Data Sharing and Data Spaces: opening new business opportunities

Elisa Rossi[1,1], Antonio Caruso[2], Mariarosaria Russo[2], Gert Breitfuss[3] and Roberto Rocca[4]
[1] Engineering Ingegneria Informatica SpA, Via Guglielmo Marconi 10, Bologna, Italy
[2] Engineering Ingegneria Informatica SpA, Centro Direzionale Torre Saverio, Napoli, Italy
[3] Know-Center GmbH, Sandgasse 36, Graz, Austria
[4] Politecnico di Milano, Via Lambruschini 4/b, Milano, Italy

Abstract
A journey through the importance of data, data sharing and Data Spaces is presented both from a technical and a business perspective, with an emphasis on how the European Commission is supporting and encouraging such technologies. The application of the Data Space concept in two EU-funded Projects (CLARUS and Circular TwAIn) is deep-dived to showcase the adaptability of the technology and its reference implementation (IDS, FIWARE) to different domains, opening new business scenarios that can be leveraged to enhance the competitiveness of the companies.

Keywords
Data Space, Data sharing, Sustainability, Circular Economy (CE), open-source

1. Introduction

The importance of data has been clear from the beginnings of digital technologies and has covered all sectors at the latest with the Big Data concept. Since that time data-driven technologies became crucial for any business as they enable better forecasting and planning by analysing Big Data. [1]. Until the first decade of 2000s, business considered mostly internal data, such as IoT and historical business data. Very few companies started to share data between a limited number of closely connected stakeholders, but always with the fear of data leakages, primary scope of not disclosing the know-how, or losing the control over proprietary data [2]. From 2010s these fears have been countered with the new concept of data sharing (i.e., data exchanged among many stakeholders even if they have no direct contact or connection) in the form of a technical approach named Data Space. The Data Space is by definition "a distributed system defined by a governance framework that enables secure and trustworthy data transactions between participants while supporting trust and data sovereignty. A Data Space is implemented by one or more infrastructures and enables one or more use cases" [3].

The European Commission was one of the first authorities to understand the breakthrough led by the Data Space concept and encouraged and supported many initiatives cross-country and cross-domain, in particular, adopted the European strategy for data [4] in 2020 which aims at creating a single market for data that will ensure Europe's global competitiveness and data sovereignty. The idea is to build Common European Data Spaces to ensure that data producers (both companies and individuals) retain control (ownership) of the data. According to the Commission Staff Working Document on Common European Data Spaces, "a common European Data Space brings together relevant data infrastructures and governance frameworks in order to facilitate data pooling and sharing" [5]. In other terms, Data Spaces will play a pivotal role in accelerating digital transformation within and across domains and supporting economic recovery plans, the future vision includes a

I-ESA 2024 12th International Conference on Interoperability for Enterprise Systems and Applications, April 10–12th, 2024, Crete, Greece
elisa.rossi@eng.it (E. Rossi); antonio.caruso@eng.it (A. Caruso); mariarosaria.russo@eng.it (M. Russo); gbreitfuss@know-center.at (G. Breitfuss); roberto.rocca@polimi.it (R. Rocca)
0000-0003-4176-4413 (R. Rocca)

European Data Space which interconnects different Data Spaces (possibly each of them domain-specific) and ensures that data are extensively shared and used while complying with EU values and regulations. In this context, it is important to mention that the Data Act [6], i.e., the regulation on harmonised rules on fair access to and use of data just entered into force in January 2024.

Many initiatives started in the past decade, the most relevant have been considered as reference for the design and implementation of the Data Space in the CLARUS project are: (i) International Data Space Association (IDSA) [7] aims to build a world in which data is easily and securely exchanged and used, to support innovation, economic growth, and people's lives improvement. The focus of IDSA is to create a secure, trusted, and interoperable platform for exchanging and utilizing data. (ii) GAIA-X [8] is a European initiative aimed at creating a sustainable and trustworthy data infrastructure. It is a cooperative effort between businesses, government agencies, and academic institutions focused on creating a data infrastructure that is secure, privacy-compliant, and trustworthy. (iii) Eclipse dataspace connector [9] implements the International Data Spaces standard (IDS) as well as relevant protocols associated with the GAIA-X project. It is free and open-source. (iv) FIWARE [10], an open-source initiative aimed at the creation of a unique cross-domain infrastructure to develop and deploy services for data exchange, sharing and analysis, lowering barriers, through the usage of common standards and languages, therefore easing the development of smart solutions and supporting organizations in their transition into smart organizations.

