FROM BIG DATA TO SMART DATA. OPPORTUNITIES FOR ENTREPRENEURS USING DATA SPACE ECOSYSTEM APPROACH

10.29073/jer.v1i2.19

RECEIVED: December 22, 2023.
ACCEPTED: December 26, 2023.
PUBLISHED: December 26, 2023.

AUTHOR: Joel Sepúlveda, Innovalia Association, Spain, jsepulveda@innovalia.org.

ABSTRACT

This paper is a theoretical and conceptual approach for entrepreneurs that would like to create new business based in the context of the European data strategy. The paradigm shift is happening, data reveals that we are passing from consuming a data volume of 33 zettabytes in 2018 to 175 zettabytes in 2025. In order to demonstrate the value of Data Space technology for entrepreneurs it has been applied a literature review methodology. For newly emerging topics (such as Data Space) the purpose is rather to create initial or preliminary conceptualization rather than review old big data concepts. One of the key insights of the present article is to confirm how data's increasing ubiquity and abundance makes is vital in every sector, and businesses of every size are becoming more dependent on data management. In the case of Data Space technology there is a clear problem between technology itself and its business application or business model and therefore we have a knowledge gap. Thus, it is necessary to spread out the Data Space concept to the entrepreneur's ecosystem so their flexibility and speed in order to adapt or create new business could help reducing the mentioned technology-knowledge gap (how to monetize or create new business models). In order to demonstrate the value of the technology and derivate business opportunities for entrepreneurs this theoretical review presents the basic concept of Data Space and its association with MDVC (multilateral data value chain) development in different sectors.

KEYWORDS: Business Innovation; Data Economy; Data Space; Entrepreneurship; Network.

1. INTRODUCTION

The principal aim of this conceptual review paper is to explain the potential business associated to a new complex technology in easy way for entrepreneurs. The main idea is to have a global context of what is happening at EU policy level and global investment view. Then the main concepts associated to Data Space Technology has been presented so the entrepreneurs could explore the necessity of new business models (to reduce the actual gap between the technology and the business).

1.1. EU CONTEXT

There is currently active a 4 to 6 billion investment in common European data spaces and European federation of cloud infrastructure services. In terms of the value of data economy (EU27) it will be an incremental growth from 301 billion (2018) to 829 billion in 2025. Even the number of data professionals will be double from 5.7 million to 10.9 million of people. So, there is no doubt about the potential market from different perspectives.

1.2. CONCEPT APPROACH AND IMPORTANCE FROM BUSINESS PERSPECTIVE

Some argue that the provision of digital infrastructure is a shared societal service in the same way as water, sanitation, education, and healthcare (Curry and Scerry, 2022). On the other hand, Data economy is defined as a global digital ecosystem in which data is gathered, organized, and exchanged by a network of companies, individuals, and institutions to create economic value (Sestino et al., 2023). So, it is evident the importance of sharing data. We

also know that the pace of technological change is much faster now than it has been in the past, and Data driven economy is a reality that is developing every day. Data driven economy is catalysing a change in data processing forms. In 2018 data management was centralised in 80% and smart connected objects where 20% while in 2025 this situation will be the opposite, there will be 80% of smart connected objects while only 20% of centralised computing facilities (EU Data Strategy 2020).

The actual entrepreneur needs to understand data environment as a sociotechnical ecosystem enabling value to be extracted from data value chains supported by interacting organizations and individuals (Fraunhofer 2017). It is important to note that economic value arises not from the data's content, but from combining different data sources in a way that meets the customer's needs. That said, profitable business models do not rely on selling access to data. In many cases, startups and SMEs are looking to disrupt markets by publishing and managing data. We are living a big digital transformation era. The data Governance Act approved by EC to boost data sharing is essential for the manufacturing industry. The 80% of industrial data is never used. We still have in our brain the misconception that set of data has only one use. Nowadays the potential of data access and sharing will help companies to create and develop new products and services.

