

Establishing the Enterprise Data Marketplace: Characteristics, Architecture, and Challenges

Rebecca Eichler
University of Stuttgart
Stuttgart, Germany
rebecca.eichler@ipvs.uni-stuttgart.de

Christoph Gröger
Robert Bosch GmbH
Stuttgart, Germany
christoph.groeger@de.bosch.com

Eva Hoos
Robert Bosch GmbH
Stuttgart, Germany
eva.hoos@de.bosch.com

Christoph Stach
University of Stuttgart
Stuttgart, Germany
christoph.stach@ipvs.uni-stuttgart.de

Holger Schwarz
University of Stuttgart
Stuttgart, Germany
holger.schwarz@ipvs.uni-stuttgart.de

ABSTRACT

Companies today have increasing amounts of data at their disposal, most of which is not used, leaving the data value unexploited. In order to leverage the data value, the data must be democratized, i.e., made available to the company employees. In this context, the use of enterprise data marketplaces, platforms for trading data within a company, are proposed. However, specifics of enterprise data marketplaces and how these can be implemented have not been investigated in literature so far. To shed light on these topics, we illustrate the characteristics of an enterprise data marketplace and highlight according marketplace requirements. We provide an enterprise data marketplace architecture, discuss how it integrates into a company's system landscape and present an enterprise data marketplace prototype. Finally, we examine organizational and technical challenges which arise when operating a marketplace in the enterprise context. In this paper, we thereby present the enterprise data marketplace as a distinct marketplace type and provide the basis for establishing it within a company.

1 INTRODUCTION

In this day and age, an enormous amount of data is generated by, for instance, the Internet of Things (IoT) or social media networks. This data contains a potential value which may lead to new insights, the discovery of new business models or the expansion into new markets. The data value can, however, only be extracted if the data is available for use. In this context, the data marketplace (DMP) is gaining in importance. Data marketplaces are electronic platforms for trading data as well as data-related services [24, 20]. A marketplace provides infrastructure for the data exchange by acting as a digital intermediary connecting data providers and data consumers [24]. Data marketplaces yield several advantageous outcomes. For instance, they stimulate innovation as consumers can acquire data which would have been unavailable and available data can initiate the improvement of products, services, or processes or also the development of new business models [14].

Data marketplaces are mainly considered for the exchange of data and services between organizations or private individuals. There are, however, also other relevant application scenarios for data marketplaces, such as their deployment within a company. Studies show that approximately two thirds of data goes unused

within companies [30]. In this context the FAIR principle, i.e., making data findable, accessible, interoperable and reusable [19, 38], as well as data democratization are discussed in literature. Data democratization has the objective to motivate and empower the majority of company employees to find, understand, access, use, and share data within the company, in consideration of data security and compliance [22, 3]. Lefebvre et al. [22] define four data democratization dimensions. The first describes the enablement of broader access to data and tools for users with varying skill-sets, the second signifies the development of data-related and analytic skills such as data cleaning. The third dimension covers collaborative knowledge-sharing between employees, and the fourth entails the promotion of data value like communicating the importance of data assets. In this context, it has been proposed to employ the data marketplace within a company in order to address data democratization and the corresponding dimensions [5]. In the company internal context, the marketplace is referred to as an enterprise data marketplace (EDMP) [12, 37] or an internal data marketplace [9].

In extension of Wells' [36], we propose the following definition: The enterprise data marketplace is a type of data marketplace for the exchange of data and data-related services between company employees, and optionally invited guests. It has the objective to democratize data within the company and thus offers the majority of a company's data. This includes data from different domains, data in varying processing degrees, and also data insights such as reports or machine learning models. This also entails that the enterprise data marketplace offers data from both operational systems like enterprise resource planning (ERP) systems and analytical systems like data lakes or data warehouses.

While there are first prototypes of these marketplaces in companies [5], the enterprise data marketplace has been studied very little in literature, and topics such as how it differs from other types of marketplaces, how such a marketplace is built, or what challenges arise in this marketplace context have not been discussed in detail. Therefore, we make the following contributions: Besides the provided definition, (1) we position the enterprise data marketplace in a classification framework differentiating it from other marketplaces and thereby provide a type distinction in Section 2. In Section 3 (2) we present requirements for data marketplaces and highlight which are specific to an enterprise data marketplace. Based on these



requirements (3) we provide an enterprise data marketplace platform architecture in Section 4. In extension, (4) we also discuss how the enterprise data marketplace integrates in the existent enterprise system and storage landscape in Section 5. To demonstrate how the presented concepts can be realized, (5) we showcase an enterprise data marketplace prototype in Section 6. Lastly, (6) we illustrate enterprise data marketplace challenges based on the idiosyncrasies of this marketplace type in Section 7. Finally, related work is presented in Section 8 and Section 9 concludes this paper.

2 CLASSIFYING THE ENTERPRISE DATA MARKETPLACE

In order to identify the distinguishing characteristics of the enterprise data marketplace, we position it in a classification framework for data marketplaces. The framework is presented in Section 2.1 and the identified characteristics are discussed in the following Section 2.2. By highlighting these distinct features, we introduce the enterprise data marketplace as a marketplace type.

2.1 The Marketplace Classification Framework

To classify the enterprise marketplace, we studied marketplace characteristics provided through various research articles such as [10, 33, 29, 35, 24, 20, 32]. The characteristics range from aspects like marketplace ownership over the value proposition, data access methods, monetization aspects to the underlying architecture. Some characteristics such as the marketplace ownership [33] are relevant for the distinction of the enterprise data marketplace, whereas other characteristics like the offering of pre-purchase testability [33] are not. We found that the classification frameworks provided by Meisel and Spiekermann [24] and Spiekermann [32] contain most of the relevant characteristics for the distinction of the enterprise data marketplace. The framework in [24] is a composition of characteristics provided through various research articles including [20, 34, 17] and the framework provided in [32] is based on a taxonomy developed explicitly for classifying data marketplaces based on their business models. We developed the marketplace classification framework as displayed in Figure 1 by combining both of these frameworks. We extended the resulting framework with the attribute *consumer* for the sake of completeness and renamed a few attributes and corresponding characteristics. These include the characteristic *company*, which is called “commercial” in the original source. As the term commercial signifies both a business interest and cash flow, yet the cash flow does not represent the participant, we renamed it company which complements the characteristics *private individual* and *public institution* in this section. Also the attribute “market positioning” [32] is replaced through the more expansive attribute *ownership* of [24] and the attribute “integration” [32] is renamed to *data offering*. By grouping the attributes, we receive five dimensions based on which an enterprise data marketplace can be classified: the *market participants*, the *market position*, the *market offering*, *monetization* and *technical aspects*. The characteristics that apply to the enterprise data marketplace in these dimensions are highlighted in two shades of light grey.

DIMENSION	ATTRIBUTE	CHARACTERISTIC			
MARKET PARTICIPANTS	PROVIDER	Company	Private Individual	Public Institution	Black Market
	CONSUMER	Company	Private Individual	Public Institution	Black Market
MARKET POSITION	OWNERSHIP	Private	Consortium		Independent
	MATCHING	One-to-One	One-to-Many	Many-to-One	Many-to-Many
	MARKET ACCESS	Open	Closed	Hybrid	
MARKET OFFERING	VALUE PROPOSITION	Transaction-centric		Data-centric	
	DATA OFFERING	Domain-unspecific		Domain-specific	
	TRANSFORMATION	Raw data	Normaliz.	Aggregat.	Quality Assurance
MARKET MONETIZATION	PRICE MODEL	Free	Fixed price	Pay-per-use	...
	REVENUE MODEL	Free	Flat rate	Fee	...
TECHNICAL ASPECTS	ARCHITECTURE	Central	Dezentral	Hybrid	

Characteristic applies to EDMP
 Characteristic applies when viewing EDMP participants as departments/employees
 Optional
 Not relevant for EDMP classification

Figure 1: Marketplace Classification Framework Highlighting the Characteristics of the Enterprise Data Marketplace.

2.2 Enterprise Data Marketplace Characteristics

For the attributes defined in the framework, one, several or none of the characteristics may apply to the enterprise data marketplace.

