GOBIERNO DE ESPAÑA
MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL

Oficina del Dato

Activity plan for deploying data spaces

v1.1 – March 2024

National Data Office

Data is, without a doubt, the main protagonist and the active backbone of digital transformation. Its essential role in the development of disruptive technologies such as Artificial Intelligence represents the difference factor for an industrial and technological revolution that will allow us to consolidate a broader, more fair and more inclusive digital economy.

Spain, in line with the objectives set in the Plan Digital Spain 2026 – specifically in its strategic axis number 4 "Data Economy and Artificial Intelligence" – seeks to position itself as a reference country in the construction of a Data Economy that has a real and effective impact on the well-being and security of citizens, as well as on economic and social progress.

The *National Data Office,* dependent on the State Secretariat for Digitalization and Artificial Intelligence, of the Ministry for Digital Transformation and the Public Service, is the unit in charge of stimulating the sharing, management and use of data throughout all productive sectors of the Spanish economy and society.

The order creating the *National Data Office* (Order ETD/803/2020) establishes among its functions the definition of methodologies and best practices that ensure the development of technological skills and the tools necessary for making decisions based on data. . The publication of this guide is developed within the aforementioned strategic function, and the dynamization and dissemination of organizational, technological and functional issues around data.

The Spanish National Data Office is the unit in charge of stimulating the sharing, management and use of data throughout all productive sectors within the Spanish economy and society.

Table of Contents

Introduction and context for the document

Data has become the great transforming power of society. Its ability to generate knowledge, drive innovation, and empower individuals and communities is undeniable. This document presents a comprehensive activity plan for the deployment of national data spaces, a key initiative to promote a data-based economy ("Data Economy") in accordance with the European principles of sovereignty, privacy, transparency, cybersecurity and competition fair.

Data spaces are collaborative environments where data sets and other data resources can be voluntarily made available, managed, exchanged and exploited, confidently, in an efficient and cyber-secure manner. They have the potential to generate very significant added value, both in economic and social terms, and this is due to their federated and distributed design, which allows capitalizing on the latent value in certain data sets and services, without losing control. on them, and without affecting the complexity or design of the rest of the information architecture of an organization.

They are an innovative paradigm, of a purely European invention, that allows solving problems in a collaborative and synergistic way, which would not be possible otherwise. For this reason, it is necessary to understand well the paradigm that gives them meaning, the context in which they are framed, the benefits they provide, and the challenges they must overcome. The ultimate objective is to promote access and exploitation of data in a capillary and cohesive way, guaranteeing equity and sustainability in the distribution of the value of data between companies, organizations, citizens and public administrations.

The document is structured in three parts:

- Part I: The definition of a data space is delved into, its principles and dimensions are detailed, and the importance of governance of these scenarios is discussed. Various models of data spaces are also described, comparing ways of organizing services and activities among their participants, and the relevant actors and technological solutions in the European panorama.
- Part II : The strategic axes of a plan for the deployment of data spaces in Spain are presented, which guide their development based on the established strategic objectives. The four axes cover the enabling

elements of viability and sustainability, reliable data governance, driving efforts in key industrial sectors, and the determining role that public administrations can play.

- Part III: Key heuristic considerations for the practical deployment of data spaces are explored, including their initial conception and the generation of the necessary conditions, the minimum functionalities that must be provided, and those governance, quality and interconnection considerations that serve to equip them. useful. In addition, perspectives regarding privacy are discussed and a list of requirements to take into account is provided.

This document is a useful tool for all those interested in the implementation of data spaces, and how they can contribute to the creation of value in the aforementioned sovereign Data Economy. Throughout this document, we hope to provide strategic and practical guidance to address an effective deployment of data spaces in Spain, outlining the potential of data towards a more efficient, fruitful and resilient production model.

Likewise, as the needs, problems and unknowns related to the design and operationalization of data spaces are explored in greater depth, new questions should arise, which may even lead to pivoting or slightly reestablishing previously agreed foundations. However, beyond the general issues applicable to all data spaces, the development of each of them will inevitably demonstrate different characteristics based on the specific needs of its community, its infrastructures, interest groups, actors and practices in the exchange and exploitation of data. For all these reasons, this document should not be considered a definitive guide, but rather a living document, open to feedback from the implementation of real initiatives.

Additionally, this document aims to serve as reference material for the **'Industrial Data Spaces Program'** of the *Secretary of State for Digitalization and Artificial Intelligence*, which seeks to promote the development of data space projects in strategic sectors of the economy. national. The program is part of investment 1 of Component 12 of the *Recovery, Transformation and Resilience Plan* , financed through the Recovery and Resilience Mechanism (MRR), and is made tangible in the form of aid that aims to partially subsidize the development of cybersecure spaces for sharing and exploiting data resources, interoperable between the main strategic sectors of the Spanish economy, and also aligned with similar European initiatives. The objectives of this strategic plan for the national deployment of data spaces will be addressed, as we have seen, in Part II of this document.

Likewise, the Program includes among its measures the **call for aid for experimental development projects** (*Order TDF/1461/2023 , of December 29*), which will be structured in two differentiated but complementary lines of action: the creation of demonstrator centers, and the promotion of use cases. Fundable projects, aligned with the strategy defined in the Digital Spain Agenda , will also find accommodation under this guide, whose design responds to facilitating the deployment of such initiatives. This plan relates methodologies, artifacts and various considerations - with which to set up data spaces - with those aspects of the projects presented that are susceptible to technical evaluation (according to article *10* of the regulatory bases approved in the aforementioned ministerial order), so that it **serves as prescriptive and support material for those organizations that apply to the call**.

PART I. Concept of a data space

This first part of the document addresses the concept of data space, defining its meaning, framing it in the European data strategy, establishing its principles and detailing its different dimensions (business, legal, technological, functional and operational), addressing its governance and examining how theoretical models materialize in different architectures, technological solutions and in the practice of data space itself.

1. Data space as a value-generation ecosystem

Data has become the great transforming power of society. Beyond a purely mercantilist vision, their ability to generate and exploit knowledge, promote innovation and empower those individuals and communities that choose to work with them is undeniable. In fact, it constitutes a resource with which to address, from a disruptive perspective, the great environmental, industrial, administrative, and social and health challenges, facilitating collaboration between actors, promoting innovation and improving the necessary accountability to which any democratic government aspires.

Data, and its essential role in the development of disruptive technologies such as Artificial Intelligence, represent the differentiating factor of an industrial and technological revolution that will allow us to consolidate a high-impact digital economy, fairer and more inclusive, and aligned with the United Nations Sustainable Development Goals. It shapes a Data Economy with dynamism and with the vocation to nurture, with all the necessary guarantees, the articulation of two key and strategic processes for Spain: the digital transformation and the ecological transition.

The *European Data Strategy* [1] aims to strengthen and promote the Single Digital Market [2], encouraging the creation of data ecosystems that promote collaboration and avoid the concentration of power in the market. The strategy focuses on developing innovative capabilities and taking advantage of the potential of data, establishing the appropriate connections between it and cloud services, as well as protecting European principles and digital rights. Common European data spaces are a fundamental piece in the deployment of the aforementioned strategy, as enablers of the Data Economy.

Following European guidelines, the challenge at national level is to encourage the circulation of data for common benefit, creating common and interoperable data spaces that serve to logically (not necessarily physically) consolidate data sets relevant to key socioeconomic sectors. These data spaces are therefore a key piece for the development of the long-awaited Data Economy, by enabling its access, exchange and legitimate reuse, positioning data as a non-rival resource with high added value, and whose usefulness grows as that its use becomes widespread (in a clear example of the network effect).

Thus, the aim is to move towards a collectivization of the value generated in data ecosystems. Ecosystems that have as pillars the creation of community, transparency and trust, *data-driven* innovation, and the ability to scale the synergistic generation of benefits (which logically should be shared). In contrast to platform models — more common today, and in which a large part of the value is retained in intermediation — we are committed to a more horizontal model that allows participants to maintain their autonomy, collaborating in point-to-point transactions. ready under a model of common minimum rules.

1.1. The ‘data space’ concept

A data space is an ecosystem where the voluntary sharing of its participants' data materializes within an environment of sovereignty, trust and security, established through integrated governance, legal, organizational, regulatory mechanisms. and technological. Data spaces seek to ensure high availability of

data sets and resources, for use in the economy and society, while those who generate them, custody them, or simply own their rights (companies, organizations, administrations public, and natural persons) maintain a high degree of control over them.

Thus, data spaces are conceived as (technologically) cyber-secure, (digitally) sovereign and (functionally) interoperable environments for sharing and exploiting data, where these can flow both within the different socioeconomic sectors, as well as between them, respecting in all common European standards and frameworks at the moment. The guiding principles of data spaces, in line with European principles and values, seek respect for fundamental rights under a citizen-centered approach, and influencing aspects such as cybersecurity, transparency and trust, innovation and sustainability, fair competition, fair distribution of value, and data sovereignty. Therefore, the development of data spaces not only requires the implementation of appropriate technological infrastructures, but also the design of conducive governance frameworks, something that generates important challenges from a legal point of view, especially regarding to ownership and access to data, and to liability and regulatory compliance.

Within the concept that we use to refer to a data space (and since it is a polysemic concept) there is room for so many centralized environments for agglutinating information and generating value-added services, with or without economic compensation; as innovative federated environments for sharing and exploiting data sets and resources (including computational resources, something not natively associated with the data). In fact, the former can be seen as fundamental pieces of the construction of the latter, and interoperability must be sought between the solutions deployed. Additionally, its constitution would also include those dynamics and initiatives aimed at generating community and reusable public goods ("digital public goods"), which cover both common use semantics, open data provided by administrations or even by private organizations, as different constructive technological components in repositories of free access and reuse.

In the *European Data Strategy*, the European Commission is committed to promoting the development of common and interoperable data spaces, in strategic economic sectors and areas of public interest, and always in collaboration with its most relevant actors [1]. Thus, there is talk of the conception of common European data spaces in a wide range of sectors and domains. Specifically, the industrial and manufacturing, the environmental (under the Green Pact), the mobility, the health, the financial, the energy, the agriculture, the public administrations, the data space of the skills and learning, language, cultural heritage, media, construction and tourism, as well as the European Open Science Cloud. Each of these spaces seeks to exploit the potential of data to drive significant advances in their respective disciplines, promoting transparency, competitiveness and sustainability, while ensuring compliance with relevant legislation in each sector.

There are different initiatives, architectures and reference models for the creation and operation of data spaces in Europe. Organizations like International Data Spaces Association (IDSA), the FIWARE Foundation, the Gaia-X European Association for Data & Cloud, the iSHARE Foundation, and the Big Data Value Association are in the process of confluence since the formalization in autumn 2021 of the Data Spaces Business Alliance (DSBA). Likewise, the process of technological convergence is in full development both based on actions with public means (e.g. through the Simpl initiative, or the financing of the Data Spaces Support Center) and privately (based on efforts deployed in the within the aforementioned DSBA), and seeks to bring together the principles and architecture of Gaia-X with the architectural model of IDSA, and with the pre-existing building blocks of the FIWARE community.

Therefore, the pace of deployment of the different data spaces, as well as the interoperability between them (as a proxy for the realization of a prosperous and sovereign Data Economy), will depend on the agility of those initiatives and organizations involved in the design and delivery of a transversal and consistent

definition of *data space*, which -along with a pragmatic framework of action- serves to take them from theory to practice. However, the challenge is not simple: the perspectives and approaches from which to address them are diverse, which is why it is critical to continue working on adequate consolidation, in order to, taking advantage of the current technological, economic, political and social moment, achieve an economy oriented up to date, sustainable, inclusive, and generating social value.

1.2. Benefits from data spaces

Data spaces are fundamental catalysts for the growth and competitiveness of productive sectors, based on the correct performance of digital capabilities and innovations, on which the use of data provides numerous and attractive advantages. Faced with an individual *modus operandi* and often in isolation, these platforms allow the creation of strategic alliances and the cogeneration of innovative digital business models, thanks to the exchange of information between different actors, which would not otherwise be feasible.

The collection and analysis of data along value chains can translate into optimization of processes in these chains, identifying opportunities, as well as eliminating inefficiencies. This leads to higher quality and added value services, both for end customers (B2C) and between the organizations themselves (B2B), allowing the anticipation of needs, improvements, personalization of products and services, or the creation of new models. and systems. For example, data spaces serve to enhance demand prediction, by allowing a better understanding of consumption patterns, facilitating more precise and efficient planning of production, inventory management, or distribution. By offering access to a wide range of (baseline resources with which to generate better) information, data spaces improve the productivity and competitiveness of organizations and companies. These can generate competitive advantages by using data to identify areas for improvement and make informed decisions.