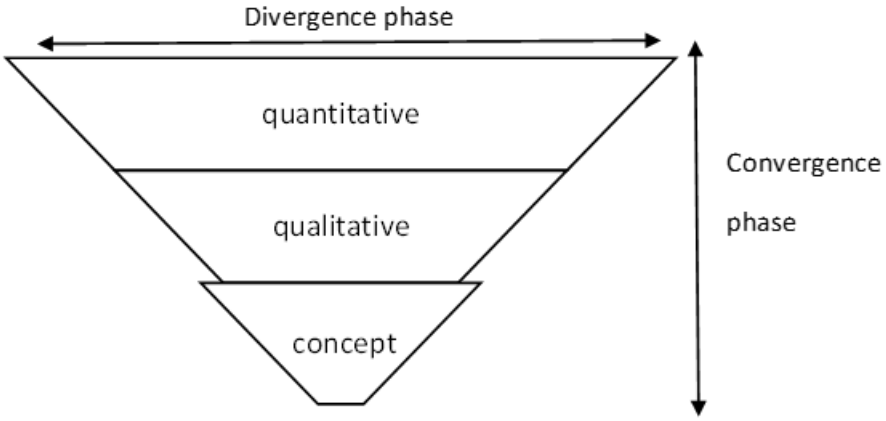
1.3. THE CASE OF MANUFACTURING INDUSTRY

The potential value of manufacturing industry data sharing has been estimated at €83 billion just in the field of process optimisation and better leveraging machine-generated data can lead to up to 20% improvements in material resource efficiency. The total value that companies can create in five key areas of data sharing is estimated to be more than $100 billion according to World Economic Forum white paper. (WEF, 2020). Conceptually the value for stakeholders derived from data sharing services in manufacturing industry will be in asset resilience, asset energy efficiency, optimised OEE (machine efficiency) and asset lifetime.

2. METHODOLOGY

This paper has been developed within the context of literature review methodology. The method has been dived in two main stages (See figure 1). The stage number one is called divergence and consists in doing an extensive literature review of the covered topic-issue (research papers, web references, white papers, or any valuable reports as result of European research projects). This stage it is more related with the classical systematic qualitative and quantitative approach (divergency phase).

FIGURE 1: Literature review methodology.



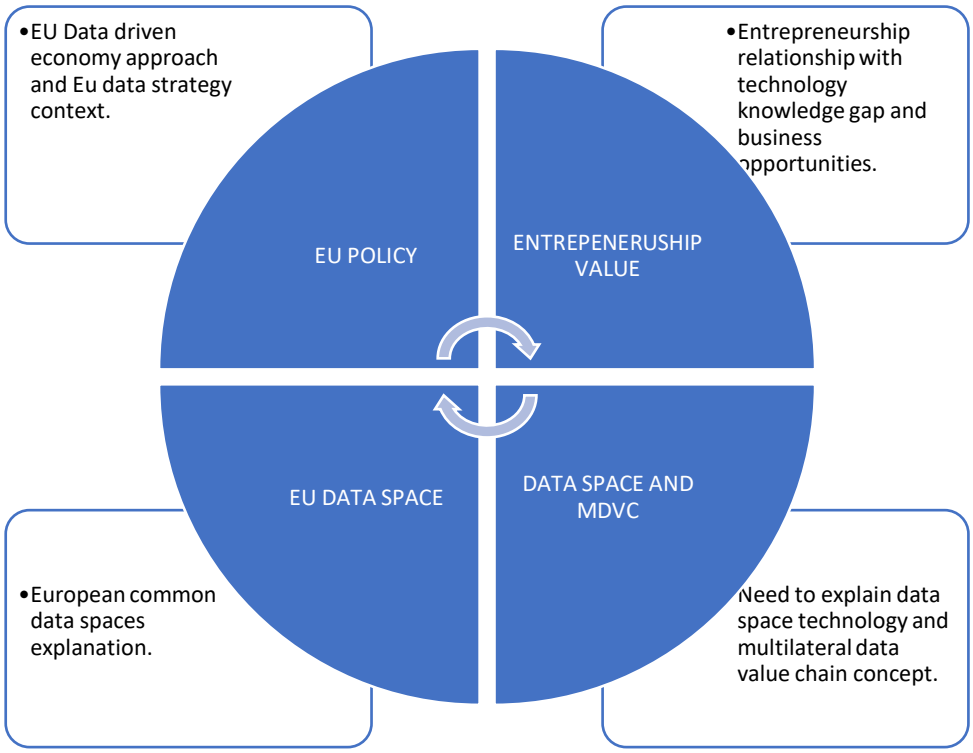
The second stage is known as the integrative theoretic framework. The main characteristic of this stage is to make valuable insight summary of the previously analysed literature (main ideas and concepts associated to the covered

