2024

Unlocking Value: Towards a Comprehensive Taxonomy for Commercializing Open Data in the Business Context

Eric Börner
TU Dresden, Germany, eric.boerner1@tu-dresden.de

Unlocking Value: Towards a Comprehensive Taxonomy for Commercializing Open Data in the Business Context

Research Paper

Eric Börner[1]

[1] TU Dresden, Chair of Business Information Systems,
esp. Information Systems in Industry and Trade, Dresden, Germany
eric.boerner1@tu-dresden.de

Abstract. The increasing availability of open data promises transparency and economic growth. Although several studies have highlighted the value of open data for businesses, the private sector has been slow to adopt it, falling short of expectations. This is due to a lack of knowledge about the real benefits of using open data as a business and few use cases for commercial reuse of open data. Against this backdrop, our study aims to shed light on the commercial use of open data by developing a comprehensive taxonomy consisting of 14 dimensions and deriving five commercial open data user archetypes. Our study contributes to the literature by providing a nuanced understanding of the commercial use of open data, as well as a conceptual framework that structures successful open data use cases, which can be used by businesses and practitioners to identify opportunities for potential open data use scenarios in their respective organizations.

Keywords: Open Data, Use Cases, Taxonomy Development

1 Introduction

Data is essential in our knowledge-driven world, driving both societal and economic shifts. Its value increases with reuse (Hossain et al., 2016), making sharing and opening datasets crucial to unlocking its full potential. In this context, the concept of open data has become increasingly popular in recent years. An expanding quantity of data is being made available as open data to promote transparency and stimulate both new businesses and economic growth (Janssen and Zuiderwijk, 2014). The open data movement originated with the release of public data as open data by governments with the goal of enhancing governmental transparency, fostering citizen participation and collaboration, and facilitating economic growth (Janssen and Zuiderwijk, 2014; Davies et al., 2019). Open data refers to "data that is freely available, and can be used as well as republished by everyone without restrictions from copyright or patents" (Braunschweig et al., 2012, p. 1). Institutional policies and pressures have encouraged both public and private entities to provide access to their information (Hossain et al., 2016; Rößler, 2022). Open Data has the potential to improve governance and democratic processes, empower citizens, adress public problems and create economic opportunities (Ahmadi Zeleti and

19th International Conference on Wirtschaftsinformatik,
September 2024, Würzburg, Germany

Kerins, 2023). In terms of economic potential, open data has been the subject of high expectations, being seen as a "gold mine" of business opportunities (The Economist, 2013). According to Manyika et al. (2013), the economic value of open data ranges from $3.2 to $5.4 trillion annually worldwide. Open data has the potential to enhance growth and prosperity across various industries, including transportation, consumer goods, electricity, finance, agriculture, urban development, and the social sector, as evidenced by prior research (Davies et al., 2019; Gruen et al., 2014; Manyika et al., 2013). Despite the widespread enthusiasm for the vast potential of open data and the increasing availability of open datasets, the private sector has lagged in its adoption of open data, and the use of open data by businesses is below expectations (Krasikov et al., 2020; Zuiderwijk et al., 2015). While research on the expected benefits of open data provides initial insights, little is known about its potential benefits in the private sector (Enders et al., 2021). There is a lack of understanding regarding the precise variables, parameters, and pathways through which open data translates into growth and opportunities (Verhulst and Caplan, 2015) and the literature lacks a specific perspective on the use of open data in a business context (Krasikov et al., 2021). Furthermore, there is a shortage of open data use cases that can serve as a foundation for purposeful and precise consumption of open data (Krasikov et al., 2019).

This study seeks to examine how economic value can be created through open data in a business environment by viewing this issue from the perspective of enterprises. The research objective is to create a taxonomy for the commercial use of open data by businesses. This will synthesize the relevant knowledge on this topic and expand comprehension of commercial open data reuse in the private sector by improving the still insufficient understanding of open data value creation. The developed taxonomy can assist researchers and practitioners in objectively assessing key facets of the potential business applications of open data. The taxonomy provides various perspectives on the commercial reuse of open data and contains major dimensions that determine the possible use scenarios of open data for businesses. The taxonomy was developed following the approach specified by Nickerson et al. (2013). Throughout the process of four iterations, diverse sources such as research literature, whitepapers, and use cases were examined. Based on the developed taxonomy, we derived five archetypes of commercial open data users. This paper is organized as follows. The second section introduces open data and value creation. In the following two sections, we describe our methodology and the results. In the fifth section, we discuss the contributions and limitations.