Also, in terms of the circular economy, data spaces facilitate product traceability, waste management and the supply chain, thus promoting reuse and recycling instead of disposal. Likewise, these spaces allow greater resilience in sectors sensitive to economic and health changes, such as health, tourism, or logistics. Access to up-to-date data and real-time analysis enables a better understanding of trends and risks, facilitating agile, data-driven decision-making to minimize negative impacts and maximize opportunities in changing situations. Likewise, these spaces promote energy efficiency and resource optimization, by allowing the monitoring and analysis of energy consumption and the use of resources in industrial processes. This translates into a more modern and sustainable industry, which adopts innovative technologies and respects the environment.

Finally, data spaces foster international collaboration and leadership in strategic sectors. By sharing data and knowledge, organizations can anticipate and strengthen their position in the international market, increasing their capacity for innovation and facing challenges together.

2. Data context and diagnosis

2.1. Forecast for the Data Economy

The Data Economy extends beyond the mere supply and demand of data sets; It also focuses on quantifying the comprehensive impact that the data market has on the Economy as a whole. This holistic approach encompasses not only those tangible elements such as the generation, collection, distribution, storage, processing and transformation, analysis, and/or delivery of products made from data, but also considers the products and services necessary for their exploitation. But even further, its scope transcends these technological operations to encompass all the influence that data has throughout all sectors and activities,

from strategic decision making to operational performance, through innovation or development. of new business opportunities. Ultimately, the Data Economy evaluates how data – as a modern form of capital – can transform and propel the economy in all its dimensions and aspects.

In this sense, the latest figures collected in the *European Data Market Study 2021-2023* [3], a recurring analysis by the European Commission, estimate that the value of the Data Economy for the EU27 reached more than €455 billion in 2021 and €496 billion in 2022, with an annual growth rate of 8.9% in 2022. The participation of the Data Economy in the GDP of the EU27 will range between 3.7% in 2021 and 5.4% in 2030. And an important contribution to this growth of the Data Economy are the reforms brought about by the programs national recovery, transformation, and resilience measures, as well as investment measures channeled through NextGenerationEU funds, which will allow Member States to deploy a series of resources and capabilities in key technologies such as cloud-to-edge computing, 5G, smart devices and the *Internet of Things* (IoT), Artificial Intelligence, or data spaces.

By 2025, European data initiatives (including data regulation) are expected to be producing significant results: a strong increase in the value of the Data Economy as part of the overall economy is estimated, with the impact of the former increasing by 1 % in terms of GDP. Likewise, in 2030, the Data Economy (EU27 + United Kingdom) is expected to exceed one trillion euros (€1,000,000M), with an annual growth rate between 2025 and 2030 of 5.6%. On a national level, Spain closed 2021 as the fifth strongest data market in the EU, with revenues of €6,022M and a growth of 4.4% compared to 2021. In the future, it is expected that our data market will have a growth of 7.4% for the period 2025-2030, with revenues of €8,368M in 2025 and €11,939M in 2030.

2.2. European and national strategies and initiatives

The European Data Strategy [1]seeks to make the Union a leader in a data-driven society. The creation of a single data market will allow data to flow freely both across the continent and between sectors, for the benefit of companies and organizations, researchers and public administrations, guaranteeing international competitiveness, as well as the sovereignty of the data in Europe. The common European data spaces are the catalyst for making more (and better quality) data available for use in the economy and society, keeping its participants in control, and under full respect at all times to the data. European rights and principles.

Thus, the *European Declaration on Digital Rights and Principles for the Digital Decade* [4], of the European Parliament, the Council and the Commission, influences the rights of citizens and companies in the European digital transformation process , seeking to put people and their rights at the center, support solidarity and inclusion, guarantee freedom of online choice, encourage participation in the digital public space, increase people's safety and empowerment, and promote the sustainability of the digital future. Along these same lines, we must highlight the *Digital Rights Charter* [5], prepared from the work carried out by the 'Advisory Group of Experts' constituted by the *Secretary of State for Digitalization and Artificial Intelligence*, which does not try to create new fundamental rights, but rather to outline those most relevant to the digital environment, describing rights instrumental, or auxiliary to the first.

The protection of natural persons in relation to the processing of their personal data is a fundamental right. The *General Data Protection Regulation* [6] (GDPR) is the positive embodiment of that right, offering a consistent framework for the protection of privacy in the EU and ensuring a high and uniform level of protection. The control of individuals over their personal data is a central element and, therefore, what is sought is to generate trust and legal security. The EU legal framework in the field of personal data protection is therefore a facilitating element, and not an obstacle, for the development of a Data Economy that

corresponds to the values and principles of the Union, and is the basis on which to build a European model for the capitalization of latent value in data.

In this sense, the European strategy for this purpose is also deployed throughout different legislative initiatives, among which the *Data Governance Act* [7] (DGA) and the *Data Act* [8] (DA) are of particular interest. The DGA enables a facilitating framework for data exchange, which promotes its availability and the creation of a reliable and secure environment in which to materialize new innovative services and products. Its main measures include a more extensive reuse of protected information held in the public sector (with full respect for its confidentiality); a framework for the promotion of neutral data intermediation services, as an instrument that guarantees data sovereignty; as well as mechanisms for the altruistic transfer of data (ie data donation).

The *Data Act* aims to establish harmonized rules on the equitable access and use of data, addressing imbalances in contractual relationships between providers and users regarding the ownership and use of data, also promoting its interoperability and efficient portability, as well as the guarantee of minimum conditions in data processing services.

Likewise, the application of these two regulations on general data must be understood together with the *Open Data Directive* [9], the *Digital Markets Act* [10] and the *Digital Services Act* [11], the General Data Protection Regulation [6], and the *Regulation on a framework for the free flow of non-personal data* [12], the Artificial Intelligence Act [13], cybersecurity regulations [14], the European interoperability framework [15], as well as the rest of the applicable industrial and horizontal regulations. For all these reasons, the practical application of these legislative frameworks (across the different Member States) must ensure at all times a coherent and harmonious implementation, thus avoiding fragmentation of the digital single market due to the imposition of additional requirements or restrictions.

Objetivo	Datos cubiertos	Actores regulados
Garantizar la **EQUIDAD** en la asignación del valor del dato entre los actores de la economía de datos.	Datos del sector privado, datos personales y no personales, y datos cogenerados (IoT).	Empresas, organismos del sector público, cloud y otros proveedores de servicios de procesamiento de datos.
Abordar los desequilibrios causados por el **PODER DE MERCADO** de los guardianes de datos.	Datos personales y datos del sector privado en poder de plataformas online y los originados por los usuarios.	Cloud y otros proveedores de servicios de procesamiento de datos, grandes plataformas.
Garantizar la **CONFIANZA** en las transacciones de datos.	Datos no personales públicos y privados, y datos personales puestos a disposición voluntariamente por sus titulares.	Proveedores de servicios de intermediación de datos, organismos del sector público, organizaciones de altruismo de datos.
Promover el uso de **DATOS ABIERTOS.**	Datos en un formato abierto que cualquier persona puede usar, reutilizar y compartir libremente.	Organismos del sector público, organismos de derecho público, empresas públicas, universidades.
Garantizar el **FLUJO LIBRE DE DATOS** distintos de los datos personales.	Datos no personales.	Estados miembros, autoridades competentes, usuarios profesionales.
Garantizar un alto nivel de **PROTECCIÓN DE DATOS** y la circulación de datos personales.	Datos personales.	Controlador de datos, procesador de datos, propietario de datos, DPO, autoridades de control, EDPB.
Promover un mercado competitivo o con las reglas **ESPECÍFICAS DEL SECTOR.**	Datos personales y no personales.	Particulares y sector público y privado.

Fig 1. European regulatory framework quick summary (based on DG-CNECT, European Commission).

Regarding initiatives and operational instruments within the aforementioned data strategy, the European Commission is launching a common middleware infrastructure through the Simpl initiative , through which it will acquire and deploy cloud-to-edge technical capabilities on which implement shared data spaces. Additionally, through the Digital Europe Program (DIGITAL) and other funding programs such as Horizon Europe , Health 4EU , the Galileo and Copernicus space programs , the European Commission will contract, develop and deploy common and interoperable data spaces of European level in all strategic sectors: mobility, energy, industry, health, agriculture, etc. In DIGITAL alone, it is planned to invest €2.1 billion between 2021-2027 in data spaces and Artificial Intelligence. As an example of these DIGITAL investments, CSA ("Coordination and Support Action") type projects in data spaces aim to support and promote collaboration between different interest groups (such as academic institutions, researchers, technology and R&D&i centers, companies, governments, and/or third sector organizations) in the analysis of future data spaces of the different sectors of the economy and society. Some of these CSA projects are DATES [16], DSFT (*Data Space For Tourism*) [17], DS4Skills (*Data Space For Skills*) [18], PrepDSpace4mobility [19], AGRIDATASPACE [20], GREAT (*Green Deal Data Space*) [21], or Data Space 4.0 for Manufacturing [22].

At the Spanish level, decisive work is also being done on the creation of a legal, political and financing environment conducive to the promotion and establishment of the national Data Economy, through the different initiatives detailed in the *Digital Spain 2026 strategy.* [23], and which are deployed, among others, through the National Artificial Intelligence Strategy, in the Plan for the Digitalisation of Spain's Public Administration, in the SME Digitalisation Plan, in the Digital Infrastructures and Connectivity Plan, and in the Strategy to Promote 5G technology .

All these priorities are part of an ambitious *Recovery, Transformation and Resilience Plan*, which takes advantage of the opportunity provided by the *NextGenEU funds* for its undertaking. Thus, in axis 4 of *Digital Spain 2026* there are measures focused on moving towards a Data Economy, taking advantage of opportunities with the aim that by 2026 at least 25% of companies use Big Data and Artificial Intelligence. And more specifically, *Measure 16* seeks to strengthen supercomputing technological infrastructures, and the creation of shared data infrastructures at the European level (i.e. data spaces), with which - among other ambitious goals, too - develop AI in languages. officers of our nation.

Within this list of actions, the creation of the *National Data Office within* the Government of Spain (framed within the *Secretariat of State for Digitalization and Artificial Intelligence*) should also be mentioned, which was born with a facilitating and coordinating mission, focused on strategic development and conceptual analysis of infrastructures and processes related to data and information generation, based on generalized methodologies that are easy to transfer across different sectors. The Office was formally established with the rank of 'Division' in mid-2020 (ministerial order ETD/803/2020), combining its external vision of promoting and supporting industrial sectors, with the internal vision of reinforcing the digital transformation of the Administration, always with the motivation of promoting digital strategic autonomy .

Likewise, there are also other important initiatives already deployed that contribute to laying the foundations on which to promote a transversal, fair and inclusive Data Economy. Among them, the Digital Kit can be highlighted, aimed at increasing the digital maturity of Spanish small and medium-sized companies; the National Digital Skills Plan, which seeks to guarantee the training and digital inclusion of workers and all citizens in order to promote the creation of quality jobs, reduce unemployment, increase productivity and contribute to closing gaps in gender, social and territorial; the Startup Law, focused on promoting the creation and relocation of emerging companies in Spain, and attracting talent and capital; the formulation of the *Charter of Digital Rights* [5], as a framework of reference to guarantee the rights of citizens in the new digital reality; the Quantum Spain project; as well as the various initiatives in the general framework for the

deployment of Artificial Intelligence (AI Missions, Integration of AI in value chains, the AI Cathedras, the Spain Agency for AI Supervision, or the AI Regulation Sandbox).

For all these reasons, the national development of data spaces, covered by this plan, must consider convergence with common European data spaces and infrastructures, as well as the possible complementarity and synergies with the various programs and projects underway both at European and national level, relying on them to overcome the challenges to data sharing.

2.3. Challenges for data sharing, and potential solutions

Business innovation through data sharing at scale poses a still immature business model, with significant challenges and barriers. Furthermore, there is a lack of certainty about what are the key elements that activate said data sharing. The *Support Center for Data Sharing* [24], an initiative of the European Commission[1], was one of the pioneers in this sense, characterizing various projects [25], and even developing early versions of various resources to facilitate sharing processes. Of course, there is currently no clear model that defines the success factors, but it is understood that the generalized consolidation of data from different sources is not a sustainable or scalable solution, to the extent that these sources are considered of high quality. organizational value.