The following section focuses on the characterization of the Data Space technology from a business perspective as outcome of a workshop conducted within the CLARUS Project[2] Consortium. Furthermore, how the Data Space concept has been leveraged and developed in two EU-funded Projects, CLARUS and Circular TwAIn[3] as references for two core applications of the Data Space technology in the European landscape will be detailed. The two approaches proposed are made up on the same Engineering asset, the TRUE (TRUsted Engineering) Connector, an open-source component for the IDS ecosystem enabling data exchanges through different communication protocols (i.e., a multi-protocol connector).The paper concludes with reflections on possible future applications of the Data Space technology (including currently ongoing experimentations in Circular TwAIn), with a view also on the European initiatives (e.g., regulation and legislation) to support industries, governments, citizen and IT solutions providers in the adoption of the Data Space technology.

2. The Data Space as an enabler for new Circular Economy Business Scenarios

A collaborative workshop has been conducted among the CLARUS Partners, with representatives of the industries (Pilots), technology providers and business development experts. The idea was to characterize the Data Space from a wider perspective, discussing the actual problems the technology aims to solve, the alternatives and the barriers to its adoption in order to define its Unique Selling Point (USP). The relevance of data acquisition, processing and analysis is commonly considered crucial, as well as the potential of data sharing among many stakeholders, thus some concerns both societal (i.e., human factors and organizational culture to be incentivized and supported to get trust in the technologies and the ecosystem), technical (e.g., interoperability, trusted data exchange, usage and access control) and regulatory (lack of common legislation) arose: the willingness of making available proprietary data (coming from sensors, cameras but also from product and process management) hits with the multiplicity of data sources and formats (i.e., digital data are complex and fragmented) deeming to the need to be homogenized through standard ontologies, with the need of keeping the sovereignty over those data and privacy issues on sensitive data, leading to the need to prevent data being unproperly used. Further, the data sharing mechanisms should be improved from both a technical and a cybersecurity point of view, considering also the starting point of the industries, often lacking an already based IT-culture and IT-experts, thus an initial huge investment for digitalization has to be supported by a full-awareness of the importance of this step and a full alignment with other involved stakeholders in the definition of the pathway. The USP of Data Space technology has been defined as follows: “Data Space allows the implementation of secure and reliable

[2] https://clarus-project.eu/
[3] https://www.circular-twain-project.eu/

data sharing mechanisms suitable for the creation of interconnected decentralized ecosystems. Eases the collaboration between different participants, domains, and ecosystems through the implementation of standard protocols".

