What Characterizes Data Spaces in Industry 4.0? Towards a Better Understanding

Oliver Werth,[1[0000-0002-6767-5905]] Christian Koldewey,[2[0000-0001-7992-6399]] Mathias Uslar,[1[0000-0002-8451-1574]] and Julian Zerbin[2]

[1] OFFIS – Institute for Information Technology, Escherweg 2, 26121 Oldenburg, Germany
oliver.werth@offis.de; mathias.uslar@offis.de

[2] Advanced Systems Engineering, Heinz Nixdorf Institute, Paderborn University, Paderborn, Germany
christian.koldewey@hni.uni-paderborn.de; julian.zerbin@hni.uni-paderborn.de

Abstract. Data spaces are receiving an emerging interest in Information Systems Research and industry practice. They are central to many European research initiatives and shape the data economy in Industry 4.0. Generally, they aim to create secure environments for cross-organizational data management and sharing. Currently, there is considerable interest in developing new data spaces in Industry 4.0, also accelerated through regulatory changes. However, key questions about what precisely characterizes a data space in Industry 4.0 remain unresolved. Against this backdrop, we build a taxonomy of data spaces in the Industry 4.0 context. We identified nine distinctive dimensions and 40 corresponding characteristics among the 19 data spaces analyzed. The taxonomy enables clearer classification and nomenclature of data spaces in this context. This short paper will ignite planned further research on data spaces in Industry 4.0 and contribute to a conceptualization of a taxonomic theory for interested researchers.

Keywords: Industry 4.0, Taxonomy, Data Spaces, Characterization

1 Introduction and Problem Statement

In today's interconnected world, data and data ecosystems are essential enablers of digital innovation. This is particularly true for the industrial sector, where the concept of Industry 4.0 has emerged. Industry 4.0 is defined by the real-time synchronization of information flows and the ability to produce more personalized, smart products [1]. Effective data handling and exchange across organizational boundaries are central to this development, yet they present considerable challenges. Transparency, monitoring of indirect suppliers, regulations, and enhanced collaboration are all key requirements that organizations increasingly face [2, 3].

To address these challenges, industrial data ecosystems are emerging. Within them, data spaces facilitate secure, cross-organizational data sharing while preserving data sovereignty. According to Möller et al. [3], data spaces are inter-organizational information systems that implement both technical and organizational mechanisms to support trusted data exchange. They also play an intermediary role, helping companies collaborate without losing control of proprietary information. From a technical standpoint, data spaces are decentralized, do not rely on a shared database, are nested and overlapping, and support data redundancy. These characteristics align well with the distributed and even more data-driven nature of Industry 4.0. Furthermore, data spaces promote interoperability and standardization, allowing systems across organizations to integrate efficiently and operate seamlessly [4, 5]. Moreover, data spaces enable open innovation by allowing businesses to co-develop products, optimize processes, and respond quickly to market demands from an economic standpoint. Access to real-time, high-quality data allows for better strategic decision-making, enabling companies to refine production workflows, improve product quality, and gain a competitive advantage [6]. Despite their potential, data spaces also raise concerns, particularly around data security and privacy. These concerns can hinder trust among participating entities. Therefore, strong governance frameworks, technical standards, and access controls are critical to ensuring secure and ethical data sharing [4].

While the concept of data spaces has been increasingly studied in information systems (IS) research, the literature remains fragmented, and a comprehensive taxonomy specific to Industry 4.0 is lacking [3]. Prior studies by Gieß et al. [7, 8] have developed taxonomies for general data spaces and connectors, identifying relevant dimensions and characteristics. However, they do not focus explicitly on the Industry 4.0 context. On the other hand, existing taxonomies in Industry 4.0, such as those by Weking et al. [9] and Azkan et al. [10], focus on business models or data-driven services, but do not address data spaces. In addition, small to medium enterprises in Industry 4.0 refrain from participating in data spaces based on restricted knowledge capabilities and personnel resources despite the chances that data spaces deliver [11–13]. As a result, a clear, structured view of potential data space configurations and characterizations in Industry 4.0 remains missing but could be of interest for practitioners and researchers of data spaces in Industry 4.0. To date, such an overview is not present by previous literature. A taxonomy will reduce the complexity of the market, structure knowledge, and support both theoretical development and practical applications within a particular business sector [14, 15]. We pose the following research question (RQ): *What are common characteristics and dimensions among data spaces in an Industry 4.0 context?*