This is why the 'data space' paradigm is invented; however, one of its main impediments is the lack of consensus on what a data space really is, how to set it up, and to what extent it can provide a competitive advantage. It is therefore crucial to show how this approach contributes to business improvement, which we suggest doing based on specific metrics. Also representing an obstacle are fears about a hypothetical loss of market position, if the organizations' own data were to be shared, as well as the lack of knowledge of the value that this data could have. It is therefore critical that the organization's senior management promotes adherence to projects of creation and participation in data spaces, given that only in this way will strategic traction be ensured, organizationally permeabilizing downwards within a context of open innovation that looks beyond the short term, and yet seeks early capitalization of results.

In this sense, and as a lever to mitigate uncertainty, it is also interesting to have demonstrator use cases that allow the native concepts to be specified in data spaces, usually abstract. These should illustrate the process of creating a future data space, accompanied by an attractive story that facilitates understanding. It is also important to reduce entry barriers for organizations that are less mature in data processing and exploitation. To achieve this, it is advisable to encourage the deployment of intermediaries that facilitate work with use cases and data quality.

The need to formalize methodologies, guides and/or templates on how to develop and operate a data space is equally evident, as is the relevance of having a mature reference architectural model. These *data spaces* must also deploy affordable and simple tools, with which to facilitate the entry of SMEs, and as an instrument to capitalize on the very extensive footprint of this type of companies in the national productive fabric.

Given that the problem does not seem to be purely technological, but rather one of cultural change, the most important barriers include lack of knowledge about the opportunities in data spaces, their possible business models, and doubts about the different risks (commercial, reputational, regulatory, and/or ethical) derived from data sharing, as well as the difficulty in testing these paradigms before undertaking investments. This is why public-private collaboration is configured as an essential activation lever, seeking an adequate balance between market, trust, financing and emotion.

[1] Apparently already closed, and whose website (*eudatasharing.eu*) is not accessible.

3. Principles for data spaces

Following the analysis of the political positioning present in the document of the *Joint Research Center* of the European Commission 'European Data Spaces – Scientific Insights into Data Sharing and Utilization at Scale' [26]and the abundant bibliography on the matter [27] [28] [29][30] [31] [32] [33], it is possible to identify the principles with which to define a data space as an environment of sovereignty, trust and cybersecurity on which to enable data sharing, going beyond the technological solution and covering governance, organizational, regulatory and technical aspects.

Every data space must ensure compliance with an intrinsic set of principles:

3.1. Guiding principles and legislation

1) The data space **enables the access and exchange of data** (considering the multiplicity of concurrent use of the same data set based on its *non-rival nature* , and not necessarily in reference to the exchange of its property), and promotes the **creation continuous and sustainable development of new products and services** based on the use of data, where the data space completes its transformative value chain, promoting fair competition, reducing market entry barriers, promoting innovation, and generating impact and benefit in society.
2) The data space supports **compliance with European Union legislation**, in particular personal data protection, consumer protection legislation and competition laws, as well as regulation relating to horizontal and industrial data processing. that is applicable to you.
3) The data space, as a sharing ecosystem, guarantees the **sovereignty of participants over their data**. The participant decides with whom they exchange information and under what contractual conditions they do so, emphasizing the prevention of unauthorized uses by third parties.
4) The data space fosters the **trust of participants** beyond regulatory compliance and the monitoring of European values, also generating a fair and responsible data regime, and enabling ethics in data processing around the values of objectivity, cooperation, participation, connection, integrity, transparency, social responsibility, equity and sustainability.
5) The data space pays special attention to all aspects related to its **cybersecurity**, influencing it from the design. In this sense, participant identification systems are established, as well as the suitability of the software components used through homologation or certification mechanisms.

3.2. Participants and roles

6) The data space is conceptually **open to the participation** of any actor in the ecosystem committed to complying with its lines of operation, from the public or private sector to the third sector. In particular, a specific data space will guide its operation by an explicit governance code, guaranteeing its greater transparency, with special attention to preserving the equity of participants and their non-discrimination, as well as its sustainability over time.
7) The data space allows the various participants to **perform different roles** within it, whether as data producers, consumers, or providers and/or consumers of data services, component developers, operators of essential services, facilitators, or guarantors of the governance of the system. The data space also facilitates the role of 'data intermediary', which facilitates the appropriate intermediation services, and in particular will enable the creation of the figure of the **trusted third party** where - if necessary - the processing of information can be carried out according to the commitment agreed upon by the parties involved.

3.3. Governance

8) The governance code of a data space will enable the establishment of **access and use policies** for data resources, through corresponding agreements. It may establish economic compensation or other types of incentives, taking into account in any case its proportionality, in accordance with the provisions present in the recent *Data Act*. Electronic processing will also take precedence over traditional paper contracts.

9) The operation of the data space as a whole, as well as the use given to the data by the different participants, may **be subject to external audit,** in particular in everything related to the verification of compliance with the incorporation policies and the agreements of access and use of data resources. To do this, timely traces of the operation should be generated, guaranteeing the access of each participant to their corresponding part of information.

10) Cases of conflict may be managed through **conflict resolution mechanisms** stipulated in the governance code itself.

11) The data space will make it easier to find, access and use data under *FAIR principles* [34] (findable, accessible, interchangeable, reusable). To do this, it will sufficiently describe the data sets involved and their use restrictions, the data structures used, the vocabularies and taxonomies, as well as the technical means of access. Said description will be made, to the extent possible, following the instructions of the standardization bodies or, in their absence, in accordance with the common criteria dictated by the European institutions for interoperability.

12) The data space will promote intensive use of **open data sources** (both those provided by the public sector and by private initiatives) in application of the most typical general licensing conditions, or other specific contractual conditions. In particular, the potential of High-Value Data Sets (*HVDS*) will be considered.

13) The data space will have mechanisms to ensure transparency regarding the level of **quality of the data** provided by participants.

14) The data space will facilitate and promote the **development and use of common semantics and vocabularies**, facilitators of exchange and its exploitation. The maintenance of these semantics must be observed by the operating consortium.

15) The data space will ensure **interoperability and synergies with other** national and/or European industrial **data spaces**.

3.4. Technology

16) The data space must aim for **independence from the underlying technological solution**, allowing its portability and deployment across different infrastructures. In this way, the generation of an innovative ecosystem is encouraged, promoting creativity and competition between different providers, thus contributing to the emergence of new ideas, technological improvements, and a greater offer of solutions and services in the data space.

17) The enabling software components of each data space (in the sense of "operating systems" of each *data space*) will be created under licensing schemes that promote their knowledge and reuse. They can be obtained, whenever possible, from **open source code repositories** and any adaptation made to them must be - back - accessible to the development community.

18) If it is deemed useful, the data space will make **use of innovative technologies**. Specifically, the deployment of the following technologies will be valued:

- *Advanced analytics services and tools* (descriptive, predictive, prescriptive) with which to obtain additional value from the data exchanged.
- Privacy-Enhancing Technologies (*PET*) to guarantee the appropriate treatment of protected or confidential information.
- Innovative *federated learning models***,** facilitating the execution of knowledge extraction algorithms from different participants, and without requiring direct access or copying of raw data.
- *Distributed ledger technologies* to guarantee both the ownership of the data itself and its access rights, seeking to avoid the dilution of its value.
- Smart contract technologies and other similar ones, to carry out the acceptance of the contract *of adhesion to the* online data space, as well as the negotiation and compliance with the conditions of use of the data sets and resources.
- *"Digital wallet" type technologies* (or *Personal Information Management Systems*) where people, both natural and legal, will be able to identify themselves digitally and decide on the use made of the information provided.

4. Data spaces' dimensions

When talking about data spaces, it is essential to keep in mind that its implementation, and above all its operation, scalability and sustainability, goes far beyond the technological solution, also involving various organizational and business aspects. To address them in a comprehensive manner, ensuring that the contour needs are met to carry out a correct sharing of data sets and resources between participants, it is necessary to analyze these needs along the business, legal, operational, functional dimensions. , and technological.

The *BLOFT Model* (Business, Legal, Operational, Functional, Technical) of the Dutch Data Sharing Coalition allows the aforementioned dimensions to be covered, having been developed based on previous experience in creating trusted data sharing frameworks [35]. Each dimension thus represents a set of specific considerations to take into account in the design and management of a data space. It is important, however, to highlight that, in practice, the separation between these dimensions is not always perfect, due to the overlap that exists between them, since it is possible to approach them from different perspectives or modeling schemes.

And since there are various approaches available, it is pertinent to establish a single methodology as a fundamental reference point to guide the conversation and, at the same time, to provide usefulness in the practical configuration of data spaces. Below, what should be addressed in each of the aforementioned dimensions is indicated, detailing the typical products present in them.

- **Business dimension** (*What for?*). In this dimension, questions arise regarding how the data space is going to create value, who will be the actors involved and how they will interact, or what business models will be followed by both the promoter of the data space and the actors involved in it. same. It focuses on formulating the general objectives of the data space, creating value and identifying the paths to achieve it. In addition, topics such as ethical principles, definition of roles and responsibilities, risk assessment and definition of acceptable thresholds, financing models, or brand image are addressed.

  Typical products present in this dimension will include:
  - Analysis of the industrial state of the art.
  - Data space strategic plan (includes purpose and vision of the data space).
  - Economic and commercial management plan.

- Sustainability and business scalability plan.
- Risk management plan.
- Social responsibility plan.
- Communication and brand management plan.
- Role definition document.
- Plan to monitor and control the value optimization process.

- **Legal or regulatory dimension** (*With what regulation or policies?*). In this dimension, questions arise about what regulatory aspects are relevant in the configuration and use of the data space, or how to ensure that the data space takes into account European values and those of the ecosystem that will make use of it. It focuses on the legal and regulatory framework in which the data space operates, and contemplates compliance with applicable laws, standards and regulations, legal risk management, preservation of competition law, the definition of contractual principles and responsibilities. associated in relation to collaboration between interested parties, the governance of who can or cannot access, and how, or certifications of processes and technologies, as well as the framework of sanctions to be applied.

  Typical products present in this dimension will include:
  - Report on applicable legislation.
  - Data space governance bodies.
  - Data space constitution act.
  - Data sovereignty and access governance plan.
  - Code of conduct and ethical principles.
  - Document of contractual principles and responsibilities.
  - Modeling contracts between the parties.
  - Cybersecurity plan.
  - Report on detection and management of legal risks.

- **Operational dimension** *(How is it managed?*). This dimension answers questions about what the operations framework of the data space is, including what day-to-day activities and processes are essential to maintain and evolve it. Its scope therefore covers those practical and daily aspects of the operation of the data space, and includes compliance with established standards, monitoring and control of operational activities (process management, incident management, change management, conflicts), the allocation of resources and talent management, the realization of the business model defined for the data space, as well as the management of service level agreements between the parties. This dimension includes project management actions and the use of support tools.

  Typical products present in this dimension will include:
  - Policy and management plan for the service offered by the data space:
    - Service level agreements.
    - Responsibility assignment matrix (RASCI).
    - Communication management matrix.
    - Risk control matrix.
    - Change control matrix.
    - Incident control matrix.
    - Agreement management plan.
    - Certification management plan.
  - Resource management plan (both material and human).
  - Billing.

- Reports on effective operation of the data space, including verification of service level agreements.
- Interoperability evaluation reports.
- Deployment of support tools.
- Communication management matrix.
- Support tools user manual.
- Report on lessons learned.

- **Functional dimension** (*What services or functionalities are offered?*). This dimension seeks to answer questions about what functionality the data space should offer, what services should be deployed, and what its functional and non-functional requirements are, satisfying the conditions of the business and regulatory dimensions. The services deployed include those accessible to the final participant, and those necessary for the correct performance of the data space, as support for the former. Each service therefore has its own definition, objectives, processes and metrics, and contribute to the achievement of the general objectives of the data space.

  Typical products present in this dimension will include:
  - Definition of functional scope.
  - List of data space requirements.
  - Logical model of interaction between roles.
  - Reports on the value of the data and the value generated.
  - Functional document on interoperability.
  - Functional document on data sovereignty and reuse conditions.
  - Functional document on security and privacy assurance.
  - Semantics, vocabularies and data models
  - Cataloging and metadata management.
  - UX interaction model in data space.
  - data consumption capabilities.
  - Interconnection plan between data spaces.