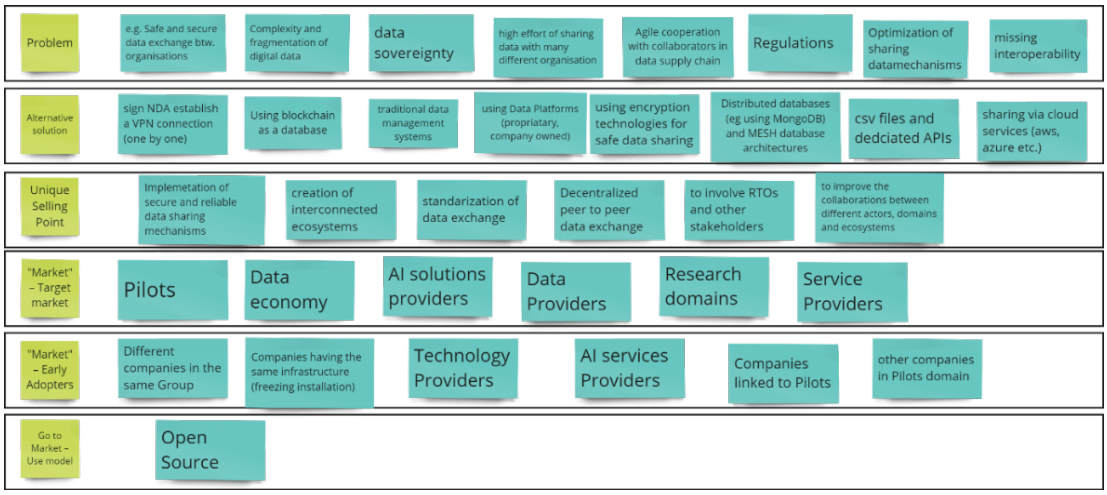

Figure 1: Collaborative board reporting on CLARUS workshop outcomes

Finally, workshop participants were requested to define the 'Target Market' and the 'Early Adopters', of interest is the discussion held with respect to the latter. In fact, industry representatives revealed during interviews that they have difficulties understanding the actual effort required for an industry (regardless the company size) to adopt a Data Space and to achieve optimization objectives (energy consumption, water consumption, waste etc.) through it, along with the necessity to impact also on others businesses digitalization process (i.e., all the stakeholders of a certain business should agree in the adoption of a Data Space, and make the respective infrastructure compliant with it). The identified early adopters are Pilots and Technology providers, as a generalization, to move the first steps on an unknown field as the Data Space is from an 'Industry perspective', the 1:1 relationship customer-supplier (Industry – IT provider): such relationship is suggested to run from a VPN/cloud solution to a Data Space, and only when the trust among the parties is established and the benefit (read as higher efficiency and lower costs with respect to previously adopted solutions) of the Data Space are recognized, the ecosystem may be enriched with the other suppliers/customers of the Industry. As well, the Open-Source nature of the most promising frameworks to build a Data Space is considered crucial. The outcomes of the workshop seem fully aligned with the Data Space scenario that is under development in CLARUS project.

3. Toward the implementation of a Data Space for Circularity

A data-driven industrial platform will be the final goal of the technical effort in CLARUS project building a Data Space for manufacturing, able to improve business opportunities for value-added services relating to the industrial data of the subjects involved in the Project (not only the Pilots but also technical partners that will play the role of Data Consumer), in order to expand their offer in terms of products and services by supporting the transition towards a circular economy.

However, the transition to a data-driven platform requires strict compliance with data processing rules and regulations, on their use as well as the ability to have full control over digital data management, i.e., the hardware and software on which a data-driven digital eco-system is built. The EU is helping European companies in this area by launching several initiatives, which aim to support them in gaining full control of the critical infrastructures for data management. The Digital Markets Act (DMA) [11] and the Digital Services Act (DSA) [12] are initiatives that provide guidance on how to create a FAIRer [13] and safer digital space for businesses, outlining the guidelines and regulations for the protection of user rights and create a level playing field to allow companies to innovate, grow and compete.

IDSA [7] is a not-for-profit association of more than 140 organisations and is the main promoter of the creation of standards for sharing data in Data Spaces, including reference models and architecture (IDS RAM), constituting standard de-facto in European data sharing landscape. Among the components that constitute a Data Space, what is pivotal to share data and to establish secured and trusted communication is a Data Space Connector. Data Space Connectors serve as instrumental tools in connecting various data endpoints, and by connecting multiple connectors, participants are allowed to freely exchange data while upholding principles of data sovereignty, transparency, and fairness through adherence to specified rules [14]. In essence, Connectors address two critical aspects: (i) offer Data Exchange Services, serving as an Application Programming Interface (API) to facilitate interoperability among participants in a Data Space. (ii) act as a trustworthy component, managing data through the implementation of policy enforcement mechanisms and a standardized baseline for cybersecurity.

In the dynamic landscape of the data-driven economy, Data Space Connectors foster interoperability, breaking down silos and creating a collaborative, trust and secure ecosystem where data flows seamlessly among stakeholders from different domains and of varied nature (as per IDS Glossary, "roles" [7]). Data Space Connectors exponentially increase the pool of available data opening avenues for businesses to access a wealth of information, fostering innovation and informed decision-making, creating new revenue streams, partnerships, and collaborations fueled by the efficient flow of information. As the inherent design of Data Space Connectors ensures data sovereignty not only meets regulatory requirements but also builds a foundation for businesses to confidently navigate the complex landscape of data governance.