2 Methodology and Data Collection for this Short Paper

For this short paper, we make use of the structured approach to taxonomy development originally introduced by Nickerson et al. [14]. The procedure is suitable for sorting a specific field and carving out characterizations and similarities. As a first step, we determine the meta-characteristic of our taxonomy [15]. We define the meta-characteristic of our objects as “documented data spaces designed for an Industry 4.0 context from the perspective of its intended stakeholders.” In the second part, we define to follow six objective and all subjective ending conditions as proposed by Nickerson et al. [14]. We make use of three iterations, one conceptual-to-empirical (C2E) approach, and two empirical-to-conceptual (E2C) approaches. and started with a C2E approach to receive a preliminary taxonomic structure that served as a suitable ignition for orientation. For this, we make use of specific literature on data spaces and the Industry 4.0 context [7, 10]. Our first taxonomy contained data space generic and Industry 4.0 specific dimensions and characteristics, like “Underlying Framework“ or “Application Area“, without sorting any objects. To receive a meaningful list of data spaces in the Industry 4.0 context for the E2C iterations, we make use of two main data sources: (1) a publicly available dataset on use cases and data spaces in Industry 4.0, which was published within a third-party-funded project context in the Gaia-X universe [16] and (2) the Data Spaces Radar Version 4 from March 2024 [17]. We searched both data sources for data spaces with a clear context and purpose for Industry 4.0. As a result, we collected and received a set of 19 objects to classify. Within two consecutive E2C iterations, we sorted the identified data space into the taxonomic structure that we received from our first C2E iteration and refined it. We selected three data spaces (n=3) for the first E2C iteration to check the robustness of the preliminary structure derived from the first iteration. We continued the taxonomic process since it violates several ending conditions. Then, in the second E2C iteration, we sorted the remaining 16 (n=16) data spaces into the taxonomy. The E2C iterations were conducted in a manner that more dimensions and characteristics could be identified, merged, or added after screening more and more data spaces that were investigated. To receive meaningful information about the dimensions and characteristics, we read white papers, scientific papers, project homepages, and/or case study reports about the respective data spaces, and constantly discussed the results for appropriateness. After these three iterations, we identified that the taxonomy met all objective and subjective ending conditions as chosen before. Figure 1 depicted our research process described.