Market participants involve both the data and service providers as well as the consumers in the marketplace. In the case of the enterprise data marketplace the participants in both categories are employees within the same *company*, this is not immediately apparent through the classification framework. In some cases, an enterprise may choose to open their marketplace to selected business partners [36], which also classify as a company.

The *market position* signifies who *owns* or operates the marketplace, the *matching*, i.e., the number of parties involved, together with the service orientation among these, as well as the *accessibility* of the marketplace. As the enterprise data marketplace mainly contains enterprise internal data, including classified and personal data, it is usually owned and operated by the same company, hence is *private*. Considering not the entire company, but its departments or employees as participants, it can be argued that it is either a *consortium* or *independent* marketplace depending on whether the department operating the marketplace is an active participant. Therefore, all three characteristics are highlighted. In the same sense, it is a *one-to-one* matching, considering the entire company exchanging data and services with itself, or a *one-to-many* or *many-to-one* matching, if business partners are involved and the company is either sharing with or receiving data and services from them. The *many-to-many* matching refers to the company’s departments or employees trading data amongst each other. Depending on whether the enterprise

data marketplace is accessible only to the company employees or also to invited guests, it is *closed* or *hybrid* respectively.

The dimension *market offering* constitutes the *value proposition*, *data offering* and *transformation functionality* in the marketplace. The enterprise data marketplace’s value proposition is *transaction-centric* as its core offering is the switching function of data and services, i.e., bringing data providers and consumers together. It only forwards the consumer to tools for data analysis, visualization and preparation and does not incorporate this functionality and is therefore not data-centric, according to [32]. The scope of offered data spans across all company data, hence, the data offering is *domain-unspecific*. According to Spiekermann [32] transformation refers to the marketplace’s ability to transform raw data into a normalized or an aggregated state or assure data quality. While we argued in [4] that a marketplace does not offer functionality to process data, e.g., aggregate it, the marketplace can offer data in various transformation states, e.g., data stored in data lake zones in varying processing degrees. These characteristics are undefined, as they are not relevant for classifying the enterprise data marketplace.

As *monetization* of data offerings would hinder the enterprise data marketplace’s goal of democratizing data within a company, the *price model* for most offerings is *free*. There may be instances in which a cash flow between separate business units is required for legal reasons, or if data is sold to a business partner, therefore, the enterprise data marketplace may support any other form of price model as well. The *revenue model* signifies under which monetary conditions participants can use the marketplace. As a revenue model would be a barrier for employees to use the marketplace, and therefore hinder data democratization, the revenue model is *free* in the enterprise data marketplace.

With the goal of democratizing most enterprise data, it is feasible to retain data in the source systems, as opposed to storing it redundantly in a centralized marketplace repository. Therefore, it has a *decentralized* data storage architecture. However, to support the registration of, e.g., a single report or file which should not be stored in any other storage system a *hybrid* approach with both a centralized and decentralized repository can be chosen. Concluding, a data marketplace that meets these criteria is classified as an enterprise data marketplace and is subject to specific requirements.

3 ENTERPRISE DATA MARKETPLACE REQUIREMENTS

Having identified an enterprise data marketplace’s characteristics, we now specify requirements concerning the marketplace’s offering in terms of data and services, functionality and as this marketplace is operated within an company, requirements to how the marketplace should integrate with the existent enterprise system landscape. The requirements are derived from existing literature on data marketplaces and data democratization, complemented by a case study on a large industrial company in order to include an enterprise-practice point of view. The company in question, is a globally active manufacturer, striving to become a data-driven Industry 4.0 company and is therefore building a tool landscape including a data marketplace (for details on the case study see [6]). In the following Sections 3.1-3.3, we highlight and explain which requirements are specific to the enterprise data marketplace and

Table 1: Relevance of Requirements in the DMP and EDMP.

Requirement		DMP	EDMP
Offerings	Data-as-a-Service	+	+
	Infrastructure-as-a-Service	o	+
	Software-as-a-Service	o	+
	Professional Services	o	+
Functionality	Consumer-Side	+	+
	Provider-Side	+	+
	Administration-Side	+	+
	Metadata-Management	+	++
	Privacy & Security	+	++
Enterprise Integration		-	++

(-) irrelevant (o) not specifically relevant (+) relevant (++) specifically relevant

which are relevant for data marketplaces in general as shown in Table 1.

3.1 Required Marketplace Offerings

As mentioned in the introduction, it is the objective of an enterprise data marketplace to address data democratization, which implicitly sets the baseline for the required offerings.

In order to facilitate the data democratization dimension of broader access to data [22], all kinds of data has to be made available within the company [12]. Therefore, the data marketplace’s main offer must be *Data-as-a-Service* [37]. Ultimately, the marketplace should make all corporate data available. This includes data from operational systems such as ERP systems as well as analytical systems like data lakes. Both internal company and externally acquired data are included in this. Likewise, raw data, data in various processing degrees as well as ready-to-use data and data insights such as machine learning models or reports, belong into this scope. As explained in Section 2.2, the data is not limited to a domain such as finance or manufacturing.

The definitions of data democratization also specify that the data must be made available to all kinds of users, i.e. also non-specialist users [12, 3]. This type of user may lack the skills for setting up the required environment or only have skills to work with data in specific tools. Hence, the marketplace must also offer *Infrastructure-as-a-Service*, and *Software-as-a-Service* in combination with the data. For instance, a user may order data with infrastructure like a virtual machine or have it provided directly in a tool such as a Tableau¹ or Microsoft Power BI² instance. Thereby, the marketplace supports self-service consumption of data. Any marketplace can offer these services, yet they are relevant in the enterprise data marketplace to achieve the first data democratization dimension.

The development and sharing of data skills is part of the second data democratization dimension [22, 3]. Hence, the marketplace should also offer *Professional Services*. These are services offered by users with specific skills and can, for example, involve courses to acquire skills for processing data, dashboarding or data preparation.

¹www.tableau.com

²powerbi.microsoft.com

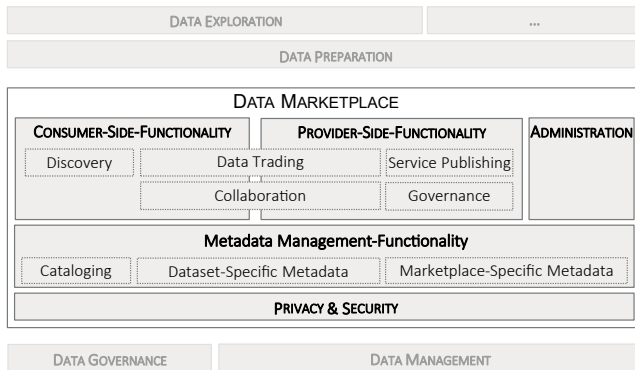


Figure 2: Data Marketplace Functionality Framework. Summarized and Adapted from [4].

While all these offerings are not exclusive to an enterprise data marketplace, they are relevant for it because of the democratization objective of this type of marketplace.

3.2 Required Functionality

Based on the general functionality framework for data marketplaces we present in [4] there is role-based functionality for the *consumer*, *provider* and the *administration*. In addition, marketplaces offer cross-sectional functionality which includes *metadata management* as well as handling issues of *privacy and security*. A condensed version of the framework is illustrated in Figure 2. Besides depicting the functionality within a data marketplace it also shows which functionality is not part of the marketplace. This involves *data governance and management* topics, as these concern the management as opposed to exchange of data, as well as all topics which follow the acquisition of data, such as *data preparation* as these are beyond the exchange of data. From our point of view the data marketplace is merely a broker which offers data and can provide a stepping-stone to data-related tasks through courses or by providing infrastructure. While most of the functionality listed is also required in other data marketplaces, we point out that specific aspects like the metadata management as well as privacy and security, may take on a broader scope in the enterprise data marketplace.

The role-based functionality is not necessarily specific to enterprise data marketplaces, yet also required therein. The consumer requires *discovery* features such as a search function and detailed description of the offerings. They also need access to *collaborational* features to, e.g., rate or comment on data. Both the consumer and provider need *data trading* features. These includes features like service access management, e.g., to request or provide access to data, or subscription and order management through which consumers can manage their acquired data and services, and the provider can manage the running subscriptions on their offered data and services. The provider also requires features for *service publishing*, such as a service registration, e.g., for registering data in the marketplace, or data import features for uploading data. *Governance* features are required for the provider to retain data sovereignty and offer the data compliantly. Administration requires features to manage users and offerings in the marketplace.