issue in the paper). The synthesizing process is called the convergence phase and in this stage is essential the critique and potential modelization framework.

3. RESULTS

After the divergency phase there has been read 35 research papers, 12 online references, 4 EU projects (digital Europe) and 10 reports. To be able to sum up the findings it has been decided to make a concept map in these has been the base to develop the sections of the paper regarding data space technology. The next figure (figure 2) is summary of the main ideas or results derived from the literature review.

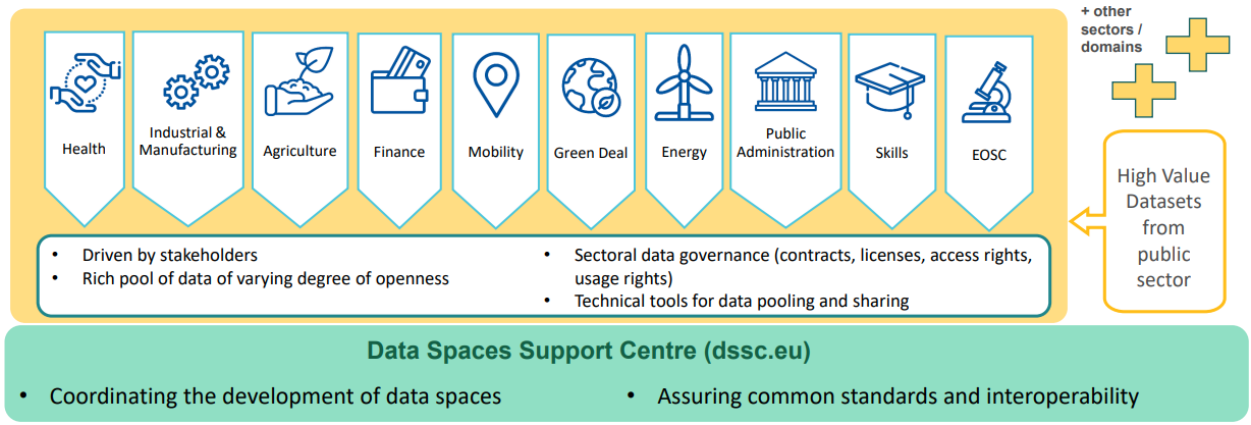
FIGURE 2: Mind map of the literature review.



3.1. COMMON EUROPEAN DATA SPACES

The creation of EU-wide common, interoperable data spaces in strategic sectors (see figure 3) aims at overcoming legal and technical barriers to data sharing by combining the necessary tools and infrastructures and addressing issues of trust by way of common rules. A common European data space brings together relevant data infrastructures and governance frameworks in order to facilitate data pooling and sharing (EC, 2002).

FIGURE 1: European Data Spaces Strategic sectors.



SOURCE: EU data strategy.

In the Commission's data strategy, diverse data spaces were proposed, ranging from industry to energy, and from health to the European Green Deal. They will, for example, contribute to the green transition by improving the management of energy consumption, make delivery of personalised medicine a reality, and facilitate access to public services.

