Sectorial Data Ecosystems as Enablers for Data-driven Services for SMEs

Henrique Diogo Silva[1,2]
Ana Chaves[1]
António Lucas Soares[1,2]
César Toscano[1]
Filipa Rente Ramalho[1,2]

[1] INESC TEC, Porto, Portugal
{henrique.d.silva}@inesctec.pt
[2] University of Porto, Porto, Portugal

Abstract. SMEs are vital to the global economy, yet their adoption of state-of-the-art technologies is hindered by ad-hoc data management practices and limited access to high-quality datasets. This paper proposes a structured approach to overcoming these barriers through sector-specific data ecosystems. The approach includes a data maturity upgrading service designed to enhance SMEs data capabilities and an integrated infrastructure with governance mechanisms to support ecosystem development. By promoting collaboration and scalability, these ecosystems facilitate SME access to data-driven services, fostering digital transformation. The study highlights the importance of business, social, and technical considerations in enabling SMEs to harness the potential of data ecosystems for sustainable technological advancement.

Keywords: Data ecosystems · Digital platform ecosystems · Data Spaces

1 Introduction

Small and medium enterprises (SMEs) are pivotal in the global economy, representing over 90% of businesses worldwide. Their ability to adopt advanced technologies such as Artificial Intelligence and Machine Learning is often hindered by limitations in data accessibility, management, and quality because the effectiveness of these technologies is intrinsically dependent on the quality, volume, and available context-relevant data to train and customize these systems. While larger organizations often have mature data management (DM) processes to support these requirements, SMEs face significant challenges in implementing such processes and systems [2, 7]. These barriers stem from the immaturity of DM processes, limited access to high-quality datasets, and resource constraints that hinder the development of scalable data-drive solutions such as AI [1, 2, 4]. Consequently, the data-dependent nature of AI systems disproportionately benefits organizations with more advanced digital infrastructures, potentially widening

the gap between larger corporations and smaller enterprises, exacerbating digital inequalities across sectors and regions.

In this paper, we explore how platform-provided services integrating sectorial data ecosystems - can bridge this gap by offering SMEs access to contextually relevant data through standard robust processes. By providing the technological and business infrastructure for shared ecosystems of data-rich services while ensuring factors such as data sovereignty and trust between actors, the platform can provide the tools as services to address the resource and process deficiencies of smaller businesses. By leveraging these services, SMEs can more effectively deploy data-driven tools, further driving the gaining of insights for operational efficiency and innovation.

We outline the framework for developing sectorial data ecosystems and detail how this framework is being explored. The insights presented are expected to inform academic discourse and practical implementations, paving the way for future research into the development of sectorial, data-driven services. In conclusion, the proposed framework aims to provide access to data-driven tools for SMEs and serve as a catalyst for innovation, ensuring sustainable growth and resilience in a rapidly digitalizing economy.

2 State of the Art

The European data strategy aims to position the EU as a leading, secure, and dynamic data-driven economy. To support this goal, IDSA introduced a federated technical architecture for trusted data exchange—Data Spaces (DS)—ensuring data security and sovereignty for all providers [11]. Originating in computer science in 2005, DS was conceptualized as an alternative to traditional databases, allowing data coexistence without full physical or semantic integration [8]. Over time, DS evolved into a collaboration-focused framework for achieving shared goals [11].

DS operates without a central data store; instead, data is exchanged directly between Providers and Consumers. The Federator plays a key intermediary role, offering services such as data cataloging, brokering, and sovereignty enforcement to ensure trust among participants [11].

From a technical perspective DS rely on three building blocks: data interoperability, data sovereignty, and data value creation [3]. Materialized in components such as network protocols or middleware components, these elements enable the implementation of the technical architecture of a data space and enable the integration of different systems.

These developments notwithstanding, without proper framework, DM becomes uncoordinated, leading to inconsistencies in data handling and an inability to retain the value of constructs such as data spaces. Inadequate Data Governance can disclose organizations to non-existent standards and policies, exposing them to data breaches resulting in financial wastage [12].

