic Affairs and Energy as part of the scientific assistance for the 'Smart Data Economy' technology programme

Smarte Datenwirtschaft

IMPRINT

Study commissioned by the Federal Ministry for Economic Affairs and Energy as part of the scientific assistance for the 'Smart Data Economy' technology programme

AUTHORS

Maximilian Lindner
Sebastian Straub LL.M.
Dr. Bettina Kühne

EDITOR

Peter Gabriel
Begleitforschung Smarte Datenwirtschaft
Institut für Innovation und Technik (iit)
in der VDI / VDE Innovation + Technik GmbH
Steinplatz 1
10623 Berlin
gabriel@iit-berlin.de

PUBLICATION

April 2021

LAYOUT

LoeschHundLiepold Kommunikation GmbH
Hauptstraße 28
10827 Berlin

PICTURES

pickup (title, Pugun & Photo Studio (p. 39),
ipopba (p. 42), Urupong (p. 45), Urupong (p. 49),
sdecoret (p. 55) – stock.adobe.com

CONTENT

DATA EXCHANGE BETWEEN COMPANIES HAS AN ENORMOUS POTENTIAL. THIS IS MADE POSSIBLE VIA DATA SHARING PLATFORMS.

BUT WHAT ARE DATA SHARING PLATFORMS CONTRIBUTING IN DETAIL AND WHAT CHALLENGES DO PLATFORM OPERATORS FACE? A PRACTICAL ANALYSIS GIVES ANSWERS.

EXECUTIVE SUMMARY

Data is the basis for the development of new products and services and the prerequisite for self-learning systems in digital value chains. Companies can use data as an essential resource to achieve significant competitive advantages. A special kind of added value is created when data from different sources is brought together, aggregated and evaluated.

The exchange of data between organizations is hence becoming increasingly important, even if some organizations are still reluctant to transfer data to a greater extent due to a lack of technical understanding and expertise or doubts about the quality of the available data and fears of being disadvantaged.

Multi-sided data platforms are generally an ideal opportunity to organize cross-organization data exchange between data providers and users. Although first implementations are now already in place, development is still in its infancy. This study will take a look at current practice and derive first suggestions for the design of current and future platforms.

For this purpose, 24 existing platforms were analysed and their operators interviewed. Taking this as a basis, the study presents the basic functions of data sharing platforms and their legal framework.

It then looks at how today's active offerings are shaping up in terms of value proposition, architecture, access, understanding of the operator's role and data transformation, if any. Both domain-specific platforms in the areas of industry, trade, health and logistics as well as domain-unspecific offerings are taken into account. The study finds that there are two types of platforms emerging today. On the one hand, there are the real data sharing platforms where nothing but data sharing takes place. On the other hand, marketplaces are emerging where data is also traded and prices are formed as a basis for purchase transactions on the platform. At least today, data sharing platforms focus strongly on single domains, such as the health sector, while marketplaces and hybrid forms strive to work across domains.

The study concludes with first lessons learned from today's practice for a successful design of platform-based data exchange. A key success factor is the focus on stakeholder demands, which specifically relate to the availability and quality of relevant data and platform security. Agile software and product development supports this focus with a view to possible adaptation and the addition of new features. A clear definition of the respective platform operator's role and its communications with the outside world has also proven to be advantageous when setting up data platforms in order to support the creation of trust among potential participants.

01

1 INTRODUCTION

Data is an essential resource in digital value chains. Depending on the respective application, data that is generated, for example, when operating production machines, in online commerce or while examining patients, forms the basis for new products and services. Data is also often used to develop self-learning systems and artificial intelligence. In many sectors of the economy, the collection and analysis of data can contribute to productivity gains, more efficient use of resources or more evidence-based decision-making within a company's organization. It is hence not surprising that the number of organizations that use data has increased significantly over the last four years according to the results of the European Data Market Monitoring Tool (see European Commission 2020). However, added value is typically only created when data from different sources is exchanged, aggregated and evaluated since the data available in an organization is insufficient or inadequate in terms of volume or quality. Organization-spanning data exchange is proving to be a key factor for the development of innovative business models (see Otto et al. 2016). One possible solution to realize this exchange is the model of multi-sided data platforms that organize interaction between data providers, data users and a platform operator. On the other hand, many organizations are either not yet sufficiently aware of the potential of data use and its exchange or are opposed to opening up due to important obstacles. The most important factor is fear of unauthorized access to sensitive organization data (see KPMG 2020). This is because the exchange of data can also go hand in hand with serious disadvantages – especially for data providers – for example when it comes to machine data and company information that is of interest to competitors or the personal data of end customers, users or patients. It is therefore of vital importance for data providers to retain control over the use of the data provided. That being said, sharing must also pay off for data providers: The collection, processing and provision of data requires an effort that must be rewarded by commercial benefits. From the data user's point of view, data quality is particularly important, but often cannot be verified directly. This requires trust structures to assure that the data products offered to users meet certain criteria, such as completeness, correctness and up-to-dateness. For platform operators, it is particularly important to find a business model that justifies operating the platform from an economic point of view and offers the greatest possible flexibility in the design of the service portfolio. This can lead to conflicts of interest between data providers and data users, for example, if the platform operator uses its natural information advantage to compete with data providers with its own data products. When designing the operator role, it is therefore important to create trust among stakeholders that the platform is unbiased and fair. Appropriate organizational or legal instruments must be created for this purpose.

— ORGANIZATION-SPANNING DATA EXCHANGE IS A KEY FACTOR FOR THE DEVELOPMENT OF INNOVATIVE BUSINESS MODELS. —

Several concepts were developed in the past to show how data sovereignty can be legally regulated. Discussions included the establishment of a 'right to data ownership' or the creation of new ancillary copyrights. Existing laws do not foresee the concept of a property-like right to data, nor have approaches to create such a right been pursued further so far. Furthermore, existing property rights, such as database protection under copyright law, do not guarantee comprehensive protection and can only be used partially to guarantee data sovereignty. In practice, the rules for data access are therefore agreed individually for each B2B platform and laid down in agreements between platform operators and participants. The rights and obligations of the contracting parties, including remuneration, are also determined there. This opens up a very wide range of design options.

— ESTABLISHING TRUST OF THE PARTICIPANTS IN THE NEUTRALITY AND FAIRNESS OF THE PLATFORM PLAYS AN IMPORTANT ROLE. —

Digital multi-sided B2B data platforms are still in a rather early development phase compared to their B2C counterparts. However, there is still no up-to-date and broad-based overview of established B2B data platforms with a systematic overview of their governance structures and identifiable business model elements. Although some overview studies address the topic of identifying data marketplaces and data sharing platforms (see Spiekermann, 2019), these are either out of date, purely theory-forming or strongly focused on a specific economic sector (see Otto et al. 2020). Another service that is of particular practical relevance for data users is the Datarade.ai directory website which supports searches for data sources with suitable content and hence also corresponding platform offers. Other players include, for instance, the European Commission's Support Centre for Data Sharing (https://eudatasharing.eu) which attempts to structure the confusing landscape of offers. Besides informing about the legal framework of sharing or monetizing data and presenting technical implementation options, the focus of interest is primarily on researching and documenting successful data sharing models. Although exciting examples are given here, the presentation is rather episodic and neither clearly structured nor particularly comprehensive.

In the Smart Data Economy programme, organizations and research institutions design and test innovative data products and data services in 21 application-oriented joint projects. Several of these projects are also working on data sharing platforms. The study is intended to offer the partners in the Smart Data Economy programme, but also other interested parties, practice-oriented support on how cross-organization data exchange via digital platforms can benefit all participants in a commercially sensible and legally secure manner, what examples of data sharing platforms currently exist and what recommendations for action can already be derived today. In order to narrow down the thematic focus, the study does not deal with the technical implementation of the platforms in detail.

The study was designed as concept-driven baseline research. Based on a literature and internet search, more than 100 potentially relevant data sharing platforms or data marketplaces were identified in Europe and North America. In the next step, those platforms and marketplaces were analyzed which

- actually include data sharing as a core function,
- already have a real operating history – even in pilot projects,
- address the priority sectors in the Smart Industry programme, i.e., logistics, trade and the health sector or which have a cross-domain orientation and
- are open to German market players.

The outcome of this process is 24 practical examples of existing data marketplaces or data sharing platforms. Information sources are the presentations of the operators' own websites and – if available – press articles or individual scientific case studies as well as a series of qualifying guideline interviews with some of the representatives of the respective platform. If available, information regarding the number of data sets, transactions, prices and sales is provided. However, this information is not provided throughout. This is a clear sign that the market for data sharing platforms is still at an early stage.

The study deals with data sharing in the form of data exchange via platforms. Two terms are often used simultaneously in this context that have related, but not identical meanings, i.e., data marketplaces and data sharing platforms.

These two terms differ with a view to their underlying transaction mechanisms. Prices on data marketplaces are determined directly by balancing supply and demand, whereas prices on data sharing platforms are not charged directly by the data provider and access to the platform may be subject to a fee. Monetization of the data then takes place indirectly via the platform operator. In the interest of better readability, the term 'data sharing platform' will be used for both forms in this study.

Chapter 2 of the study first deals with the basic commercial benefits of data sharing platforms and their obstacles and uses these as a basis for deriving their fundamental functionalities. Chapter 3 addresses the legal challenges of sharing data via sharing platforms. Chapter 4 describes the organizational structures and business models that currently prevail, while chapter 5 presents the platforms that are relevant today in detail. Chapter 6 draws a preliminary conclusion: What are the main points of focus of today's platforms, and what lessons learned are already emerging?

We would like to thank the experts who kindly made themselves available for interviews:

- Rolf Apweiler (Director, EMBL-EBI/European COVID 19 Data Platform)
- Jürgen Bretfeld (Advaneo Data Marketplace)
- David Knight (CEO, TERBINE)
- Prof. Ulrich L. Manz (CEO, IfCC GmbH/VTH-eData-Pool)
- Sebastian Wiemann (Product Manager, Telekom Data Intelligence Hub)

Another three interviewees who took part in the confidential interviews did not wish to be named.

02

2 DATA SHARING PLATFORMS: BASIC CONCEPTS

The term 'data sharing' describes a process where data is exchanged between different players. The term is typically used in conjunction with the exchange of data between businesses (B2B). The European Commission uses the term 'data sharing' in this context, referring to all conceivable models of data exchange between businesses (B2B), from provision and use right through to the exploitation of data. According to the European Commission's definition, this must be distinguished from the exchange of data between businesses and governments (B2G) or data sharing between businesses and consumers (B2C) (see Arnaut et al 2018). The present study pursues a pragmatic approach by focusing on B2B platforms, however, without excluding other systems, such as Otonomo (see section 5.1.3) and Dateva (see section 5.4.4), which also cover B2C and B2G.

2.1 Benefits of data sharing

As the digital transformation of the business world increases, high-quality and available data as a digital image of product or process-relevant information represents a central resource for an organization's success. This information even forms the basis for new data-driven business models or changed service offerings. Accordingly, the use and processing of internal and external data significantly determines the innovative strength and economic success of companies (see Ringel et al. 2017). The drastic increase in potential data sources due to greater use of mobile devices and applications, the growing digital transformation of objects in the Internet of Things (see Fricker et al. 2017) and developments in the area of Open Data are reinforcing this trend. Due to the strong increase in data sources and the premise of ever-more complex value creation processes, however, there is also an greater probability of data relevant to organizations not being available within their own sphere of influence, so that data must also be exchanged with third parties. According to some experts, the ability to exchange data within value creation networks will even be one of the key factors for the growth and success of organizations (see Niederee 2019).

New business models and changed service offerings. With increasing digitalization, both the market environment and customer needs are changing dramatically. The availability of high-quality data can change an organization's future orientation in two ways: In the first step, data forms the basis for the further development of existing product or service portfolios supplemented by digital features. In the second step, active data collection and its use can also lead to the development and establishment of purely digital business models (see Nentwig et al. 2019). The availability of data and the increase in data sources – especially developments in the field of the IoT – are creating the opportunity for both digital start-ups and traditional companies from traditional sectors to develop from pure producers or service providers into data-driven software or service companies (see Ringel et al. 2020).

More efficient performance and delivery processes. Using external data sources combined with opening up internal data silos across functional units of the respective organization can also make a significant contribution towards optimizing business processes and supply chains (see Meisel and Spiekermann 2019, INFORM 2017). This specifically applies to cases where the provision of services or the supply of products is embedded in complex network structures where the exchange of data beyond company boundaries can be the key to success, while a corresponding lack of exchange can lead to serious information deficits that can severely disrupt service delivery and thus have a massive negative impact on business success (see Linnartz et al. 2020).

Improved service provision and more evidence-based decisions. Improved availability of data through data sharing can also help organizations to improve analogue products and services and to develop them in a more demand-driven and customized way for their customers. An improved product portfolio as well as productivity and profitability increases can generate considerable competitive advantages by using internal and external data resources (see Brynjolfsson 2016). The systematic use of data also opens up a range of opportunities for organizations to gain business-relevant insights and to incorporate these into strategic or operational decisions. Besides descriptions of customer groups or competitor analyses, companies can make forecasts with the help of predictive analyses and derive recommendations for future decisions through prescriptive analyses (see Wölfl et al. 2018).

Data monetization. Intangible assets, such as human capital, product-specific and process-related knowledge and the quality of customer relationships increasingly determine the competitiveness of companies, and even data and its use are now also classified as intangible assets. A greater understanding of the added value of data also increases its genuine value. This makes it possible to monetize data assets (see Krotova et al. 2019) on data marketplaces or in direct exchange with interested organizations.