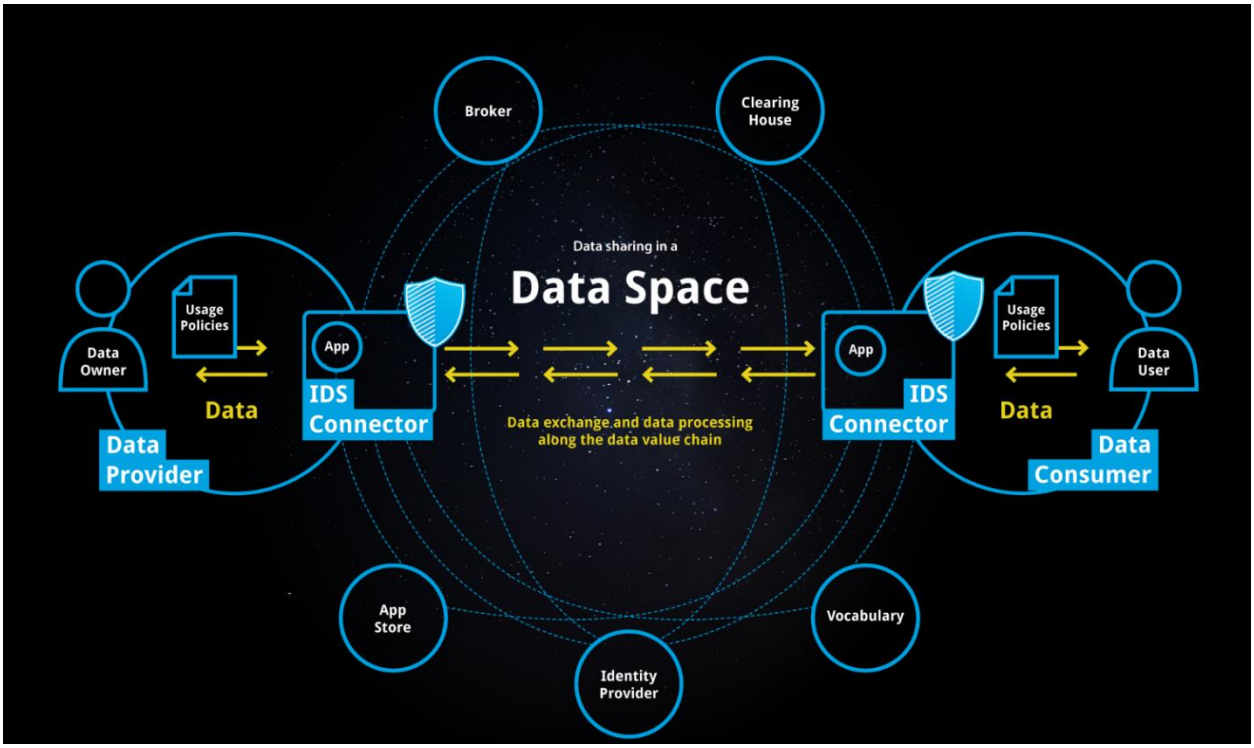
Common European data spaces will allow data from across the EU, both from the public sector and businesses, to be exchanged in a trustworthy manner and at a lower cost, thereby boosting the development of new data-driven products and services. Data spaces are composed of both the secure technological infrastructure and the governance mechanisms.

3.2. THE DATA SPACE CONCEPT

As we look to the evolution of data-driven innovation, Data, in particular large quantities of high-quality data, is critical to creating competitive smart solutions for companies. With few exceptions, our current large-scale data infrastructures are beyond the reach of small organizations that cannot deal with the complexity of data management and the high costs associated with data infrastructure.

The modern economy is developing based in a technology-concept called Data Space and therefore data driven entrepreneurship and data driven business opportunities are inevitably associated to it. A common European data space will ensure that economies and societies have access to more data while preserving the power of communities and people to create. In the actual socioeconomic context, the big industrial players are transforming their business into a full vertical and horizontal integration of systems and components (Rojko et al., 2017). Applying data space concepts in companies will make them more competitive and flexible in the face of any change in the market. It is important to have innovative services such as predictive maintenance or process autonomy. In order to clarify what a data space is, see figure 4 (data space components and framework).

FIGURE 4: Data Space components.



SOURCE: IDSA (International Data Space Association).