The literature on DM maturity is vast, providing methodologies that focus on Data Governance, Data Quality, and many others. Moreover, scholars consider

that technology maturity assessments should consider the organization's strategy, emphasizing digital and DM strategies, which can profit from the definition of tailor-made roadmaps [13].

In summary, by using DS, companies, and institutions could take advantage of the benefits of digitization without increasing their risks. It accomplishes this by leveraging a trustworthy architecture for DM with standards for sovereignty and secure data exchange [11].

3 A Platform for Sectorial Ecosystem Services

Data has, for a long time been the lifeblood of enterprises. At the core of every business process, every business relies on data to conduct its day-to-day activities and keep up with the current (often required) pace of innovation. However, the reality is that only a small percentage of industrial data is currently used in ways that add value to businesses, especially in medium to long time frames[6, 15].

The vision developed in this paper supports the development of sectorial data ecosystems to facilitate enterprise data exchanges in secure, trusted, and semantically interoperable environments. We adopt a perspective in which the elevation of a sector's data management maturity is crucial for the generation of value for the individual organizations and, as a consequence, the sector as a whole. In this sense, we define a bottom-up approach where supporting organizations in adapting internal processes and information systems to allow for an active (and value-added) participation in data ecosystems from both business, social and technical perspectives.

While this bottom-up approach tackles the individual organizations that compose the different sectorial ecosystems, on a long-term perspective and including a top-down strategy, it will generate a network of federated sectorial ecosystems, managed and orchestrated by Ecosystem Owners that enable a more intelligent and trusted exchange of data among organizations.

In figure 1, in line with the literature highlighted in section 2, we distill the necessary development dimensions of our vision into: (1) the business and social elements, corresponding to the development of data maturity advancing services; (2) the governance characteristics of the sectorial ecosystems; and (3) the technical elements that bring the previous two together to provide the infrastructure for the ecosystem. In the remainder of this section we detail the activities in these three dimensions.

3.1 Data Management as Driver for Ecosystem Adoption

Organizations need help in making sense of the increasing volume of industrial data. In our bottom-up approach is important to note that the actions highlighted bellow are to be initially implemented on an individual enterprise scale, but with an intent on having and effect on the entire sector [1]. In this sense, for the characterization of each step, we highlight the importance for both of these perspectives.

The first step in our vision is to *characterize the company* and its position in its activity sectors. At an individual level, a characterization that encompasses current information (intensive) systems and business processes accompanied by an assessment of the maturity of the associated data is crucial for adapting governance models aligned with the industry's needs [10]. From a sectorial perspective, the characterization and assessment of collective challenges, the strategic and regulatory context and data-sharing requirements allow for the development of structures that facilitate the integration of the various actors.

When using the ecosystem metaphor for business contexts keystone organizations are a crucial building block. As [9, 1] explain keystone actors have the potential of increasing an ecosystem's productivity by simplifying the complex task of connecting network participants to one another or making the creation of new products by third parties more efficient. This contextualization of an organization when relative the an ecosystem may support developers in ensuring a fit for a number of other organizations in the same ecosystem and serve as support for its growth.

After the initial characterization, next step involve the definition of *governance and data quality requirements* [1, 7]. Defining how data is accessed, managed and shared, to ensure security and regulatory compliance [1, 7] is a the core of data management processes and systems that enable organizations to extract value from their data sustainably and over time [7]. In contrast, at the sector level, shared standards, interoperability guidelines and trust rules should be implemented to encourage collaboration [1]. Ensuring data quality - through accuracy, reliability and consistency - increases confidence in data-driven decision-making. Reference models, such as DAMA-DMBOK, provide structured methodologies that SMEs can adapt to sector-specific governance needs, reinforcing standardization and interoperability in sector ecosystems [5].