THE INCREASE OF DATA SOURCES ENABLES THE DEVELOPMENT OF DIGITAL START-UPS AS WELL AS TRADITIONAL COMPANIES FROM SOLELY BEING A PRODUCER OR SERVICE-PROVIDER TOWARDS A DATA-DRIVEN SOFTWARE OR SERVICE COMPANY.

2.2 Obstacles

However, the potential benefits of increased data exchange are also offset by factors inhibiting the exchange of data between organizations, especially on the part of data providers.

Ignorance of the added value of data sharing. The recognition of data as a valuable resource in economic competition is still relatively new: At the beginning of digital transformation processes, data was generally perceived more as a by-product of the transition from analogue to digital information exchange (see Niederee 2019). This view has not yet been accepted by all organizations and/or practical implementation is insufficient. Only a minute share of organizations are making full use of their own data resources or are even part of a loose digital network or an established ecosystem where the exchange of data is pursued proactively and as a natural element (see Otto et al. 2019). However, data sharing is not an end in itself, even for the identified pioneers, but only the realization of one or more of the above-mentioned benefits of sharing triggers a corresponding commitment.

Lack of expertise and inadequate infrastructure. However, the prerequisites for realizing the advantages described above are often lacking. Although some developments can already be seen, there is still a certain degree of immaturity on the part of organizations when it comes to using and, in particular, providing data for third parties across the board. This applies to all areas of application, even highly digitalized sectors, such as the manufacturing industry (see Cattaneo et al. 2020). The reasons for this include a lack of human resources, but primarily insufficient technical expertise and inadequate technical conditions (see Nentwig et al. 2019 and Otto et al. 2019). Moreover, uncertainty exists with regard to the legally sound implementation of data exchange, especially with regard to data protection and liability issues when providing data.

Doubts regarding the quality of the available data. In order to realize the added value potential of the systematic processing and use of external data sources in organizations, the data provided must also fulfil basic requirements. The actual benefit depends on the validity of the data, its timeliness and completeness as well as its comprehensibility and accessibility (see Clarke 2016). The special challenge is hence to ensure the availability of data that is accurate, useful, understandable and easily accessible (see Wölfl et al. 2018). Especially in highly competitive markets, companies have doubts about the actual quality of data provided by third parties. However, trust is a basic prerequisite for the success of data sharing platforms and data marketplaces.

Fear of overreaching. Despite the increasingly strong desire to cooperate with competitors in certain industries (see Marx 2020), the decision to open up in-house data silos to external parties is fraught with fear of being overcharged by direct competitors or through additional competition from established technology companies. This applies, in particular, to highly sensitive business data, such as customer master data, purchase and sales prices as well as detailed product or process data (see INFORM 2017). The fear of overreaching, which exists with a view to the sharing of business-relevant data, also leads to a fundamental need on the part of potential data providers for maximum data security and protection against unauthorized data access on the respective exchange platform (see Niederee 2019 and Guggenberger et al. 2020). Establishing a trusted platform architecture and, above all, a culture of trust is hence one of the biggest challenges for exchange platform operators.

2.3 Basic functions of data sharing platforms

Data sharing platforms must leverage the added values outlined above and remove the obstacles. This raises the question regarding the basic functionalities of a data sharing platform. Building on initial systematization attempts (see Fricker et al. 2017 and Stahl et al. 2016), it is particularly worthwhile to take a look at the reference model differentiated by Meisel and Spiekermann (2019) which presents individual functions within several function groups that can be adopted by data sharing platforms in different combinations. This model is the basis for this study, albeit with some minor modifications on the basis of the practical examples and the interviews (see Table 1).

FUNCTION GROUPS	FUNCTIONS			
Transaction infrastructure	Data discovery	Rulemaking and access management	Transaction execution	Storage
Interfaces and security	Interfaces	Data security	Profile security	
Data integration	Data import	Data transformation	Metadata management	
Data services	Data analysis	Data enrichment	Data cleansing	Data-based consulting
Platform administration	Data history management	User administration	Conflict management	Network maintenance
Complementary infrastructure services	Computing power	Storage space	Software for internal data processing	

Table 1 The functions of data sharing platforms based on Meisel and Spiekermann's reference model (2019)

Based on these fundamental functions of data sharing platforms and the associated function groups, the specific offer of individual features can be fleshed out in detail for each platform in different possible combinations. The implementation of a transaction infrastructure for the exchange of data forms the basis for all other function groups. The following sections explain the six function groups of the reference model and their components. A detailed presentation of specific features within the individual functions can be found in Meisel and Spiekermann (2019) in the Annex.

Transaction infrastructure. The first function group comprises the features of the transaction infrastructure as the fundamental technical and organizational preconditions for exchanging data between several parties. They include all elements that make it possible to find individual data sets in the sense of data discovery, including features such as search functions, bundles of data sets whose contents are curated by the platform operator or the implementation of subscribable data streams. Other elements to be mentioned in this context are the central storage of data by the platform operator or the provision of a decentralized storage infrastructure using, for instance, blockchain solutions. All functions for setting rules for data exchange on the platform and for managing data access are also part of this complex. Lastly, the transaction execution features should be mentioned here. They include the provision of a communication and exchange infrastructure and, when necessary, the implementation of tools for data providers to monetize the data and to guarantee the payment flows that arise in the process. This is the fundamental function group within the reference model since the functions and features in this group enable the exchange of data in the sense of a bridge between data providers and data users (see BMWi, 2020).

Interfaces and security. The second function group serves the purely technical implementation of features to ensure that data exchange is smooth and secure at the same time. This includes all interface functionalities for the provision and use of data as well as potential data-related services of platform operators or third-party providers, as well as interfaces for external notification and communication channels. In terms of security, all measures can be listed that both guarantee the security and authenticity of the available data and protect the profiles of platform participants. With regard to data security, the possibility of restoring corrupted or lost data sets is of central importance in addition to the technical protection of data against unauthorized or improper access by third parties. Further components are technical tools for encrypting or anonymizing data to increase transaction security. In order to ensure the integrity of platform providers and users, profile security can also be increased through authentication and certification mechanisms (see Meisel and Spiekermann 2019).

§ PRACTICAL NOTES

Data protection and data security. The legal system provides for a variety of data protection and data security obligations in different laws. When processing personal data, the controller and processor must ensure the security of processing (article 32 GDPR). To this end, suitable technical, organizational and other measures must be taken in order to ensure a level of protection that is appropriate to the risk. In this regard, the GDPR provides specific assessment criteria, such as the nature, scope, circumstances and purpose of processing, as well as the likelihood and severity of the risk to the rights and freedoms of natural persons. Security measures include pseudonymization and encryption of personal data (section 64 (2) of the Federal Data Protection Act (BDSG, Bundesdatenschutzgesetz)). State-of-the-art technical and organizational means must be selected. The term 'state of the art' is not defined by law. Furthermore, it is also a dynamic variable that can change during the course of technological development. Therefore, platforms must regularly review and, when necessary, improve their protective measures within the framework of a risk analysis.

In addition, the Telemedia and Telecommunications Act contains further provisions relevant to IT security. Operators of data sharing platforms, as providers of 'telemedia', must take technical and organizational precautions in order to ensure that users can use telemedia in a way that is protected from disclosure to third parties (section 13 (4) of the Telemedia Act (TMG, Telemediengesetz)). In this case too, state-of-the-art precautions must be selected. A risk analysis and appropriate protective measures are necessary here as well.

Furthermore, the Act on the Federal Office for Information Security (BSIG, Gesetz über das Bundesamt für Sicherheit in der Informationstechnik) stipulates IT security and reporting obligations for providers of digital services. In particular, online marketplaces, online search engines and cloud computing services are considered providers of digital services (section 2 (11) BSIG). Data sharing platforms that fall under these categories must hence take appropriate and proportionate technical and organizational measures to manage risks to the security of network and information systems. Providers of digital services must additionally report significant security incidents to the Federal Office for Information Security (section 8c (3) BSIG).

THE TRANSACTIONAL INFRASTRUCTURE FOR THE DATA EXCHANGE FORMS THE BASIS FOR THE FIVE ADDITIONAL FUNCTIONAL GROUPS OF DATA SHARING PLATFORMS.

Data integration. The third function group comprises all possible features and measures of a platform that enable the general usability of the available data through meaningful integration and additionally increase the added value of the data by further processing. This includes the definition of uniform metadata and control of their compliance, as well as tools for transforming available data sets according to the preferences of the respective data user. In essence, however, it is primarily about the functions to ensure the smooth import and export of data via the platform in order to guarantee its integrity and user-friendliness, such as first and foremost the implementation of different input and output formats as well as basic plausibility checks for data sets provided.

Data services. Building on this, a fourth function group can be identified: The platform operator itself or, if the platform is open enough, also third-party providers can offer services to data users and providers. Besides tools for independent data analysis in the sense of further quality control, for cleaning up data errors or for comparing with other data providers, this also involves the provision of independently expandable data models or tools for visualising results – if possible in real time. Further aspects worth mentioning here are services or tools for enriching and linking different data sets according to specific needs of data users. In addition, data-based consulting services also belong to this function group, which, given the required consent, can be offered either by platform operators directly or by third-party providers, especially for data providers.

Platform administration. Although often seen as self-evident, the elements of the platform administration function group can contribute significantly towards sustainable use of the respective platform. In detail, this includes, for instance, management of user accounts, including the addition of individual usage and notification preferences, as well as adaptation of the scope of the platform functions used (see Meisel and Spiekermann 2019). Furthermore, the data history can be mapped with regard to origin, usage behaviour or, if necessary, changes. The provision of a support service for platform users and the establishment of a conflict management system are further elements in this context. Active network maintenance to maintain and/or increase the number of parties providing and using data is also part of this function group. This includes forums, newsletters, interfaces with social media or the organization of events, such as user meetings or hackathons.

Complementary technical infrastructure services. Finally, complementary technical infrastructure services can be listed in the sixth function group. This group includes the provision of basic technical requirements for data processing, such as computing power and storage space, as well as the provision of software for internal data processing on the part of data providers or users by platform operators.

03

3 LEGAL FRAMEWORK

Access and exchange of data via platforms raise a number of legal issues. Despite the economic importance of data, the legal system does not currently provide for a property right or a comparable absolute right to data. Data is not a corporeal object (within the meaning of section 90 of the German Civil Code (BGB, Bürgerliches Gesetzbuch). The property law norms of the German Civil Code are not applicable to data. The allocation of data is hence in principle purely fact-based: Whoever has the power of disposal over data may also use it (see Froese and Straub, 2020). Although this does not give 'data owners' the position of 'beneficial owners', they do acquire a position similar to beneficial ownership, according to which he can decide whether, how and to what extent data is collected, used or disclosed (see Fezer 2018). In order to regulate data access and use outside of this de-facto power of disposal, contractual agreements are needed. The principle of private autonomy leaves the contracting parties considerable discretionary freedom. However, limits exist when the granting of data access and use rights violates mandatory law. Such restrictions may arise in particular from copyright law, data protection law or rules for the protection of business secrets.

3.1 Copyright protection of database works and databases

Copyright serves to protect creative works and grants the holder an exclusive right of exploitation to their work for a limited period of time. The originator may grant third parties exploitation rights or, in the absence of consent, prohibit the unauthorized act of exploitation. However, copyright protection requires a personal intellectual creation (section 2 (2) of the Act on Copyright and Related Rights (UrhG, Gesetz über Urheberrecht und verwandte Schutzrechte). A precondition for this is a human creative process. This means that machine-generated data cannot claim copyright protection. The same applies to individual data since the necessary minimum degree of individuality (the so-called level of creation) is lacking here. However, the protection of data in the form of a database work is possible and also expressly provided for by law. A database work is a collection whose elements are arranged systematically or methodically and the individual elements are individually accessible by electronic or other means (section 4(2) UrhG). However, a work can only be protected if there is a personal intellectual creation due to the selection or arrangement of the elements (section 4 (1) UrhG). This means that the essential structural features of the database must be determined by a human being. The database must also have a certain originality.

— IN ORDER TO REGULATE DATA ACCESS AND USE, CONTRACTUAL AGREEMENTS ARE NEEDED. —

If the characteristic of personal intellectual creation is lacking, a database may also be protected by the right of the database producer (sections 87a et seqq. UrhG.). This case also requires a systematic or methodical arrangement of data which must be accessible by electronic or other means. Another crucial requirement is that the procurement, verification or presentation must be associated with a significant investment in terms of type or scope. This is the core of ancillary copyright: It is not the individual pieces of information included in the database that are protected, "but the database as a totality of the contents collected, arranged and made individually accessible with a substantial investment of time and effort as an immaterial good" (Cologne Higher Regional Court (Oberlandesgericht (OLG) Köln, judgement of 15 December 2006, case 6 U 229/05). The producer of a database, i.e., the person who has

THE BASIC PRINCIPLE OF PRIVATE AUTONOMY LEAVES CONTRACTING PARTIES MUCH SCOPE OF ACTION.

HOWEVER, LIMITS EXIST WHEN THE GRANTING OF DATA ACCESS AND USE RIGHTS VIOLATES MANDATORY LAW.

made the essential investment, benefits from the ancillary copyright (section 87a (2) UrhG). If the preconditions for protection are met, the database producer receives the exclusive right to exploit the database. This means that they can grant to third parties the right to reproduce, distribute or communicate to the public or - in the case of an unauthorized act of exploitation - prohibit it. The exploitation right relates to the database in its entirety or to a substantial part of the database in terms of its nature or size.