The main idea is that a data space is defined as a federated ecosystem based on shares policies and rules. Basically, it is a decentralized infrastructure that enables diverse actors to share and use data in a secure, reliable, and trustworthy manner, following common governance, organizational, regulatory, and technical mechanisms (JRC, 2020).

Correlated to data space technology there is a new approach to value chain. The concept for companies to adopt the data perspective and to optimize data usage, can be found in the field of Big Data by setting up sophisticated data value chains (DVC) (Jony et al., 2016). SMEs need to understand the concept of the data value chain as early as possible and start playing the game as soon as possible. One promising embryonic data space project for Spanish and Portuguese companies is BAIDATA project. The BAIDATA Association is helping to drive the development of data sovereignty and the data economy. Founded in collaboration with the International Data Spaces Association (IDSA), BAIDATA implements research, development and training activities to help build the public-private data ecosystem at Iberian Peninsula scale.

On the other hand, it is important to note the DVC approach from strategic point of view. It helps in value creation (Faroukhi et al., 2020) and integrates all steps that affect data starting from data generation and collection and ending with the possibility of decision-making based on data output (Kasim et al., 2012; Miller et al., 2013.).

The amount of data from various distributed sources is growing rapidly which creates significant opportunities to gain valuable insights. Data ecosystems can create conditions for market competition between participants or enable mutually beneficial cooperation between different and interdependent participants. Dataspaces can provide a clear framework to support data sharing within a data ecosystem.

3.4. BUSINESS OPPORTUNITIES FOR ENTREPRENEURSHIP-STARTUP ECOSYSTEM

The fall in the number of start-ups accelerated in 2021, falling by 60% in the European Union, the United Kingdom and the United States (EC, 2022). However, the number of start-ups rose somewhat in 2021 in Germany in knowledge-intensive services sectors, after a trend decline since the early 2000s (Deutsche Bundesbank, 2022).

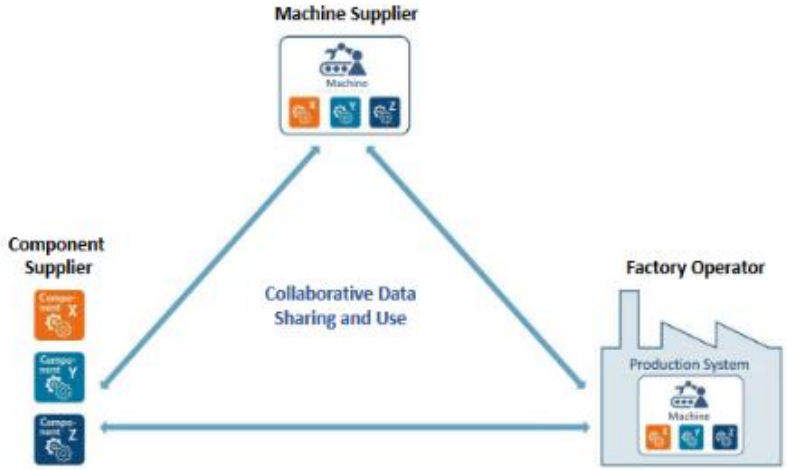
There is gap and an opportunity to recover the innovation path for the rest of European countries. The missing start-ups can worsen long-term productivity developments. Historically severe recessions have often been associated with the loss of a generation of startups or poor growth performance of those that survived. The impact is expected to be small in the short term but since start-ups make up only a small part of the business their absence has a long-term impact on productivity as these start-ups often play a key role in competitive innovation. (Kolev et al., 2022) and/or in diffusing new technologies and business models (Criscuolo, Gal and Menon, 2016). Some start-ups also have strong growth potential, with related economic benefits (OECD, 2021).

Nowadays Industrial companies are faced with challenges with the increasing complexity of their production processes. There is a good example of a complexity driver is the ever-increasing demand for products in the customized market and short life cycles and delivery times. Due to competitors companies make more products to remain more attractive in their market (Gottman, 2019).