- **Technological dimension** (*With what technologies and tools?*). This dimension seeks to answer questions about what technological specifications satisfy the conditions of the rest of the dimensions, what international standards and reference frameworks must be taken into account, what technical construction blocks are required and how they are materialized in the appropriate infrastructure. hardware and software, with special attention to security considerations, scalability, sustainability and technological interoperability.

  Typical products to be developed in this dimension will include:
  - Definition of the selected digital systems and infrastructures.
  - Manual for the implementation and deployment of technologies (infrastructure and systems).
  - Evolutionary development planning of the data space (development *roadmaps).*
  - Administration procedures and technological support of the data space.

Harmonious progress in all these dimensions allows us to create a robust and collaborative ecosystem, where data is shared reliably, used in an innovative way to drive informed decision-making, as well as encourage the creation of new products and services. Each dimension addresses key aspects that must be addressed to ensure the success of a data space, while the products mentioned are examples of typical deliverables in each dimension.

5. Agreements needed for making a data space happen: the trust framework

5.1. *Trust framework* for a data space

Data spaces are complex systems that can have a wide variety of actors and roles that perform a wide spectrum of activities that, in turn, are subject to specific boundary conditions. Both the activities and their conditions must be reflected in a set of agreements that, ultimately, constitute the true trust framework of the data space because they define the scope of action of each of the participants in the data space.

From this point of view, the *trust model* or *trust framework* of a data space is the set of agreements or decisions that define the data space as a playing field for the development of the Data Economy therein. The term *agreement or decision* in this context applies to any element of the data space that ultimately allows and guarantees interoperability, trust, data sovereignty, and the capitalization of the value of the data among its participants: from technological standards to the definition of roles and business models, including the functionalities offered to potential participants, the operating conditions, the legal framework or the chosen governance model.

Establishing these decisions requires carrying out a prior analysis of the data space that you want to create, so that its features are defined from different perspectives. These definitions will ultimately take shape in the decisions and agreements that constitute the aforementioned trust framework. To carry out this analysis task, it is necessary to use tools that provide a comprehensive understanding of the system from multiple angles. One of the most powerful models, and the one used in this document, is the BLOFT dimension model.

In this section, a general review will be made of the nature of the decisions that affect the functioning of a data space, based on the application of the BLOFT model. This model, widely used in this area, provides a holistic view of a data space as a service from multiple perspectives, making it suitable for the analysis and collection of the characteristics of the agreements that will ultimately be reflected in the *trust framework*.

5.2. Building blocks for data spaces, according to the BLOFT model

As we have seen in the section above, the BLOFT model contemplates 5 dimensions or analysis perspectives for data spaces:

a) Business
b) Legal
c) Operational
d) Functional
e) Technological

The approach offered by the BLOFT model is comprehensive, so the result of its application is a complete understanding of the multiple facets of a data space. These facets can also be modeled at a more granular level of detail, based on conceptual components linked to functionalities or capabilities in a data space, and as atomic and orthogonal to each other as possible, so that their rationalization can be undertaken independently (without generating an interrelation of choice between them).

Therefore, in this document we call these atomic (or quasi-atomic) conceptualizations composable with each other a “building block”, and we propose the following as an evolution of the OpenDEI model [36]. The proposal is the result of a consolidation exercise of several previous references from European experts and interest groups, including the 'Data-Sharing canvas' of the *Data Sharing Coalition* [27], the characterization provided by the *i4Trust initiative* [30], the report of the Geonovum committee of the *National Spatial Data Infrastructure* (NSDI) of the Netherlands [37], version 0.5 of the *Blueprint for data spaces* of the DSSC [38], or the work of the *European Open Science Cloud* (EOSC).

Our characterization of *building blocks* strictly follows the dimensions of the BLOFT model, since we believe that this perspective provides a simplified understanding, and that they are also the result of pioneering practical experiences [35]. However, it should be noted that our proposal is only an additional perspective on those already available throughout the most recent European literature. We also understand that any model in this sense is subject to the publication of version 1.0 of the 'blueprint' for data spaces of the *Data Spaces Support Centre*, expected in spring 2024.

Fig. 2. Building blocks according to the BLOFT dimensional model.

5.2.1. Business building blocks

This dimension refers to the business models around a data space, which we can characterize among those of the different participants around data-driven use cases, those that make data products and services available, and those that intermediate. But there is also the perspective and feasibility analysis of the data space itself as a final service (as a complete set), which is complementary.

Fig 3. Building business blocks, according to the DSSC Blueprint v0.5.

- *Data space service business model:* It is crucial for the development and sustainability of the data space, as it defines how it creates and offers value to its participants, as well as how it captures value to maintain itself in the long term and even scale and evolve. The *promoter of the data space will be responsible for* its start-up and operation, and its business model may be based on periodic contributions from the rest of the participants or on a percentage of transactions, for example.

- *Use cases development:* It is a strategic approach to amplify the value of a data space by encouraging the creation, support and scaling of use cases. Use cases depend on the data space to solve business challenges or create value for one or more participants. Use cases generate demand and leads for the data space, and data space facilitates the profitable implementation of those use cases. Each of them connects more participants and data products to the data space, thus generating crucial synergies for different business models. A user role in this development of use cases will be the consumer of data and services, whose business model will be based on the benefits that the specific data sharing use case brings. Sometimes, the use of a data service is not subject to economic compensation, although this is one of the possible business models.
- *Provision of data products and services:* This role provides the data products and services themselves. Their business model is very dependent on the use case. It may be based on contributions from consumers interested in the product or service, but there must not necessarily be economic compensation; This is, for example, the case of consortia of providers and consumers of data products that design and execute sharing use cases to obtain a derived benefit[2], in which case, the costs associated with the provision of the service are covered by indirect benefits obtained from the sharing. Therefore, this role will also be closely linked to the development of use cases.
- *Data intermediation:* Intermediation service providers provide value- added services necessary for the operation of data space, but which are not data services themselves. Examples of the former include a service catalog, or a vocabularies repository and description hub. Typically, the data intermediation business model will be based on contributions from the participants who use said services, or on contracts signed with the promoter of the data space, who would be interested in this type of services being offered efficiently and efficiently. Ultimately, brokers can specialize in delivering the same services to multiple data spaces.

The possible range of roles is not restricted to those mentioned. There may be many others, such as technology providers, support officers, or development communities. In any case, it is the governance of the data space that is responsible for defining the roles based on their purpose, and analyzing where the business interest lies for each of them to ensure that the conditions of their viability are met.

5.2.2. Legal building blocks

Fig. 4. Building Legal blocks according to the DSSC Blueprint v0.5.

Legal Building Blocks offer guidance and resources to ensure regulatory compliance in and of data spaces, establishing a robust contractual framework. This selection includes the following 2 blocks:

- *Regulatory Compliance*: The goal is to clarify legal requirements in data spaces, identify relationships between different frameworks (i.e., provide guidance when multiple legal frameworks overlap or

[2] This is the case, for example, of the consortium of manufacturers and suppliers of products and services to the automotive industry that makes up Catena-X.

intersect), and ensure that responsible data sharing and innovation is achieved. articulated around data spaces is done based on coherent legal frameworks. This identification and application of current regulations regarding data sharing can be divided into two areas:

- Global data regulations: It includes the regulations applicable to all types of data, such as the data protection regulations (GDPR), the provisions of the Data Act, Data Governance Act, eIDAS, or the regulations for the free circulation of non-personal data in the European Union. It also includes general regulations such as antitrust laws, or regulations applicable to information society services.
- Sector regulations: Covers the specific regulations that apply to the sector to which the data corresponds, such as specific regulations for health or mobility data.

- *Contractual framework*: This relates to the contractual framework that underpins the data spaces, as well as the contractual resources made available to participants to facilitate transactions within the data space. This includes the contractual relationships between the participants, whatever their role and the reason for said relationship, covering both those between the users of the data space (data providers and consumers), as well as those existing between them with the data service providers. added value (e.g., catalogues, services to guarantee trust, or any other activity), and also of these with those responsible for the data space.

Therefore, under this legal dimension, in addition to the purely regulatory building blocks, aspects that are related to the governance of the data space are also included, in the sense that they transversally affect several of the dimensions such as legal, or the operational one. This is reflected in the organizational building blocks of the Data Spaces Blueprint v0.5 [38], and also in the 'data-sharing canvas' of the Data-Sharing Coalition [27].

Fig 5. Governance building blocks according to DSSC Blueprint v0.5.

Fig 6. Governance within the 'Legal' scope in the DS Canvas.

- *Organizational Governance:* We call this the governance of the entire data space, understood as a service whose main value proposition is to guarantee interoperability between its participants, as a mechanism to facilitate the sharing of data resources. This therefore covers key points in decision-making for the establishment and effective operation of a data space. Organizational governance involves coordination, decision-making, and goal-setting within data spaces that promote collaboration among multiple stakeholders.

  Data Space Authority' must be established, which will act as the entity responsible for operationalizing and maintaining such governance. One of the first steps after defining a representative authority is for it to establish, maintain and enforce a governance framework that defines the scope, rules, roles and responsibilities of the data space, providing the opportunity for interested parties review and contribute to it. Once established, these rules must be documented in a book, for operational use (it is important to note that the management and operation of this organizational governance - already defined - will fall under the scope of the operational dimension). The data space authority will also intervene in the management of disputes between participants.

- *Peer-to-peer data sharing governance:* Well-defined rules for data sharing help participants from different backgrounds exchange data with complete trust. To this end, it is relevant that participants can

maintain control over who accesses their data, for what purposes and under what terms. There are several ways to organize data exchange within a data space, and therefore the governance authority will need to choose how the different services that allow the exchange of data sets and services will be established. These rules may include requirements to verify participants (e.g., secure identification) or establish requirements for the products and services available therein (e.g., language, data formats, etc.). The governance authority must balance reducing barriers to entry (i.e., having flexible rules) and promoting interoperability and data quality (strict rules). As in the previous block, once these rules are established, we catalog the activities of its management and daily operations under the operational dimension.

The BLOFT model [27] distinguishes between two types of phases for the implementation of governance:

i. Development of the trust framework[3] (Legal dimension). It consists of the definition itself of the governance agreements. These agreements are also called governance rules, since - once approved - they become the context of policies and rules of the data space itself.

ii. Management of the governance framework (Operational dimension). Once a framework of trust has been established, it is necessary to ensure its compliance and evolution, and this is done from bodies that must establish a neutral position and that will also be in charge of mediating any disputes that may arise ('Data space Authority').

More detailed information on governance in data spaces can also be found in chap. 6.

5.2.3. Operational building blocks

A data space is a service that offers consumers and providers an environment in which to consume and offer data resources. Like any other service, it is not exempt from specifying a set of agreements (with the different actors and roles, which must be in line with the defined governance framework) that reflect the conditions under which the service is provided, in order to supervise it and manage it.

Fig 7. Governance within the 'Operational' scope in the Data-Sharing Canvas.

Furthermore, the service itself also needs control and management that guarantees its correct operation, maintenance and - if necessary - escalation. In this sense, *operational governance* covers the areas of execution, control and monitoring of the regulatory and governance policies defined (for the data space) set out in the previous section. Also defining other specific actions in line with the day-to-day operations of the service, the following actions can be listed (not exclusive):

- User life cycle management. It involves defining procedures for registration, cancellation, modification of profiles and conditions of permanence in the data space.
- Service level agreements for the services offered by the technological components.

[3] Not to be confused with the concept of "trust framework" used by the Gaia-X initiative, which has some added technological connotations over the definition used here, which is more general, and speaks of the generation of optimal conditions to facilitate data sharing scenarios. data.

- Operation and support to the end user.
- Maintenance of service components.
- Management of technical incidents.
- Risk management.
- Change management, including the evolution of technological components.

In practice, operation and support, maintenance, incident management, *Service Level Agreements* (SLAs), or change management, fundamentally apply to the services offered through the technological components that support the functionalities of the service space. data (described in the *building blocks* of section 5.2.4), which together constitute the service offered by the data space itself. These functionalities can be offered by the promoter of the data space, or by authorized entities to which the promoter delegates.

Configuration Items

The last key element in establishing the trust framework is found in the management of the configuration of the data space, understanding this as the process of controlling and monitoring the elements that shape it. In turn, Configuration Items (CI) are those individually identifiable components or resources that are part of the infrastructure[4] necessary to offer the data space service (that is, the materialization of the *building blocks*) and also its maintenance and scaling.