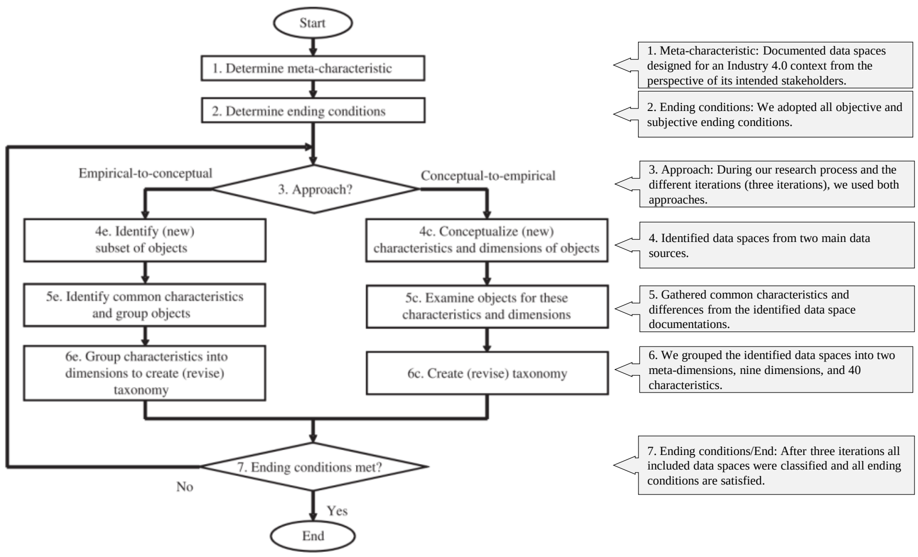

Fig. 1. Research Process

We developed an online appendix, where the taxonomy, a check of the subjective and objective ending conditions, definitions for each dimension, and data spaces are documented. It can be found here[1] to provide transparency for the reader since the iterations were briefly documented within this short paper, based on page restrictions. The final taxonomy is presented in Figure 2. Numbers in brackets indicate the number of occurrences identified in the 19 data spaces in the Industry 4.0 context. Bold characteristics represent the most observed occurrence among a dimension.

3 Results Observed from the Taxonomy and Discussions

To answer our RQ in this short paper, "What are common characteristics and dimensions among data spaces in an Industry 4.0 context?" we can make several observations from our taxonomy: First, we identified nine **dimensions** among 40 distinct *characteristics*. We sorted the nine dimensions into two meta-dimensions, namely, an economic and technical one. This delimitation allows interpreting the data spaces in different directions through their characteristics as identified.

Economic aspects. Generally, we identified five economic dimensions that characterize the analyzed data spaces in Industry 4.0. We found a clear tendency in *publicly-funded* Industry 4.0 data spaces, e.g., Manufacturing-X (n=18). Also, most of them have a mainly *economic* **value proposition**, e.g., CollMi (n=13). The **key purposes** of the data spaces are quite diverse, including *predictive maintenance, manufacturing optimization*, and *supply chain optimization*; however, we found a small tendency (n=5) for *multiple* key purposes, e.g., Boost 4.0. Only six data spaces (n=6) explicitly mention that they are currently *operational* in the real world (dimension: **real-world operation**) based on our observations, e.g., FLEX4RES, while some are purely *conceptual*, e.g., CoLoDaS (n=6) or have an *unknown* status, e.g., Pressious Data Space (n=7).

Various **delivery channels**, i.e., *software solutions*, of the data spaces are used. Here, we observe different solutions like *applications*, i.e., Metal Domain Data Space (n=1), *portals*, i.e., Sm4rtenance (n=1), and *platforms*, e.g., DV4CUL (6), while the majority do *not explicitly mention* their delivery channel, e.g., Manuspace (n=7). We refrain here from the term "data space connector" (for an overview of definitions see Gieß et al. [7]) since we interpreted the documentation more like a delivery channel "only" on how to carry the data space service, e.g., through a software frontend, to the participants without explaining technical details and backend functionalities. *Economic* benefits are the most frequently cited, e.g., UCIMU (n=13), but combinations like *economic + environmental value*, e.g., e-CMR Hub (n=2) are also present. Our observed **value propositions** are somewhat in line with the already discussed one by Möller et al. [3].