In contrast, the metadata management functionality is distinctive in enterprise data marketplaces. Data marketplaces are metadata-driven platforms, therefore the handling of metadata is a central aspect within these. It includes building a data *catalog* with a inventory of data and services offered, the collection of *metadata specific to these datasets* such as descriptions, quality metrics, the data model etc., and also storing *marketplace-specific metadata* on the marketplaces internal processes like the purchase and search history therein. Companies already have infrastructure that collects and manages a wide variety of metadata, for instance, with tools such as data catalogs or business glossaries [6]. In the company, the marketplace thus has significantly more metadata at its disposal. Furthermore, the enterprise data marketplace can be tailored to reflect enterprise idiosyncrasies. For example, companies often have a company internal “language”, i.e., specific vocabulary, which is maintained through tools like business glossaries. By way of example, a company may refer to an end product as “material”. Yet normally the term “material” refers to a product’s elements. In an enterprise data marketplace this vocabulary can be incorporated in the description of the dataset. In this sense, the enterprise data marketplace is more flexible than other data marketplaces, which cannot, for instance, support a “customized” language across various companies.

Like metadata management, privacy and security aspects are especially relevant in the enterprise data marketplace. While security aspects like ensuring data confidentiality, integrity, availability, and authenticity are relevant in all types of marketplaces, the issue of ensuring data privacy is more challenging in the enterprise data marketplace. For instance, selected datasets are traded across companies, rarely including personal data, and if so regulations such as the general data protection regulation (GDPR) require the consent of the data subject for this exchange [8]. In contrast, the enterprise data marketplace’s data includes most of the personal data in the company, which was collected and approved for certain purposes. Therefore, the marketplace has to ensure that it is used and shared for these purposes only. That is, some parties may access the entire datasets, other parties may access an anonymized version of the data, and some may not be allowed to know that this data exist. Therefore, issues of remaining compliant with legal regulations like GDPR may be more challenging and significant in the enterprise data marketplace. Due to limited space, metadata aspects will be discussed in more detail in the following, whereas security and privacy aspects will not be the focus in this paper and are subject to future work.

3.3 Enterprise Integration Requirements

Unlike marketplaces for trading data between organizations, which are usually stand-alone marketplaces, an enterprise data marketplace can tightly integrate with a company’s system landscape and incorporate existent functionality, data and metadata. In this sense, we present the following set of integration requirements.

To begin with, it should *integrate with existing data management and storage systems*. This may include operational systems like ERP systems as well as analytical systems like data warehouses or data lakes. The ability to reference data in various data management systems is not per se specific to an enterprise data marketplace.

An enterprise data marketplace should, however, be able to reflect peculiarities of such a system or reflect data in a customized way according to the source system. For instance, it could reflect a data lakes customized zone architecture such as [11] and reference the data accordingly throughout the zones.

As mentioned previously, there are a variety of metadata management tools that are used to manage data and the understanding thereof within a company. These tools include data catalogs, business glossaries, and model repositories. Some of these tools provide functionality which is required in a marketplace. The data catalog, for example, contains a data inventory, which is also required within a data marketplace. The business glossary and other tools contain metadata which is relevant for finding, understanding and consequently choosing data for use. This information can be reused within a marketplace. Therefore, the enterprise data marketplace should tightly *integrate with the existent metadata management tool landscape*, build on existing functionality and incorporate the existing relevant metadata.

There are also administrative systems in companies such as identity management systems for managing company employees, or systems that deal with the corresponding employee rights. By *integrating with administrative tools* single sign-on and authorization management across sources systems, including the enterprise data marketplace is possible. The marketplace can then also access existent information in the user profiles such as an employee's clearance level and reuse this, e.g., to filter appropriate search results.

As the relevance of all the individual integration requirements are the same for the data marketplace and enterprise data marketplace, these have been consolidated in Table 1.

4 PLATFORM ARCHITECTURE

Marketplace architectures presented in literature thus far provide various perspectives on required components and the component-interactions. These include architectures that illustrate how marketplaces can be implemented with blockchain [31, 26], architectures that position the marketplace in IoT ecosystems [28], or architectures that focus on matching supply and demand through a so called data market management system [9]. So far, the presented architectures have not considered the special features and requirements of an enterprise data marketplace.

Therefore, we present a platform architecture that reflects the components of an enterprise data marketplace, displayed in Figure 3. Components that are potentially distinctive in the enterprise data marketplace, e.g., in regard to implementation aspects, are highlighted in grey. How this marketplace platform integrates into the existent system landscape and how the components interact therein is discussed in the following Section 5.

The architecture distinguishes *frontend* and *backend* components. The frontend is responsible for offering functionality to the marketplace participants and the backend for implementing this functionality through a variety of *services*. The frontend and backend components communicate via REST through an *API Gateway*. In addition, there are storage components for metadata and data. Components labeled as tools or platforms may already exist as standalone solutions within an enterprise. This is a unique characteristic within the enterprise and can be exploited by tightly

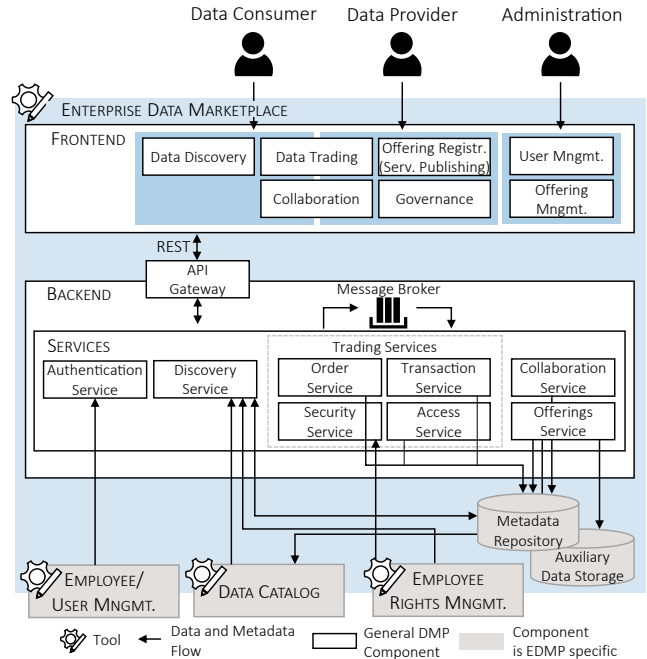


Figure 3: Enterprise Data Marketplace Architecture Featuring a Component Overview.

integrating the enterprise data marketplace with the existent solutions as specified in Section 3.3. The architecture also indicates which services depend on which of these potentially outsourced components. Alternatively, the features of these components can also be implemented within the according backend services yielding a self-sufficient data marketplace.

4.1 Frontend

The marketplace functionality is available to the roles, data consumer, data provider, and administrator in the frontend through a graphical user interface as well as an API. It includes the functionality as described in Section 3.2 and as listed in the functionality framework [4]. Namely, this is *data discovery*, *data trading*, and *collaboration* functionality for the data consumer, and complementary, *offerings registration* and *governance* functionality for the data provider, as well as *user* and *offerings management* for the administrators. Since the functionality from the cross-sectional areas, i.e., metadata management and privacy, security and compliance, is not directly accessible to users, it is not represented in the frontend. These are addressed indirectly throughout the backend services.

4.2 Backend

The backend provides a variety of services according to the functionality offered through the frontend. The services partially build on each other and communicate via a *message broker*. There are services for *authentication*, *discovery*, *order*, *security*, *transaction*, *access*, *offerings* and *collaboration* functionality. The authentication service is responsible for managing user access to the marketplace and in this sense handles the registration and login. Search functionality

together with a detailed view on offerings is provided through the discovery service. To facilitate trading, several services are required. The creation, monitoring and management of orders and subscriptions is handled through the order service. The security service deals with permission and provision approvals for the orders. This entails topics such as the verification whether a user has appropriate access rights for data with a higher security class. If any form of monetary transaction is called for, this is dealt with by the transaction service, and the access service is accountable for creating and managing access methods such as data base access, or access-links to data. The offerings service is responsible for the registration of any kind of service as described in Section 3.1, i.e., data, courses etc. It adds the data offerings to the data catalog which maintains a data inventory, with according metadata relevant for finding and understanding data and stores additional metadata which is not associated with the catalog, e.g., metadata for accessing the offerings, in the metadata repository. Lastly, the collaboration service takes care of any form of interaction on the offerings such as comments, use-case-documentations or ratings.