The final step integrates the previous steps by *intervening in business processes* to create-value. At the individual company level, this step implements the integration of data governance into decision-making workflows and operational

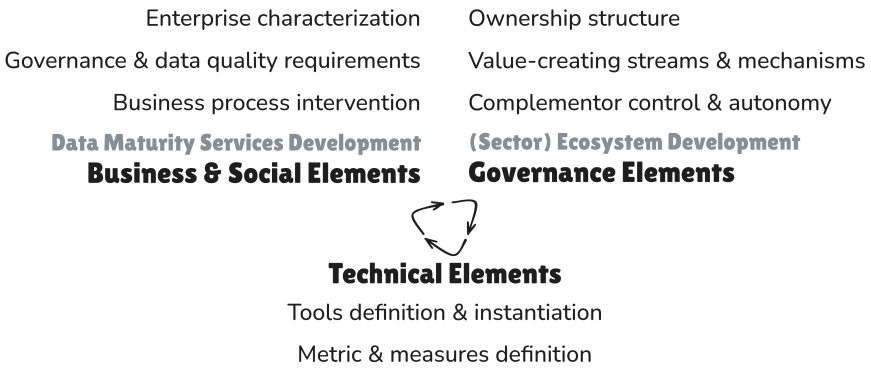

Fig. 1. Ecosystem Design Elements

processes. From a sectorial perspective, there is expected to be greater alignment in inter-company business processes since the actors have updated their processes to incorporate governance and data quality requirements. The updated processes following the good practices and sectorial characterization and context previously surveyed and analyzed. In short, these actions ultimately strive to upgrade data management maturity to a degree where active business processes can actively benefit from the ingestion and supplying of data from participating ecosystem's.

3.2 Ecosystem Development

Adopting the building blocks of platform bound ecosystems defined in section 2, we envision a ecosystem where (1) activities include sharing and consuming of sector relevant datasets and the creation and use of data-centric and added-value services that leverage the ecosystem's user base and technical elements; (2) actors become the companies that comprise each sector and other elements on the sector's value-chain, such as RTOs, cluster management structures, or relevant service providers; and (3) the DS becomes the data-level architectural framework that enables the envisioned functionalities.

Governance in the DS reference architecture revolves around *ownership*, *control structures*, and *value-creation mechanisms*.

The DS reference architecture, built on a federated model, places the Federator as provider of complementary services of varying importance [11], and sets the tone for the *ownership structure* of the ecosystem data. Ownership is crucial in governance [14], ranging from decentralized models that empower participants to centralized, rigidly controlled federated systems.

Value-creation strategies and mechanisms are equally critical. While the ecosystem facilitates data transactions, how complementary services interact with this core value proposition is crucial to enable it as a source of innovation [16]. Well-structured boundary resources, such as APIs, are essential, along with services ensuring semantic interoperability across industrial systems.

Control mechanisms shape the platform's initial setup and evolution. Owners traditionally exert control through gatekeeping, process regulation, and performance metrics [14]. Additionally, informal mechanisms—such as shared norms and values—help influence ecosystem behavior.

3.3 Technical Elements

Configuring and instantiating technical elements is essential for the ecosystem's foundational infrastructure. *Technical specifications* must align with DM requirements and the governance model. Key considerations include (1) designing the core platform with its functionalities and services and (2) integrating boundary resources to facilitate interactions among participants.

Metrics also play a crucial role in platform and ecosystem evolution over the medium and long-term time frames. [14] describe how metrics such as "synergy", "plasticity", "mutation" and "envelopment" highlight how the focus on

both internal and external environment are crucial to ensure that platform and ecosystem remain consistent over time in ensuring that offered functionalities are aligned with actors' expectations and the ecosystems' evolving needs.

4 Use Cases

The development of the vision detailed in the previous sections is enabled by ongoing work in several use cases, from different research projects across various industries. In this section we describe the ongoing work on three such projects that implement DS as their core ecosystem infrastructure: "TexP@ct" and "tExtended" for the textile sector, highlighting how the business, social, governance and technical perspective was essential for the tackling of the sector's digitization challenges; and "CircThread", for the iterative nature of the business, governance and technical elements of our vision. The findings are summarized in table 1.