2 Theoretical Background

2.1 Open Data

Open data is a widely used term and can be defined as "data that is freely available, and can be used as well as republished by everyone without restrictions from copyright or patents" (Braunschweig et al., 2012, p. 1). Although there are many definitions of open data, there is no universal approach to determine whether data is open or not. According to Verhulst and Caplan (2015), some authors view open data as part of a spectrum of

openness. In this regard, Lindman et al. (2013) define openness as including technical openness of data (e.g., through the use of interfaces and standards), legal openness (including copyright and licensing), and commercial openness (allowing for commercial use of the data). Open data can be provided by both government and private organizations, resulting in the subsets of Open Government Data (OGD) and Open Private Data (Berends et al., 2017). Open government data can be defined as "data collected and held by a government and its agencies, shared and made freely available to anyone in order to pursue open government objectives" (Marmier and Mettler, 2020, p. 3). Not all published data can be considered 'pure' open data, as they do not always meet strict criteria, such as the Sebastopol Open Government Data Principles (Verhulst and Caplan, 2015). Shared corporate data, i.e. data voluntarily published by companies that were previously internal to the company, is considered by many authors to be a form of open data, although it is often subject to a number of restrictions and is therefore not open data according to a strict interpretation (Verhulst and Caplan, 2015). This work considers open data to be content that is freely accessible, in a machine-readable format, freely usable without licensing restrictions, and free of charge.

2.2 Value Creation with Open Data

The literature proposes various concepts to comprehend how value can be created from data or its use. One of the most prominent is the data value chain, which explains how raw data is transformed into actionable information. Open Data Watch (2018) has proposed a data value chain with a focus on open data, which is described as follows. The data value chain shows the process that data goes through, from identification of data needs to the impact of data use, as shown in Figure 1. There are four main stages in the data value chain: collection, publication, uptake, and impact, which are further broken down into sub-stages. First, in the *collection* phase, it is determined what data will be collected and how it will be used before the data is collected, for example, through surveys or by retrieving or using sensors. The data is then processed to ensure its correctness and stored in formats that allow further use. In the *publication* phase, the data must be published in a way that the intended users can access it. This includes analyzing the data to extract useful information and visualize it, releasing the data and associated metadata to make it accessible online and offline, and disseminating it to potential users through appropriate channels such as information intermediaries. The *uptake* phase comprises the substages connect, incentivize, and influence. In this phase, connections are established with end users, and the data can be reprocessed to gain new insights. Users are encouraged to utilize the data in their decision-making process and are influenced to value the data. In the final phase (*impact*), the data is used to understand problems or make decisions, change outcomes or improve situations, and finally, the data can be reused by combining it with other data. ("The Data Value Chain," 2018). The focus of the taxonomy developed within this paper is mainly on the impact phase.

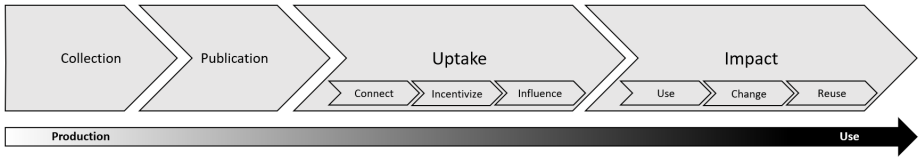

Figure 1. Data Value Chain, adapted from Open Data Watch (2018)

3 Research Method: Taxonomy Development

The objective of this work is to develop a taxonomy that describes different perspectives on the commercial (re)use of open data, due to the absence of suitable frameworks for its commercial use. Taxonomies comprising pertinent dimensions and corresponding attributes can assist researchers in organizing knowledge (Kundisch et al., 2021) and can be employed for ex post theory building (Bapna et al., 2004). For a rigorous classification of commercial open data use scenarios, we chose the approach proposed by Nickerson et al. (2013) due to its widespread acceptance as a de facto standard in IS research (Kundisch et al., 2021). Following Nickerson et al. (2013), we first specified the meta-characteristics and the ending conditions for the taxonomy development process. As ending conditions, we applied the objective and subjective ending conditions proposed by Nickerson et al. (2013). Next, we developed the Taxonomy with four iterations. During the iterative development process of the taxonomy, new dimensions were added when they were deemed suitable for distinguishing different commercial open data use scenarios. These dimensions were integrated into the taxonomy when they did not overlap with existing dimensions. Dimensions were excluded if they were not suitable for categorizing commercial open data use scenarios because they were too generic or could not cover the entire spectrum of use scenarios. The taxonomy development process is illustrated by Figure 2. The four employed iterations are described in the following.

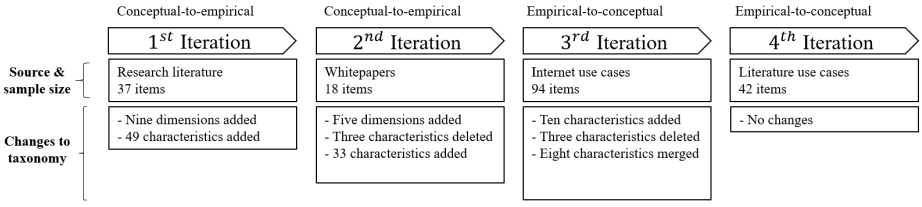

Figure 2. Summary of the taxonomy development process

First Iteration (Research Literature): In the initial iteration, we employed a conceptual-to-empirical approach, conducting a systematic literature review following the guidelines of Webster and Watson (2002). This approach enabled us to synthesize relevant information on the commercial use of open data in the literature and extract key concepts for the taxonomy. In accordance with the recommendations set forth by Webster and Watson (2002), a four-step process was employed, commencing with an initial search in the 51 key information systems journals on Scopus via litbaskets.io. This was