4.3 Enterprise Data Marketplace Specific Components

The components highlighted in gray in Figure 3 are required in all marketplaces, but can be specifically adapted to the enterprise setting, and are therefore termed as enterprise data marketplace specific components. For instance, the components marked as tools can be implemented as part of the marketplace, producing a stand-alone solution which could be used in an external context. These components can, however, already exist within an enterprise setting, and could therefore also be reused and integrated in the marketplace.

The component *employee/user management* is responsible for the identity management and authentication of users, meaning, enterprise employees and invited guests that have access to the enterprise data marketplace. Essentially this is the user database. In terms of the data democratization goal, getting access through, e.g., a user account, should be easy and attainable for the employees. As mentioned previously, companies usually have tools to manage information on their employees, such as Employee Database Software³ which offers a directory of employee profiles and functionality to structure and secure employee data including personal information, qualifications, skills and so on. As the marketplace will require an extract of exactly this metadata, it can be built on such an existing tool instead of recording the same information twice.

Closely related is the component *employee rights management*, which handles the users authorization, meaning rights, e.g., for various tools and platforms and potentially specific actions therein. Through it users can apply for, attain and manage these rights. Like before, there are tools for this on the market that are already used within the enterprises such as Access Rights Manager⁴ and could be integrated into the marketplace.

A *data catalog* such as Alation⁵ is a tool for maintaining a data inventory and amongst others, offers discovery, administration and

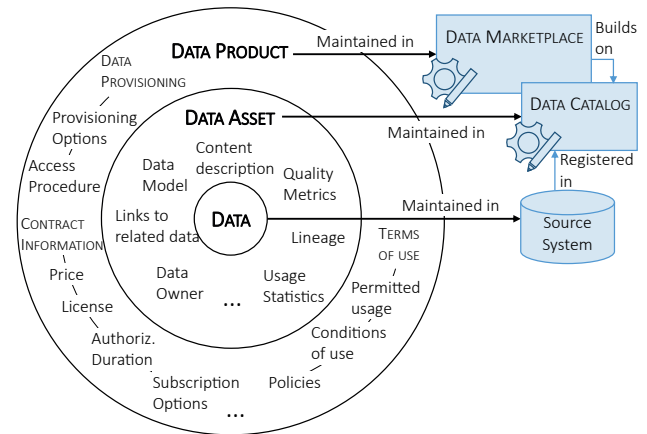


Figure 4: The Distinction of Data Assets and Data Products, Adapted from [4].

governance functionality [19, 39]. Within the marketplace this inventory would reflect the offered data and services like courses, with according metadata like a content description, the owner, who may access and use it and so on. This inventory can be maintained as part of the marketplace’s metadata repository, or could be maintained within an external tool. As companies are in the process of building and maintaining data catalogs [6] the stored information could be reused within the marketplace as opposed to doubling the inventory with collected metadata and functionality. As a marketplace requires more metadata for data trading than is normally collected within data catalogs this requires a distinction of data which is registered in the data catalog, i.e., a data asset, and data which is explicitly registered in the marketplace, i.e., a data product [4]. This distinction is illustrated in Figure 4. The product merely provides an extended set of metadata to the asset that explicitly enables the exchange of data, such as information on the license, price, or access options. As discussed in [4] this differentiation of data assets and products and the integration of an existent catalog supports and relieves the data provider, who has the potentially laborious task of making data known and providing provisioning options.

The *metadata repository* stores the metadata which is relevant for operating the data marketplace. As data marketplaces are metadata driven platforms [12] this is an essential component. What metadata is maintained in the enterprise data marketplaces varies depending on whether the above mentioned tools are integrated in the marketplace, or if it is implemented as a stand-alone solution. Besides metadata for cataloging the offerings, user information and access rights, the metadata repository may store metadata on, e.g., the order process, the purchase history, transaction history, or search history.

As explained in Section 2.2 an enterprise data marketplace may have a hybrid architecture with both a centralized and decentralized data storage. Most of the offered data should be referenced in the according storage systems, in order to support the scope of most enterprise data, and is therefore, part of the decentralized storage. However, if there is no storage system that can be referenced for

³www.scnsoft.com/software-development/databases/employee

⁴www.solarwinds.com/de/access-rights-manager

⁵www.alation.com

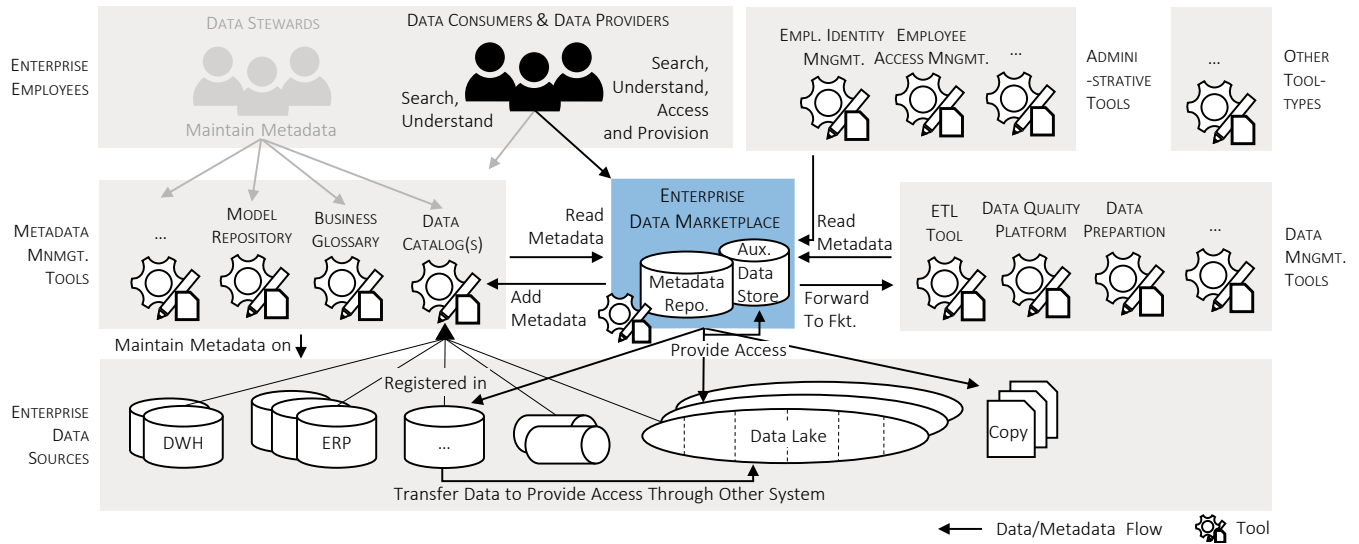


Figure 5: An Illustration of how the Enterprise Data Marketplace Integrates with a Company’s Existing System Landscape, i.e., Data Sources and Tools.

certain data, there is the option of loading the data directly into the integrated *auxiliary data storage* of the marketplace. This data storage may be omitted if such data can be loaded into and provided through an external system like a data lake.

The extent to which the marketplace distinctive components constitute an independent tool or have to be implemented in the marketplace also depends on the existing system and tool landscape in the company which we discuss in the following section.

5 ENTERPRISE INTEGRATION

In this section, we explain how the enterprise data marketplace can integrate into a company’s existent system and tool landscape, as depicted in Figure 5, and how this integration can be advantageous. This is distinctive for the enterprise data marketplace, as stand-alone marketplaces, for instance, for trading data between companies are usually not connected with the various data management systems within the participating companies. For one, this would be challenging for reasons of data security and privacy, but also in view of the fact that the participating organizations have a wide variety of system landscapes that the marketplace would have to be able to reflect. The typical integration scenarios are derived from our previous work in [6, 5, 4].