In recent years Engineering[4] has developed its IDS compliant Connector: the TRUE (TRUsted Engineering) Connector. The TRUE Connector is an open-source implementation, part of the IDS open-source catalogue (i.e., compliant with IDS reference architecture RAMI3.0) and of the FIWARE Catalogue (i.e., the integration of existing FIWARE ecosystems is guaranteed by the dedicated Data APP, enabling the IDS-based interaction in a plug-and-play way). Engineering is promoting the adoption of Data Space technology supporting the scenarios of various projects through its TRUE Connector.

In the context of the CLARUS Project, the Data Space will therefore be a key element in the creation of a data economy that, involving various partners, will enable the identification of new business opportunities and new areas of application for the obtained results. In Figure 2 the graphical representation of CLARUS Data Space. The CLARUS Data Space represents a virtual bridge among all the Project stakeholders, allowing the identification of roles each plays in the generation, sharing, and utilization of data. The two pilots act as Data Providers making their datasets available to Data Consumers (Universities and Research Centers) to fuel and enhance AI algorithms, logistics support systems and, in general, decision support solutions based on AI techniques. To overcome the entirely legitimate adversity to sharing their data, which represents both value for the company and a tempting opportunity for competitors to access sensitive data, the component selected to guarantee the trust and secure data exchange within the Data Space is the IDS TRUE Connector: it ensures that data sovereignty principles are respected, and data access is governed by internal policies established among the various participants.

[4] Engineering Ingegneria Informatica SpA is a digital transformation company in Italy www.eng.it

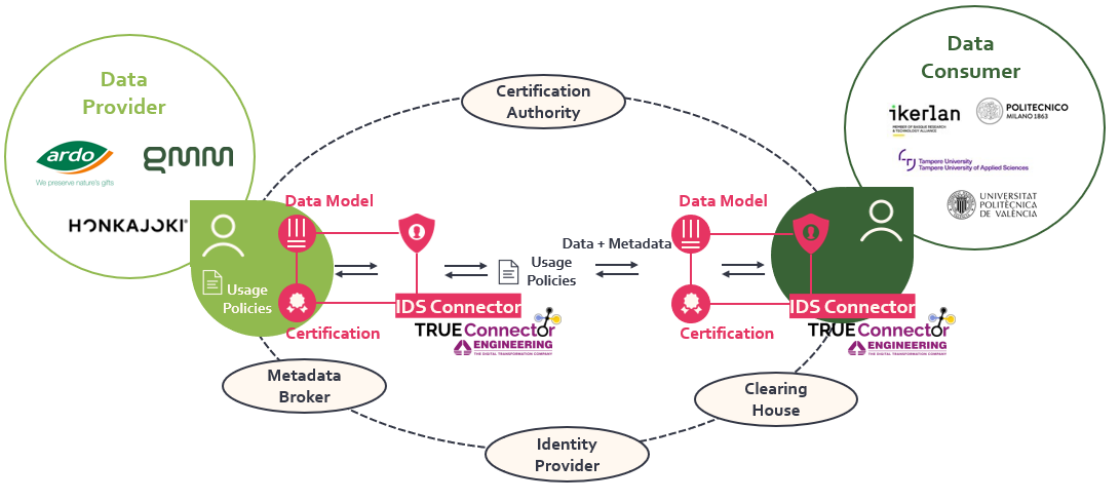

Figure 2: CLARUS Data Space implementation

CLARUS Primary food processing Pilot. The Data Space will be instrumental in achieving the main objective of adopting AI models to detect anomalies in water consumption and optimizing production and logistics operations considering energy consumption as main parameter of interest. In this scenario, various stakeholders, including water utilities, maintenance teams, production units, and logistics planners, should contribute relevant data to a decentralized Data Space. The Data Space acts as a collaborative platform where real-time and historical data on water and energy consumption, maintenance records, production schedules, and logistics operations will be integrated. In addition, the Data Space will extend beyond water-related data to include information on other energetic resources and logistics planning. By leveraging a Data Space in this manner, organizations can enhance collaboration, make informed decisions based on integrated data, and achieve more sustainable and efficient resource management, not only for water but also for other energetic resources.

CLARUS Food By-product processing Pilot. The Data Space will be the fundamental tool in enhancing the efficiency and effectiveness of the logistics operations for the company. In this scenario, the logistics data collected from slaughterhouses, trucks, and other relevant sources (including weather forecasts) will be aggregated in a Data Space facilitating real-time monitoring and allowing stakeholders to have a visual representation of container statuses in each slaughterhouse and the current positions of trucks. The AI systems, fed by the data available through the Data Space, will implement decision support systems generating optimal suggestions for the combination of collecting containers from slaughterhouses taking into account factors such as freshness (critical for the final quality of the product), transportation efficiency, and logistics optimization.