Data Space technology will help supply chain analytics to describe the use of new data sources and analytical techniques to help companies design and run smarter, cheaper and more flexible supply processes. One of the most significant benefits is that with data space concept, companies could gain access to better data-driven predictions of obstacles in their supply chain to potentially avoid disruption scenarios altogether. Supply chain disruptions include natural disasters, labour shortages, cyberattacks, and global crises like COVID-19.

In order to increase new market potential opportunities there needs to be a change in mindset from classic bilateral data exchange to holistic, standardised and multilateral sharing of data from multiple stakeholders (value creation). Business related to the management of MDVC (multilateral data value chain) will increase by 60 in EU market (EU data strategy, 2020). The MDVC concept refers to the multilateral data sharing and can be explained within the CCM three-point fractal approach that is, (figure 5) the smallest possible unit for multilateral data sharing (Plattform industrie 4.0, 2022). The main idea is that any multilateral cooperation has a minimum of 3 components: factory, it provider and machine or process provider with the aim of generating economic added value for all parties involved.

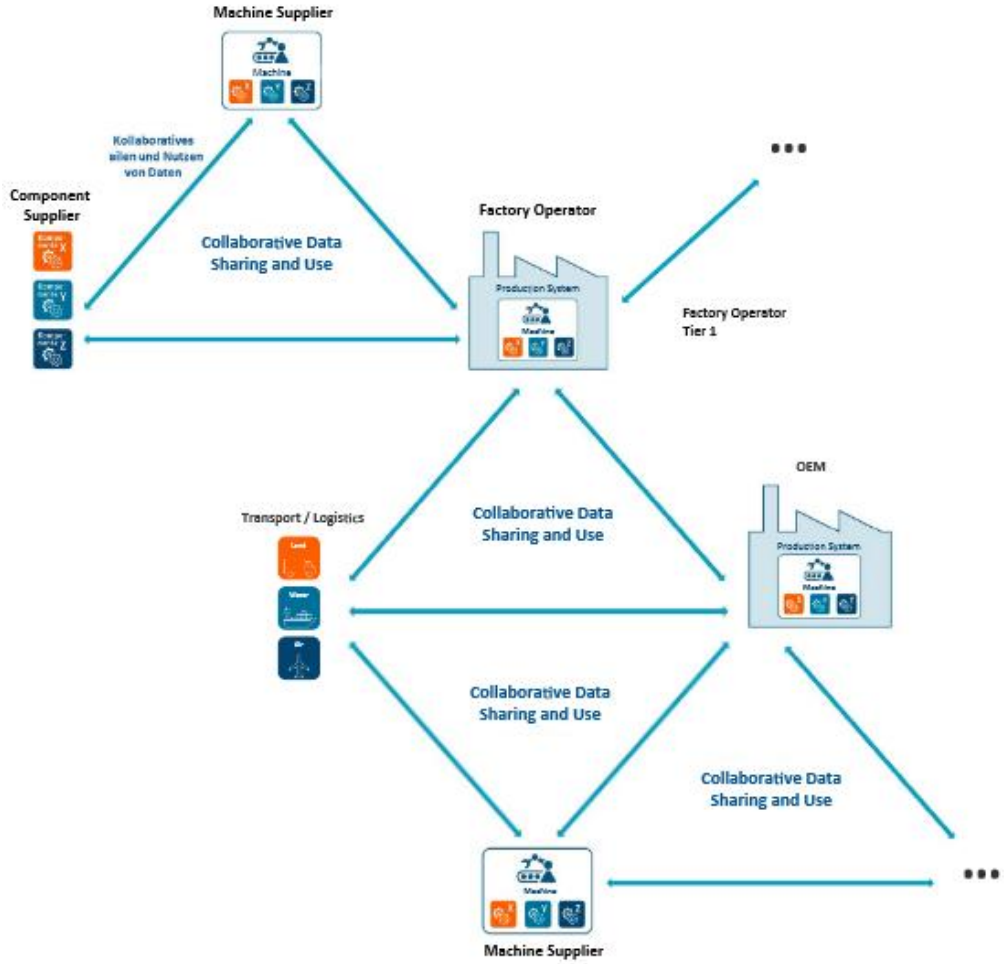
FIGURE 5: Multilateral data sharing reference model for industry 4.0.