[1] https://osf.io/ruzg4/?view_only=8c50397946a042219d3ee0616bd9a2c9

Meta Dimension	Dimension	Characteristics			
Economic	Funding	**Public (18)**	Public + Private (1)		
Economic	Key Purpose	(Predictive) Maintenance (4)	Optimized Manufacturing (2)	E-Signature (1)	Optimization (Supply Chain) (1)
Economic	Key Purpose	Data Sharing / Data Exchange (3)	Monitoring (3)	**Multiple (5)**	
Economic	Real-World Operation	Yes (6)	No, Conceptual (6)	**Unknown (7)**	
Economic	Delivery Channel	Application(s) (1)	Portal (1)	Platform (6)	
Economic	Delivery Channel	Online Marketplace (1)	Multiple (3)	**Not Explicitly Mentioned (7)**	
Economic	Value Proposition	Environmental (2)	**Economic (13)**	Economic + Environmental (2)	
Economic	Value Proposition	Economic + Environmental + New Business Model (1)	Unclear (1)		
Technical	Data Space Access	**Membership / Project Participation (18)**	Fee (0)	Unknown (1)	
Technical	Underlying Model	**IDS Reference Architecture Model (18)**	Not Explicitly Mentioned (1)		
Technical	Data Type Shared	Repository Data (2)	Resilience Data (1)	Process Control Data (3)	
Technical	Data Type Shared	Historical Data (1)	**Not Specified (12)**		
Technical	Technical Enabler	Artificial Intelligence (4)	Blockchain (1)	Digital Twin (2)	Big Data Analytics (1)
Technical	Technical Enabler	Data Space (1)	Multiple (2)	**Not Explicitly Mentioned (8)**	

Fig 2. Taxonomy of Data Spaces in Industry 4.0 (n=19)

Technical aspects. We found four technical dimensions in our corpus of data spaces in Industry 4.0 that are of particular interest. Data space access is mainly based on *membership/project participation* within the dimension of **data space access**, e.g., GlobShare (18). In contrast to Gieß et al. [8], we did not find any *fee access model* (n=0) showing a "structural hole" in our taxonomy. Also, we did not observe any pricing models. This is somewhat in line with difficulties of data value assessment or the distinction between data spaces and so-called data marketplaces, which are intended to help companies trade data and data assets, as identified by previous literature [18, 19]. As a result, we see this issue as suggestion for further research in terms of data spaces in Industry 4.0. The *IDS Reference Architecture Model* is the predominant **underlying model**, e.g., within the data space Underpin (mentioned 18 times). This observation is not surprising since the majority of investigated data spaces in Industry 4.0 are dedicated to the standards provided by the International Data Space Organization.

Various **data types** that are shared and processed in the corresponding data spaces were found. Here, we observed data types such as *historical data*, i.e., Pressious Data Space (n=1) and *process control data*, e.g., UCIMU (n=3), though "*Not Specified*" appears most frequently, e.g., OSME (12 times), indicating no further information on the processed data. This fuzziness is somewhat surprising since data space participants should be aware of what data types can be shared in a data space. However, our taxonomy shows a smaller step towards clarity on data types than the two characteristics for the same dimension examined by Gieß et al. [8]. Data spaces investigated articulated (further) **technical enablers** that are used within the space. Commonly cited technologies include *artificial intelligence*, e.g., Underpin (n=4), *blockchain*, i.e., CollMi (n=1), and *big data analytics*, i.e., Boost 4.0 (n=1). "*Not Explicitly Mentioned*" appears most frequently, e.g., CoLoDaS (8 times). Also, we found one data space that named the *data space itself* as an enabler, i.e., the Metal Domain Data Space (n=1). As a result, we identified a diverse mix of technological enablers without any explicit tendency.

Discussions and first contributions. As a theoretical contribution from this short paper, currently, data space or data ecosystem researchers in general and Industry 4.0 researchers in particular, interested in the dimensions and characterizations, can use the taxonomy to receive an initial look at this phenomenon. An overview of data spaces in Industry 4.0 remains missing with this short paper we close the call for further investigations on the fragmentation of the data space market in this particular sector [3].

In this short paper, we make use of the process of taxonomy development by Nickerson et al. [14]. As an implication for research, we can conclude that the process was suitable to some extent to provide meaningful dimensions and characteristics in the context of data spaces and Industry 4.0. However, most problems came up with the diverse quality and depth of information provided by the data spaces, resulting in characteristics like “Not Explicitly Mentioned” or “Unclear” which are sometimes the most common ones inside the dimensions provided. We were able to achieve mutually exclusiveness within the dimensions and characteristics provided by adding a “multiple” characteristic to some dimensions. For example, the value proposition provided by the usage of the data pace was mentioned in a fuzzy way by the data spaces and within the documents. Economic values were proposed with no concrete measurement. Therefore, no further refinement through more explicit characteristics was possible. As a limitation, validity of the taxonomy is somewhat restricted based on the information found and further interpreted.