followed by an expanded search on popular databases, including AIS eLibrary, ScienceDirect, and EBSCO. A backward search and a forward search were then conducted. For the database searches, the query "Open Data AND Use OR Re-use OR Reuse" was used. Following the literature extraction, the abstracts and full texts of the identified studies were carefully read to identify patterns of commercial use of open data. The review sample was limited to articles written in English or German and clearly related to the use or consumption of open data by private organizations. Publications that focused on open data disclosure, open data policy, and the use of open data by governments, nonprofits, and citizens were also excluded. After excluding ineligible articles, the final sample contained 37 items. For the literature synthesis, the identified articles were scanned for concepts related to commercial open data reuse. After this scanning, nine concepts relevant to the taxonomy were identified. The complete literature sample list can be found in section A of the online Appendix[1].

Second Iteration (Whitepapers): In the second iteration of the taxonomy development process, grey literature, especially whitepapers from companies and public organizations, were extracted and analyzed. The literature was identified through online and forward-backward searches based on the research literature already identified in the first iteration. The selection and analysis were performed similarly to that of the research literature in the first iteration. The executive summaries and full texts of identified publications were carefully examined and inappropriate papers were removed. From an initial 25 publications identified, seven were removed due to exclusion criteria, resulting in a final sample size of 18 whitepapers. The identified publications were scanned for relevant concepts related to the commercial reuse of open data. Additional six concepts relevant to the taxonomy were identified in this phase. Section B of the online Appendix[1] contains a table with all the analyzed whitepapers.

Third Iteration (Internet Use Cases): For the third iteration, we chose an empirical-to-conceptual approach by examining use cases. Several online sources were identified that contain open data use cases related to the commercial use of open data. For use case identification and extraction, the sources used include European Data Portal, Open Data 500 Project, Open Data Impact Map, Open data means business by ODI.

The open data use cases were limited to scenarios from private, profit-seeking organizations with an active online presence. In addition, there had to be a clear indication of open data use and an accurate description of how open data was being used. The information from the mentioned sources and the respective companies' websites were analyzed and compiled into an Excel sheet. In total, 94 cases from online sources were included in the sample. All cases extracted were grouped according to the existing dimensions of the taxonomy. During this phase of grouping and categorizing cases, there arises a need for new characteristics, while some of the existing ones are proving inadequate.

Fourth Iteration (Literature Use Cases): In the fourth iteration, we shifted the focus to validating our taxonomy using an empirical-to-conceptual approach and applying the taxonomy to a number of use cases extracted from literature sources. As in the previous

[1] https://figshare.com/s/fc52e7f2468d479dc58e

iteration, the use cases were limited to scenarios from private, profit-oriented companies, with a clear description of the actual usage of open data. Various whitepapers and research papers were used to extract a total of 42 open data use cases. The extracted cases were categorized based on the existing dimensions of the taxonomy. During this phase, there was no need for adaptation or change of the existing dimensions and characteristics of the taxonomy. Therefore, the taxonomy development process concludes with this iteration, as no changes were made and both the subjective and objective ending conditions were met. Section C of the online Appendix[1] contains the sample of all use cases examined in the last two iterations.

Identification of Archetypes: To obtain appropriate archetypes for commercial users of open data, we applied our developed taxonomy to all previously identified open data use cases. We categorized a total of 136 use cases based on the dimensions and related characteristics of our taxonomy. This categorization allowed us to compare the frequencies of each dimension's characteristics with those of every other dimension in the taxonomy to identify common relationships. Through a meticulous comparison of the frequencies of characteristics, we identified several characteristic patterns that resulted in the creation of five distinct archetypes.

4 Results

4.1 Taxonomy for commercial Open Data use

This section presents the taxonomy with its 14 dimensions and related characteristics, as depicted in Figure 3.