3.2 Data protection law

With the General Data Protection Regulation (GDPR) in force in the European Union since May 2018, data protection has become considerably more important. The provisions of the Regulation apply to the (automatic) processing of personal data. Personal data means any information relating to an identified or identifiable natural person (article 4 (1) GDPR). A person is deemed to be identified if the identity results directly from the information itself, such as name, address and date of birth (see Kühling and Buchner-2018, Klar, article 4 (1) GDPR, para. 18). However, it is also sufficient if the person is identifiable. This is the case if the information in itself cannot already be linked to a person, but if the person can be identified when additional information is added. A person is therefore identifiable if he or she can be identified directly or indirectly, for example, by means of association with an online identifier or location data (article 4 (1) GDPR). Anonymous or anonymized data, on the other hand, does not fall within the scope of the GDPR (recital 26 GDPR).

§ PRACTICAL NOTES

Personal reference of data. A clear demarcation between personal and anonymous data is not always possible. In order to determine whether a person is identifiable, account must be taken of all the means which are, in all probability, likely to be used to identify the person directly or indirectly (see recital 26 of the GDPR). This means that the assessment of personal relevance depends very much on the capabilities and resources of the processor. These characteristics can vary from case to case, so that the question of the existence of a personal reference cannot be answered schematically (see BMWi 2020b).

The term ‘data processing’ is to be understood broadly and includes almost every form of data handling, from data collection, storage and adaptation right through to data transmission, provision and erasure (article 4 (2) GDPR). The consequence is that data protection law must be observed at all stages of the value chain (see Bussche, 2020).

Both data marketplaces and data sharing platforms are hence subject to the requirements of data protection law for all processes with relevance under data protection law. This has implications for the contractual arrangements between platform operators and users, but also for the granting of possible exclusive or access rights (see Specht and Kerber, 2017). The principles standardized in article 5 GDPR must be additionally taken into account. These include the principles of

- lawfulness, processing in good faith, transparency
- purpose limitation
- data minimization
- accuracy
- storage limitation
- integrity and confidentiality

The principle of purpose limitation plays a central role in the context of data-based value creation systems since this can significantly limit potential (further) processing options. According to this principle, personal data must be collected for specified, explicit and legitimate purposes (article 5(1)(b) GDPR). Processing operations that deviate from the original purpose are only permitted if the further processing is compatible with the original purpose (so-called purpose compatibility). This requires substantive proximity between the original purpose and the purpose of the further processing (see Kühling and Buchner 2018, Herbst, article 5 GDPR, para. 24). Determining the purpose already at the time of data collection means that future data uses must be anticipated when developing data-based business models. This restricts the corridor of permissible data use. It is hence not permitted to collect personal data in advance without concrete objectives. Furthermore, the principle of transparency of data processing requires that the data subject be informed about the intended purposes of processing. Only when he knows what happens to 'his' data can he make a decision, for instance, whether to give his consent or to use a service.

The principle of lawfulness also requires that the data processing be based on a sufficient legal basis. Article 6 GDPR lists further legal grounds for permission in addition to consent. These include, for example, data processing for the performance of a contract or processing to protect the legitimate interests of the controller or a third party. Only in exceptional cases should platforms that depend heavily on the processing of personal data rely on consent as a basis for legitimacy. On the one hand, there are high requirements for consent to be effective. Consent must, in particular, be given in an informed manner, i.e., the data subject must be informed in advance about all the circumstances of the data processing. Moreover, the lawfulness of data processing depends on the existence of consent. Consent can be withdrawn at any time with effect for the future. This usually leads to inadmissibility of further data processing. If there is no other legal basis for further use, data affected by withdrawal must be erased (article 17 (1) (b) GDPR).

In addition, responsibility under data protection law for data sharing platforms or data marketplaces is of central importance. The controller is the person who determines the purposes and means of the data processing (article 4 (7) GDPR). The controller must ensure data protection. This includes, but is not limited to, guaranteed compliance with data protection law (see above), but also the implementation of technical and organizational measures for the protection of personal data (article 25 GDPR). The controller must also be able to prove compliance with data protection obligations (so-called accountability) and is the addressee for fines and claims for damages in the event of data protection violations. In platform-based value networks, it is possible that the means and purpose of data processing are determined not by a single, but by several (joint) controllers (article 26 GDPR). In this case, the joint controllers are obliged to agree on who is to implement which GDPR-specific obligation. This agreement must be recorded. So-called processing (on behalf of a controller) (article 28 GDPR) must be distinguished from joint responsibility. A processor is anyone who acts on behalf of the controller. Unlike the controller, the processor does not determine the means and purposes of the data processing, but acts so to speak as an 'extended arm' on the instructions of the controller. Agreement is also required with regard to processing, whereby the processor is obliged to process the data in accordance with data protection law (so-called processing contract).

3.3 Protection of business secrets

The exchange of data via platforms, especially in the B2B sector, can also affect sensitive company information, such as machine data that allows conclusions to be drawn about internal production processes. If this information in the form of data falls within the scope of application of the Business Secrets Protection Act (GeschGehG, Gesetz zum Schutz von Geschäftsgeheimnissen), use of data may be restricted. A business secret is information that is secret (and hence of economic value) and protected by appropriate confidentiality measures and for which there is a legitimate interest in keeping it secret (section 2 (1) GeschGehG). The owner of a business secret is the person who has lawful control over a business secret (section 2 (2) GeschGehG). If business secrets were obtained, used or disclosed without authorization, the damaged party is entitled to defence claims against the damaging party (sections 6 et seqq. GeschGehG). It should be noted, however, that the provisions of the Business Secrets Protection Act do not establish an exclusive right with regard to data. This means that the owner of a business secret does not have a positive right of disposal, but can only assert defence claims in the event of infringement. This initial situation means that organizations that provide data must ensure that the data provided by them is protected by appropriate confidentiality measures on the part of the data recipient. These can be technical, organizational or legal measures. Organizations can obtain contractual assurance of compliance with these obligations, for example, in the form of non-disclosure agreements. If business secrets include personal data, it must be ensured that the confidentiality obligations are in line with the technical and organizational data protection obligations.

3.4 Contract law

As should have become clear, sovereignty over data based on civil law can only be derived in part from the current legal framework. However, copyright protection of databases, data protection law or the protection of business secrets may influence data use and disclosure. However, given these limitations, stakeholders in platform-based value networks are free to regulate data transfer and use in contracts. The principle of freedom of contract opens up the possibility of making data the subject of a contract under the law of obligations. Rights and obligations regarding data transfer and use can be negotiated on a case-by-case basis according to the parties' respective interests. The contractual allocation of data offers contracting parties a high degree of flexibility. Nevertheless, further restrictions still exist in addition to the special legal rights of protection and defence discussed earlier. For instance, general legal provisions exist to protect the contracting party that is typically inferior and/or weaker in certain constellations. One example worth mentioning in this context is section 138 of the German Civil Code pursuant to which an immoral legal transaction is null and void, as well as the restrictions of the law on standard business terms. Standard business terms are "[...] all contract terms pre-formulated for more than two contracts which one party to the contract (the user) presents to the other party upon the entering into of the contract" (sec-

tion 305 BGB). If standard business terms are effectively made part of a contract, the provisions contained therein are subject to the test of reasonableness of contents (section 307 BGB). As a consequence, provisions are ineffective if they unreasonably disadvantage the other party contrary to the requirements of good faith. However, the so-called indeterminate legal terms, such as 'unreasonable disadvantage' or 'good customs' contained in the above-mentioned norms only become fully effective when they are fleshed out by the courts.

3.5 Contract design

If data is made the subject of a contract, it must first be distinguished at what level and between which parties the transfer or use of data is to be regulated. Data marketplaces and data sharing platforms are typically designed to pass on existing data to third parties. These are also referred to as secondary data markets (see Specht and Kerber 2017). If the party receiving the data is to be granted the right to use the data transferred, corresponding data transfer or data licensing contracts can be concluded. The circumstances of the data transfer must then be determined between the parties on a case-by-case basis. From a legal point of view, the type of performance or consideration owed is important because this determines the type of applicable law: If data is to be transferred permanently and payment of a purchase price has been agreed upon, the provisions of the law on the sale of goods apply first and foremost. If, on the other hand, data is to be transferred temporarily, the provisions of lease law are more appropriate. A clear assignment to any of the contract types provided for by law is not always possible nor necessary because the law of obligations, due to the principle of freedom of contract, does not force the contracting parties to choose a specific type of contract determined by law (so-called numerus clausus under the laws of property). However, the assignment of the contract to one of the legally standardized models is nevertheless highly relevant since the different types of contract under the German Civil Code provide for different legal consequences with regard to defective performance.

In addition to the contractual performance obligations, the type and scope of intended data use can also be regulated. Similar to the licensing of intangible property, two variants can be considered. First, the recipient of data can obtain an exclusive right of use, i.e., he can use the contractual data for himself and additionally grant third parties a sublicense to use such data. Second, the data recipient can be granted a simple right of use which only entitles him to use the data for his own purposes. The duty under the law of obligations to provide data can also be supplemented by additional duties which can also be enforced in the event of a dispute. In this context, it should be clearly defined in which format, via which interface and at which intervals data should be made available. When exchanging sensitive information, a supplementary non-disclosure agreement may be necessary. The same applies to the processing of personal data for which a joint responsibility agreement or a processing contract can be considered.

04

4 ORGANIZATIONAL STRUCTURES AND BUSINESS MODELS IN CURRENT PRACTICE

Building on the legal framework of data sharing described in chapter 3, the next step is now to shed light on how organizational structures and business models are developing on today's data sharing platforms. The basis for this is the 24 practical examples that are presented in detail in the next chapter. With minor changes, the presentation follows the taxonomy proposed by Markus Spiekermann (2019), which in turn builds on the work by Engelhardt et al. (2017) (see Table 2).

CATEGORIES	CHARACTERISTICS		
Value proposition	Transaction centricity		Data centricity
Transaction mechanism	Data marketplace		Sharing platform
Role	Neutral		Active
Platform access	Openness	Semi-open	Closed
Data integration	Domain-specific		Domain-unspecific
Data transformation	Passive provision		Active editing
Platform architecture	Centralized	Hybrid	Decentralized
Business model	Focused		Diversified
Financing structure	One-dimensional		Multi-dimensional

Table 2 Taxonomy for classifying data sharing platforms, based on Spiekermann (2019)

Value proposition and added value – transaction or data centricity? The first differentiation criterion is based on the dichotomy between two different value propositions of a data sharing platform:

Transaction-centric platforms place the role of neutral intermediary at the centre of their value proposition. The platform provides the technical basis for exchanging or selling data and thereby brings together supply by data providers and demand of data users, enabling trusted processing of individual transactions through technical and contractual security and control mechanisms. This also includes a well-functioning and user-friendly search engine in order to efficiently identify relevant data sets and the associated data providers as possible cooperation partners, as well as the provision of model contracts and/or contract principles for the secure and simple execution of transactions. Furthermore, transaction-centric platforms establish arbitration procedures which increase the integrity of the platform as a neutral anchor of trust through the moderated resolution of disputes between platform participants.

Data sharing platforms with a general data-centric orientation mark the other end of the range. Unlike transaction-centric platforms, they intervene more strongly in the respective totality of available data. This is reflected, among other things, in the fact that platform operators of data-centric platforms, for example, make data sets of their own organization available on the platform or actively search for freely available data sets – especially from the public sector –, process them and then make them available on their own platform. Both serve to increase the attractiveness of the offering. Furthermore, establishing technical interoperability of the individual platform

participants is the focus of interest, which is why data-centric platforms often offer consulting and even technical services to improve data quality and the infrastructure of data exchange as an additional feature. Finally, data-centric platforms often offer a set of tools for analyzing and visualizing the available data.

The analysis of practical cases shows that this theoretically ideal distinction does not hold in its pure form, but must rather be understood as the complementary focus of individual platform operators. Nevertheless, understanding this distinction is crucial, especially when it comes to addressing the platform's respective target group(s). Depending on the technical and cultural background of (potential) platform participants, a successful approach depends on the specific focus in the communication of the value proposition.

Transaction mechanism – data marketplace or sharing platform? The transaction mechanism used can also be used as a further distinguishing feature which can be assigned either to the type of data marketplace or the type of sharing platform. In the first case, pricing for the available data sets takes place directly as a function of supply and demand of market participants without the platform operator's involvement. Data can also be made available openly and free of charge here. In the case of sharing platforms, there is no direct pricing between supply and demand of participants, but the availability of data within the network of the platform is subject to membership status. Membership can be either free of charge or subject to a participation or subscription fee.

UNLIKE WITH DATA MARKETPLACES, THE ACCESS TO SHARING PLATFORMS IS NOT GRANTED BY DIRECT PRICING, BUT INDIRECTLY BY MEMBERSHIP.

Role of the platform operator – neutral or active? The third characteristic is the understanding of the role of the respective platform operator. A distinction must be made between a neutral and an actively involved position. The latter is characterized by participation of the platform operator on the demand and/or supply side of the data platform. The platform operator's role is neutral if it neither supplies nor demands data. Depending on the nature of the competition of the respective target group and its cultural openness towards the concept of data sharing, ensuring neutrality of the platform has a prominent role to play. The practical projects show that all large organizations, in particular, are using their position and the existing wealth of their data to build their own – sometimes only unilaterally open – platforms. If smaller companies acknowledge that a platform with a neutral operator can be trusted, they then tend more towards these platforms. This does not necessarily have to involve a single operator, i.e., operation of the platform can also be organized as a cooperative operating model with appropriate corporate structures. A basic distinction can be made between loose cooperation agreements between largely independent parties and the joint establishment of an independent company as a joint venture (see Niederee 2019).