In a data space, these configuration elements are what allow users (providers and consumers of data resources) to carry out their activities. Which specific components are CIs of a specific data space will depend on the trust framework and the agreements adopted therein. Typically, it is at least those components intended to articulate *building blocks* of sovereignty and trust, but they can also be value-added services such as catalogs or traceability services. The resources themselves (products and services) offered by suppliers do not constitute CIs, but only those components intended for the use cases to be executed.

Configuration management involves performing the following tasks:

i. Identify configuration elements (hardware, software, personnel, facilities, etc.).
ii. Assembly of said elements.
iii. Configuration record, including element status, version, location, relationship between components, and any other relevant attributes.

5.2.4. (Technical-) Functional building blocks

The fourth dimension of the BLOFT model for the characterization of a data space is the functional one, which defines the functionalities of the data space. Like other BLOFT dimensions, it is not completely independent of the others, but is largely determined by the characteristics of the others. Specifically, the business and legal dimensions, which establish the business models, as well as the regulatory and contractual framework, are determining factors for the essential functionalities that a data space must offer.

For the description of the functional dimension, we take as reference the taxonomy proposed by the *Data Spaces Support Center* (DSSC). This proposes a grouping of functionalities in what it calls technical *building blocks* and which it defines as "*basic units or components that can be implemented and combined with other building blocks to build data spaces*". In practice, *building blocks* are groupings of functionalities that can be implemented through software components.

[4] In a more general sense than computing infrastructure.

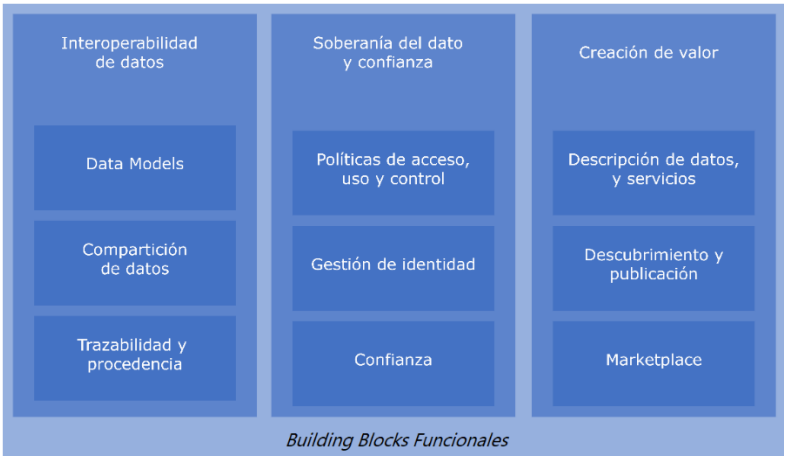

Fig. 8. Technical Building Blocks, in DSSC terminology.

In turn, the DSSC groups the technical *building blocks* into three thematic pillars:

1) Data interoperability pillar.
2) Data sovereignty and trust pillar.
3) Value creation pillar.

Each of the pillars is described in more detail below.

5.2.4.1. Data interoperability

This pillar groups together the technical *building blocks* that provide the necessary functionalities so that participants can effectively share data with each other. On this pillar, participants will build and execute use cases and, on the use cases, services that add value. Three *building blocks* are distinguished that correspond to three groups of functionalities:

- *Data models*

  This *building block* groups together the set of functionalities offered for the semantic modeling of the data that is going to be shared in order to describe it. These functionalities allow, therefore, a data provider to define and express the characteristics of the data it makes available so that potential consumers obtain a clear idea of the *dataset* they are going to download.

  It must be taken into account that data is the final product provided in a data space and that, in dynamic ecosystem-type environments, there is potentially an infinite variety of types of data sets, so there must be an agreement between participants on how to define and express the nature and characteristics of the data being shared. A set of data that cannot be described or whose description cannot be communicated because there is no previously agreed upon common language is of no use.

  The definition of the data model answers three challenges:

  - Ensure that the different actors and systems in a data space are able to define and interpret data sets consistently, accurately and without ambiguity.
  - Ensure that data can be interpreted between actors belonging to different industrial domains.
  - Manage and maintain the semantic models created.

  To do this, you must have the following capabilities:

  - Vocabularies. Vocabularies are the semantics of the data space itself. They are made up of the definition of the entities that can be part of a *dataset* along with their ontology. Vocabularies are

defined by specifying data models, schemas that allow the specified model to be reflected, and open APIs that enable participants to programmatically describe the data.

- Processes and APIs for vocabulary management. The provision and management of vocabularies in a data space is a value-added service that is in effect an intermediation service provided by a vocabulary provider. The vocabulary provider will be responsible for creating, maintaining and publishing one or more vocabularies. Typically, vocabularies have an industrial scope and are defined by communities, so the vocabulary provider is usually a delegated entity.
- Vocabulary hub. It is a technical component itself (software and hardware) that physically houses the vocabularies and offers APIs for consultation and management.

- *Data sharing*

  This *building block* provides mechanisms for effective data sharing between participants. This requires the adoption by the parties involved of a standard API that implements a protocol with capabilities to address the functionality. The functions of this protocol go beyond the mere transfer of data between provider and consumer, distinguishing in its implementation two levels of activities that correspond to the two phases of which the data sharing activity consists of: the control phase and the transfer phase.

  The control plane: responsible for the activities of the control phase. In this phase, the participants identify themselves, authenticate themselves, agree on the assets (data) to be exchanged and negotiate the conditions of sharing, including consideration, conditions of use and how the transfer will be carried out (plan of the data).

  The data plane: responsible for the data transfer itself. It begins once the control plane phase has finished. The execution conditions of this phase are agreed during the control plane phase; For example, during that phase it may be agreed that the transfer (data plane) is carried out using protocols such as NGSI-LD.

  The technical components (hardware and software) responsible for executing the API that implements the functionality are typically the connectors (*Data Space Connectors*) through which both data providers and consumers access the data space.

- *Traceability and provenance*

  This *building block* groups the functionalities necessary to record evidence of the data transactions carried out. It requires the definition of semantics that allow defining and recording transactions and an API for recording the transaction in a recording service offered by a third party. This service is typically another intermediary service within the data space. This functionality is typically implemented in the connectors and in the component in charge of offering the value-added service.

5.2.4.2. Data sovereignty and trust

This pillar of technical *building blocks* groups together the functionalities that guarantee trust between participants from the following aspects:

- *Identity management*

  The identity management *building block functionalities* provide data space participants with capabilities for self-identification and verification of the identity of other participants (authentication). Logically, these capabilities constitute the cornerstone of trust in any information system, and also in a data space.

  From an implementation point of view, identity management functionality must be implemented in all components of the data space that act on behalf of a participant, be it a data resource provider, consumer or intermediary service provider.

- *Definition of data use and access control policies, as well as compliance assurance*

  The functionalities of this *building block* make it possible to guarantee data sovereignty in the data space, this concept understood as the ability of the data owner to define who is authorized to use the data and under what conditions. This *building block* is made up of two components:

  a. A language that allows the definition of use and access policies by data owners.
  b. Mechanisms that guarantee compliance with established policies (*enforcement*).

  An example of a language for defining access and use policies is ODRL, promoted by the DSSC. The interpretation of established policies and their application typically resides in the logic of the connectors through which data consumers access data resources.

- *Trust in data space*

  This *building block* groups together the functionalities to guarantee trust, understood as the ability to make and verify *statements* [5] by participants that go beyond their own identity.

  Identification management has been previously described as one of the technical *building blocks* considered by the DSSC. Identification management is based on participants making statements about who they are (identification) in such a way that the rest of the participants can verify these statements about their identity (authentication). In this sense, identification and authentication are a particular case of declaration and verification of declarations.

  However, to ensure trust in a data space in a broader sense, it is necessary to be able to make and verify many other statements regarding the characteristics of the services, the datasets and the profile of the participants themselves. These capabilities require a specific infrastructure whose functionalities are included in this technical *building block*. It is called a trust *building block* because it enables participants in the data space to trust assertions or statements made about entities in the data space, many of which may even have legal implications. For example, a service provider may describe its data service by stating that the data resides in the EU; This fact, which has legal implications because it involves applying European regulations on, for example, data protection, must be able to be verified by the consumer of the data.

  Although the ideal would be that any participant in the data space could make any declaration and that any declaration could be verified, in practice, in order to guarantee interoperability, it is necessary to establish a common criterion on which declarations and about which entities it is possible to state a declaration, which implies the definition of a complex information model that allows sharing a common vision of the aforementioned entities and their characteristics and declaration validation mechanisms. Typically, the entities about which claims are made are the shared data, the services offered, and the participants in the data space. An initiative to determine this common criterion and the modeling of entities about which to make statements is the so-called *Trust Framework* of the Gaia-X association.

  From a technological point of view, the implementation of trust mechanisms such as the one defined by the verifiable credentials standard (*Verifiable Credentials*) and distributed identity systems (SSI) defined by the W3C is a trend. The paradigm of verifiable credentials, the credentials are created by the holder of the declaration and signed by a recognized authority (trust anchor).

[5] *Claim* is translated as *a statement* in the sense of affirmation or claim. A statement in this context is any statement that provides information about an entity such as a service, a data set, or a participant.

The implementation of the trust *building block* resides in the connectors themselves and in specific components of the trust service provider.

5.2.4.3. Data value creation

One of the fundamental objectives of a data space is to generate value for different participants through data exchange. The data model and data trust and sovereignty pillars provide basic functionality for trusted peer *-to-peer data sharing*, but data sharing goes beyond having capabilities to interpret and transfer data. data ('Data Interoperability' pillar) and to ensure trust. In fact, data is made available by providers in the form of products and services that must be described along with their conditions of use and their possible consideration, for example. In other words, the final products and services go beyond the mere transfer of data (for example, they may offer, together with the data, *dataset* cleaning services or the application of privacy technologies -PET-). At the same time, the search and location of the offer of products and services available in the data space must be made possible (for example, through catalogues) together with their descriptions and there must also be mechanisms to, for example, financially compensate suppliers.

In other words, the first two pillars provide basic functionalities for describing and sharing data with confidence, but an additional pillar is necessary to integrate that activity into an environment of providers and consumers of products and services through additional functionalities. This task is what provides the value creation pillar, which is based on the functionalities of the first two pillars and on them (data description, statements, etc.) provides these additional functionalities that turn the data sharing activity into an environment for creating and consuming data products and services. The value creation pillar consists of three *building blocks*:

- *Data Descriptions and Service Offerings*

  This pillar provides tools to describe data products and services in a way that is understandable by all participants in the data space along with data use and access policies and how the described products and services can be accessed. It is therefore a set of functionalities offered through the modeling of resources in the data space.

- *Publication and discovery*

  This *building block* allows the publication of the service offers described through the functionalities of the previous one, so that they can be located and consulted by consumers.

- *Marketplace*

  This *building block* provides functionality for peer-to-peer marketing of data products and services.

The publishing and discovery functionalities, as well as those of *the Marketplace*, are typically provided by intermediary service providers, this role being understood as that of a facilitator for the provision of data products and services that does not directly supply these products and services, but rather that provides services (a catalog, a payment gateway) that allow the provision of the former.

***Building blocks* and components**

The three pillars of functionality must be implemented in software and hardware technological components. This entails the choice of technological standards, for which the following subsection outlines some considerations. Likewise, section 6.6.1points out how the different functionalities (enabled based on technical components) also serve to model interoperability between different data spaces.

5.2.5. Technological building blocks

The technological perspective (sometimes also called 'technical') refers to the establishment of specifications, standards and technologies that will be used for the implementation of the *building blocks*

mentioned in the previous section. The *Data Spaces Support Center* offers an overview of standards and specifications that are classified according to the pillar of *building blocks* in which they are applicable. A list of these standards can be found at [39]. It should be taken into account that the data space does not have to support all the standards included in the DSSC list, but it is advisable, to the extent possible, to reuse existing standards instead of developing new ones.

Sometimes, the specifications that are part of the technical perspective are in turn sets of specifications that, together, allow the implementation of a building block. This is the case, for example, with the Gaia-X specifications. Below, we describe in greater detail the specifications applicable to each *building block pillar*.

- *Specifications, standards and technologies for data interoperability*

  Following the taxonomy of *building blocks* proposed by the DSSC for this pillar, it is possible to distinguish three types of specifications:
  - Data modeling. These include vocabularies and ontologies, as well as languages to express them (for example JSON-LD, Schema.org) and related standards.
  - Data exchange: specifications, standards and technologies for the transfer of data between participants, such as APIs for data sharing (NGSI-LD) or standards for the definition of APIS (GraphQL, OpenAPI) used for sharing.
  - Traceability and provenance. Specifications for this type of activity such as, for example, the PROV-O specification (PROV Ontology -W3C) intended to express data models that define the traceability and provenance of the data.