Role of data in value proposition	Final good itself	Key ingredient of product/service	Marginal ingredient of product/service	No ingredient of product/service				
Process of data use	Data-to-fact	Data-to-data	Data-to-information	Data-to-interface	Data-to-service			
Purpose of data use	Improving internal processes						Product/service development	
	Data management			Intelligence & analytics	Benchmarking	Supply chain management	Risk management	Enhancement of existing products/services	Creation of new products/services
	Enrichment	Validation	Cleansing						
Data source	Government				Corporates	Academic institutions		
	International institutions	National government	Regional government	Municipal government				
Type of data	Geodata	Environmental data	Economic & business data	Traffic & transport data	Agricultural data	Legal data	Demographic data	Science data
Role of enterprise in value chain	Enabler	Intermediary					Service provider	
		Aggregator			Developer	Enricher		
Role of data in enterprise	Bread and butter for the organization	Bread and butter for a product/service	Addition to existing product/service	Addition for organization				
Type of data elaboration	Aggregation & mash-up	Structuring & classification	Geo-referencing	Validation	Analytics			
Products & services	Tangible products	Services			Digital products			
		Consulting services	Other services	Aggregated information	Open Data repositories	Service platforms	Search- and comparison platforms	
Price mechanism	Premium	Freemium	Free					
Customer segment	Business-to-business (B2B)	Business-to-consumer (B2C)	Business-to-government (B2G)					
Stage of enterprise	Entrepreneurial	Established						
Sector of enterprise	Agriculture	Extractives	Transport	Infrastructure	Consumer products	Finance & insurance	IT & technology	Business services	Health & pharmaceuticals
Business impact	Innovation		Harmonized data	Improved processes	Improved decision making	Reduced costs		
	New products/services	Enhanced products/services						

Figure 3. Taxonomy for commercial open data use

Role of Data in Value Proposition: Based on Ferro and Osella (2012, 2013), the concept of open data in the value proposition is an important categorization for distinguishing between different types of open data use in the enterprise context. It describes the extent to which open data is a component of marketable products or services. In our taxonomy, we distinguish between the characteristics “final good itself”, “key ingredient of product/service”, “marginal ingredient of product/service” and “no ingredient of product/service”. This categorization can be interpreted as a continuum ranging from maximal to minimal contribution of open data to the enterprise offering. For instance, in the case of the online portal *TrendEconomy*, which provides aggregated open macroeconomic data with visualizations, open data can be considered as "final good itself". If companies use open data solely for internal purposes, such as insurance companies incorporating open data to enhance their risk models, the open data is considered as "no ingredient of product/service".

Process of Data Use: In the realm of open data exploitation processes, various types can be differentiated. Davies (2010) initially introduced a distinction between data-to-fact, data-to-data, data-to-information, data-to-interface, and data-to-service. Data-to-fact refers to the extraction of factual information from data for further use, such as in bureaucratic processes (Davies, 2010). Possible applications of the data-to-fact approach include advocacy and consultancy (Magalhaes and Roseira, 2020). Data-to-data involves manipulating data to create new datasets that are more suitable for use by other users (Davies, 2010; Magalhaes and Roseira, 2020). This process is useful for refining and structuring data. Data-to-information covers creating visualizations, infographics, and reports to represent and interpret data (Davies, 2010). Data-to-interface refers to creating interactive access to data through interfaces, resulting in data platforms and open data portals (Davies, 2010; Magalhaes and Roseira, 2020). Open data portals are repositories for open data, while data platforms are more interactive and allow users to combine open data with other types of data for deeper analysis (Magalhaes and Roseira, 2020). When data is used as input for services, it is referred to as data-to-service (Kaasenbrood et al., 2015). This can involve integrating data into existing services or products, or using data to create new services (Davies, 2010).

Purpose of Data Use: The purpose of the data use dimension is to describe the specific purpose for which open data is being utilized. We distinguish between improving internal processes, which includes data management, intelligence and analytics, benchmarking, supply chain management, risk management, and product/service development, which involves enhancing existing products/services and creating new ones. According to Krasikov et al. (2019), data management can be distinguished into three usage categories: enrichment, validation, and cleansing.

Data Source: The data source dimension describes the origin of the open data according to the type of data provider. In our taxonomy, we differentiate between open data provided by governments and non-governmental actors such as corporations and academic institutions. In terms of Open Governmental Data, providers can be categorized into international institutions, such as the European Union, and national, regional, and municipal governments.

Type of Data: The type of data dimension distinguishes between different types of open data. The eight most common types of open data with the highest potential for commercial use have been included: geospatial data, environmental data, economic and

business data, traffic and transportation data, agricultural data, legal data, demographic data, and scientific data. As indicated by Carrara et al. (2015), the most valuable types of information for commercial reuse of open data are geographical, environmental, and economic and business information. Geospatial data encompasses a range of resources, including cadastral and topographic maps, aerial photos, administrative boundaries, and hydrographical maps. Open geospatial data is a commonly utilized resource by insurance companies, as evidenced by the British company *Aviva*'s incorporation of geospatial data into its modeling of floods and other property-related issues, which subsequently informs its pricing models (Ordnance Survey, 2020). Economic and business data includes price data, economic trends, trade data, growth projections, industry profiles, market shares, and other pertinent information. This data can be utilized by companies across a diverse array of domains.