All of these entities in the global value network of production have a variety of different bilateral relationships with their partners in the value chain. The complexity of these value networks can be modelled by scaling the three-point

fractal. So, the final idea consists in having a data space where all the MDVC could exchange data in order to create economic value (figure 6).

FIGURE 6: Cross-company collaboration in a production value network.



The data space and MDVC combination will let companies to access knowledge and innovation networks faster. Accessing knowledge networks is critical for SMEs to innovate and transform. Firms seldom innovate in isolation and networks of innovation involving multiple actors are the rule rather than the exception (DeBresson, 1996). Collaborative firms, even smaller ones, tend to be more innovative than non-collaborative ones, even larger firms (Eurostat, 2022). This is because innovation results from the accumulation of increasingly specialised knowledge and knowledge-based capital that calls for co-operating and opening innovation to gain efficiency and reduce time to market (Chesbrough, 2003). Indeed, networks are increasingly seen as an innovation asset (Corrado et al., 2005; OECD/Eurostat, 2018). The companies within the MDVC have the goal of commercialising data from the processes for creating goods and services and will define corresponding pricing models for this purpose. Another important data is that only 39% of European businesses share data with other businesses (Scaria, 2018). In table 1 it is possible to see the market segmentation by opportunity.

TABLE 1: Market segments by opportunity.

MARKET SEGMENT ANALYSIS	
TECHNOLOGY OPPORTUNITY	**INDUSTRY OPPORTUNITY**
▪ Cloud Computing ▪ Big Data and Analytics ▪ Cybersecurity ▪ Artificial Intelligence ▪ Internet of Things ▪ Others (Block chain & Business Intelligence)	▪ Banking ▪ Financial Services and Insurance (BFSI) ▪ Manufacturing ▪ IT and Telecommunications ▪ Retail and Consumer Goods ▪ Healthcare ▪ Education ▪ Transportation and Logistics ▪ Government ▪ Environment ▪ Others

In the upcoming years (2023–2025), open standards and interoperability will be the basis for new business models and a stronger focus on the customer experience. This includes the monetization of operational data as well as the co-development of products and their joint use.

5. DISCUSSION

This paper provides an introduction to a data space concept and multilateral data sharing concept, including its advantages and requirements. The article has analysed the European context of the data economy. It has also presented at a conceptual level the importance and significance of the data space technology and which are the sectors of socio-economic activity involved. We are talking about a new technology, therefore as any innovation process, it will need a period of testing and maturity based on the adoption by the companies. The article is intended as an introductory guide to the subject for entrepreneurs who want to get into the subject as well as for SME's that have some IT service that can join the technological wave of the European Data Spaces. It has been argued how the systematic collection and processing of usage data for multilateral data sharing therefore has significant economic potential.

6. CONCLUSION

Data Space is a key element of the digital future. These offer new business opportunities and strategic roles in a data-driven environment. Although its potential value is considered high and various use cases are working nowadays (BAIDATA; MANUFACTURING-X, DAWEX...), the underlying business model considerations are not widely communicated. For example, there is a lack of literature and must be a s future research step to consider resilience in the lifecycle of a value chain as an output of Data Space applications.

The paper provided an original conceptual assessment on the actual role played by space-based data at EU digital transformation scale for new businesses, but further studies are recommended to examine how to monetize Data Space environments.

REFERENCES

Chesbrough, H. (2003). *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business School Press.

Corrado, C., et al. (2005). *Measuring Capital in the New Economy*. University of Chicago Press. https://www.nber.org/books-and-chapters/measuring-capital-new-economy

Criscuolo, C., Gal, P., & Menon, C. (2016). Do micro start-ups fuel job creation? Cross-country evidence from the DynEmp Express database. *Small Business Economics, 48*(2), 393–412. https://doi.org/10.1007/s11187-016-9778-x

Curry, E., Scerri, E., Tuikka, T. (Eds.) (2022). *Data Spaces Design, Deployment and Future Directions*. ISBN 978-3-030-98635-3, ISBN 978-3-030-98636-0 (eBook). https://doi.org/10.1007/978-3-030-98636-0

DeBresson, C. (1996). *Economic Interdependence and Innovative Activity*. Edward Elgar Publishing.