- *Specifications, standards and technologies for sovereignty and trust*

  In this group are those specifications for the implementation of trust and sovereignty functionalities. Specific:
  - Definition and control of access and use policies. Standards in this category are, for example, XACML or SAML. For the definition of use policies, the ODRL (W3C) standard is being imposed.
  - Identity management. Particularly relevant are OAUth, OpenID, X.509 certificates, the standard for decentralized identities DID (w3C) or the electronic wallet architecture promoted by the European Commission (ARF).
  - Trust. This section includes standards for APIs querying blockchain networks such as EBSI (for the network of the same name) or the verifiable credentials (VC) model defined by the W3C.

- *Specifications, standards and technologies for value creation services*

  This group of standards includes those standards necessary for the definition and access to data products and services. Not to be confused with standards intended to describe data sets; The latter are lower level and are part of the *building blocks pillar* defined as *data interoperability*.

  Specifications, standards and technologies of this type are, for example, those intended to define catalogs of data sets (for example, DCAT), or service offerings (Gaia-X).

5.3. Complementarity between different structural models

As we said, this structuring around atomic dimensions and functionalities (*building blocks*) is an explanatory exercise to simplify a complex map of relationships between functions, roles, stakeholders, artifacts, and methodologies.

In the following figure, from the Joint Research Center document [26], *various* aspects to be considered when establishing the governance framework are synthesized and ordered in the form of a mental map, highlighting the strategic and regulatory context; the organizational part of the system as a whole based on each of the parts; the necessary institutional coordination and regulatory supervision; the actors present; and those more technical facets. It should be noted that this JRC governance framework is not incompatible with our model above, since essentially similar elements are included, modeled in a different way.

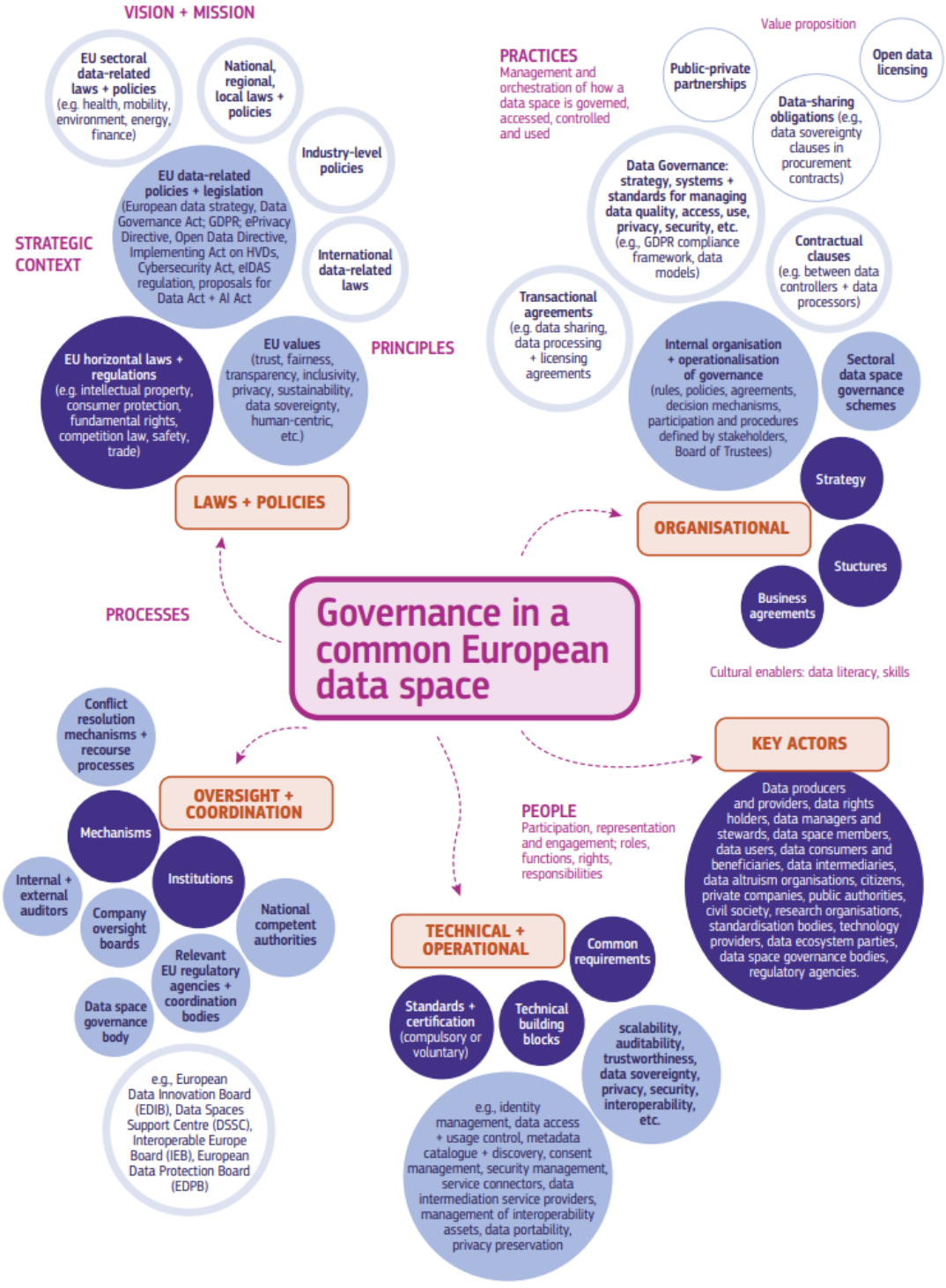

Fig 9. Governance areas within a data space (JRC – European Commission).

6. Comprehensive governance for interoperable data spaces

In this chapter, an analysis of the comprehensive governance for interoperable data spaces is carried out, understanding as such those data spaces that are not isolated, but rather allow their participants, to some extent, to offer total or partial services to data consumers. other data spaces and, likewise, consume services from providers of other data spaces.

6.1. Scope and structure for governance of data spaces

6.1.1. Governance of a data space

A data space is, in short, a service consisting of providing an environment suitable for the sharing of data between data providers and consumers and for the provision of related services such as, for example, data processing and exploitation applications or computational resources (disk space, processing capacity) necessary for the exploitation of data in the cloud. In this document, the set of services offered by the providers of a data space will be called *data resources*, whether they are data sets themselves or related services such as data exploitation applications or computational resources.

For such an environment to exist, a framework is required for governance or decision making that must be adopted during the design, creation, maintenance (and evolution), and operation of the data space service, following a certain model. In addition, there must be a recognized authority that exercises governance and guarantees compliance with the decisions made. Thus, according to the definition of the *Data Spaces Support Center* (DSSC) [40], the main task of governance of a data space is the creation of a framework made up of "*a set of principles, standards, policies (including regulations) and practices that apply to the creation, administration and operation of a data space with a certain scope, that has mechanisms to enforce compliance and to resolve conflicts"*.

Therefore, governance of a data space should be understood not as the management of the data space itself, but as the definition, maintenance, enforcement, and (potentially) evolution of a framework composed of decisions relating to:

- Data space scope
- To its operating principles (for example, that it works according to the rules of an ecosystem)
- To the application standards and specifications (for example, those adopted by the informational model of the Gaia-X architecture, or the semantic vocabularies that are applicable in a specific sector)
- To the rules and regulations for participating, including the roles and scope of each person, as well as their responsibilities.
- And to management practices (for example, how the operation, support and maintenance of the data space will be carried out).

Governance is completed with the existence of one or several bodies or authorities in charge of maintaining the aforementioned framework, supervising its compliance, enforcing decisions, and resolving potential conflicts.

6.1.2. Governance and interoperability: the European Interoperability Framework (EIF)

Logically, the governance characteristics of a service are closely linked to its very nature. In the case of a data space, it is a service offered to providers and consumers so that they can provide and consume data resources between them, through peer *-to-peer relationships* and autonomously.

It is important to consider that, if interoperability is key in any information service, it can be stated that interoperability is, in itself, the service of a data space. That is, the main mission of a data space promoter is

to guarantee an interoperable environment in which participants can offer and consume data resources through *peer-to-peer relationships*. The autonomy of the participants should be understood as the freedom to define, create and consume data resources with no other requirement than complying with interoperability rules established by the governance of the data space in which they participate. In other words: the data space promoter acts as a facilitator for the sharing of data and services, but does not provide the data resources, but merely provides the suitable environment; The resources are provided autonomously by the dataspace participants. In such an environment, it is easier to understand why interoperability is a central point. Likewise, and in line with what has been mentioned, it is the responsibility of the governance of the data space to decide what interoperability standards will apply, and to create a framework accepted and followed by all participants.

For the above reason, the development of the governance framework of a data space will be strongly guided by the set of decisions necessary for the creation of the interoperability framework that makes the service inherent to the data space possible. This makes the interoperability framework serve as a beacon of the governance framework, according to several analysts [26]: given the nature of the data space as a service, through the decisions to define its interoperability, practically all of the decisions that make up its framework will be covered. governance.

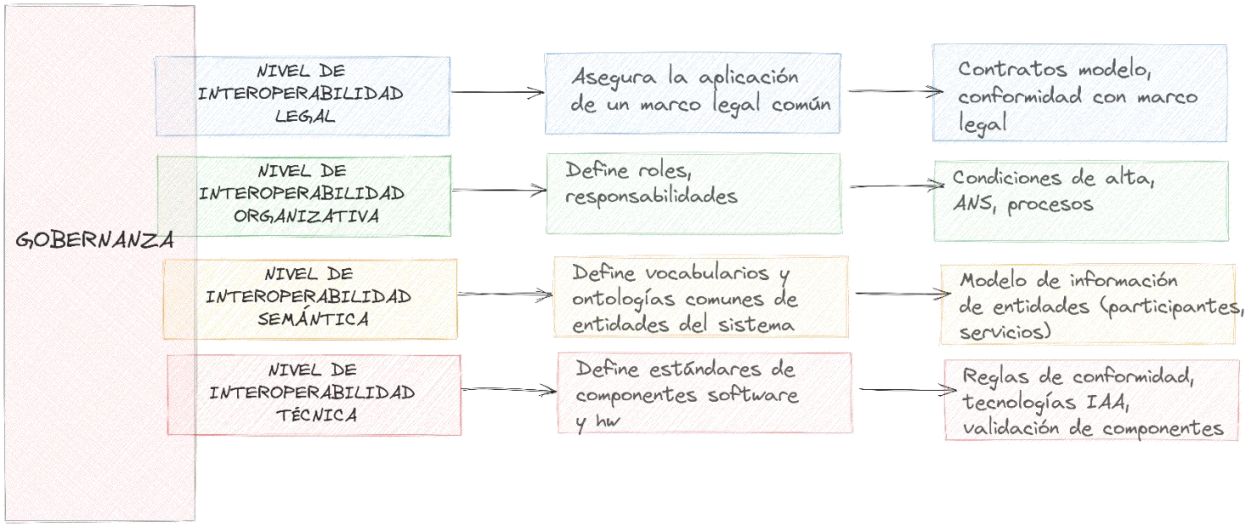

Fig 10. EIF interoperability levels.

In this sense, the criticality of interoperability - as a central objective within the data space governance framework - allows taking advantage of some of the existing tools for the analysis of the interoperability of ICT services such as, for example, the *European Interoperability Framework* (EIF [15]), which is in fact the one used in the available literature and the one followed in this document. Thus, the governance framework of a data space will consider decisions related to interoperability across 4 levels: legal, organizational, semantic, and technical.

6.1.3. Governance body structuring

The maximum degree of interoperability is offered by those data spaces that guarantee inter- and intra-operability, these terms understood as:

- Intra-operability: internal interoperability to a singular data space.
- Inter-operability: interoperability of a data space with other data spaces.

Guaranteeing both types of interoperability requires having a governance structure at 2 levels, with different but mutually dependent areas of action, each of which is conceptually governed by a separate body:

a. A (single) global governance body in charge of harmonizing interoperability between different data spaces. Generally this governance model is usually light or basic.

b. Multiple governance bodies for each of the unique data spaces, covering interoperability aspects not covered by the overarching governance body, which will typically include many more aspects or prescriptions.