Role of Enterprise in Value Chain: This dimension describes where the organization is positioned in the data value chain. Our taxonomy includes five roles: enabler, aggregator, developer, enricher, and service provider. According to Jaakkola et al. (2014), aggregators, developers, and enrichers can be grouped as open data intermediaries. Enablers facilitate the supply or use of open data, including storage and hosting, retrieval, categorization, and advisory services (Ferro and Osella, 2013; Jaakkola et al., 2014). Aggregators collect and combine multiple open data sources and process them to produce insights that are useful to their users (Jaakkola et al., 2014; Janssen and Zuiderwijk, 2014). Developers design and build applications around open data (Immonen et al., 2014; Jaakkola et al., 2014; Janssen and Zuiderwijk, 2014; Hjalmarsson et al., 2015). Enrichers utilize open data to enable or enhance existing products/services or improve internal processes (Jaakkola et al., 2014; Janssen and Zuiderwijk, 2014).

Role of Data in the Enterprise: This dimension emphasises the importance of open data to the organization, as described by Ferro and Osella (2012, 2013) and Kaasenbrood et al. (2015). If open data is crucial to the organization's operations to the extent that it could not exist without it, it is referred to as the “bread and butter for the organization”. Open data can be a crucial component in certain products or services, which is commonly referred to as the “bread and butter for a product or service”. Furthermore, open data can be an addition to preexisting products/services or to the organization to strengthen internal capabilities and processes.

Type of Data Elaboration: The type of data elaboration is the process or method used to manipulate or analyze the data. Based on Ferro and Osella (2012, 2013), the most important techniques in context of commercial open data reuse, such as data aggregation & mash-up, structuring & classification, geo-referencing, validation, and analysis, have been included in our taxonomy.

Products & Services: The product & service dimension encompasses the offerings marketed by the organization. Our taxonomy distinguishes between tangible products, services, and digital products. In terms of services, we differentiate between consulting services and other services. Digital products include aggregated information, open data repositories, service platforms, and search and comparison platforms.

Price Mechanism: The price mechanism describes how monetary profit is generated by offering products and services. Our taxonomy includes the three price mechanisms proposed by Ferro and Osella (2013): premium, freemium, and free.

Customer Segment: The Customer segment refers to the group of individuals or organizations that the company intends to target with its products or services. Our taxonomy includes the business-to-business (B2B), business-to-consumer (B2C), and business-to-government (B2G) segments.
Stage of Enterprise: We added a dimension to our taxonomy, named "stage of enterprise", to assess the enterprise's maturity level. We differentiated between entrepreneurial and established enterprises.
Sector of Enterprise: The sector of the enterprise describes the industry in which the enterprise operates. Our taxonomy includes the nine most important sectors for commercial open data reuse: agriculture, extractives, transport, infrastructure, consumer products, finance & insurance, IT & technology, business services, and health & pharmaceuticals. In our sample of commercial open data use cases, business services, IT & technology, and finance & insurance were the most prevalent sectors of companies using open data.
Business Impact: The dimension of business impact describes the immediate effect of utilizing open data on an organization's business. In our taxonomy, we included innovation (including new or enhanced products/services), harmonized data, improved processes, improved decision-making, and reduced costs. As indicated by Berends et al. (2020), the most significant benefit of utilizing open data for businesses is innovation, specifically the introduction of new products and services. This is followed by reduced costs and data harmonization.

4.2 Commercial Open Data User Archetypes

Using our proposed taxonomy, we have identified five common archetypes of commercial open data users. These archetypes are based on open data use cases and represent the most frequent combinations of characteristics from the introduced taxonomy. Although most of them relate to one of the identified roles within the value chain, it is the combination of characteristics that distinguishes the archetypes from the identified roles of the enterprise in the value chain.
Established Enrichers: The established Enricher archetype is a frequently encountered user of open data that enhances its current business model by integrating open datasets to enhance internal processes such as data management. Most established enrichers are large, mature companies in data-intensive industries such as finance and insurance or real estate. Insurance companies, in particular, utilize open data for risk management and improving their predictive models. For example, the lending company *OnDeck* uses open financial data to evaluate the financial health of businesses (Verhulst and Caplan, 2015). Big retailers often use open data to improve store operations, such as identifying optimal store locations and optimizing store assortments using demographic, weather, and consumer data. For example, supermarket *Tesco* used open weather data and sales records to create accurate demand models, reducing overstocking and spoilage (Manyika et al., 2013). The primary distinction between the established enricher and other archetypes is that the established enrichers exclusively utilize open data to enhance internal issues rather than to improve products or services or create new ones. The offering of this open data archetype includes tangible products or

services like insurances, financial services, or distribution electricity. The business impact of open data use is reliant on the intended purpose of its utilization. If open data is utilized for Data Management, it is probable that the outcome will be harmonized data, resulting in a positive business impact. This archetype could potentially achieve harmonized data, improved processes, enhanced decision making, and cost reduction through open data usage.