Also, it was surprisingly difficult to verify if a data space was productive in an industry environment or not, by the documentation of the data space provided. Here, as an implication for data space providers, we argue that a clearer and more standardized communication of use cases and values, e.g., economic ones are necessary for outreach and external communication.

Taxonomies order the complexity and serve as a foundation that can lead to future research and serve as a nomenclature on a specific phenomenon [14]. Therefore, IS researchers can use the taxonomy provided as inspiration for their research inquiries and clearer classification of data spaces in Industry 4.0. Practitioners, like small and medium enterprises in Industry 4.0, can use this taxonomy as a general and initial background for discussing patterns and similarities of data spaces to participate in them (or not) with data space providers or other sector representatives.

4 Next Steps and Conclusions

We plan to conduct the following next research steps towards ICSOB and for the ultimate goal of a journal article: First, our objects were chosen from two databases, and the data space market is constantly changing. For some data spaces mentioned in the two databases, no information was available anymore. An update of possibly more data spaces that will emerge is necessary. As the topic of data spaces is a vital research field, we will first revise the taxonomy based on completeness and potential new knowledge for dimensions and characteristics from recent literature [20] after the conference in November 2025, since the taxonomy provided in this short paper was developed in July 2025.

Second, as Design Science Research artefacts [14], taxonomies should be further evaluated in terms of their completeness and appropriateness for their intended purposes as solution-building instruments [15]. Examined classifications, e.g., through clustering techniques, can lead to higher theorizing, and will allow us to provide a more unified view of the data space context in Industry 4.0 [3, 21]. We expect that classifications and provided clusters would carve out interrelations and connections between services bundles within data space. This will show a more nuanced overview on the data spaces market and what is currently offered to interested participants in Industry 4.0. Here, we will recent developments on adequate clustering methods designed for taxonomy data [22] and suggestions for taxonomy evaluation, e.g., through interviews with data space providers [15].

Third, with this more comprehensive overview and clusters of other data space researchers will have a foundation for stronger theoretical contributions to the conceptualization of data spaces toward a taxonomic theory. This theory is useful in developing testable propositions and relationships between dimensions and characteristics provided, researching the phenomenon towards a theory for analysis. It will equip IS researchers to describe novel Information Technology-related phenomena with eloquent knowledge [21, 23]. To the best of our knowledge, such an investigation and proposition of a taxonomic theory has not been conducted so far for data spaces. Our package of planned research outputs will be contributed to a subsequent journal article.

5 Acknowledgments and Use of Generative AI

This research was conducted within the project URANOS-X. It was funded by the Federal Ministry of Education and Research of Germany under the subsidy indicator 02J21D001. The authors would like to thank all people involved who provided feedback on the manuscript in subsequent discussions. We used Generative AI for proofreading some parts of the manuscript.

References

1. Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., Hoffmann, M.: Industry 4.0. Bus Inf Syst Eng, vol. 6, 239–242 (2014)
2. Leckel, A., Linnartz, M.: Towards the Internet of Production – How to increase data sharing for successful supply chain collaboration. Journal of Production Systems and Logistics, vol. (2023)
3. Möller, F., Jussen, I., Springer, V., Gieß, A., Schweihoff, J.C., Gelhaar, J., Guggenberger, T., Otto, B.: Industrial data ecosystems and data spaces. Electron Markets, vol. 34 (2024)
4. Otto, B., Hompel, M. ten, Wrobel, S.: Designing Data Spaces. Springer International Publishing, Cham (2022)
5. Contreras, J.D., Garcia, J.I., Diaz, J.D.: Developing of Industry 4.0 applications. Int. J. Onl. Eng., vol. 13, 30–47 (2017)
6. Czvetkó, T., Abonyi, J.: Data sharing in Industry 4.0 - AutomationML, B2MML and international data spaces-based solutions. Journal of Industrial Information Integration, vol. 33, 100438 (2023)