The separation of governance at these two levels serves to differentiate areas of competence, and thus avoid data silos, guaranteeing interoperability between different data spaces, but at the same time allowing the creation of specialized data spaces autonomously governed by their own rules. in everything to which global governance does not apply. This multi-level structure is in fact covered by the *Regulation on harmonized standards for fair access to and use of data* (Data Act [8]), and is recommended for European data spaces.

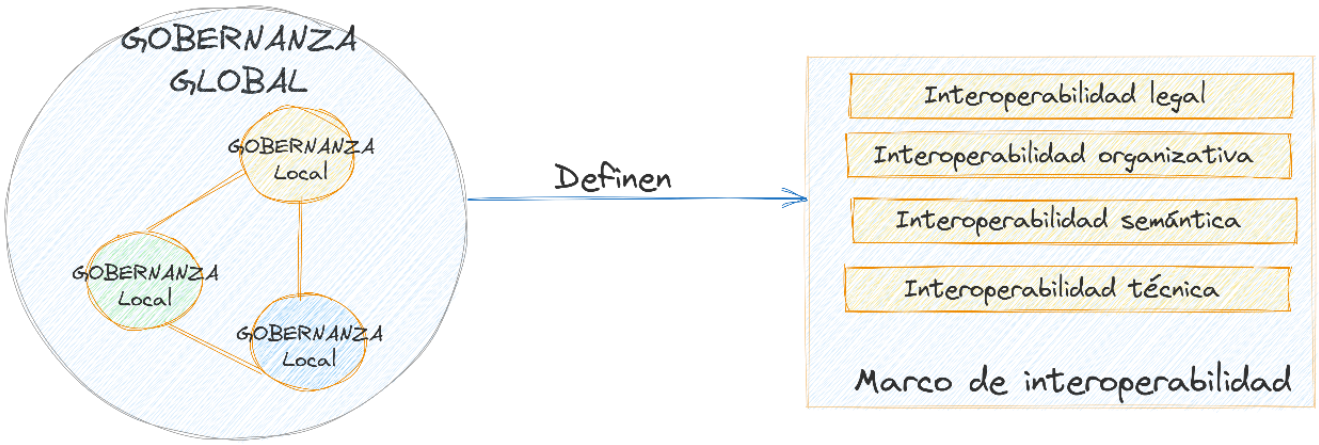

Fig. 11. A tiered interoperability structure.

6.2. Relationship between *building blocks* and interoperability levels according to the EIF

In the description of a data space, two conceptual models have been used as analysis tools: the BLOFT dimensions and the European Interoperability Framework (EIF). In turn, a *blueprint* of its own has been proposed to describe the functional dimension of a data space. With the intention of closing the circle between these three conceptual models, a correspondence between the proposed *building blocks*, following the BLOFT model, and the assets for interoperability according to the EIF framework, is described below.

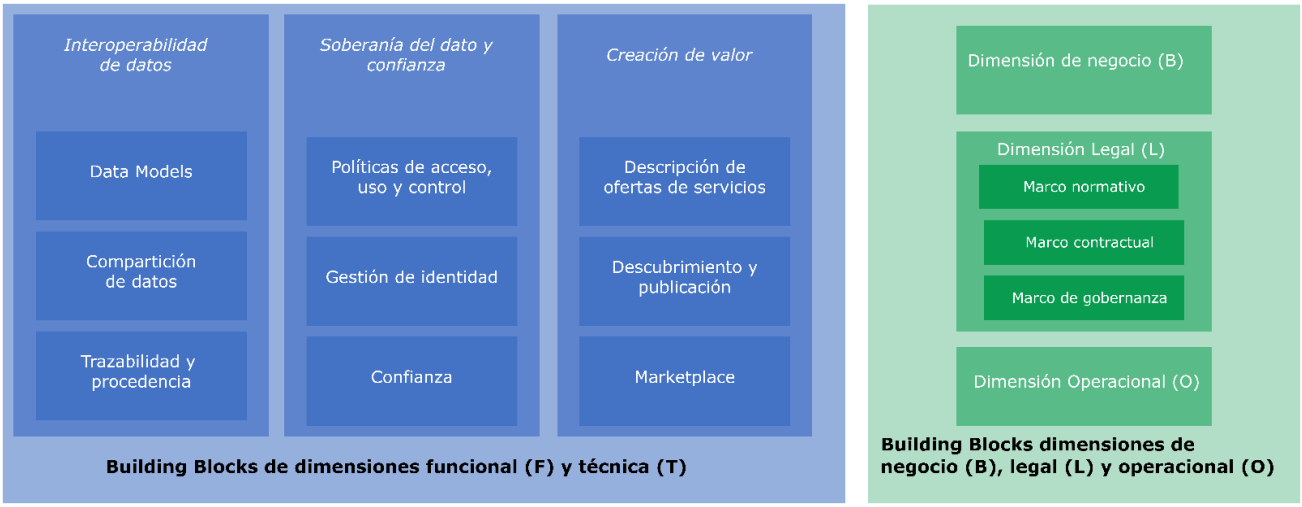

Fig. 12. Building Blocks and dimensions of the BLOFT model.

The following table shows the correspondence between *building blocks* and EIF interoperability levels:

Building Block		*EIF Interoperability Level*
Building functional and technical blocks	Data Models	Semantic
	Data sharing	Semantic and technological
	Traceability and provenance	Semantic and technological
	Access, use and control policies	Semantic and technological
	Identity management	Semantic and technological
	Trust	Semantic and technological
	Service offering descriptions	Semantic
	Discovery and publication	Technological
	Marketplace	Semantic and technological
Building business, legal and organizational blocks	Business	Organizational
	Regulatory framework	Legal
	Contractual framework	Organizational and legal
	Governance framework	Organizational
	Operations	Organizational

Table 1. Map between buildings blocks and EIF interoperability levels.

6.3. Roles within a data space

Assuming governance at two levels, the roles involved in a data space are described below:

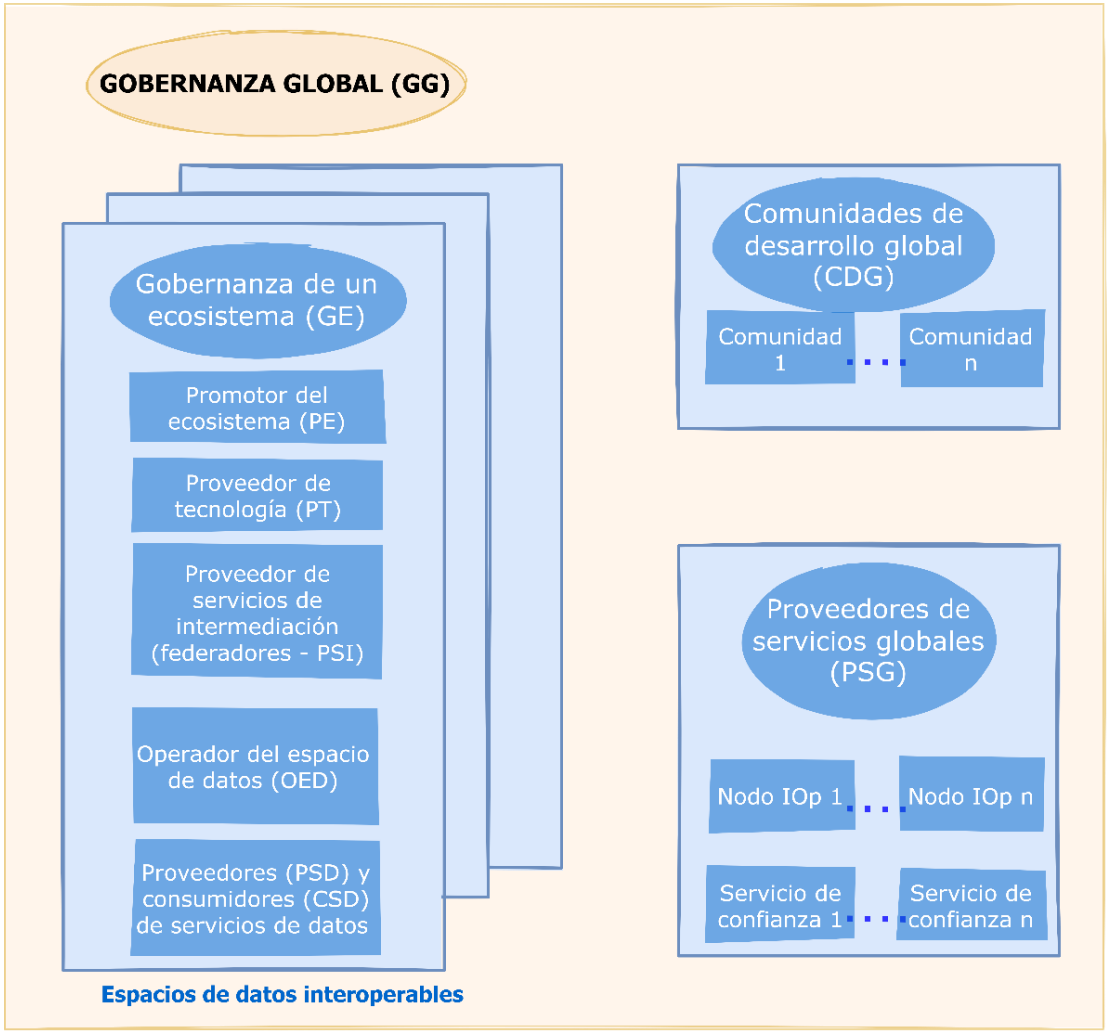

Fig. 13. Roles involved in the governance of an interoperable data space.

i. *Global Governance* (GG): is the entity responsible for defining interoperability standards between different data spaces. Its competencies cover the definition of specifications, the adoption of standards, the validation of components in accordance with their specifications and the authorization of global interoperability and trust service providers. There can only be one global governance body and its existence is mandatory. Global Governance also manages and monitors the certification process, and is in charge of the evaluating entities.

ii. *Evaluating Entity* (EE): Evaluates component implementations.

iii. *Global development communities* (CDG): *open-source* communities in charge of developing components for the deployment of ecosystems in accordance with the standards defined by global governance. These components are optional for data space instances, so the existence of these component communities is not mandatory either. In any case, the creation of this type of initiatives is considered a good practice because they facilitate the creation of data spaces that comply with the interoperability rules defined by global governance. In the event that communities of this type exist, the developments provided, as they apply to the interoperability standards of global governance, require validation by the GG.

iv. *Providers of interoperability and trust global services* (PSG). Although each ecosystem is autonomous, interoperability between them normally requires the provision of minimum interoperability and trust services. Interoperability services are understood to be those intended to prove compliance with interoperability rules and are provided by interoperability nodes. Trust services are understood as common services that provide a common identification, authentication and authorization (IAA) service between data spaces. In both cases, the services would be provided under the authorization and compliance of Global Governance. The specific services that are mandatory to be provided are determined by the GG. There may be multiple providers of interoperability and trust services at the discretion of the GG.

v. *Governance of an ecosystem* (EG). Apart from the common decisions regarding the establishment and compliance of interoperability rules between different ecosystems, the rest of the decisions fall under the responsibility of the governance of each of the instances of interoperable domain ecosystems. In fact, it is normal for ecosystems to enjoy great autonomy to determine their operation, their business rules and their technology in everything that does not have to do with interoperability with other data spaces. The roles belonging to each of the instances are the following:

   - Ecosystem promoter (PE). He is the driver of the ecosystem and the person in charge of its operation. As such, it is also responsible for the governance of the instance and for defining and executing the registration and deletion processes in the system and its supervision, among others, either directly or by granting work to a third party.
   - Technology provider (PT). He is in charge of integrating the technological solution that allows the ecosystem to function on behalf of the ecosystem promoter. You may have to carry out specific development of components for the final solution in the event that you do not want to use those supplied by the global development community (CDG), but in any case the technological solution they provide must be compatible with the defined interoperability standards. by the GG.
   - Intermediation service providers (*federators*) (PSI). They offer services that are necessary for the functioning of the ecosystem, such as a service catalog, activity registry or any other that the governance of the data space has established as part of the ecosystem.
   - Data Space Operator (OED). Like any other system, data space requires operation to function. This work covers all those activities that are not data resources and that are not covered by the intermediation services, such as the supervision of the subsystems or the registration, deletion and modification of profiles. If there were any intermediary services that the developer wanted to reserve for itself (for example, data catalog), the data space would be included among the operator's jobs.
   - Providers (PSD) and consumers (CSD) of data resources (PSD). They offer and consume these according to the scope of the ecosystem defined by the promoter.