Only a few architectures presented in literature consider the marketplace in the context of a company’s internal system or tool landscape. Gröger [12] presents the core elements of a data ecosystem with an enterprise data marketplace, yet states that implementation and integration aspects are yet to be investigated. Wells [36] roughly highlights which technologies are needed within the marketplace components, i.e., data lake management, data pipeline management, data catalogs and data preparation. How the marketplaces interact with existing tools that implement these technologies is not discussed. Therefore, we address this topic in this section.

5.1 Integration with Data Sources

To begin with, we would like to illustrate how the marketplace will be integrated with or reference data within the enterprise source systems. This does not concern the integration of data, but the exchange between the marketplace and these systems. As can be seen in Figure 5, a wide variety of data sources, such as operational systems, e.g., ERP systems, and analytical systems, such as data lakes, are registered in a data catalog, as currently set up and maintained in many companies [6]. The marketplace references these systems via the data catalog. As discussed previously, only data that cannot be referenced in any external system is loaded and stored in the marketplace. If data cannot be provided in the source systems, there is also the option that these are transferred into another system such as a data lake. The marketplace can then grant access to this new system.

5.2 Integration with Tools

As stated previously, many companies have a variety of tools that provide functionality which is partly required in the marketplace. This includes functionality in tools for managing data and metadata, or administrative tools. Figure 5 indicates how the marketplace interacts with these tool groups.

As the enterprise data marketplace is a metadata-driven tool [12] most of its functionality is based on metadata. An example of this is the data inventory, which consists of metadata listing available datasets with information such as the storage location. Apart from the auxiliary data store, the marketplace does not interact with the actual data, only with the according metadata. As can be seen in Figure 5, metadata is collected and maintained in the company through *metadata management tools* such as data catalogs, business glossaries, for defining business terms and term relations [13], model repositories with semantic data models which are integrated

with the business glossary [6] and so on. These metadata are relevant in the selection process of a dataset. As described in [5], the distribution of metadata across a wide range of tools is a challenge for data consumers in the process of finding relevant data. For this reason, the marketplace requests the metadata from these tools and provides it in an integrated view. This is a read-only process on these tools. The data catalog is an exception in this context. Since an inventory of data records is already maintained in the catalog, the marketplace builds on this inventory, i.e. when new data is registered in the marketplace, it creates an entry in the existing data catalog for the new dataset, and thus performs a write operation. Although the marketplace extracts metadata from these tools, it is important to note that the metadata will continue to be maintained by the employees within the respective tools. The exception being the data catalog, which metadata is maintained through both the marketplace and catalog. Therefore, the introduction of the marketplace does not change the entire metadata management workflow and the marketplace does not need to provide the functionality of all these different tools. Also, while a consumer can find an integrated version of the metadata in the marketplace, it is still possible to view this metadata in the individual tools.

There are also *data management tools* that collect metadata. These include for instance, ETL tools that can reflect data lineage, or quality management platforms that amongst other things collect quality metrics such as a datasets completeness. As with the other tools, the marketplace can extract metadata from these tools and provide it in the integrated view if these are of interest in the data selection process. Furthermore, as explained in Section 3.2, the enterprise data marketplace is a broker for data between consumers and producers, and does not provide functionality for processing data. It can however, provide the data within an instance of such a tool, e.g., in Tableau, or transfer the consumer to tools with required functionality like data preparation after data acquisition.

In addition to the data and metadata management tools, the marketplace is integrated with *administrative tools* for, e.g., identity management. Thereby, employees only need to acquire the rights to access the marketplace, and the marketplace can then extract employee information from these tools. Based on the extracted information it can for instance, display only those records that match the employee’s clearance level.

5.3 Enterprise Integration Advantages

Integrating the enterprise data marketplace in the enterprise system and tool landscape has several advantages. For one, *existent functionality is reused*. By building on the existent tools, the marketplace does not double functionality such as rights management which also avoids the marketplace becoming a jack of all trades monolithic application. Also, there is a *comprehensive view on metadata*. If metadata collected throughout various tools is displayed in an integrated view in the marketplace this provides holistic information on the data. It is, however, important to note that integrating the marketplace with metadata management tools, as well as the integration of the metadata itself is a complex topic which elicits a variety of challenges including the classic data integration problems. Another advantage of integrating the marketplace in the enterprise is a *reduction in metadata management effort and errors*.

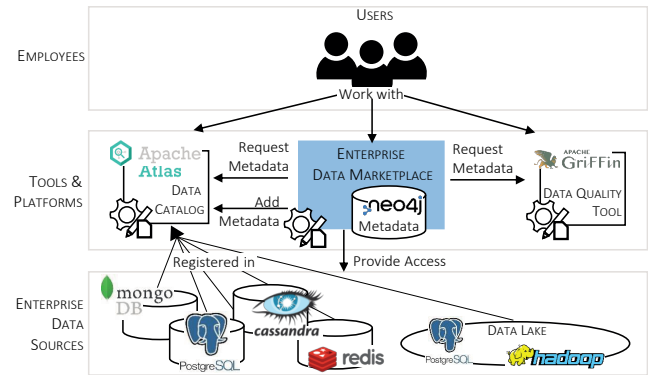


Figure 6: An Overview of the Tools Used to Implement the Enterprise Data Marketplace Prototype.

By reusing metadata already collected within other tools, there is no additional effort for maintaining a redundant set of metadata in the marketplace. This reduces the workload of the data providers that only have to maintain the metadata in one system and is also less error-prone. More information on this can be found in [4]. Finally, there is *less redundant data*. The same is true for the data, when referencing data within the data sources as opposed to uploading the data redundantly in the marketplace, there is less effort on behalf of the providers, reduced storage-cost, no synchronization-efforts and so on.

6 PROTOTYPICAL DEMONSTRATION

To evaluate the presented marketplace concepts, validate their feasibility and further examine the idiosyncrasies of marketplaces used within enterprises, we implemented an enterprise data marketplace prototype. It is a work in progress, which represents a large part, but not the full scope of the concepts presented above. An overview of the prototype is presented in Section 6.1 and Section 6.2 demonstrates how three typical data marketplace scenarios can be realized, namely adding data, searching for, and then ordering this data.

6.1 Prototype Overview

We based the choice of tools for the prototype on non-commercial and open-source tools because we want to enable free usability and customization. As depicted in Figure 6, a source system landscape is represented by a variety of database types and a data lake. The databases include the document store MongoDB⁶, the object-relational database PostgreSQL⁷, the columnar database Cassandra⁸ and the key-value database Redis⁹. These databases contain a variety of structured, semi- and unstructured sample datasets. In order to explore how a marketplace can reflect the characteristics of specific system types, we have also implemented a data lake. It is realized as a conglomeration of storage systems, including the Hadoop Distributed File System (HDFS)¹⁰ and PostgreSQL, and is

⁶www.mongodb.com

⁷www.postgresql.org

⁸www.cassandra.apache.org

⁹www.redis.io

¹⁰hadoop.apache.org

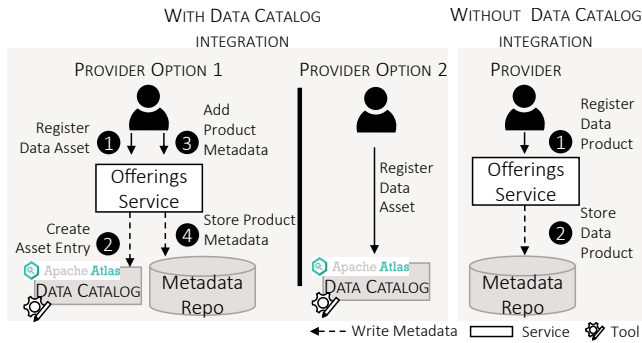


Figure 7: Data Registration Process and Possible Implementation Variants with or without a Data Catalog as well as Process Variants Concerning the User Workflow.

based on the data lake zone model by Giebler et al. [11]. Apache Airflow¹¹, a workflow management tool, is used to coordinate processes for moving the data into the appropriate zones based on three exemplary use cases.