7. Gieß, A., Hupperz, M., Schoormann, T., Möller, F.: What does it take to connect? Unveiling characteristics of data space connectors Proceedings of the 57th Hawaii International Conference on System Sciences (2024)
8. Gieß, A., Möller, F., Schoormann, T., Otto, B.: Design options for data spaces Proceedings of the 31st European Conference on Information Systems (2023)
9. Weking, J., Hein, A., Böhm, M., Krcmar, H.: A hierarchical taxonomy of business model patterns. Electron Markets, vol. 30, 447–468 (2020)
10. Azkan, C., Iggena, L., Gür, I., Möller, F., Otto, B.: A taxonomy for data-driven services in manufacturing industries Proceedings of the 24th Pacific Asia Conference on Information Systems (2020)
11. Dumitrescu, R., Riedel, O., Henke, M., Metternich, J., Ihlenfeldt, S., Koldewey, C.: Gaia-X in industriellen Wertschöpfungsnetzwerken – Erkenntnisse aus der Förderlinie InGAIA-X. VDI Verlag (2025)
12. Dorkaev, E., Werth, O., Uslar, M.: MAMOGI - A Maturity Model Towards Gaia-X Conformity in Industry 4.0 Proceedings of the 31st Americas Conference on Information Systems (2025)
13. Lick, J., Lamarz, J., Dohmann, F., Kulkarni, J.P., Zerbin, J., Koldewey, C.: Guidance on Dataspaces: Development of a Learning Environment for Industrial SMEs. In: Proceedings of the 6th International Conference on Control and Robotics (ICCR), pp. 348–354. IEEE (2024)
14. Nickerson, R.C., Varshney, U., Muntermann, J.: A method for taxonomy development and its application in information systems. European Journal of Information Systems, vol. 22, 336–359 (2013)
15. Kundisch, D., Muntermann, J., Oberländer, A.M., Rau, D., Röglinger, M., Schoormann, T., Szopinski, D.: An update for taxonomy designers. Bus Inf Syst Eng, vol. 64, 421–439 (2022)
16. Hilger, T., Schütz, J.: Data preparation on state-of-the-art and progress of European Gaia-X and data space initiatives with a focus on Industry 4.0 use cases (v1.0.0). Zenodo (2024)
17. International Data Space Association: The data spaces radar version 4 - Data spaces support centre edition (2024), https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/The-Data-Spaces-Radar-Version-4.pdf
18. Werling, M., Stuber, K., Petrik, D., Lachenmaier, J., Herzwurm, G.: Enabling Inter-organizational Data Sharing: Towards a Method for Assessing Data Assets. In: Papatheocharous, E., Farshidi, S., Jansen, S., Hyrynsalmi, S. (eds.) Software Business. Lecture Notes in Business Information Processing, vol. 539, pp. 89–96. Springer Nature Switzerland, Cham (2025)
19. Jussen, I., Möller, F., Schweihoff, J., Gieß, A., Giussani, G., Otto, B.: Issues in inter-organizational data sharing: Findings from practice and research challenges. Data & Knowledge Engineering, vol. 150, 102280 (2024)
20. Brechtel, M.: From hesitation to participation in industrial data ecosystems: Analysis of motives and incentives in the automotive industry. Electron Markets, vol. 35 (2025)
21. Gregor: The nature of theory in Information Systems. MIS Quarterly, vol. 30, 611–642 (2006)
22. Heumann, M., Kraschewski, T., Werth, O., Breitner, M.H.: Reassessing taxonomy-based data clustering: Unveiling insights and guidelines for application. Decision Support Systems, vol. 187, 114344 (2024)
23. Baier, M.-S., Berger, S., Kreuzer, T., Oberländer, A.M., Röglinger, M.: What makes digital technology? A categorization based on purpose. CAIS, vol. 52, 332–384 (2023)