§ PRACTICAL NOTES

Accountability of platforms. The question of the platform operator's liability is of central importance. Platforms play an important role in data ecosystems as a link between providers and users. In the Telemedia Act (TMG, Telemediengesetz), the legislator has laid down special liability rules for service providers. Platforms are considered to be service providers and fall within the scope of the act if they make their own or third-party telemedia available for use or provide access to use (section 2 No. 1 TMG). A distinction is made between different types of providers (see Bartels et al, 2020), whereby the provisions regarding the responsibility of content and host providers are particularly important for data sharing platforms. While the content provider provides its own content on its platform and is fully responsible for this, the host provider is generally not liable for the provision of third-party content. An exception applies only if the host provider had positive knowledge of illegal content and did not remove it despite being aware of it. Furthermore, the host provider is not obliged to monitor the information transmitted or stored by it or to investigate circumstances that indicate illegal activity. The purpose of liability relief in the Telemedia Act is to give preference to the operation of pure intermediaries. Platforms are only liable for the information they publish from third parties if they have adopted this information as their own.

Access to the platform: open, semi-open or closed? The criterion of platform access is measured by the degree of openness. Closed platforms mark one side of the range where the platform operator decides on the inclusion of data providers or data users. While the selection of partners limits the growth of the platform's network, it facilitates control over the quality of data and its use. At the opposite end, we find open platforms without any kind of access restrictions on the supply or demand side which basically facilitate scaling of the platform. The group of practical examples includes an increasing number of hybrid forms that enable the exchange of high-quality data among selected partners while at the same time allowing the marketplace to grow by adding new players subject to fulfilment of certain access requirements. With regard to the data access governance structures, different instruments can be used. Besides open provision with no access restrictions, blacklisting or whitelisting methods and, in individual cases, even individual releases can be found here.

Data integration – domain-specific or domain-unspecific? Another characteristic is the way in which the integration of concrete data content is handled. Here, the range spans between domain-specific and domain-unspecific platform types. The latter are characterized by openness with regard to the contents of the data to be made available. They do not limit the contents of the data stock and thus cover several economic or application areas. In this way, the platform becomes fundamentally interesting for a large number of data providers and especially for users due to the diversity and the larger number of relevant data sets. At the same time, the search for suitable data sets and the long-term retention of platform participants becomes more difficult. Domain-specific platforms, on the other hand, set clear boundaries for the sources and contents of the available data sets and curate them according to their fit with the chosen content orientation. A specification of the offering is possible here that can even lead to the determination of a narrowly defined niche. The focus on content generally makes it easier to address specific target groups, but also limits the attractiveness of new players who do not or only marginally identify themselves with the specification made.

§ PRACTICAL NOTES

Competition law. Platform-based business models not only facilitate networking between organizations, they also enable new forms of collaboration. This can lead to cooperation between actual or potential competitors, where admissibility of such cooperation is governed by competition law. Article 101 TFEU prohibits all agreements between undertakings, decisions by associations of undertakings and concerted practices which may affect trade [...] and which have as their [...] effect the prevention, restriction or distortion of competition. However, agreements between undertakings are only objectionable if they are restrictive of competition. Furthermore, the planned 10th amendment to the Act against Restraints of Competition (GWB, Gesetz gegen Wettbewerbsbeschränkungen) aims at stricter abuse control in order to enable more effective action against anti-competitive behaviour by large digital platforms.

Transformation of data stock – passive provision or active editing? The category of data stock transformation refers to the degree of intervention by platform operators in the data sets provided. On the one hand, platforms exist that passively collect raw data and make it available to interested users in data streams that have not been further examined or aggregated. On the other hand, there are platforms that actively curate the available data stock. The following four levels of active further editing can be distinguished in this context, which move further and further away from purely passive provision and build on each other:

1. In the first step of normalization, the data is compared with standardized data models and metadata catalogues and is standardized when deviations are identified.

2. In the aggregation step, the platform operator compiles data packages or data streams tailored to the needs of the respective target groups, which facilitate further editing and analysis through their combination.

3. The next step of active editing is to constantly ensure data quality while taking responsibility which is carried out by means of in-house consistency checks by operators.

4. The use of data by the platform operator to provide data-based consulting services constitutes another level of transformation.

§ PRACTICAL NOTES

Data quality and liability. Data can be made the subject of a contractual agreement. If this data then proves to be unsuitable for the purpose intended in the agreement, for example, because it is out of date or does not have the necessary quality, this may give rise to warranty claims on the part of the data recipient. The specific warranty rights that exist in individual cases depend on the specific performance owed. This determines which type of contract and which legal provisions apply. If, for example, sales law applies, the data recipient can demand rectification in the form of error correction, analogous to standard practice in the field of software development. If rectification fails, claims for reduction and rescission may then exist. If a loss occurs due to incorrect data, claims for damages may also be considered. In this case too, the existence and scope of claims depend on the relevant regulations, which in turn result from the underlying type of contract.

Platform architecture – centralized, decentralized or hybrid? With regard to the design of the technical platform architecture, a distinction can be made between centralized and decentralized approaches. Centrally built platform architectures collect the data sets provided in one place based on a cloud infrastructure. This facilitates access management for platform operators and enables direct further editing of the data for users. Furthermore, central storage is characterized by better accessibility because neither data providers nor data users have to create their own extensive technical and organizational requirements for data transfer. At the same time, central storage can also pose a serious obstacle to participation due to a lack of trust between platform operators and data providers. With decentralized platform architectures, the data remains with data providers and only the necessary information about its contents and nature is made available on the platform.

This approach increases the data sovereignty of data providers, but complicates the real data transfer and its further processing by users. It also requires greater expertise and better technical equipment on the part of platform participants. This is why hybrid forms of platform architecture often emerge in practice, enabling decentralized data exchange while providing complementary assistance and additional infrastructure services to facilitate the actual transaction.

Design of the business model – focused or diversified? The design of the business model is closely linked to the characteristics described earlier, and in particular to the value proposition, the understanding of the role and the transformation service provided by the respective platform operator. The findings from the practical examples suggest that the following elements are relevant in this context. Besides enabling the technical and organizational exchange of data, active quality assurance of data stocks and their enrichment as well as the provision of analysis and visualization tools are important elements of the business model. Further elements are data-based consulting and infrastructure services in the form of storage space or computing power. The proactive collection of further, freely available data stocks and their provision by platform operators should also be mentioned. Two business models can be distinguished in practice: One model focuses on one or, at most, two function groups, whilst the other group consists of models which are highly diversified and encompass a large number of the function groups described.

Design of the financing structure – one-dimensional or multi-dimensional? There is generally a close connection between the design of the business model and the value proposition of the respective platform with its specific revenue model. The optimal design of the revenue model must always consider the attractiveness of the overall platform system and the relative pricing of the individual user groups (see Engelhardt et al 2017). Asymmetric revenue models, up to and including free use of the platform or even the granting of monetary or non-monetary benefits to a user group, are therefore the rule rather than the exception in multi-sided data markets. Three funding streams are generally available to platforms for generating revenue: In addition to sponsorship by a partner network or a single network player, such as an association for operating the platform, financing by third parties in the form of public funding or financial support by private foundations should also be mentioned. However, the most widespread practical design of the revenue model involves generating revenue by pricing data use. Common pricing models are:

- Subscription or fixed-price model for a certain period of time and a fixed set of data records
- Package price model for predefined data packages and a defined period of time
- Pay-per-use model
- Memberships (free use or barter)
- Transaction fees
- Progressive pricing model depending on the intensity of use of the entire marketplace

These different price models can also appear as hybrid models in a wide variety of combinations. Here, on the one hand, freemium is a combination of free access to attract paying users, whilst the two-part tariff combines a basic fee and either an indirect (package price) or direct (pay-per-use) use-dependent share (see Stahl et al. 2015). It goes without saying that other combinations are also possible, including customized price models for individual user groups.

05

5 SHORT PORTRAITS OF DATA MARKETPLACES AND DATA SHARING PLATFORMS

For this study, 24 practical examples of data marketplaces and data sharing platforms were chosen which are particularly relevant because

- they are actually already in operation, also within the framework of a pilot project,
- they cover the industry, logistics, healthcare or trade sectors and/or are domain-neutral and
- they are open to market participants from Germany.

These platforms are presented below in brief portraits. The taxonomy follows that of chapter 4 above. A general overview of all platforms can be found in the Annex. All information is up to date as of December 2020.

5.1.1 BMW CARDATA

With CarData, the BMW Group is building a platform for telematics data for its models of the BMW and MINI brands. The use of customized or also aggregated vehicle data can be requested by registered third parties, such as workshops, insurance companies or providers of new mobility services. However, data is only exchanged if the vehicle owner expressly agrees to the use of data or actively wishes to use the provider's service on condition that data is made available. This makes it a two-sided market, whilst on the data provider side it is clearly limited to vehicle owners of BMW Group models. On the other hand, the marketplace is open to third parties who are required to register. As the platform operator and central data owner, the BMW Group monetizes the availability of the data according to a clearly structured pay-per-use model for €0.29 per data retrieval and vehicle, capped at a maximum monthly amount of €5.00 per vehicle.

WEBSITE
https://www.bmwgroup.com/de/innovation/technologie-und-mobilitaet/cardata.html

Platform operator	Bayerische Motoren Werke Aktiengesellschaft
Company location	Germany
Reach	International
Data type	Telematics data from vehicles
Value proposition	Data-centric
Transaction mechanism	Sharing platform
Role	Active (aggregation and quality)
Access	Semi-open (supply side is closed and limited to the BMW Group; demand side is open to registered third parties)
Data integration	Domain-specific (connected car data)
Data transformation	Active
Platform architecture	Centralized
Business model	Focused (sale of authorized data access for third-party providers)
Financing structure	One-dimensional (Revenues from data users via pay-per-use with cap per vehicle and month)

WEBSITE
https://www.caruso-dataplace.com

5.1.2 CARUSO DATAPLACE

The CARUSO data marketplace was launched at the beginning of 2017 as a result of an initiative by a number of independent suppliers to the independent automotive aftermarket and has since pursued the goal of supplying high-quality data from the automotive and mobility sector. It is based on openness, non-discrimination of all market participants and fairness. Neutrality of the platform is to be guaranteed by a diversified structure of more than 20 shareholders from the automotive sector. The platform provides internal vehicle data, which enables data-based services in the mobility sector, for instance, in vehicle service business, car sharing via apps or insurance companies that can offer customized policies depending on driving profiles.

A development portal is additionally offered to facilitate technical integration of the data by corporate clients, even with limited technical or human resources. The platform's business model is monetized via a four-tier membership model: In addition to a three-month trial subscription, two further levels are available which allow a higher number of data accesses for individual vehicles for €1,500 or €15,000 per year, depending on requirements. The 'On Demand' category additionally offers customized options, including exclusive access to manufacturer data or additional consulting services for technical integration or for designing data-based services.

Platform operator	Caruso GmbH
Company location	Germany
Reach	International
Data type	Vehicle data
Value proposition	Transaction-centric
Transaction mechanism	Marketplace (data providers can set prices for data packages themselves; however, one of several account variants must also be available)
Role	Neutral
Access	Open (but release options for drivers for individual data services)
Data integration	Domain-specific (connected car data)
Data transformation	Active (aggregation and quality)
Platform architecture	Centralized
Business model	Diversified (data platform, provision of a development portal, but not profit-oriented since proceeds are reinvested exclusively in the platform)
Financing structure	Two-dimensional (financial contribution to the development of ten active partner organizations; in addition, generation of revenue from data users as a hybrid form of membership and pay-per-use)

5.1.3 OTONOMO

WEBSITE
https://otonomo.io

The Israel-based company Otonomo has developed and built the cloud-based platform of the same name for exchanging connected car data and its monetization. The platform enables car manufacturers, third-party providers and drivers to exchange data within an ecosystem. Otonomo acts as a neutral player and provides third-party providers with easy access to high-quality data and on the other hand gives car manufacturers and drivers control over availability through transparent rights management. In total, data sets of more than 18 million passenger and commercial vehicles from the US, Canada, Asia and Europe are available. Otonomo is thus taking a pioneering role in the highly competitive market for vehicle and traffic data. Individual data sets are accessed either via a pay-per-use model with entry prices starting at 60 US dollars per data access or via customizable access schemes, including realtime retrievability and an extended range of data types. Applications range from the development of autonomously driving cars and the insurance industry to fleet management and service providers (see Schmid, 2020).

Platform operator	Otonomo Ltd.
Company location	Israel
Reach	International
Data type	Vehicle and traffic data
Value proposition	Transaction-centric
Transaction mechanism	Sharing platform
Role	Neutral
Access	Open (but release options for drivers for individual data services)
Data integration	Domain-specific (connected car data)
Data transformation	Active (aggregation and quality)
Platform architecture	Centralized
Business model	Diversified (data platform, services for auditing and billing)
Financing structure	One-dimensional (revenues from data users with a mix of pay-per-use or individual payment model)

5.2 LOGISTICS

WEBSITE
https://www.aviation-datahub.com

5.2.1 AVIATION DATA HUB

The AVIATION Data Hub was established by Lufthansa in 2019 as an independent company to establish an independent platform for all data generated in air transport. In order to ensure the platform's neutrality in the long term, the hub will be supported by a network of participating partners, according to its own statement on the website, and is hence open to any market player. However, no strategic partners have been announced so far. The aim is to digitize the aircraft maintenance industry and to integrate the related areas of flight operations and ground handling. To this end, data aggregation and exchange will be enabled without restrictions vis-à-vis competitors. The guiding principles are guaranteed data security, quality and integrity as well as efficient control by data providers. Airlines in particular would then be able to decide whether and to whom they want to make their data available for technical support of aircraft or for improving ground and flight operations.