Search & Comparison Platforms: This archetype utilizes open data to provide a portal or platform for consumers and businesses to find the optimal product, home, contractor, business partner or other based on their needs and preferences. There are various platforms available in different sectors such as finance & insurance and transportation. Some real estate search and comparison platforms, such as *Zillow*, provide property searches for home buyers and sellers or real estate agents, using open demographic data, price indexes, transaction data, etc. ("Open Data Impact Map," 2023). In the transportation sector, some apps and platforms help their customers find the optimal routes and connections using open data on routes and schedules. For these platforms, open data plays an important role, sometimes the platforms rely only on publicly available data. This archetype applies the data to interface and data to information approaches, as data is provided and made available to customers through an interface and may be transformed into valuable and usable information for the customer. The providers of search & comparison platforms can be considered as aggregators or developers, as they aggregate and combine different datasets into their platform and develop the actual platform itself with its inherent search and compare engines and algorithms. The data elaboration methods used by search & comparison platforms are aggregation & mash-up, structuring & classification and analytics. These platforms are mostly provided with a free or freemium price mechanism. The enterprises offering such platforms can be considered for the most parts as small, young entrepreneurial companies.

Aggregation Platforms: This archetype provides aggregated information on specific topics, such as financial reports from companies, patent data, and lawsuit data. Enterprises operating aggregation platforms create value by aggregating datasets from different sources, process them into homogenous, aggregated datasets and provide its customers access to those aggregated datasets via their platform. One example is *OpenCorporates*, which aggregates publicly available corporate data from a variety of sources into corporate and financial data in a standardized format (Davies et al., 2019). For this archetype, open data can be considered as a final good itself, processed and marketed on an online platform. Thus, the purpose of utilizing open data is to process it into a new product. The approach of data to interface and data to data is mostly used by aggregation platform providers. Businesses that provide aggregation platforms are considered aggregators in the data value chain. The data elaboration techniques utilized are aggregation & mash-up and structuring & classification to make the datasets easier to understand and discover by the users of the platform. Aggregated information can be considered as the offering of enterprises following this archteype.

Enablers: Enablers provide software and tools for open data analysis, construct open data portals and repositories, and offer assistance with open data publication and distribution. They primarily aid in the construction of repositories and provide guidance to government agencies or companies seeking to open up datasets. The enabler archetype

provides products and services, such as open data repositories and consulting services on open data disclosure. These offerings exclude the provision of open data itself. Unlike other archetypes, enablers do not use open data directly for their products. Instead, they work with open data to develop their offerings. Enablers do not elaborate open data for their own use, but provide data elaboration tools to their customers, such as data aggregation or analysis tools. Open data can be considered crucial for this archetype's organization due to its reliance on open data for their business model, even though they do not actually utilize open data in their offerings. An example of the enabler archetype is the U.S. company *3 Round Stones*, which makes a web data publishing platform used by U.S. government agencies to publish and consume data ("Open Data Impact Map," 2023).

Business Service Providers: Business service providers utilize open data in their products and services, which they offer to other businesses. They typically aggregate and analyze economic & business data or demographic data to provide useful insights to their customers. The business service provider archetype utilizes the data to service or data to information approach, converting open data into valuable information and insights for their clients. Therefore, open data is used for product/service development purposes, either for the enhancement of the existing offering or for the sake of new business services. The methods used to process this data include aggregation & mashup, analytics, and in some cases, geo-referencing or structuring & classification. The offering of the business service provider may include service platforms, consulting services, or other services. The relevant sectors in which this archetype operates are predominantly business services, with some overlap in finance, insurance, or IT & technology. *SLR Consulting* is an example of a business service provider that collects open data for research purposes and uses the insights derived from the open data in its advisory and environmental consulting services ("SLR Consulting | data.europa.eu," 2020).

5 Discussion and Conclusion

This study aims to identify the characteristics and determinants of open data usage for value creation in a private business setting. Therefore, we identified specific dimensions with respective characteristics across commercial open data use scenarios. For this purpose, we reviewed prior literature and analyzed a broad range of open data use cases from various sources. A taxonomy for creating value with open data in a business context was developed following the guidelines of Nickerson et al. (2013), and archetypes for commercial open data users were derived.

The study make several important contributions to the open data literature. Firstly, the study clarifies and describes the concept of value creation through open data by private for-profit organizations by deriving a taxonomy consisting of 14 key dimensions. We addressed the need for further research in this area, as suggested by Magalhaes and Roseira (2020). Our taxonomy expands on previous research in this area and offers a more comprehensive view of the commercial use of open data. It highlights a wide range of dimensions in which enterprises use open data to create value. Compared to previous studies on the commercial use of open data, our taxonomy offers a more de-

tailed and comprehensive understanding of the elements and relationships that characterize open data value creation in enterprises. Additionally, our taxonomy helps to create an broader understanding of value creation in the business context based on open data, which goes beyond the widespread focus on intermediaries and infomediaries. This includes scenarios with marginal open data utilization. In addition, our taxonomy includes general indicators and determinants that are not specific to open data. Including indicators such as the sector, customer segment, and enterprise stage is crucial for understanding the specific business environment and covering the complex topic of commercial open data utilization in a holistic way. The taxonomy offers a comprehensive and detailed insight into how organizations are incorporating open data into their operations and value chains covering a wide range of industries, business processes, and data sources.