4. Conclusions

The Data Space technology has been pushed a lot by the European Commission in the last few years, and the Work Programme 2023-2024 confirms such interest through funding in diverse domains (e.g., Data Space for Agricultural, Cultural Heritage, Tourism, Energy, Green Deal, Manufacturing, Health…) [15][15], and through regulatory framework [7] aimed at covering multiple aspects in some cases interplaying with existing legislation. In fact, the European Commission proposed different regulations covering data protection laws, competition laws, and regulations on intellectual property. Thus, industries and mostly technology providers are researching on the topic trying to leverage data sharing to optimize current processes and open to new business opportunities. Despite the general uncertainty and concerns coming from the industries as depicted also during the workshop conducted in CLARUS, some other projects are actively developing solutions using Data Space associated with other core technologies such as Digital Twins, AI and Digital Product Passport (DPP). An example in this sense is given by Circular TwAIn.

The Circular TwAIn's BATTERY Pilot aims to enhance the efficiency of the current circular value chain for automotive battery systems, which currently lacks optimization in recovering functionalities and materials. The existing process involves batteries being transported to authorized

dismantlers for disassembly, where battery modules or cells are subjected to recycling treatments. The Pilot introduces innovative circular economy nodes to this framework. End-of-life batteries are first assessed for reusability, and compliant batteries are directly reused in automotive or stationary applications. Batteries not suitable for direct reuse undergo disassembly and testing at the module and cell levels. Reusable units are then reassembled into second-life batteries, while those with no residual electric properties undergo recycling for raw material recovery. The Pilot is made by various stakeholders of the Circular Value Chain (dismantler, remanufacturer and recycler) cooperating to improve the circular economy capabilities of electric vehicles batteries anticipate the European Regulation on batteries, the Digital Product Passport (DPP), that will be available in 2027 [7]. The respective information gathered by each participant are shared through the Data Space (leveraging the TRUE Connector) to create and update the Digital Product Passport, the ID Card of the battery, from which the stakeholders may understand among the others the residual life (state of health), the performance over time, insights on charge cycles and on the raw material used.

Acknowledgements

This work has been supported by the projects “CLARUS” and “Circular TwAIn”, which have received funding from the European Union’s Horizon Europe research and innovation programme, respectively under grant agreement No. 101070076 and No. 101058585.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] Fessl, A. et al. (2024). Supply Chain Data Spaces–The Next Generation of Data Sharing. In: Haber, P., Lampoltshammer, T.J., Mayr, M. (eds) Data Science—Analytics and Applications. iDSC 2023. Springer, Cham. https://doi.org/10.1007/978-3-031-42171-6_11

[2] Otto, B. (2022). The Evolution of Data Spaces. In: Otto, B., ten Hompel, M., Wrobel, S. (eds) Designing Data Spaces. Springer, Cham. https://doi.org/10.1007/978-3-030-93975-5_1

[3] A. Poikola, B. Verdonck, R. Joosten, T. Guggenberger and S. Salminen, "DSSC Glossary," https://dssc.eu/space/Glossary/176553985/DSSC+Glossary+%7C+Version+2.0+%7C+September+2023, Jan. 2024

[4] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066, Jan. 2024

[5] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0232, Jan. 2024

[6] https://digital-strategy.ec.europa.eu/en/policies/data-act, Jan. 2024

[7] https://internationaldataspaces.org/, Jan. 2024

[8] https://gaia-x.eu/, Jan. 2024

[9] https://projects.eclipse.org/projects/technology.edc, Jan. 2024

[10] https://www.fiware.org/, Jan. 2024

[11] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en, Jan. 2024

[12] https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package, Jan. 2024

[13] https://www.go-fair.org/fair-principles/, Jan. 2024

[14] Giussani G., Steinbuss S., Data Connector Report, International Data Spaces Association, (8), November 2023 https://doi.org/10.5281/zenodo.10227097

[15] https://ec.europa.eu/newsroom/dae/redirection/document/100740, Jan. 2024