Platform operator	AVIATION DataHub GmbH
Company location	Germany
Reach	International
Data type	Aircraft and logistics data
Value proposition	Transaction-centric
Transaction mechanism	Sharing platform
Role	Active (normalization, aggregation and consulting)
Access	Open
Data integration	Domain-specific (logistics and product data from aircraft operations)
Data transformation	Active
Platform architecture	Centralized
Business model	Diversified (securing transaction/consulting)
Financing structure	According to its own statement, no profit-making intention and sponsorship of the partner network yet to be set up; so far financed by Lufthansa; in addition, monetization of the consulting services of the Aviatar analysis tool

5.2.2 ISHARE.ORG

WEBSITE
https://www.ishareworks.org

Strictly speaking, ISHARE is not an integrated technical platform for data exchange, but rather a coherent model of functional, technical and legal agreements and standards that are used within the Dutch transport and logistics sector to exchange data. The core of the model consists of agreements and standards focusing on identification, authentication and authorization, which are intended to standardize and thus to significantly facilitate the exchange of data between organizations. Together with a number of certified partner companies, the aim is to promote the organizational and technical integration into this network of data-sharing companies. Compliance of the model is ensured and its continuous development achieved by the independent iSHARE Foundation that was specifically established for this purpose. Another special feature is the model creation process which is based on cooperation between various government bodies and companies from the transport and logistics sector. The iSHARE model was also included in the reference architecture of the International Data Spaces Association (ISDA).

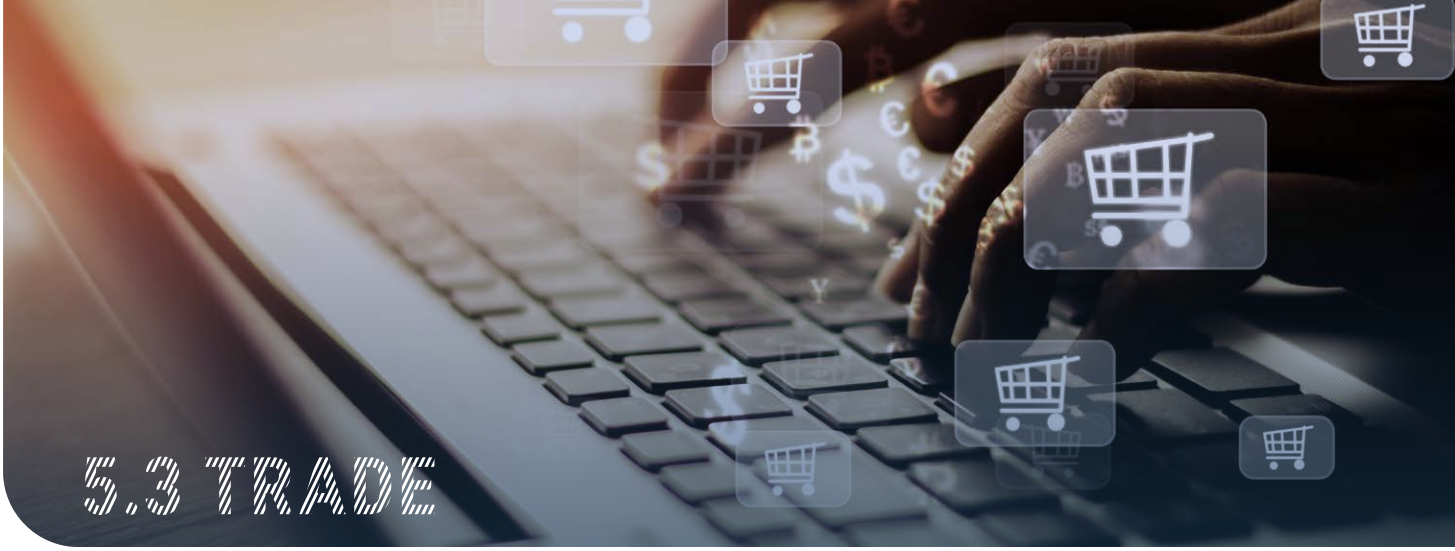
Platform operator	iSHARE Foundation
Company location	Netherlands
Reach	National/European
Data type	Logistics data
Value proposition	Transaction-centric
Transaction mechanism	Sharing platform (NLIP platform in combination with a set of sharing arrangements)
Role	Passive
Access	Semi-open (joining after verification of the organization and assignment of an iSHARE identity)
Data integration	Domain-specific (logistics data)
Data transformation	Passive
Platform architecture	Decentralized
Business model	Focused
Financing structure	Two-dimensional (funded by the iSHARE Foundation consisting of registered partners as well as by sponsors: mainly governmental institutions and business associations)

WEBSITE
https://onetransport.io

5.2.3 ONETRANSPORT DATA MARKETPLACE

The oneTRANSPORT Data Marketplace is an open, transaction-centric data marketplace for logistics and mobility data that emerged from a government-funded project in the UK from 2015 to 2017. The Marketplace encourages public and private-sector organizations and companies of all sizes to publish, share and collaborate on their mobility-related data in realtime in order to create new data-driven smart mobility solutions. The oneTRANSPORT Data Marketplace operates on a cloud-based platform in order to enable open data exchange for any organization. It is supported by a commercial framework and enables data transfer via standardized interfaces. Revenues are generated by oneTRANSPORT from monthly subscription fees which vary depending on the size of the company and the type of organization. A brokerage fee of 10% of the data purchase fee is additionally charged by data providers who make their data available for a fee. All marketplace participants also have the option of purchasing additional storage space for their data for a further fee.

Platform operator	One TRANSPORT Data Marketplace
Company location	USA/UK
Reach	International
Data type	Mobility and logistics data
Value proposition	Transaction-centric
Transaction mechanism	Data marketplace
Role	Passive
Access	Open
Data integration	Domain-specific (logistics and mobility)
Data transformation	Passive
Platform architecture	Hybrid (cloud-based, but also acceptance of other data transfer options)
Business model	Diversified (offering of infrastructure services, web-based in-house data sharing, provision of the marketplace)
Financing structure	One-dimensional (hybrid model combining a ‘subscription fee’ – varying according to company size and type of organization – for marketplace participants and a ‘brokerage fee’ of 10% of the data purchase price)



5.3.1 PEEKD

WEBSITE
https://peekd.ai

peekd provides a specially developed tool for optimizing search engine results for product offerings (SEO tool) which is used by online retail companies to analyze data sets from Google or Adobe Analytics. By using this peekd tool, retail companies become data providers for the peekd data platform. This approach makes e-commerce sales data from more than 500,000 online retailers available on the peekd platform. Using a freemium model, interested companies can thus access the sales figures of branded products sold worldwide from the electronics and clothing industry. peekd additionally offers data-based consulting and market research services.

Platform operator	peekd.ai
Company location	Germany
Reach	International
Data type	E-commerce data
Value proposition	Data-centric
Transaction mechanism	Sharing platform
Role	Active (data is released for use by the participating companies when using the in-house SEO tool and then placed on the platform)
Access	Semi-open (data providers use the in-house SEO tool, open for all data users after confirmation of the code of conduct)
Data integration	Domain-specific (retail)
Data transformation	Active (quality assurance)
Platform architecture	Centralized
Business model	Diversified (securing the transaction, supplementary technical support and consulting services regarding data handling)
Financing structure	One-dimensional (pricing of data use via a freemium model consisting of a freely available data plan and a plan that can be customized in terms of scope and price)

WEBSITE
https://www.salesforce.com/products/marketing-cloud/data-sharing

5.3.2 SALESFORCE DATA STUDIO

Salesforce Data Studio is a transaction-centric data marketplace launched in 1999 by Salesforce.com, a cloud computing solutions provider focused on customer relationship management (CRM). Salesforce Data Studio enables marketing experts to search for specific customer profiles and target groups and discover new marketing opportunities for target group identification and activation. Data providers, for their part, can use data governance tools on the platform in order to control access to their data with regard to scope, duration and purpose of data access authorization. In addition, support tools are offered to data providers to help them upgrade their own data resources.

Platform operator	Salesforce.org EMEA Limited
Company location	USA
Reach	International
Data type	E-Commerce data
Value proposition	Transaction-centric
Transaction mechanism	Data marketplace (granting the possibility of private sharing groups)
Role	Neutral
Access	Semi-open (possibility to set up closed user groups)
Data integration	Domain-specific (trade and consumer data)
Data transformation	Passive
Platform architecture	Centralized
Business model	Focused (with regard to the Data Studio as a marketplace, but further services via Salesforce as a company as a goal)
Financing structure	Not specified

5.3.3 SPOCC

WEBSITE
https://www.spocc.io

The SPOCC (Single Point of Content and Communication) platform claims to be the central data platform for the footwear and leather goods industry and is designed to enable professional data handling of marketing content. SPOCC was launched in 2019 as a collaborative project between merchandise management providers ETOS, Brandt Software-Produkte and Ariston Informatik in cooperation with the Federal Association of the German Footwear and Leather Goods Industry (HDS/L) to connect manufacturers and retailers and to make it easier for them to work with digital content. The open data platform, which is operated by the operating company SPOCC that was established in 2018, offers manufacturers the possibility to make content materials, such as photo, graphic and video files or product master data and marketing information, available to the stationary trade on a user-related basis. Rights and periods of use for the data can be managed by drag and drop. SPOCC also has interfaces with merchandise management systems in the footwear industry, so that the data made available on the platform can be integrated directly into retailers' online shops or kiosk systems. The data is stored centrally in data centres in Germany. Package prices are charged for the use of the platform, which are defined on the basis of the number of traders authorized to access the platform and the number of brands.

Platform operator	SPOOC GmbH & Co. KG
Company location	Germany
Reach	Germany
Data type	Product data and marketing content
Value proposition	Transaction-centric
Transaction mechanism	Sharing platform
Role	Active
Access	Open
Data integration	Domain-specific (product data from the footwear/textile industry and retail)
Data transformation	Active (provision, aggregation, quality assurance)
Platform architecture	Centralized
Business model	Diversified (provision, data management tools, ensuring data quality)
Financing structure	Not specified (only information on the foundation as a cooperation project between the Association of the German Footwear and Leather Goods Industry and individual trading companies as well as providers of merchandise management systems)

WEBSITE
https://vth.ifcc.de/de

5.3.4 VTH EDATA POOL

The eData Pool is a platform for exchanging data between manufacturing companies and retailers. In this context, the manufacturers are responsible for publishing and updating their product range and the associated data in terms of up-to-dateness, completeness and correctness. Retailers in turn receive high-quality product data in a standardized format. Trading companies can use standardized, defined characteristics and the corresponding enrichment of the individual data stocks as well as their final control in order to access data sets with the required attributes without having to do too much checking themselves. The data is accessed via data feeds that provide traders with the master data information of the products. Each feed requires separate approval by the respective manufacturer. According to its own account, more than 40 well-known manufacturing companies from various industries and more than 50 trading companies are already taking advantage of the services of the eData pool. Regular user meetings will further improve cooperation between data providers and data users.

Platform operator	IFCC GmbH
Company location	Germany
Reach	Germany
Data type	Product and article data
Value proposition	Transaction-centric
Transaction mechanism	Sharing platform
Role	Neutral
Access	Open
Data integration	Domain-specific (product data from trade/manufacturing)
Data transformation	Active (normalization, aggregation, quality assurance)
Platform architecture	Centralized
Business model	Diversified (securing the transaction, network expansion, consulting services and technical support)
Financing structure	Multi-dimensional (paying premium partnerships, also pricing of data usage by manufacturers and retailers, based on a hybrid pricing model of basic fee and pay-per-use)

5.4 HEALTH SECTOR

5.4.1 BURSTIQ

The starting point of burstIQ's corporate history is the provision of services and technical infrastructure for functioning intra-organizational data management. The domain-specific focus is on highly sensitive data in the health sector; however, other sectors with highly sensitive data, such as the financial sector, also use burstIQ's services. This was the basis for the further development to establish a higher-level 'network of networks' in the sense of a global inter-organizational data exchange platform that maps the respective data releases and transactions using several interconnected blockchains. This enables granular consent to data usage, from individual data points right through to huge data packages. More than 30 companies and organizations from the health sector are currently part of the platform. No information is available regarding the number of users from other areas of application.

WEBSITE
https://www.burstiq.com

Platform operator	BurstIQ Corporation
Company location	USA
Reach	International (however, focus on the US)
Data type	Health data
Value proposition	Data-centric
Transaction mechanism	Sharing platform
Role	Active
Access	Open
Data integration	Domain-specific (health data: study and patient data)
Data transformation	Active (aggregation and recruitment of new data-providing organizations)
Platform architecture	Decentralized (via blockchain)
Business model	Diversified (technical support and granting network access, network maintenance, consulting)
Financing structure	One-dimensional (basic fee for network participation, transaction fee per data retrieval, individual pricing for consulting services)

WEBSITE
https://www.covid19dataportal.org

5.4.2 COVID-19 DATA PORTAL

The COVID-19 Data Portal was launched in April 2020 in response to the Covid-19 pandemic. The aim of the portal is to bring together all relevant data sets for sharing and analysis, thereby accelerating coronavirus research. It enables researchers to upload, access and analyze Covid-19 related reference data and relevant data sets as part of the wider European COVID-19 Data Platform. The platform is run by the European Commission and the European Bioinformatics Institute of the European Molecular Biology Laboratory (EMBL-EBI) together with EU Member States and research organizations, such as the European Life Sciences Data and Information Research Infrastructure ELIXIR. It aims to enable the European and global research community to rapidly collect and extensively share available research data from a variety of sources.