Furthermore, the taxonomy is designed to categorize successful commercial use scenarios of open data, including the roles of open data reusers, sectors, and data types with high commercial potential for open data exploitation. Our taxonomy and the derived archetypes can be regarded as valuable solution space for practitioners and businesses developing open data use cases within their organizations and provide a new perspective on possible business models involving open data. Practitioners and managers can use our taxonomy and archetypes to discover the range of open data use scenarios and identify opportunities for open data exploitation.

The results and implications of our study should be interpreted with several limitations in mind. Firstly, open data use cases can vary in scale and focus, and it is possible for multiple open data use cases to exist in parallel within an organization. Secondly, the development of a taxonomy cannot be entirely objective since it involves interpretation and data classification based on the researcher's comprehension of conceptual frameworks. There may be other possible classifications, as there is no definitive or completely objective approach to this subject matter. Additionally, it is important to note that the development of taxonomies may not always include the full range of concepts or categories within a given field. This is due to limiting factors such as the need for clarity and usability, which can constrain the size of classifications. The application of the taxonomy depends on the specific situation, as not every dimension of the taxonomy is relevant in every case. As Nickerson et al. (2013) point out, a taxonomy does not need to be perfect, but it does need to be useful and meet the needs of its users.

Further research is needed due to the complexity of commercial reuse of open data and the still insufficient understanding of the business opportunities that open data offers. Detailed case studies are necessary to better understand the motivations, challenges, and success factors related to the use of commercial open data at the company level. Furthermore, additional research is necessary to quantify the economic impact of using open data on a enterprise-level. This will assist companies in accurately assessing the business potential of open data.

In conclusion, the developed taxonomy and proposed archetypes contribute to the existing open data literature by enhancing the understanding of open data use for commercial purposes and providing a conceptual basis for further research in this area. We hope that our study will stimulate additional work on the promising topic of commercial use of open data in a variety of contexts.

References

Ahmadi Zeleti, F. and Kerins, R. (2023) 'Open Data Impact Taxonomy – Analysis of Open Data Use Cases in Response to the Ireland's National Open Data Strategy 2017–2022', in K. Arai (ed.) Advances in Information and Communication. Cham: Springer Nature Switzerland (Lecture Notes in Networks and Systems), pp. 340–357.

Bapna, R., Goes, P., Gupta, A., & Jin, Y. (2004) 'User Heterogeneity and Its Impact on Electronic Auction Market Design: An Empirical Exploration', *MIS Quarterly*, 28(1), p. 21.

Berends, J., Carrara, W., Engbers, W., & Vollers, H. (2020) 'Reusing open data – A study on companies transforming open data into economic and societal value', *Publications Office of the European Union: Luxembourg*.

Berends, J., Carrara, W., & Radu, C. (2017) 'Analytical report 9: the economic benefits of open data', *Publications Office of the European Union: Luxembourg.*

Bonina, C. (2013) 'New business models and the value of open data: definitions, challenges and opportunities'.

Braunschweig, K., Eberius, J., Thiele, M., & Lehner, W. (2012) 'The State of Open Data Limits of Current Open Data Platforms'.

Carrara, W., Fischer, S., van Steenbergen, E., & San Chan, W. (2015) 'Creating value through open data – Study on the impact of re-use of public data resources', *Publications Office of the European Union: Luxembourg.*

Davies, T. (2010) 'Open data, democracy and public sector reform', *A look at open government data use from data. gov. uk*, pp. 1–47.

Davies, T., Walker, S.B., Rubinstein, M., & Perini, F. (2019) *The State of Open Data: Histories and Horizons*. African Minds and IDRC.

Enders, T., Benz, C., & Satzger, G. (2021) 'Untangling the open data value paradox: How organizations benefit from revealing data', in *Innovation Through Information Systems: Volume III: A Collection of Latest Research on Management Issues*. Springer, pp. 200–205.

Ferro, E., & Osella, M. (2012) 'Business models for PSI re-use: A multidimensional framework', in *Using Open data: Policy Modeling, Citizen Empowerment, Data Journalism Workshop, European Comission, Brussels*.

Ferro, E., & Osella, M. (2013) 'Eight Business Model Archetypes for PSI Re-Use', in *Open Data on the Web Workshop, Google Campus, Shoreditch, London*.

Gruen, N., Houghton, J., & Tooth, R. (2014) 'Open for business: How open data can help achieve the G20 growth target'.

Hjalmarsson, A., Johansson, N. and Rudmark, D. (2015) 'Mind the Gap: Exploring Stakeholders' Value with Open Data Assessment', in 2015 48th Hawaii International Conference on System Sciences. 2015 48th Hawaii International Conference on System Sciences (HICSS), HI, USA: IEEE, pp. 1314–1323.