The data sources are registered in the open source data catalog apache Atlas¹². Amongst others, it provides governance and metadata management functionality for building a catalog of data assets. Besides classic metadata such as a content description, our Atlas instance also reflects system specific metadata such as the mapping of data assets to data lake zones. Next to the data catalog we introduced Apache GriffIn¹³ into our tool landscape. It is a data quality solution which can measure data quality metrics such as the completeness, accuracy or timeliness of datasets. GriffIn tracks quality metrics on a selection of datasets in our source system landscape.

The enterprise data marketplace itself is implemented with the Spring framework¹⁴ based on a micro services architecture including an authentication, discovery, order, security, access and offerings service. The services communicate via the message broker RabbitMQ¹⁵. Marketplace specific metadata is stored in a Neo4J¹⁶ graph database and the metadata is modeled according to our metadata model HANDLE [7].

6.2 Usage Scenario Demonstration

Based on three standard scenarios in data marketplaces, derived from our previous work in [5, 4], we demonstrate how the marketplace components and enterprise tools interact with each other. In this regard, we present the scenarios of registering data in the company in Section 6.2.1, how this data can be searched for and found in Section 6.2.2 and finally ordered in Section 6.2.3. Individual steps of these scenarios are exemplified with screenshots of the prototype.

6.2.1 Scenario 1 - Registering Data: In order for a data marketplace to become effective it needs to have an assortment of offerings. The registration process is different depending on whether the

¹¹www.airflow.apache.org
¹²www.atlas.apache.org
¹³www.GriffIn.apache.org
¹⁴www.spring.io
¹⁵www.rabbitmq.com
¹⁶www.neo4j.com

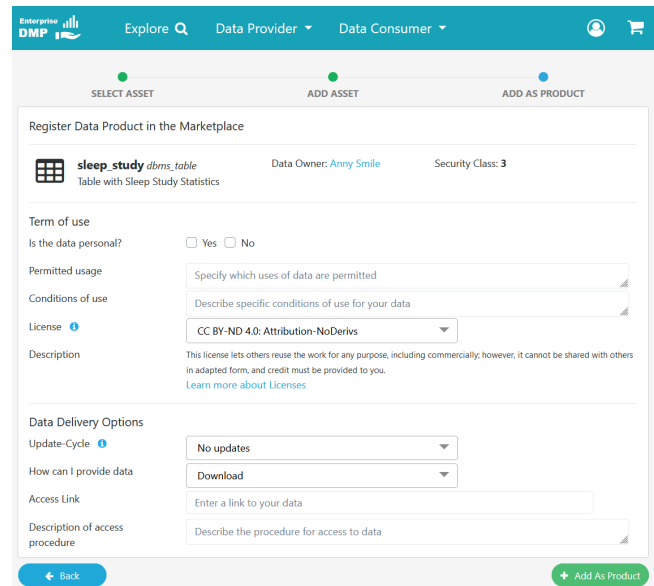


Figure 8: Prototype - Data Product Registration Wizard.

marketplace integrates with an existent data catalog or not, as illustrated in Figure 7. With a data catalog, a data provider has two options. They can register data through the marketplace as shown in Figure 7 on the left hand side. They enter asset metadata, meaning, descriptive metadata relevant for understanding the data through a form in step 1. The offerings service then creates an according entry for the data asset in the data catalog, i.e., Atlas, in step 2. At this point the provider can stop as this dataset can be found in the marketplace by potential consumers. Yet, this data is missing product metadata relevant for acquiring it [4]. This could be for instance, the allowed usage, a license, price or subscription or provisioning options. In our prototype this product metadata is added through the product registration wizard as shown in Figure 8. This constitutes step 3. The product metadata is specific to data trading and thus stored in the marketplace by the offerings service in step 4. At this point the data is ready to be ordered and provisioned to consumers.

Alternatively, the provider can register data directly in the catalog. This is illustrated as provider option 2 on the left side of Figure 7. Figure 9 depicts what a registration form for adding data to a data catalog, here Atlas, can look like. As the marketplace is integrated with the catalog this entry can be found in the marketplace, yet, once the data is requested by a consumer, the provider will be prompted to add the product metadata through the marketplace, continuing option 1 at step 3.

If the marketplace does not build on an existent data catalog there is no distinction between data assets and data products and the provider only has the option of registering the data through the marketplace, step 1, which will store it in the marketplace's inventory in the metadata repository, step 2, as shown in Figure 7 on the right.

6.2.2 Scenario 2 - Searching for Data: After registering data in the marketplace it can be found, as displayed in Figure 10. The consumer enters a request into the frontend search in step 1. Based

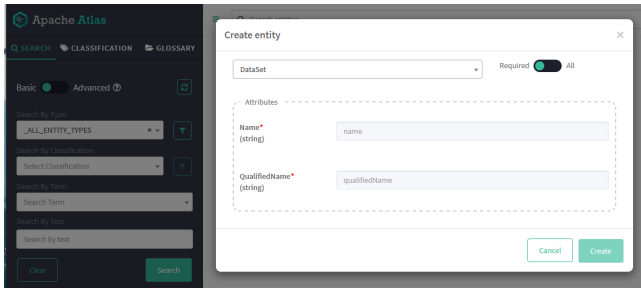


Figure 9: Prototype - Data Asset Registration in Atlas.

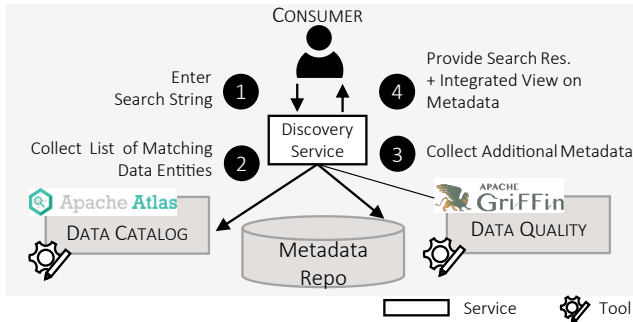


Figure 10: Search Process for Data with Involved Tools and Components.

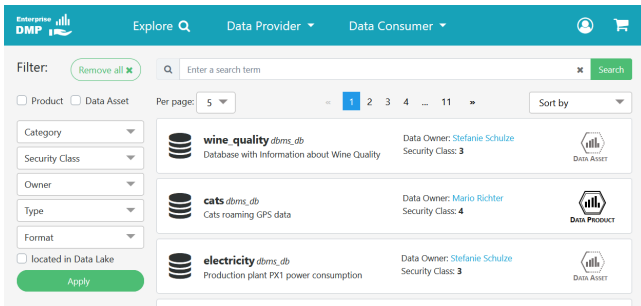


Figure 11: Prototype - Search Results View.

on the search string the discovery service collects entries from the data catalog, in our case Atlas, in step 2. Then it collects additional metadata such as product metadata from the metadata store and according metadata from other tools such as quality metadata from GriffFin in step 3. A list of search results as shown in Figure 11 is returned to the consumer in step 4. The single results can be expanded to provide a detailed integrated view on all the collected metadata.

6.2.3 *Scenario 3 - Ordering Data:* In the marketplace’s detailed-view-page on data the consumer can add the data to a shopping cart and order it to gain access which is illustrated as step 1. To issue the order the consumer also specifies the intended usage and chooses the provisioning option. Figure 12 demonstrates how once the order is submitted the order service checks if the chosen

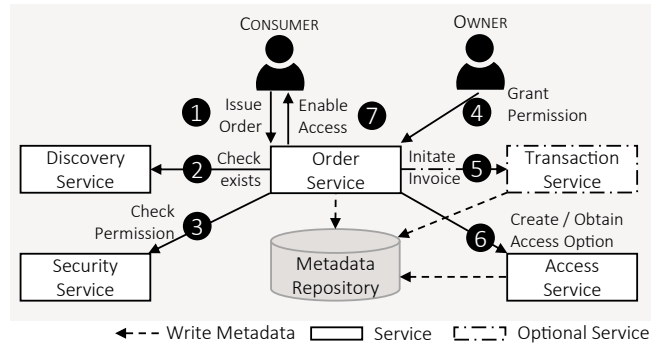


Figure 12: Data Order Process with Involved Components and Their Interaction Patterns.

dataset is valid through the discovery service in step 2. After this has been verified the order service transfers the request to the security service in step 3 though which the consumer’s permission to access this dataset is checked. For example, this includes a check if the consumer has an adequate clearance level for the dataset’s specified security class. If all is adequate, the order services notifies the data owner that they have a new access request. The owner can then grant or deny permission in step 4 based on, e.g., the specified usage information. If monetization is involved the order service initiates the transaction process in step 5 through which an invoice is sent to the consumer. When the transactions are completed the order service forwards the request to the access service as part of step 6. The access service deals with data provisioning options, for example, depending on the chosen and available provisioning options, the access service could create and store an access link which is then forwarded to the consumer in step 7 through the order service. The consumer can now access the ordered data.