Platform operator	Elixir
Company location	UK
Reach	European
Data type	Health data and study data on COVID-19
Value proposition	Data-centric
Transaction mechanism	Sharing platform
Role	Active
Access	Open
Data integration	Domain-specific (health data: genome and patient data)
Data transformation	Active (collection and aggregation)
Platform architecture	Centralized
Business model	Focused
Financing structure	One-dimensional (external basic funding from the European Commission, national health and research bodies)

5.4.3 CSDR

WEBSITE
https://www.clinicalstudydatare-quest.com

The CSDR (ClinicalStudyDataRequest) platform provides access to patient-level clinical trial data shared by 14 of the world's leading pharmaceutical companies with each other and the research community for research purposes. Seven years after the platform's establishment, data sets from more than 3,500 clinical trials are now available. The platform is operated by ideaPoint as the technical third-party provider. Access is granted upon request to registered users whose research proposals must be approved in advance by an independent reviewer panel. Furthermore, the detailed data use agreement for the CSDR platform must be signed. Most of the data is stored centrally within a secure working environment and made available for a total of 12 months after release. In the working environment, the data can be used via integrated statistical evaluation software. Direct exchange of data between data providers and users is also supported organizationally. Funding is provided by the membership fees of institutional members, foundations and public research funding institutions. There is no fee for individual use.

Platform operator	ideaPoint/Anaqua, Inc.
Company location	USA
Reach	International
Data type	Clinical trial data
Value proposition	Transaction-centric
Transaction mechanism	Sharing platform
Role	Neutral
Access	Open (for registered organizations and researchers, but individual approval procedure)
Data integration	Domain-specific (health data: clinical trials)
Data transformation	Passive
Platform architecture	Centralized
Business model	Focused
Financing structure	One-dimensional (external basic funding from public health organizations in the UK, Bill & Melinda Gates Foundation and other foundations)

WEBSITE
https://dateva.biz

5.4.4 DATEVA

Since 2016, the data-centric platform Dateva has been pooling and analyzing health data and research results from various sources and making this data available to individuals and health organizations. The aim is to contribute to the development of preventive, predictive and personalized medicine. Besides public data, hospital data and drug-related research data from clinical trials, data sources also include data provided by patients or users via apps and wearables, for instance, via the VITALx patient app or various health or fitness apps or trackers. Data is uploaded to the cloud and combined with other data sets. Following a pooled data analysis, reports are passed on to medical organizations and companies in the medical and pharmaceutical sectors as well as to insurance and financial service providers and other interested parties, such as doctors and scientists. The Dateva platform is operated by the Canadian company of the same name based in Toronto.

Platform operator	DATEVA Inc.
Company location	Canada
Reach	International
Data type	Health data
Value proposition	Data-centric
Transaction mechanism	Sharing platform
Role	Active
Access	Open
Data integration	Domain-specific (health data: study and patient data)
Data transformation	Active (collection, normalization, aggregation and visualization)
Platform architecture	Centralized
Business model	Not specified
Financing structure	Not specified

5.4.5 LONGENESIS

WEBSITE
https://longenesis.com

Longensis focuses on fostering collaboration between stakeholders in the health sector by creating a platform-based ecosystem that enables direct communication and data exchange between different stakeholders. A central component is the utilization of medical data from bio-databases, hospitals, pharmaceutical and biotech companies and patient organizations for all stakeholders in the ecosystem. In order to facilitate and at the same time legitimize the direct involvement of patients, a digital consent tool based on blockchain technology is offered in order to obtain and use individual patient data for research purposes. However, further information regarding concrete usage figures or access requirements is not publicly available. Despite numerous nominations for technology awards and recognizable pilot applications of the platform architecture, this circumstance suggests that the Hong Kong-based start-up of the same name behind the platform is still in an early development phase.

Platform operator	Longensis Inc.
Company location	Latvia/Hong Kong
Reach	International
Data type	Health data
Value proposition	Data-centric
Transaction mechanism	Sharing platform
Role	Active
Access	Open
Data integration	Domain-specific (health data: patient data)
Data transformation	Active (aggregation and recruitment of new data providers)
Platform architecture	Not specified
Business model	Diversified
Financing structure	Not specified

WEBSITE
https://vivli.org

5.4.6 VIVLI

In 2015, the Multi-Regional Clinical Trials Center (MRCT) at Brighham and Women's Hospital and Harvard University, together with partners from the pharmaceutical industry and medical research, set out to design and jointly implement an independent platform for sharing clinical trial data. In July 2018, the non-profit platform Vivli was then launched, making available data from more than 5,400 clinical trials from 28 institutional members in the pharmaceutical industry and more than three million individual data users in 2020. Vivli hence claims to be the largest data exchange platform for clinical trials worldwide. The platform additionally offers a specialized search engine and its own data analysis tools for users. Access to individual data sets is only possible for verified users whose qualified request must be confirmed by the data provider. The platform is funded by the membership fees of its institutional members and foundations. There is no fee for individual use.

Platform operator	The Multi-Regional Clinical Trials Center of Brigham and Women's Hospital and Harvard
Company location	USA
Reach	International
Data type	Clinical trial data
Value proposition	Transaction-centric
Transaction mechanism	Sharing platform
Role	Neutral
Access	Open (for registered organizations, but individual approval procedure)
Data integration	Domain-specific (health data: clinical trials)
Data transformation	Passive
Platform architecture	Hybrid (data is not stored centrally, but shared via the platform for authorized parties after release)
Business model	Focused
Financing structure	Two-dimensional (external funding by foundations as well as by participating pharmaceutical companies as partners)

5.5 DOMAIN-SPANNING

5.5.1 ADVANEO

WEBSITE
https://www.advaneo.de

Data sovereignty is the focus of the ADVANEO data marketplace, which has been active since 2020 and is operated by the German software company of the same name. The marketplace is continuously updated and extended by further data sets through regular updates and offers open, domain-unspecific data from international sources. It also offers commercial data for sale directly via the marketplace, whereby this is offered exclusively via metadata, i.e., descriptive information. Compared to other marketplaces, ADVANEO's decentralized solution is characterized by the fact that the raw data remains exclusively in the sovereignty of providers and can only be found via metadata. The data offers are provided with rights and obligations of use and encrypted transmission of the original data by certified transmission software only takes place directly between the data provider and the buyer after successful conclusion of the deal. The data can then be analyzed directly in the web application. Revenues are generated via a freemium model where monthly basic fees and transaction commission are charged in addition to the free trial version for access to the marketplace. Organizations can agree upon further individual price and usage models on request.

Platform operator	Advaneo GmbH
Company location	Germany
Reach	International
Data type	No specification
Value proposition	Transaction-centric
Transaction mechanism	Data marketplace (with most of the data available free of charge)
Role	Neutral
Access	Semi-open (certification of participants according to IDS procedure; also option to set up closed user groups)
Data integration	Domain-unspecific
Data transformation	Passive
Platform architecture	Decentralized
Business model	Diversified (ensuring transaction), data science workbench, data pricing tool (in final development phase: technical support)
Financing structure	One-dimensional (pricing of access to the platform via freemium model with several levels from free to €29.50 per month as well as individual price agreements)

WEBSITE
https://aws.amazon.com/de/data-exchange

5.5.2 AWS DATA EXCHANGE

The AWS Data Exchange offering from online retailer and cloud provider Amazon is a cloud-based data marketplace for the secure subscription and use of third-party data products. The transaction-centric marketplace specifically focuses on qualified data providers whose brands are already established. These data providers pay monthly storage fees when using the platform, which are measured in 'byte-hours' and calculated according to data size and region. Data providers, for their part, charge subscription fees for use of their data which they determine themselves along with the terms of use. Once subscribed, subscribed data can be exported via an application programming interface (API) or used via other AWS services within the AWS cloud portfolio, some of which are paid for, for example, by transferring the data to the Amazon Simple Storage Service (Amazon S3) using the AWS Data Exchange console. Additional services are offered for data use in the AWS cloud portfolio.

Platform operator	Amazon Web Services, Inc.
Company location	USA
Reach	International
Data type	No specification
Value proposition	Transaction-centric
Transaction mechanism	Data marketplace (but supplemented by sharing components through the provision of publicly available data sets)
Role	Active
Access	Open
Data integration	Domain-unspecific
Data transformation	Active (provision of open data, aggregation)
Platform architecture	Centralized
Business model	Diversified (securing the transaction, data management, offering of AWS analysis tools and machine learning services)
Financing structure	One-dimensional (pricing of data use through subscription fees, including commission for AWS (determined by the data provider) and for data storage through storage fees)

5.5.3 DATA BROKER GLOBAL

WEBSITE
https://www.databroker.global

Data Broker Global is a neutral data marketplace developed and operated by Belgian provider SettleMint. Based on preliminary work on an originally conceived IoT marketplace for sensor data, this data marketplace went into commercial operation in 2020. The Data Broker Global marketplace uses blockchain technology to ensure secure data transfer between seller and buyer. Complementary services offered on the marketplace include consulting and a free DataMatch service to help customers find suitable data and potential data partners. Data Broker is also available as a white-label Platform as a Service (PaaS) solution, so that organizations with large amounts of data can operate their own data exchange platform using Data Broker.

Platform operator	SettleMint NV
Company location	Belgium
Reach	International
Data type	No specification
Value proposition	Transaction-centric
Transaction mechanism	Data marketplace
Role	Neutral
Access	Open
Data integration	Domain-unspecific
Data transformation	Active (data search, aggregation)
Platform architecture	Decentralized (using blockchain)
Business model	Diversified (PaaS services up to white-label solution for intra-organizational data exchange, securing the transaction, consulting services and guidance for data search, active data search for users)
Financing structure	One-dimensional (transaction fees for data use, fee for infrastructure services)

WEBSITE
https://www.dawex.com/de

5.5.4 DAWEX

The French company Dawex Systems SAS offers companies IT solutions for building platforms for data exchange and operates a data marketplace itself. Dawex Global Data Marketplace has been active since 2017 as an open marketplace where all types of data can be bought or sold worldwide, such as files, APIs, raw data or processed data. Dawex verifies the identity of all stakeholders operating on the marketplace in a validation process before they can buy or exchange data in individual transactions or by subscription. A blockchain-based license agreement is generated between buyers and data providers to carry out data transactions. Various pricing models are offered for access to the marketplace, which differentiate between different functions and scopes of services, ranging from free entry to a monthly fee of around €3,000 to individual services and price agreements. One year after the marketplace, the commercial launch of Dawex's Data Exchange Platform took place. This sharing platform offers organizations the possibility to confidentially exchange and enhance their data internally and within specific groups (such as subsidiaries or partners, customers and suppliers, etc.) while using all Dawex functionalities in private mode.

Platform operator	Dawex Systems SAS
Company location	France
Reach	International
Data type	No specification
Value proposition	Transaction-centric
Transaction mechanism	Hybrid (both open data marketplace and sharing platform in 'Data Exchange Platform mode')
Role	Neutral
Access	Semi-open (possibility to set up closed user groups)
Data integration	Domain-unspecific
Data transformation	Passive
Platform architecture	Centralized
Business model	Diversified (ensuring transaction security, onboarding support, data search simplification service)
Financing structure	One-dimensional (pricing of access to the platform via freemium model with several levels from free to €3,000 per month as well as individual price agreements)

5.5.5 SENATE

WEBSITE
https://www.datarepublic.com

SENATE is the sharing platform of Australian start-up Data Republic which enables data exchange between multiple parties and thus cross-organizational data-related collaboration. Data is exchanged only in predefined groups of several parties ('multi-party groups'). SENATE supports the transactions with various tools and utilities, such as a matching tool, governance workflows and licensing tools. SENATE generates revenues through use of the platform and utilization of its services either on a project-related basis (pay-per-use) or as an annual company license fee.

Platform operator	Datarepublic
Company location	Australia
Reach	International
Data type	No specification
Value proposition	Transaction-centric
Transaction mechanism	Sharing platform/data marketplace
Role	Neutral
Access	Closed (participation in principle open, but sharing is limited to predefined 'multi-party' groups.)
Data integration	Domain-unspecific
Data transformation	Active (aggregation via SENATE Matching)
Platform architecture	Decentralized
Business model	Diversified (securing the transaction, matching tool as well as further tools for quality assurance)
Financing structure	Multi-dimensional (revenues through various public/private funding rounds; pricing of platform use via a pay-per-use model, subscription model on an annual basis or fixed partnership)

WEBSITE
https://dih.telekom.net

5.5.6 TELEKOM DATA INTELLIGENCE HUB

The Data Intelligence Hub (DIH) is a cross-sector and cross-industry sharing platform operated by Deutsche Telekom IoT. A complementary marketplace function is currently under development. The aim of DIH is to provide organizations with a holistic overview of data that is freely available for exchange or direct purchase, the use of which is intended to help organizations optimize their internal processes along the value chain or develop innovative data-driven business models. For this purpose, the operator provides various supplementary analysis tools to process and structure the organization's own data combined with the data available on the platform using machine learning methods. The Data Intelligence Hub claims to be the first platform to meet the security requirements of the International Data Spaces Association (IDSA). A data trust architecture, the choice between centralized or decentralized data storage to ensure data sovereignty as well as certification of participants serve this purpose. Revenues are generated through platform use via a freemium model that charges supplementary fees for storage, computing power and transactions.