The tExtended is a European project aiming to improve the circular economy in the textile sector by developping recycling solutions to add value to textile waste. Among these, a traceability system has been developed so as to capture and store data in circular value chains about who has done what, when, where, and why on each company involved in a supply chain. On a different scope, TexP@ct is a Portuguese project designed to drive the digital transition within the national textile and clothing sector. While working of different scales of implementation, both projects faced similar challenges in integrating digital solution in current business processes and systems, and in enabling semantic interoperability between the ecosystem.

Early in tExtended's initial stages of **company characterization and data governance definition** for supply chain actors, it was highlighted that ensuring the participation of all companies would prove challenging, particularly due to the high number of involved SMEs with lower technological maturity. To tackle these challenges, the deployed DS communication component was extended, to allow ecosystem complementors (software providers) to create services that **allows for the integration of current information systems directly with the ecosystem**, avoiding any additional implementation efforts. Similar efforts were devoted in TexP@ct, where four Portuguese industry ERP providers collaborated in a standardized framework based on the eBiz standard language. Pilot users were engaged to identify relevant processes and based on these insights, providers adapted their ERP to implement the eBiz standard. In this sense, the projects are enabling semantic interoperability across the supply chain of the textile sector.

From the appliance industry the CircThread project, aims at improving the circularity of home appliances. To this end, a per product Digital Product Passport was created to provide public information to end-users, encouraging them to contribute appliance usage data (e.g., the number of cycles in a washing machine). Supporting services such as a a Meta-Data Catalogue are being designed to support the sharing of product life-cycle data between value chain partic-

Table 1. Use case challenges and developed solutions

	Business & Social Elements	**Governance Elements**	**Technical Elements**
How to ensure all companies are able to participate in the ecosystem?	Identify systems and processes for enabling automated data exchange	Provide complementors with tools to enhance DS communication	Work with software providers to directly enable solutions for the supply chain companies
How to ensure semantic interoperability in the ecosystem?	Identify data requirements across processes and systems	Provide standard framework based on the eBiz language	Work with software providers to directly enable solutions for the supply chain companies
How to enforce access control in some data existing in the DS?	Identify data sharing requirements and establish policies	Define and enforce data sharing policies and control mechanisms	Extend base DS components to allow for usage control based on defined policies

ipants, such as recyclers and manufacturers, in order to improve data-driven decision making. During the **business and governance design phases**, relevant challenges were found: even among the supply chain participants, some information, such as product certification, and product repair manuals, must stay semi-private. To tackle this challenge, additional policies and features were introduced to the DS. These guarantee that the data owner retains complete control over its data, preventing unauthorized access and maintaining data sovereignty by controlling how long and with whom its data is shared.

5 Conclusions

As data becomes evermore crucial for powering decision-making processes and to ignite innovation activities, inefficiencies in how it is managed and missing opportunities for data use also become more costly. In this paper we highlight how sectorial platform-driven ecosystems can be designed to allow for SMEs and other organizations to derive value from available data, by working on elevating the relevant data maturity competences. We highlight how a business, social and technical-oriented approach is crucial for this development's success. From the business and social perspective, we focus on (1) the upgrading of DM processes and practices based on standard good practices in a way that allows SMEs to derive additional value from already existing data and data available on the ecosystem and (2) on the ecosystem-building activities, essential for the sustainability of the platform and the platform-provided services. For the technical components, we describe a Data Space-based (IDS) technical architecture, emphasizing the requirements for data sovereignty and trust as core elements for

designing digital platforms that can serve as catalysts for SME transformation in increasingly data-driven sectors. Finally, we highlight how this vision is already being applied in on ongoing I&D projects for Portuguese and European industries, helping in the solving of challenges faced by organization in the textile and consumer appliance industries.

Acknowledgements

This work is co-financed by Component 5 - Capitalization and Business Innovation, integrated in the Resilience Dimension of the Recovery and Resilience Plan within the scope of the Recovery and Resilience Mechanism (MRR) of the European Union (EU), framed in the Next Generation EU, for the period 2021 - 2026, within project TEXPACT, with reference 61 and the Ph.D. Grant UI/BD/152565/2022 from the Portuguese funding agency, FCT-Fundação para a Ciência e a Tecnologia.