Based on the three scenarios and the prototype, we have demonstrated how the platform architecture introduced in Section 4 can be implemented, how the components interact and how different processes unfold in the marketplace as well as how the marketplace can be integrated with external tools like Atlas and GriffFin.

7 CHALLENGES IN THE ENTERPRISE DATA MARKETPLACE

Based on an in-depth knowledge exchange with our industrial partner investigating enterprise data marketplace (see Section 3) and by implementing such a prototype we discovered technical and organizational challenges surrounding this type of marketplace.

One challenge is the *lack of incentives* for providers to share the data. By omitting monetization the main incentive for data providers to share data is removed. Initially, the provider has effort that is not compensated. Consequently, other forms of incentives are required for providing data in the enterprise data marketplace. Other researchers have suggested bonus points as an incentive [9]. Possibly, gamification in the marketplace, publicity through visibility, coupons for coffee or the awarding of data sharing titles might be interesting ideas for incentivization. This topic calls for an investigation what drives providers and their executives to promote data sharing or what would prevent them from doing so. With this

knowledge possible impediments could be removed. For example, we have discussed in [4] how the marketplace can support the provider in data sharing by minimizing the sharing effort.

How data ownership can be retained or passed to acquiring consumers constitutes another challenge. If data is acquired, processed and then made available again as a product, who will be the data owner? It might be the original data owner, who must then keep track of all processed versions of their data. Yet, how far they can feasibly track their data down the chain of processed versions of this data is questionable. Also, there may be a situation in which several datasets from different owners are processed together, providing the question, which of the data owners would be responsible in this case. Alternatively, the person processing the data or their supervisor may be the new data owner. In any case, there must be a data owner for every processed data instance to ensure compliance to legal regulations.

Furthermore, *preventing the flooding of the marketplace with unusable data* also represents a challenge. Two scenarios emanate from the data democratization goal of publishing as much data for as many users as possible. For one, there is the targeted provision of data for known use cases. In this case the relevance of the data, as well as the processing state in which this data is required are known. In the second scenario, data is provisioned without knowing if it is relevant for other participants and in what form they would need it. This bears the risk of flooding the marketplace with data that nobody needs or that is unusable for further processing. This challenge is closely related to the topic of providing data in a way that increases the consumer's utility, by considering the consumers needs, as addressed in, e.g., [9].

Finally, *integrating the enterprise data marketplace into the existing system landscape* can be challenging. Different tools support varying metadata exchange standards which must be supported or an alternative standardization for the inclusion of metadata must be provided. In addition, the marketplace must be able to display metadata dynamically, since the tools may provide a variety of different metadata per dataset. Moreover, while implementing the marketplace prototype, the inadequate documentation of the tools was an issue. In this context, the implementation of the enterprise data marketplace is more complex than those that function as stand-alone marketplaces.

These challenges are particularly pronounced in the enterprise context. Nevertheless, challenges identified for other marketplaces may also apply to the enterprise data marketplace.

8 RELATED WORK

The enterprise data marketplace is addressed in only a few research articles. Amongst others, Gröger [12] highlights the need for this specific marketplace type, Fernandez et al. [9] consider them to bring down data silos and Wells [37] defines and presents the enterprise data marketplace in a report. We also discuss the necessity and various aspects of enterprise data marketplaces in our previous research [6, 5, 4]. None of these articles, however, clearly highlight the specifics and differences to external marketplaces. Therefore, we close this gap by placing the marketplace in a marketplace classification framework and by providing the specific requirements, architecture and challenges.

There are several research articles that provide classification frameworks for data marketplaces. These include Schomm et al. [29] who provide an initial set of dimensions and Stahl et al. [33] that extend these. Meisel and Spiekermann [24] derive five classification characteristics and Spiekermann [32] provides economic and technological characteristics of marketplaces. Täuscher and Laudien [35] list key business model attributes of marketplaces, which are however not exclusive to data marketplaces and Fruhwirth et al. [10] provide a list of characteristics that are assigned to dimensions such as value capture, delivery, proposition and creation. So far, the enterprise data marketplace has not yet been classified based on any of these frameworks, hence we provide this placement in such a framework.

Requirements for data marketplaces are listed in a range of research articles. Fernandez et al. [9] introduce requirements concerning topics such as the ability to price datasets or the ability to support markets of different types like internal and external markets. Sometimes the requirements are tailored to a specific context such as trustworthiness through, e.g., blockchain [21] or marketplaces in the a IoT context [28, 18]. While requirements are often listed in a specific context such as IoT many still apply to data marketplaces in general, for example, requirements concerning scalability or security [2]. Requirements for the enterprise data marketplace could be derived from this marketplace types descriptions as supplied in, e.g., [12, 37, 5], and general requirements also partly apply to the enterprise data marketplace. It has, however, not been clarified which explicit requirements the enterprise data marketplace has and how these overlap with those of other marketplaces.

In terms of marketplace architectures, there is a variety of architectural proposals, most of which are, however, tailored to a specific context, similar to the requirements. There are marketplace architectures specific to the use of blockchain [27, 31], the IoT context [18, 2, 28], multilateral marketplace design [17], elements in decentralized marketplaces [26], personal data valuation [15], or also specific marketplace aspects like a market management system or mashup builder [9]. None of these architectures reflect the specific components of the enterprise data marketplace. In contrast Wells [37] does provide a component overview for the enterprise data marketplace, nonetheless, it is not apparent which aspects are special to the internal setting or also how the components interact. In terms of how the marketplace is embedded in a company's existing system landscape, Gröger [12] places the marketplace in the data ecosystem of an industrial enterprises and Wells [37] gives an overview of required technologies in the marketplace, both do not however explain how the marketplace and its components interacts with the other systems.

Lastly, challenges within data marketplaces are discussed in many research articles. Amongst others, these include challenges concerning the valuation and pricing of data [25, 23, 1, 29], the derivation and assurance of data quality [16, 9], the ability to combine datasets to satisfy buyers' needs and establishing trust amongst the participants [9], or issues of data procurement [40]. We highlight challenges the consumers and providers currently face when accessing and providing data within the enterprise in our previous work [5, 4], yet these are role-specific. Hence, as none of the research articles discuss challenges specific to the enterprise data marketplace we address this gap in this paper.

9 CONCLUSION

Enterprise data marketplaces for exchanging data within companies are becoming increasingly relevant as they support data democratization and consequently contribute to extracting more of a company's potential data value. In this paper we have established that the enterprise data marketplace is a distinct type of marketplace with specific characteristics. This was clarified by placing the enterprise data marketplace in a classification framework and by highlighting a set of requirements which are specific to the enterprise data marketplace. By presenting a platform architecture, discussing how this platform integrates with existent enterprise system landscapes and demonstrating these concepts through a prototype, we laid the foundations for the development of an enterprise data marketplace. The discussion has revealed that there are still a number of challenges to be addressed when using a marketplace in the enterprise internal context. One of the challenges involves integrating the enterprise data marketplace into the existing system landscape. In future, we intend to address this challenge, especially with regard to the topic of metadata management, by investigating how metadata from a variety of tools can be displayed in an integrated view in the enterprise data marketplace. Furthermore, a detailed examination how privacy and security aspects are handled in the enterprise data marketplace is also subject to future work.