Platform operator	Deutsche Telekom IoT GmbH
Company location	Germany
Reach	International
Data type	No specification
Value proposition	Hybrid: equivalent data- and transaction-centricity
Transaction mechanism	Sharing platform (but marketplace function under development)
Role	Active
Access	Semi-open (certification of participants according to IDS procedure)
Data integration	Domain-unspecific
Data transformation	Active (collection, provision, normalization, aggregation, quality assurance indirectly via tool)
Platform architecture	Hybrid (both centralized and decentralized data storage possible)
Business model	Diversified (securing of the transaction, infrastructure services, tools for quality assurance of data and for data analysis, offer for organization-internal application of the technical solution)
Financing structure	One-dimensional (pricing of use in a hybrid freemium model, including commission fees for data use, sale of additional infrastructure as well as downstream sales for internal organizational solutions; in addition, platform as a service also for public sector clients)

5.5.7 TERBINE

WEBSITE
https://terbine.com/marketplace

The basic aim of US-based start-up terbine is to enable smooth use and monetization of machine-generated data in the Internet of Things using the 5G network infrastructure. The platform offers automated harmonization of metadata and a search service for potentially relevant data streams that can be used both manually and by machines. In this way, terbine aims to act as a smart intermediate layer that enables communication between IoT devices in a targeted manner and regulates it according to the data providers' specifications. Since 5G networks are not yet available to a sufficient extent, terbine is pursuing another core business area and integrating openly available data sets and data streams, which are then compiled into individual data offerings according to user demand. The data comes mainly from various public-sector sources with content covering mainly the areas of climate, logistics and transport.

Platform operator	Terbine
Company location	USA
Reach	International (focus on the US)
Data type	Open data and IoT data
Value proposition	Data-centric
Transaction mechanism	Sharing platform
Role	Active
Access	Open
Data integration	Domain-unspecific (data sources IoT devices and open government data)
Data transformation	Active (aggregation, normalization)
Platform architecture	Centralized (at present, but decentralized architecture being planned in the long term via edge computing in 5G infrastructure)
Business model	Diversified (securing the transaction, hosting and infrastructure, set-up as 'Platform as a Service')
Financing structure	Multi-dimensional (revenue through various public/private funding rounds; pricing of platform access through freemium model for $795 per month after 30 trial days; transaction fee for individual sales; individual pricing of PaaS services)

WEBSITE
https://up42.com

5.5.8 UP42

Up42 sees itself as an open developer platform and a marketplace for exchanging geodata and analyses. High-resolution satellite imagery, drone imagery, IoT and other data are made available, as well as the infrastructure, tools and algorithms needed to use this data. Developers and data scientists can use this offering or run their own algorithms in the cloud by uploading their software code directly to the platform or via an API. The spacetech start-up Up42, founded in 2019 by the aerospace group Airbus (Airbus Defence & Space), generates revenues via a freemium package model where prices vary depending on the intensity of use.

Platform operator	Up42 GmbH
Company location	Germany
Reach	International
Data type	Geographic data
Value proposition	Data-centric
Transaction mechanism	Data sharing
Role	Neutral
Access	Open
Data integration	Domain-unspecific (but data type-specific: geospatial data)
Data transformation	Active (provision of open data, aggregation, etc.)
Platform architecture	Centralized
Business model	Diversified (ensuring data availability and transfer, analytics tool – especially image processing, infrastructure and computing power; development tools)
Financing structure	Two-dimensional (monetization of partnerships as a basic prerequisite for making data or own data processing tools available; pricing of the intensity of use according to credits with: 10,00 credits free, 50,000 for €500, then quantity discount)

§ PRACTICAL NOTES

Public-sector data. For some time now, public-sector data has been recognized to have a high value creation potential (Kuzev, 2016). National legislation regarding the use of 'open data' was very heterogeneous within the EU for a long time, which is why the EU with its PSI Directive established uniform framework conditions for the re-use of public-sector data. The Directive stipulates, amongst other things, that data subject to its scope may in principle be re-used free of charge. The prerequisite is that the data concerned has already been made accessible or that a corresponding right of access exists. If these conditions are met, data from public bodies may be made available and shared via platforms (also for commercial use).

06

6 INTERMEDIATE CONCLUSION

Most data sharing platforms are still at the very beginning of their development. In the years to come, more players are expected to enter the market with new data sets and new business models. In the long term, however, strong consolidation of providers is to be expected due to the network effects known from online platforms. It is hence still far too early to draw any final conclusions about what successful data sharing platforms will look like in the future. However, an analysis of today's platform landscape is ventured and first experience gained by platform operators is also presented in order to show which factors deserve special attention when setting up new data sharing platforms.

6.1 Positioning of today's platforms

With regard to the value propositions of the individual platforms, the share of transaction-centric offerings is somewhat higher, whereby almost all of the identifiable value propositions contain elements of both ideal-typical extremes.

This suggests that even platforms designed purely as intermediaries in terms of their underlying concept are often interested in the active further development of their own data offerings because the number of data providers is still too small. This impression is supported by the large preponderance of platforms which themselves actively transform the available data sets. With regard to the understanding of the role of the platform operator as an active provider or active user of the available data, on the other hand, the analysis of the practical examples suggests a rather neutral general orientation. At the same time, some platform operators also take a very active role. This applies both to the provision of their own data sets and the collection of openly available data sets as well as to the use of available data. Access to the platforms is almost exclusively open, apart

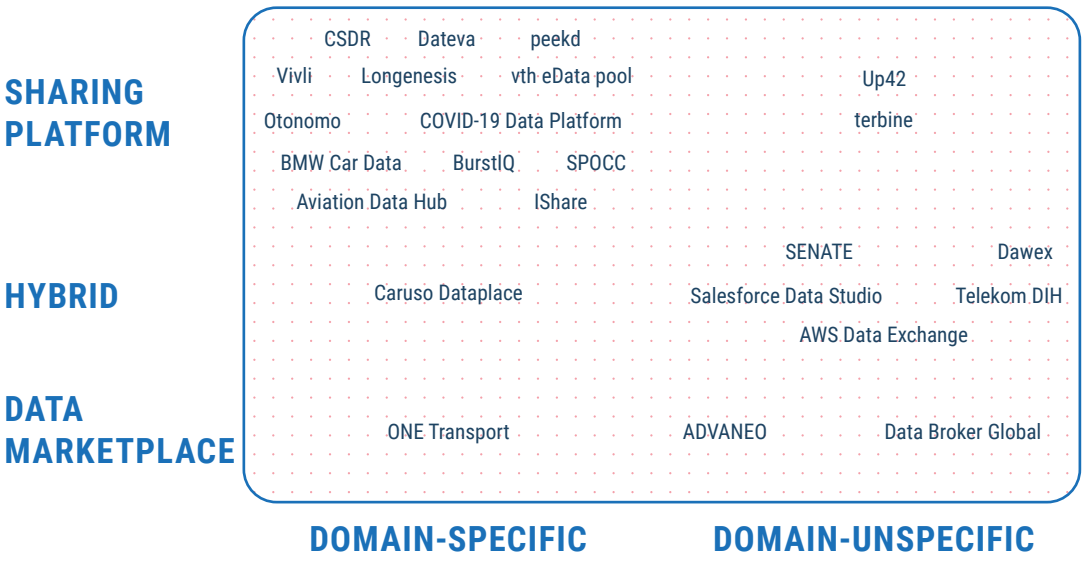


Figure 1: Focus of the 24 practical examples: domain-specific sharing platforms, but domain-unspecific data marketplaces and hybrid platforms.

from registration or authentication. Exceptions mainly exist with regard to semi-open platforms whose providers make access more difficult or even block it entirely. This holds particularly true when a platform operator is itself active in the respective industry. The fact that almost all domain-specific case studies belong to the area of sharing platforms with partially limited access and usage rules confirms the necessary level of trust for sharing data in competitive situations. In contrast, free and open sales via data marketplaces or hybrid forms that support both transaction mechanisms are found almost exclusively in the area of domain-unspecific platforms and to a much lesser extent (see Fig. 1).

Most platform architectures feature a centralized structure. Data is both stored and accessed directly via the platform. Decentralized solutions are less common and function either via a blockchain solution or via organized metadata matching. With regard to the recognizable business models, it can be said that almost all platform operators offer a – sometimes strongly – diversified portfolio of services for data providers and data users which go far beyond the pure realization of data transactions. In this sense, the additional instruments provided are mainly tools for improving the quality of data or its analysis and subsequent visualization. Supplementary technical infrastructure services and implementation support as well as partly data-based advisory services are offered in addition. The financing structure is designed almost exclusively one-dimensionally through the pricing of the services provided by the platform operator and, of course, primarily through data use. This is typically based on hybrid models consisting of a basic fee or membership combined with individual pay-per-use arrangements. However, a few one-dimensional models also exist where data is exchanged within a thematically linked group of stakeholders via a jointly funded consortium structure.

The research results as a whole show a dynamic market environment of available data marketplaces and data sharing platforms. At the beginning of the 2010s, a number of data sharing platforms or data marketplaces were already developed in a first wave – some of them also by large technology companies – that are no longer available today or can no longer be classified as data sharing platforms or data marketplaces in the core of their business model (see BMWi 2020a). However, in the last four to five years at the most, a second attempt oby a diverse group of players can be seen. Both start-ups and established companies are determined to advance the establishment and operation of data sharing platforms. However, it was not (yet) possible in this study to clearly identify data marketplaces and data sharing platforms that feature permanently high traffic and are economically sustainable. Instead, the majority of the case studies are still in the set-up or even only in the extended concept development phase. This is not surprising considering the fact that several interviewees confirmed that they only expect sufficient usage intensity to establish marketplace profitability after at least five to seven years. This is consistent with the results of studies with a similar design (see Meisel and Spiekermann 2019). For many platforms, sparse publicly available information about how the service works and the lack of willingness to respond or participate in the study suggest that these platforms are still at an early stage of development. One exception here are the platforms for clinical trial data from the health sector which have established themselves as an instrument for cross-organization data exchange within a fixed network.

6.2 Lessons learned and success factors for platform development

Building on the analysis of the case studies and on the basis of the expert interviews conducted with representatives from a total of eight of the 24 platforms presented above, the following success factors for and obstacles to the development of data marketplaces and data sharing platforms were identified. A detailed valid review of individual service offerings as well as business and funding models is not possible since the platform landscape is still at the development stage. However, factors could be identified that can promote the successful launch and scaling of the platform as quickly as possible, or that represent fundamental sources of mistakes.

Focus on stakeholder needs. A key success factor is a clear focus on the needs of the (potential) stakeholders involved, i.e., data providers and data users. Just because an exchange of data is technically possible, a sufficient number of stakeholders necessary for the platform to function still have to actively decide to participate. Accordingly, all experts consider a purely technical focus to be a major mistake in platform development. The central needs here relate primarily to the available data with high relevance for the respective users as well as their individual preferences with regard to the usable formats. Furthermore, specific stakeholder groups have specific demands regarding platform security, the procedure for admitting new stakeholders and access and usage rights management. This is especially true for domain-specific platforms with their narrow content focus. It is precisely here that the creation of initially closed group structures can help, including a jointly defined regulatory framework for the inclusion of new stakeholders. With regard to the familiar issues that can always be observed when setting up platforms, i.e., economies of scale and network effects which are initially difficult to trigger and which, in the area of data sharing platforms, are primarily due to the lack of a basic stock of relevant data, the platform operator can assume an active role and thereby contribute to the solution. The active provision of data sets from the platform operator's own area of responsibility and/or the active collection of openly available data sets – especially the various public sector sources – and their provision create a potentially relevant stock of data that can attract both data users and, with growing interest on the demand side, data providers too.

Community building. The emphasis placed by potential platform operators on integrated community-building measures is closely related to the focus on the content demanded by stakeholders, but can still be listed as a single success factor. Even before the platform is launched, it can be helpful here to identify central stakeholders in the respective application area and to address them directly or via network players, such as associations, and to involve them in the development of the platform. In this way, potential strategic partners can be found at an early stage, whose par-

ticipation can not only increase the neutrality of the platform through distributed sponsorship, but also ensure the availability of a sufficient amount of relevant data. Active and successful community building can also include the establishment of a trust-building conflict management system and the continuous further development of the metadata catalogue based on feedback from participating stakeholders. Another central building block is constant proactive communication of successful examples of data exchange that was only possible through the platform. Finally, community events, such as hackathons or user meetings, can be used to address new target groups and, in particular, to bind existing stakeholders to the platform.

Flexible and agile software and product development. Following the focus on stakeholder needs, agile software and product development are then the logical conclusion as an essential success factor. This approach is based on constant feedback from (potential) data providers and data users which should be obtained at an early stage of the platform if possible. Many case studies show that the starting point of plans for platform development is based on the basic idea of implementing a pure transaction infrastructure. During the further course, however, new features were added to the platform to improve the overall experience for data providers and data users. These include, above all, tools to ensure data quality, but also complementary infrastructure services. In this way, the business model becomes more diverse and opportunities arise to tap into new funding streams.

Clear definition of the platform operator's role. The clear-cut definition of the respective platform operator's role and its external communication can also make a significant contribution towards building trust with potential platform participants and especially with data providers. This holds particularly true when a platform operator wishes to actively use the data itself. This case generally involves a higher risk of uncertainty on the part of platform participants and should only be implemented in combination with clearly defined and communicated rules and a readily visible added value for participants. In contrast, taking a more neutral position as an anchor of trust and communicating this is more conducive to success. Here, the operator's own security standards as well as tools for authenticating participants and ensuring data quality can be cited.

07

7 LITERATURE

Arnaut, Catarina/Pont, Marta/Scaria, Elizabeth/Berghmans, Arnaud/Leconte, Sophie (2018): Study on data sharing between companies in Europe, Luxemburg: Europäische Kommission.

Bartels, Karsten U./Beck, Susanne/Buchholz, Birgit/Bürger, Matthias/Straub, Sebastian (2020): Leitfaden Kollaborative Wertschöpfungssysteme in der Industrie, hrsg. Begleitforschung PAiCE, iit – Institut für Innovation und Technik in der VDI/VDE Innovation + Technik.