Hossain, M.A., Dwivedi, Y.K., & Rana, N.P. (2016) 'State-of-the-art in open data research: Insights from existing literature and a research agenda', *Journal of Organizational Computing and Electronic Commerce*, 26(1–2), pp. 14–40.

Immonen, A., Palviainen, M. and Ovaska, E. (2014) 'Requirements of an Open Data Based Business Ecosystem', IEEE Access, 2, pp. 88–103.

Jaakkola, H., Mäkinen, T., & Eteläaho, A. (2014) 'Open data: opportunities and challenges', in *Proceedings of the 15th International Conference on Computer Systems and Technologies*. *CompSysTech'14: 15th International Conference on Computer Systems and Technologies*, Ruse Bulgaria: ACM, pp. 25–39.

Janssen, M., & Zuiderwijk, A. (2014) 'Infomediary Business Models for Connecting Open Data Providers and Users', *Social Science Computer Review*, 32(5), pp. 694–711. Available at: https://doi.org/10.1177/0894439314525902.

Kaasenbrood, M., Zuiderwijk, A., Janssen, M., de Jong, M., & Bharosa, N. (2015) 'Exploring the Factors Influencing the Adoption of Open Government Data by Private Organisations', *International Journal of Public Administration in the Digital Age*, 2(2), pp. 75–92.

Krasikov, P., Harbich, M., Legner, C., & Eurich, M. (2019) *Open Data Use Cases - Framework for the generation and documentation of open data use cases*.

Krasikov, P., Obrecht, T., Legner, C., & Eurich, M. (2020) 'Is Open Data Ready for Use by Enterprises? Learnings from Corporate Registers.', in *DATA*, pp. 109–120.

Krasikov, P., Obrecht, T., Legner, C., & Eurich, M. (2021) 'Open Data in the Enterprise Context: Assessing Open Corporate Data's Readiness for Use', in *Data Management Technologies and Applications: 9th International Conference, DATA 2020, Virtual Event, July 7–9, 2020, Revised Selected Papers 9*. Springer, pp. 80–100.

Kundisch, D., Muntermann, J., Oberländer, A. M., Rau, D., Röglinger, M., Schoormann, T., & Szopinski, D. (2021) 'An update for taxonomy designers: methodological guidance from information systems research', *Business & Information Systems Engineering*, pp. 1–19.

Lindman, J. (2014) 'Similarities of Open Data and Open Source: Impacts on Business', *Journal of theoretical and applied electronic commerce research*, 9(3), pp. 59–70.

Lindman, J., Rossi, M., & Tuunainen, V.K. (2013) 'Open Data Services: Research Agenda', in *2013 46th Hawaii International Conference on System Sciences*. *2013 46th Hawaii International Conference on System Sciences (HICSS)*, Wailea, HI, USA: IEEE, pp. 1239–1246.

Magalhaes, G., & Roseira, C. (2020) 'Open government data and the private sector: An empirical view on business models and value creation', *Government Information Quarterly*, 37(3), p. 101248.

Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., & Doshi, E. A. (2013) *Open Data: Unlocking Innovation and Performance with Liquid Information*. McKinsey Global Institute.

Marmier, A., & Mettler, T. (2020) 'Different shades of perception: How do public managers comprehend the re-use potential of open government data?'

Nickerson, R.C., Varshney, U., & Muntermann, J. (2013) 'A method for taxonomy development and its application in information systems', *European Journal of Information Systems*, 22(3), pp. 336–359.

Open Data Impact Map (n.d.), https://opendataimpactmap.org/map#DataView. Accessed: 14 December 2023.

Ordnance Survey (2020), https://www.ordnancesurvey.co.uk/customers/case-studies/aviva-geographic-intelligence-underwriting. Accessed: 15 December 2023.

Rößler, J. (2022) 'Sharing is Caring: Using Open Data To Improve Targeting Policies'.

SLR Consulting | data.europa.eu (2020), https://data.europa.eu/en/publications/use-cases/slr-consulting. Accessed: 15 December 2023.

'The Data Value Chain: Moving from Production to Impact' (2018), https://opendatawatch.com/publications/the-data-value-chain-moving-from-production-to-impact/. Accessed: 19 November 2023.

The Economist (2013): 'A new goldmine', *The Economist*, https://www.economist.com/business/2013/05/18/a-new-goldmine. Accessed: 17 November 2023.

Verhulst, S., & Caplan, R. (2015) 'Open Data: A Twenty-First-Century Asset for Small and Medium-Sized Enterprises', *SSRN Electronic Journal*

Webster, J., & Watson, R.T. (2002) 'Analyzing the past to prepare for the future: Writing a literature review', *MIS quarterly*, pp. xiii–xxiii.

Zuiderwijk, A., Janssen, M., Poulis, K., & van de Kaa, G. (2015) 'Open data for competitive advantage: insights from open data use by companies', in *Proceedings of the 16th Annual International Conference on Digital Government Research*. *dg.o 2015: 16th Annual International Digital Government Research Conference*, Phoenix Arizona: ACM, pp. 79–88.