REFERENCES

- [1] Anish Agarwal, Munther Dahleh, and Tuhin Sarkar. 2019. A Marketplace for Data: An Algorithmic Solution. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC '19)*, 701–726.
- [2] Ahmed Salim Alrawahi, Kevin Lee, and Ahmad Lotfi. 2020. AMACoT: A Marketplace Architecture for Trading Cloud of Things Resources. *IEEE Internet of Things Journal*, 7, 3, 2483–2495.
- [3] Pranjal Awasthi and Jordana George. 2020. A case for Data Democratization. In *Proceedings of the 26th Americas Conference on Information Systems (AMCIS '20)*, 23:1–23:10.
- [4] Rebecca Eichler, Christoph Gröger, Eva Hoos, and Holger Schwarz. 2022. From Data Asset to Data Product – The Role of the Data Provider in the Enterprise Data Marketplace. In *Proceedings of the 16th Symposium and Summer School On Service-Oriented Computing (SummerSoc '22)*, 1–20.
- [5] Rebecca Eichler et al. 2022. Data Shopping – How an Enterprise Data Marketplace supports Data Democratization in Companies. In *Proceedings of the 34th Conference on Advanced Information Systems Engineering (CAISE '22)*, 19–26.
- [6] Rebecca Eichler et al. 2021. Enterprise-Wide Metadata Management: An Industry Case on the Current State and Challenges. In *Proceedings of the 24th International Conference on Business Information Systems (BIS '21)*, 269–279.
- [7] Rebecca Eichler et al. 2021. Modeling Metadata in Data Lakes – A Generic Model. *Data & Knowledge Engineering*, 136, C, 101931.
- [8] European Parliament and Council of the European Union. 2016. Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (Data Protection Directive). Legislative Acts L119. Official Journal of the European Union, (Apr. 27, 2016).
- [9] Raul Castro Fernandez, Pranav Subramaniam, and Michael J. Franklin. 2020. Data Market Platforms: Trading Data Assets to Solve Data Problems. *Proceedings of the VLDB Endowment*, 13, 12, 1933–1947.
- [10] Michael Fruhwirth, Michael Rachinger, and Emina Prlja. 2020. Discovering business models of data marketplaces. In *Proceedings of the 53rd Hawaii International Conference on System Sciences (HICSS '20)*, 5738–5747.
- [11] Corinna Giebler et al. 2020. A Zone Reference Model for Enterprise-Grade Data Lake Management. In *Proceedings of the 24th IEEE Enterprise Computing Conference (EDOC '20)*, 57–66.
- [12] Christoph Gröger. 2021. There is No AI without Data. *Communications of the ACM*, 64, 11, 98–108.
- [13] Deborah Henderson, Susan Earley, and Laura Sebastian-Coleman, eds. 2017. *DAMA-DMBOK: Data Management Body of Knowledge*. Technics Publications, Basking Ridge, NJ, USA.
- [14] Marijn Janssen, Yannis Charalabidis, and Anneke Zuiderwijk. 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29, 4, 258–268.
- [15] Yonas Mitike Kassa et al. 2016. Your Data in the Eyes of the Beholders: Design of a Unified Data Valuation Portal to Estimate Value of Personal Information from Market Perspective. In *Proceedings of the 2016 11th International Conference on Availability, Reliability and Security (ARES '16)*, 701–705.
- [16] Pantelis Koutroumpis, Aija Leiponen, and Llewellyn D. W. Thomas. 2020. Markets for data. *Industrial and Corporate Change*, 29, 3, 645–660.
- [17] Pantelis Koutroumpis, Aija Leiponen, and Llewellyn D. W. Thomas. 2017. The (Unfulfilled) Potential of Data Marketplaces. ETLA Working Papers 53. The Research Institute of the Finnish Economy (ETLA), (Sept. 29, 2017).
- [18] Bhaskar Krishnamachari, Jerry Power, Seon Ho Kim, and Cyrus Shahabi. 2018. I3: An IoT marketplace for smart communities. In *Proceedings of the 16th ACM International Conference on Mobile Systems, Applications, and Services (MobiSys '18)*, 498–499.
- [19] Clement Labadie, Christine Legner, Markus Eurich, and Martin Fadler. 2020. FAIR Enough? Enhancing the Usage of Enterprise Data with Data Catalogs. In *Proceedings of the IEEE 22nd Conference on Business Informatics (CBI '20)*, 201–210.
- [20] Juliane Lange, Florian Stahl, and Gottfried Vossen. 2018. Datenmarktplätze in verschiedenen Forschungsdisziplinen: Eine übersicht. *Informatik-Spektrum*, 41, 170–180.
- [21] Sebastian Lawrenz, Priyanka Sharma, and Andreas Rausch. 2019. Blockchain Technology as an Approach for Data Marketplaces. In *Proceedings of the 2019 International Conference on Blockchain Technology (ICBCT '19)*, 55–59.
- [22] Hippolyte Lefebvre, Christine Legner, and Martin Fadler. 2021. Data democratization : toward a deeper understanding. In *Proceedings of the 2021 International Conference on Information Systems (ICIS '21)*, 2106.
- [23] Jinfei Liu et al. 2021. Dealer: An End-to-End Model Marketplace with Differential Privacy. *Proceedings of the VLDB Endowment*, 14, 6, 957–969.
- [24] Lukas Meisel and Markus Spiekermann. 2019. Datenmarktplätze – Plattformen für Datenaustausch und Datenmonetarisierung in der Data Economy. ISST-Bericht. Fraunhofer ISST, (Feb. 18, 2019).
- [25] Jian Pei. 2020. A Survey on Data Pricing: from Economics to Data Science. *IEEE Transactions on Knowledge and Data Engineering*, 1–1.
- [26] Gowri Sankar Ramachandran, Rahul Radhakrishnan, and Bhaskar Krishnamachari. 2018. Towards a Decentralized Data Marketplace for Smart Cities. In *Proceedings of the 2018 IEEE International Smart Cities Conference (ISC2 '19)*, 1–8.
- [27] Dumitru Roman and Gatti Stefano. 2016. Towards a reference architecture for trusted data marketplaces: The credit scoring perspective. In *Proceedings of the 2nd International Conference on Open and Big Data (OBD '16)*, 95–101.
- [28] Stefan Schmid et al. 2017. An architecture for interoperable IoT Ecosystems. In *Proceedings of the 2nd International Workshop on Interoperability and Open-Source Solutions for the Internet of Things (InterOSS-IOI '17)*, 39–55.
- [29] Fabian Schomm, Florian Stahl, and Gottfried Vossen. 2013. Marketplaces for data: An initial survey. *ACM SIGMOD Record*, 42, 1, 15–26.
- [30] Seagate Technology. 2020. Rethink Data – Bessere Nutzung von mehr Unternehmensdaten – vom Netzwerkrand bis hin zur Cloud. Report. Seagate.
- [31] Priyanka Sharma, Sebastian Lawrenz, and Andreas Rausch. 2020. Towards Trustworthy and Independent Data Marketplaces. In *Proceedings of the 2020 2nd International Conference on Blockchain Technology (ICBCT '20)*, 39–45.
- [32] Markus Spiekermann. 2019. Data Marketplaces: Trends and Monetisation of Data Goods. *Intereconomics*, 54, 208–216.
- [33] Florian Stahl, Fabian Schomm, Lara Vomfell, and Gottfried Vossen. 2017. Marketplaces for Digital Data: Quo Vadis? *Computer and Information Science*, 10, 4, 22–37.
- [34] Florian Stahl, Fabian Schomm, Gottfried Vossen, and Lara Vomfell. 2016. A classification framework for data marketplaces. *Vietnam Journal of Computer Science*, 3, 137–143.
- [35] Karl Täuscher and Sven M. Laudien. 2018. Understanding platform business models: A mixed methods study of marketplaces. *European Management Journal*, 36, 3, 319–329.
- [36] Dave Wells. 2018. Dynamic Data Marketplace: Fast Data for Fast Business. Report. Eckerson Group, (Mar. 2018).
- [37] Dave Wells. 2017. The Rise of the Data Marketplace: Data as a Service. Report. Eckerson Group, (Mar. 2017).
- [38] Mark D. Wilkinson et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018.
- [39] Ehtisham Zaidi, Guido De Simoni, Roxane Edjlali, and Alan D Duncan. 2017. Data Catalogs Are the New Black in Data Management and Analytics. Gartner Research. Gartner, (Dec. 13, 2017).
- [40] Zhenzhe Zheng, Weichao Mao, Fan Wu, and Guihai Chen. 2019. Challenges and opportunities in IoT data markets. In *Proceedings of the 4th International Workshop on Social Sensing (SocialSense '19)*, 1–2.