Brynjolfsson, Erik und Kristina McElheran: The Rapid Adoption of Data-Driven Decision Making, in: American Economic Review, 106(5), S. 133-139, 2016.

Bundesministerium für Wirtschaft und Energie (2020a): Datenmarktplätze in Produktionsnetzwerken – Impulspapier der Plattform Industrie 4.0.

Bundesministerium für Wirtschaft und Energie (2020b): Plattform Industrie 4.0, Diskussionspapier, Anonymisierung im Datenschutz als Chance für Wirtschaft und Innovationen.

Bussche, Axel Freiherr von dem (2020): Datenschutz 4.0, in Frenz, Walter (Hg.): Handbuch Industrie 4.0: Recht, Technik, Gesellschaft, Berlin/Heidelberg: Springer.

Cattaneo, Gabriela und Chiara Francalanci (2020): Story 9 – Scaling up data-driven innovation: European industry requirements and the role of European data spaces, Luxemburg, Europäische Kommission.

Clarke, Roger: Big Data, Big Risks (2016), in: Information Systems Journal, 26(1), S. 77-90.

Engelhardt, Sebastian von/Wangler, Leo/Wischmann, Steffen (2017): Eigenschaften und Erfolgsfaktoren digitaler Plattformen, Begleitforschung AUTONOMIK für Industrie 4.0, iit – Institut für Innovation und Technik in der VDI/VDE Innovation + Technik GmbH, hrsg. von Alfons Botthof.

Europäische Kommission (2020): The European Data Market Monitoring Tool, Brüssel.

Fezer, Karl-Heinz (2016): Repräsentatives Dateneigentum – Ein zivilgesellschaftliches Bürgerrecht, Sankt Augustin/Berlin: Konrad-Adenauer-Stiftung.

Fricker, Samuel und Yuliyan Maksimov (2017): Pricing of Data Products in Data Marketplaces, in: Proc. ICSOB International Conference of Software Business, Cham: Springer.

Froese, Julia und Sebastian Straub (2020): Wem gehören die Daten? – Rechtliche Aspekte der digitalen Souveränität in der Wirtschaft: in: Ernst A. Hartmann (Hg.): Digitalisierung souverän gestalten – Innovative Impulse im Maschinenbau, Heidelberg: Springer Vieweg.

Guggenberger, Tobias et al. (2020): Data Sharing in industrial ecosystems – Driving value across entire production lines, McKinsey & Company/Fraunhofer ISST.

INFORM (2017): Data Sharing als wichtiger Faktor für die Supply Chain, https://www.inform-software.de/news/news-details/news/data-sharing-als-wichtiger-erfolgsfaktor-fuer-die-supply-chain.

KPMG: Cloud-Monitor 2020 (2020) – Die Integrationsfähigkeit und Interoperabilität der Cloud stärken, Unternehmensumfrage in Zusammenarbeit mit Bitkom Research.

Krotova, Alevtina/Rusche, Christian/Spiekermann, Markus (2019): Die ökonomische Bewertung von Daten, IW-Analyse, Nr. 129.

Kühling, Jürgen und Benedikt Buchner (2018): Datenschutz-Grundverordnung/BDSG – Kommentar, 2. Auflage, München: Beck.

Kuzev, Pencho/Dapp, Marcus M./Balta, Dian/Palmetshofer, Walter/Krcmar, Helmut (Hg.) (2016): Open Data. The Benefits. Das volkswirtschaftliche Potential für Deutschland. St. Augustin/Berlin: Konrad-Adenauer-Stiftung.

Linnartz, Maria und Anja Leckel (2020): Data Sharing im Supply-Chain-Management – Ein Assistenzsystem für die Berwertung und Gestaltung der Offenheit und des Vertrauensniveaus zwischen Supply-Chain-Partnern, ZWF – Zeitschrift für wirtschaftlichen Fabrikbetrieb, S. 563-566.

Marx, Uwe (2020): Der Maschinenbau hinkt digital hinterher, FAZ, 18.09.2020.

Meisel, Lukas und Markus Spiekermann (2019): Datenmarktplätze – Plattformen für Datenaustausch und Datenmonetarisierung in der Data Economy, ISST-Bericht, hrsg. v. Boris Otto und Jakob Rehof, Dortmund: Fraunhofer ISST.

Nentwig, Stefan/Saft, Danilo/Taphorn, Christoph (2019): Mehrwerte aus Daten – Potenziale und Handlungsoptionen für den Mittelstand, Mittelstand 4.0-Kompetenzzentrum Dortmund.

Niederee, Michael (201): Data Sharing – Die Zukunft der Wertschöpfung in der Datenökonomie, Whitepaper, KPMG.

Otto, Boris et al (2020): Datenmarktplätze in Produktionsnetzwerken, Impulspapier der AG 6 „Digitale Geschäftsmodelle" der Plattform Industrie 4.0, hrsg. von Bundesministerium für Wirtschaft und Energie (BMWi).

Otto, Boris et al. (2019): Data Economy – Status Quo der deutschen Wirtschaft & Handlungsfelder in der Data Economy, White Paper, DEMAND-Projekt.

Otto, Boris und Hubert Österle (2016): Corporate Data Quality – Voraussetzungen erfolgreicher Geschäftsmodelle, Wiesbaden: Springer Gabler.

Ringel, Michael/Taylor, Andrew/Zablit, Hadi (2017): The Most Innovative Companies 2016 – Getting Past Not Invented Here, Boston Consulting Group.

Ringel, Michael et al. (2020): The Most Innovative Companies 2020 – The Serial Innovation Imperative, Boston Consulting Group.

Schmid, Theresa: Startup otonomo (2020): Mit Connected-Car-Austauschplattform zum Erfolg, abgerufen über: https://emobilitaetblog.de/startup-otonomo-mit-connected-car-austauschplattform-erfolg/.

Specht, Louisa (2017): Datenverwertungsverträge zwischen Datenschutz und Vertragsfreiheit – Eckpfeiler eines neuen Datenschuldrechts, DGRI Jahrbuch.

Specht, Louisa und Wolfgang Kerber (2017): Datenrecht – Eine Rechts- und Sozialwissenschaftliche Analyse im Vergleich Deutschland – USA, ABIDA-Projekt.

Spiekermann, Markus: Data Marketplaces (2019): Trends an Monetisation of Data Goods, Intereconomics, 2019/4, ZBW – Leibniz Information Centre for Economics, S. 208-216.

Stahl, Florian/Löser, Alexander/Vossen, Gottfried (2015): Preismodelle für Datenmarktplätze, in: Informatik_Spektrum 38/2, S. 133-141.

Stahl, Florian/Schomm, Fabian/Vossen, Gottfried/Vomfell, Lara (2016): A classification Framework for data marketplaces, in: Vietnam Journal of Computer Science, 3/3, S. 137-143.

Wölfl, Steffen/Leischnig, Alexander/Ivens, Björn/Hein, Daniel (2018): From Big Data to Smart Data – Problemfelder der systematischen Nutzung von Daten in Unternehmen, in: Geschäftsmodelle in der digitalen Welt, S. 213-231.

APPENDIX

OVERVIEW OF THE 24 PRACTICAL EXAMPLES

The analysis of the 24 practical examples was conducted based on publicly available company-information as well as on single expert interviews. In case of insufficient data, the particular category was accordingly marked with "not specified".

PLATFORM	PLATFORM OPERATOR	COMPANY LOCATION	DOMAIN	DATA TYPE	REACH	VALUE PROPOSITION
Advaneo	Advaneo GmbH	Germany	Domain-spanning	No specification	International	Transaction-centric
Aviation Data Hub	AVIATION DataHub GmbH	Germany	Logistik	Aircraft and logistics data	International	Transaction-centric
AWS Data Exchange	Amazon Web Services, Inc.	USA	Domain-spanning	No specification	International	Transaction-centric
BMW CarData	Bayerische Motoren Werke Aktienge-sellschaft	Germany	Industry	Telematics data from vehicles	International	Data-centric
burstIQ	BurstIQ Corporation	USA	Health sector	Health data	International (however, focus on the US)	Data-centric
Caruso Dataplace	Caruso GmbH	Germany	Industry	Vehicle data	International	Transaction-centric
COVID 19 Data Portal	Elixir	UK	Health sector	Health data and study data on COVID-19	European	Data-centric
CSDR	ideaPoint/Anaqua, Inc.	USA	Health sector	Clinical trial data	International	Transaction-centric
Data Broker Global	SettleMint NV	Belgium	Domain-spanning	No specification	International	Transaction-centric
Dateva	DATEVA Inc.	Canada	Health sector	Health data	International	Data-centric
Dawex	Dawex Systems SAS	France	Domain-spanning	No specification	International	Transaction-centric
IShare.org	iSHARE Foundation	Netherlands	Logistics	Logistics data	National/European	Transaction-centric
Longenesis	Longensis Inc.	Latvia / Hong Kong	Health sector	Health data	International	Data-centric
One TRANSPORT Data Marketplace	Chordant	USA/UK	Logistics	Mobility and logistics data	International	Transaction-centric
Otonomo	Otonomo Ltd.	Israel	Industry	Vehicle and traffic data	International	Transaction-centric
peekd	peekd.ai	Germany	Trade	E-commerce data	International	Data-centric
Salesforce Data Studio	Salesforce.org EMEA Limited	USA	Trade	E-commerce data	International	Transaction-centric
SENATE	Datarepublic	Australia	Domain-spanning	No specification	International	Transaction-centric
SPOCC	SPOOC GmbH & Co. KG	Germany	Trade	Product data and marketing content	International	Transaction-centric
Telekom DIH	Deutsche Telekom IoT GmbH	Germany	Domain-spanning	No specification	International	
terbine	terbine	USA	Domain-spanning	Open data and IoT data	International (focus on the US)	Data-centric
Up42	Up42 GmbH	Germany	Domain-spanning	Geographic data	International	Data-centric
Vivli	The Multi-Regional Clinical Trials Center of Brigham and Women's Hospital and Harvard	USA	Health sector	Clinical trial data	International	Transaction-centric
vth edata pool	IFCC GmbH	Germany	Trade	Product and article data	Germany	Transaction-centric

Chart 1 The functions of data sharing platforms based on the reference model by Meisel and Spiekermann (2019)

TRANSACTION MECHANISM	ROLE	ACCESS	DATA INTEGRATION	DATA TRANS-FORMATION	PLATFORM ARCHITEC-TURE	BUSINESS MODEL	FINANCING STRUCTURE	PLATFORM
Data marketplace	Neutral	Semi-open	Domain-unspecific	Passive	Decentralized	Diversified	One-dimensional	Advaneo
Sharing platform	Active	Open	Domain-specific	Active	Centralized	Diversified	Not specified	Aviation Data Hub
Data marketplace	Active	Open	Domain-unspecific	Active	Centralized	Diversified	One-dimensional	AWS Data Exchange
Sharing platform	Active	Semi-open	Domain-specific	Active	Centralized	Focused	One-dimensional	BMW CarData
Sharing platform	Active	Open	Domain-specific	Active	Decentralized	Diversified	One-dimensional	burstIQ
Marketplace	Neutral	Open	Domain-specific	Active	Centralized	Diversified	Two-dimensional	Caruso Dataplace
Sharing platform	Active	Open	Domain-specific	Active	Centralized	Focused	One-dimensional	COVID 19 Data Portal
Sharing platform	Neutral	Open	Domain-specific	Passive	Centralized	Focused	One-dimensional	CSDR
Data marketplace	Neutral	Open	Domain-unspecific	Active	Decentralized	Diversified	One-dimensional	Data Broker Global
Sharing platform	Active	Open	Domain-specific (health data: study and patient data)	Active (collection, normalization, aggregation & visualization)	Centralized	Not specified	Not specified	Dateva
Hybrid	Neutral	Semi-open	Domain-unspecific	Passive	Centralized	Diversified	One-dimensional	Dawex
Sharing platform	Passive	Semi-open	Domain-specific	Passive	Decentralized	Focused	Two-dimensional	IShare.org
Sharing platform	Active	Open	Domain-specific	Active	Not specified	Diversified	Not specified	Longenesis
Data marketplace	Passive	Open	Domain-specific	Passive	Hybrid	Diversified	One-dimensional	One TRANSPORT Data Marketplace
Sharing platform	Neutral	Open	Domain-specific (connected car data)	Active	Centralized	Diversified	One-dimensional	Otonomo
Sharing platform	Active	Semi-open	Domain-specific		Centralized	Diversified	One-dimensional	peekd
Data marketplace	Neutral	Semi-open	Domain-specific	Passive	Centralized	Focused	Not specified	Salesforce Data Studio
Sharing platform/ Data marketplace	Neutral	Closed	Domain-unspecific	Active	Decentralized	Diversified	Multi-dimensional	SENATE
Sharing platform	Active	Open	Domain-specific	Active	Centralized	Diversified	Not specified	SPOCC
Sharing platform	Active	Semi-open	Domain-unspecific	Active	Hybrid	Diversified	One-dimensional	Telekom DIH
Sharing platform	Active	Open	Domain-unspecific	Active	Centralized	Diversified	Multi-dimensional	terbine
Data Sharing	Neutral	Open	Domain-unspecific	Active	Centralized	Diversified	Two-dimensional	Up42
Sharing platform	Neutral	Open	Domain-specific	Passive	Hybrid	Focused	Two-dimensional	Vivli
Sharing platform	Neutral	Open	Domain-specific	Active	Centralized	Diversified	Multi-dimensional	vth edata pool