References

1. Abraham, R., Schneider, J., Vom Brocke, J.: Data governance: A conceptual framework, structured review, and research agenda. International Journal of Information Management **49**, 424–438 (Dec 2019). https://doi.org/10.1016/j.ijinfomgt.2019.07.008
2. Alhassan, I., Sammon, D., Daly, M.: Data governance activities: A comparison between scientific and practice-oriented literature. Journal of Enterprise Information Management **31**(2), 300–316 (Mar 2018). https://doi.org/10.1108/JEIM-01-2017-0007
3. Association, I.D.S.: Technical Building Blocks. https://dssc.eu/space/BBE/178422228/Technical+Building+Blocks (Jan 2025)
4. Attard, J., Brennan, R.: Challenges in Value-Driven Data Governance. In: Panetto, H., Debruyne, C., Proper, H.A., Ardagna, C.A., Roman, D., Meersman, R. (eds.) On the Move to Meaningful Internet Systems. OTM 2018 Conferences, vol. 11230, pp. 546–554. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-030-02671-4_33
5. Barrenechea, O., Mendieta, A., Armas, J., Madrid, J.M.: Data Governance Reference Model to streamline the supply chain process in SMEs. In: 2019 IEEE XXVI International Conference on Electronics, Electrical Engineering and Computing (INTERCON). pp. 1–4. IEEE, Lima, Peru (Aug 2019). https://doi.org/10.1109/INTERCON.2019.8853634
6. Bharadwaj, A., El Sawy, O.A., Pavlou, P.A., Venkatraman, N.: Digital Business Strategy: Toward a Next Generation of Insights. MIS Quarterly **37**(2), 471–482 (2013)
7. Earley, S., Henderson, D., Association, D.M. (eds.): DAMA-DMBOK: Data Management Body of Knowledge. Technics Publications, Basking Ridge, New Jersey, 2nd edition edn. (2017)
8. Franklin, M., Halevy, A., Maier, D.: From databases to dataspaces: A new abstraction for information management. SIGMOD Rec. **34**(4), 27–33 (Dec 2005). https://doi.org/10.1145/1107499.1107502

9. Iansiti, M., Levien, R.: The Keystone Advantage: What the New Dynamics of Business Ecosystems Mean for Strategy, Innovation, and Sustainability. Harvard Business Press (2004)
10. Karkošková, S.: Data Governance Model To Enhance Data Quality In Financial Institutions. Information Systems Management **40**(1), 90–110 (Jan 2023). https://doi.org/10.1080/10580530.2022.2042628
11. Otto, B.: The Evolution of Data Spaces. In: Otto, B., ten Hompel, M., Wrobel, S. (eds.) Designing Data Spaces : The Ecosystem Approach to Competitive Advantage, pp. 3–15. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-030-93975-5_1
12. Sarsfield, S.: The Data Governance Imperative. IT Governance Publishing (2009)
13. Sassanelli, C., Rossi, M., Terzi, S.: Evaluating the smart maturity of manufacturing companies along the product development process to set a PLM project roadmap. International Journal of Product Lifecycle Management **12**(3), 185–209 (Jan 2020). https://doi.org/10.1504/IJPLM.2020.109789
14. Tiwana, A.: Platform Ecosystems: Aligning Architecture, Governance, and Strategy. Newnes (2013)
15. Wang, J., Xu, C., Zhang, J., Zhong, R.: Big data analytics for intelligent manufacturing systems: A review. Journal of Manufacturing Systems **62**, 738–752 (Jan 2022). https://doi.org/10.1016/j.jmsy.2021.03.005
16. Yoo, Y., Boland, R.J., Lyytinen, K., Majchrzak, A.: Organizing for Innovation in the Digitized World. Organization Science **23**(5), 1398–1408 (2012)