Deutsche Bundesbank (2022). Productivity effects of reallocation in the corporate sector during the COVID-19 crisis. *Monthly Report*, *64*(9).

EC (2022). *Annual Report on European SMEs 2021/2022*. European Commission.

EC Staff Working Document (2022). *Staff working document on data spaces*.

EU Data Strategy (2020). *European data strategy—European Commission*.

Eurostat (2022). *Community Innovation Survey 2020 (CIS2020)* https://ec.europa.eu/eurostat/web/science-technology-innovation/data/database

Faroukhi, A. Z., et al. (2020). Big data monetization throughout Big Data Value Chain: a comprehensive review. *Journal of Big Data*, *7*(1).

Fraunhofer (2017). *Reference Architecture Model for the Industrial Data Space*. https://www.fraunhofer.de/content/dam/zv/de/Forschungsfelder/industrial-data-space/Industrial-Data-Space_ReferenceArchitecture-Model-2017.pdf

Gottmann, J. (2019). *Produktions controlling*. Springer Fachmedien Wiesbaden.

Jony, I. R., et al. (2016). Big Data Characteristics, Value Chain and Challenges. *1st International Conference*.

JRC (2020). *Definitions—JRC Data Spaces Knowledge Base—EC Public Wiki*.

Kasim, H., Hung, T., & Li, X. (2012). Data Value Chain as a Service Framework: For Enabling Data Handling, Data Security and Data Analysis in the Cloud. In *2012 IEEE 18th International Conference on Parallel and Distributed Systems (ICPADS)*. IEEE.

Kolev, J., et al. (2022). Of academics and creative destruction: Startup advantage in the process of innovation. *National Bureau of Economic Research*. https://doi.org/10.3386/w30362

Miller, H. G., & Mork, P. (2013). From Data to Decisions: A Value Chain for Big Data. *IT Professional, 15*(1), 57–59.

Multilateral data sharing in industry (2022). *Plattform industrie 4.0*.

ODI (2017). The Data Spectrum helps you understand the language of data. *Open Data Institute*. https://theodi.org/data-spectrum

OECD (2021). *Understanding Firm Growth: Helping SMEs Scale Up*. OECD Studies on SMEs and Entrepreneurship, OECD Publishing. https://doi.org/10.1787/fc60b04c-en

OECD iLibrary (n.d.). https://www.oecd-ilibrary.org/

OECD SME and Entrepreneurship Outlook Biennial (n.d.). OECD iLibrary. ISSN: 29599504 (online). https://doi.org/10.1787/8d707502-en

OECD/Eurostat (2018). *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Edition*. OECD Publishing/Eurostat. https://doi.org/10.1787/9789264304604-en

Plattform Industrie 4.0 (2022). Federal Ministry for Economic Affairs and Climate Action (BMWK).

Rojko, N. (2017). Industry 4.0 Concept: Background and Overview. *Int. J. Interact. Mob. Technol., 11*(5), 77. doi:10.3991/ijim.v11i5.7072.

Scaria, E., Berghmans, A., Pont, M., et al. (2018). *Study on data sharing between companies in Europe Final report*. Publications Office. https://data.europa.eu/doi/10.2759/354943

Sestino, A., Kahlawi, A., & De Mauro, A. (2023). Decoding the data economy: a literature review of its impact on business, society and digital transformation. *European Journal of Innovation Management*. doi:10.1108/EJIM-01-2023-0078.

WEF (2020). *World economic forum Share to Gain: Unlocking Data Value in Manufacturing*.

ETHICAL STATEMENT

CONFLICT OF INTEREST: Nothing to declare. **FUNDING:** Nothing to declare. **PEER REVIEW:** Double-blind peer review.