The interaction between these roles determines how the data space works in practice. Thus, a data provider, formally incorporated into a data space instance, will make its data products and their conditions of use accessible through the data catalog managed directly by the data space operator or by authorized data intermediaries.

When a participant in a data space wishes to access a set of data, he or she will look to see if said information is already available in a data space, proceeding to study the conditions of access and use, as well as the suitability of its semantics and vocabulary, made explicit in the appropriate data catalog (which represents an intermediation service). If the characteristics detailed in the data catalog meet your expectations, you will

proceed to establish the appropriate dialogue with the data provider, establishing effective communication between provider and consumer of the data in accordance with the conditions explained in the data catalog (technological, and business). The activities of authentication, exchange of credentials and negotiation of the conditions of services are also known as "*mutual recognition*" between participants.

The performance of the ecosystem will be monitored by the data space operator (OED) by appointment of the service promoter. The technology provider must have designed the service in such a way that all data federation and exchange carried out generates the appropriate audit trail, being available for conflict resolution, as well as for the preparation of performance and effectiveness reports. or for hypothetical audits. Ultimately, it will be the governing body of the data space that guarantees and verifies the correct functioning of the data space.

6.4. Instruments for the governance of interoperable data spaces

Given the roles that intervene in the data spaces, it remains to define what actions each of the roles are entrusted with in a two-level governance scheme. The actions that correspond to them can be broken down into areas, understood as areas of action that can and must be carried out in the aforementioned governance framework, depending on the role of each participant. The areas for the application of said framework, considered in this document are:

- **Standardization and specification**: Formalization of tasks, processes, guides. Definition of mandatory technologies and standards. The specification of the labeling of service offers (*labels*) is also included. The labels aim to provide additional assurance about the conditions under which the service is provided in terms of, for example, data localization or compliance with ethical codes. That is, it can go beyond the strictly technological.

  Standardization and specification tasks not only apply to the interoperability of data spaces, but are also necessary within the framework of each unique data space in what has to do with the corresponding industrial standards (health, logistics, mobility, etc.) For example, a health data space is affected by a specific regulation that is applicable only in those data spaces of that type.

  For their part, the service offering labels apply to all data spaces, but each of them could include additional labels related to the sector of the data space that provide additional trust in their sector. That is, there is an industrial specification task that applies to each data space depending on the sector in which it operates.

  It is conceivable that certain industrial spaces (for example, health) are affected by common legal, organizational, semantic and technical rules that would require harmonization, but this harmonization goes beyond the scope of data spaces as such that is addressed in this document.

- **Certification**: Validation of formally established standards and specifications. The certification can be divided into:
  - Technical certification of components: validation of components according to defined standards and specifications.
  - Organizational certification: validation of the organizational and legal requirements of the participants according to the defined standards and specifications.
  - Labeling of services according to trust standards (*labels* mentioned in the previous section).
- **Development**: Component software development activities, which must subsequently be certified.

- **Operation**: Services exploitation activities. A distinction must be made between the operation of the systems and the use given to them by the end user (which does not appear in these instruments since it does not govern the data space). Operation activities can be certified organizationally.
- **Communication**: Actions to disseminate information on the activities carried out, both to publicize defined standards and specifications, as well as marketing actions.

Likewise, looking at governance from the point of view of the building blocks of a data space, its responsibility applies to said functionalities and the software components that support them. In other words, the instruments for the governance of a data space apply to the building blocks of the data space. Specific:

Governance instruments	Application to building blocks
Standardization and specification	Applicable to software processes and components to guarantee interoperability.
Certification	Applicable to the validation of the processes and software components used.
Development	Applicable to the design and construction of components and processes.
Operation	Applicable to the operation of data spaces.
Communication	Applicable to initiatives to disseminate information.

Table 2. Relationship between a data space's governance instruments and their application to building blocks and dimensions.

6.5. Decision-making to ensure interoperability

The initial deployment of a data space requires asking and resolving specific questions about its governance. Thus, action protocols, roles, decision rights and responsibilities must be established to ensure precision, consistency, security and regulatory compliance, support common objectives, and facilitate mutual trust.

Furthermore, the governance of a data space is not something static, immovable over time, but rather it must be understood as a program in continuous evolution, where, in parallel, the adaptation to new conditions and the maturity of the installed system must be addressed. always seeking to optimize the effective performance of the data space and the fulfillment of its business objective.

In this sense, the following table summarizes the three concepts mentioned above:

- The levels of interoperability (legal, organizational, semantic and technical). For each level, some of the most significant assets have been indicated to facilitate interoperability.
- The instruments to exercise effective governance of data spaces (standardization, certification, development, operation, and communication).
- The roles involved in the operation of data spaces (see section 6.3).

	Description	Standardization	Certification	Development	Operation	Communication
Legal level						
Legal agreements	Model legal agreements between data space participants, to which they must adhere. It applies to both the inter- and intra-data space levels, so both Global Governance (GG) and that of the specific data space (GE) are involved in this activity.	G.G., G.E.		PE	G.E., G.G.	G.E., G.G.
Verification of the legal status of participants in a data transaction	In certain scenarios it is necessary to identify the participants exchanging data. Therefore, mechanisms are required to verify the characteristics of the actors, as a service that requires development and operation.	G.G., G.E.		P.T.	PSG, PSI	G.G., G.E.
Semantic level						
Common semantic model	Semantic model for interoperability between different data spaces. Covers the definition of entities involved in interoperability, such as participants and service offerings.	GG	G.G., E.E.	CDG		GG
Specific semantic model	Semantic model specific to each data space. Its standardization, certification, development and communication are the responsibility of the governance of each data space. The implementation corresponds to the PT.	G.E.	OED, E.E.	OED, P.T.		OED
Organizational level						
Registration and certification of users according to their role	User registration and creation of credentials. It includes the definition of standardized	G.E.	G.E., E.E.	CDG, P.T.	OED	G.E.

	processes, as well as the development of components for this purpose. Participants register in a local data space, so this task corresponds to local governance, but also applies to global governance because it affects the interoperability of trust. Therefore, the global development community can collaborate by developing components.					
Service level agreements and quality control	Definition of service levels and supervision of their compliance. The GE is responsible for establishing the implementation framework.	G.E.	OED		OED, P.T.	OED
Operations and processes	Definition and execution of processes and operations.	PE			OED	OED
Technological level						
Federation services	Operational services offered within the data space. They are the responsibility of local governments.	G.E.	G.E., E.E.	CDG, P.T.	PSI	G.E.
Identification and authorization technologies and mechanisms in a data space	Although the ideal would be for the IAA mechanisms of each data space and the common ones to coincide, it is a decision that corresponds to the promoter of the data space to determine this aspect in its domain.	G.E.	G.E., E.E.	P.T.	OED	G.E.
Definition of common access and use policies	Definition of types of access and use policies for resources that establish the way in which consumers can use them.	G.G., G.E.	G.E.			G.E.
Data space access clients (connector)	It is assumed that there must be a data space access client (connector) that will be used to access interoperable data spaces, therefore it must comply with global interoperability standards, but must also comply with those specific to the data space. Therefore, the responsibility for	GG, G.E.	GG, GE, E.E.	P.T., C.D.G.		GG, G.E.

	this activity falls on global and local governance, each in its own area.					
Global interoperability and trust services	Services that allow global interoperability of data spaces. Basically, these are services that allow global identification, authentication and authorization (mandatory) and services that enhance trust between the parties (interoperability nodes for compliance services and service labeling).	GG	G.G., E.E.	CDG	PSG	GG
Common identification and authorization technologies and mechanisms (inter-data spaces)	Valid technologies for IAA, processes and systems responsible for this task must be specified to ensure trust between different data spaces. Includes systems certification.	GG	G.G., E.E.	GDG	PSG	GG

Table 3. Roles making decisions about the different governance instruments, for each interoperability asset.

6.6. Architectural and software components

The *Data Spaces Support Center* (DSSC) is currently the main initiative for the creation of data space *blueprint.* In its most recent version (v.0.5) [38], the decomposition into building blocks of the common functionalities previously proposed by the *OpenDEI project* is adopted and updated [36]. The DSSC defines *building blocks* as "*basic units or components that can be implemented and combined with other building blocks to construct data spaces*". In practice, each of the building blocks constitutes a grouping of functionalities that can be implemented through software components. The way in which the functionalities are implemented depends on the specific architecture of the data space, but the DSSC also proposes a general implementation model, which considers the following types of components:

a) **Connectors** (*Data Space Connectors*). In the model prescribed by the DSSC, consumers and service providers access the data space through connectors through which services are provided on a peer-to-peer *basis*. Apart from serving as a means of access to such services and data sets, these connectors uniquely identify participants. It is the element that mediates the access and provision of data space resources, and is therefore key, as it is involved in all the functionalities incorporated into all the pillars of *building blocks*.

b) **Data Space Registry**. It is responsible for implementing trusted functionalities; That is, it provides mechanisms so that actors can make statements *about* themselves, and about the products and services they offer.

c) **Components of federation services**. Strictly speaking, and given that the services in a data space are provided in *peer-to-peer mode*, at the time of the provision and consumption of a data product or service, only the participants (provider and consumer) are directly involved. through its connectors. However, for the correct functioning of the data space, additional services are also necessary that implement the functionalities of the 'Value Creation' stack of *building blocks.*

These functionalities (seen in section 5.2.4.3) articulate the data space, adding value and providing capabilities without which it would not be possible for participants to locate each other or know the characteristics of the services that providers make available, to give two examples. Value creation functionalities are provided through so-called federation services. Typically, each service requires at least one software component.

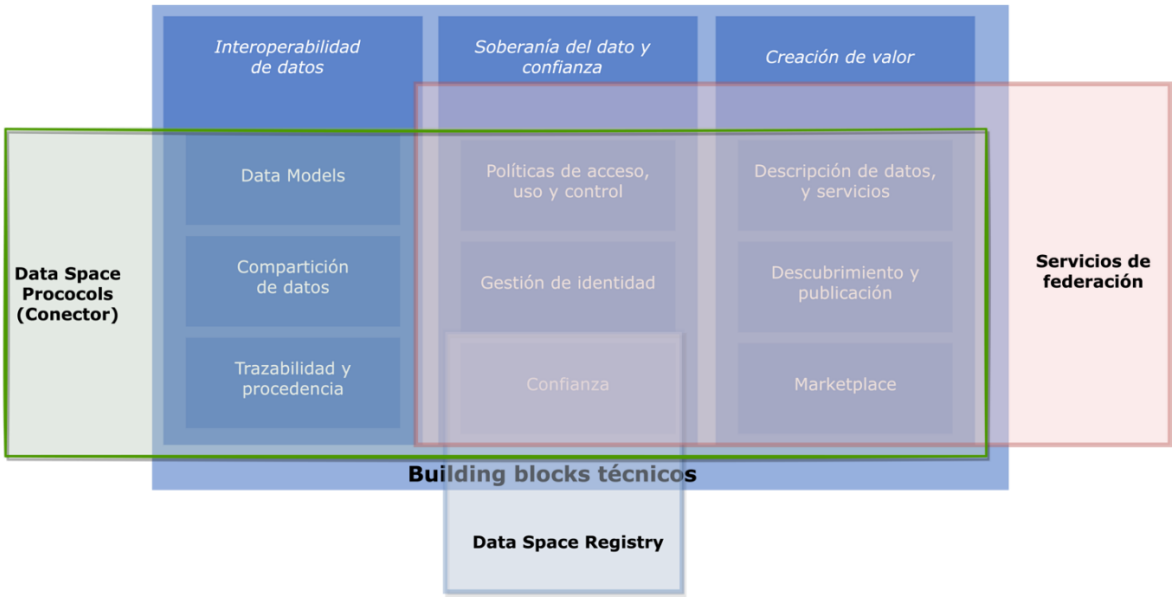

Fig. 14. Mapping of building blocks to technical components in a data space (DSSC).

To give another example, we have seen in section 5.2.4.1how the 'Data Interoperability' building block stack offers semantic capabilities for the description of data sets and APIs for effective sharing. But these capabilities do not allow describing the data product or service itself, which goes beyond the description of the data, to cover e.g. service level agreements, update frequency, monetary consideration (if any), or additional information such as the territory in which the data resides. The description of the service offered requires additional semantic capabilities and a catalog that can be consulted by potential consumers, thus making the suppliers of the products and services traceable. This, and any other services beyond data sharing activities, are provided through federation services.

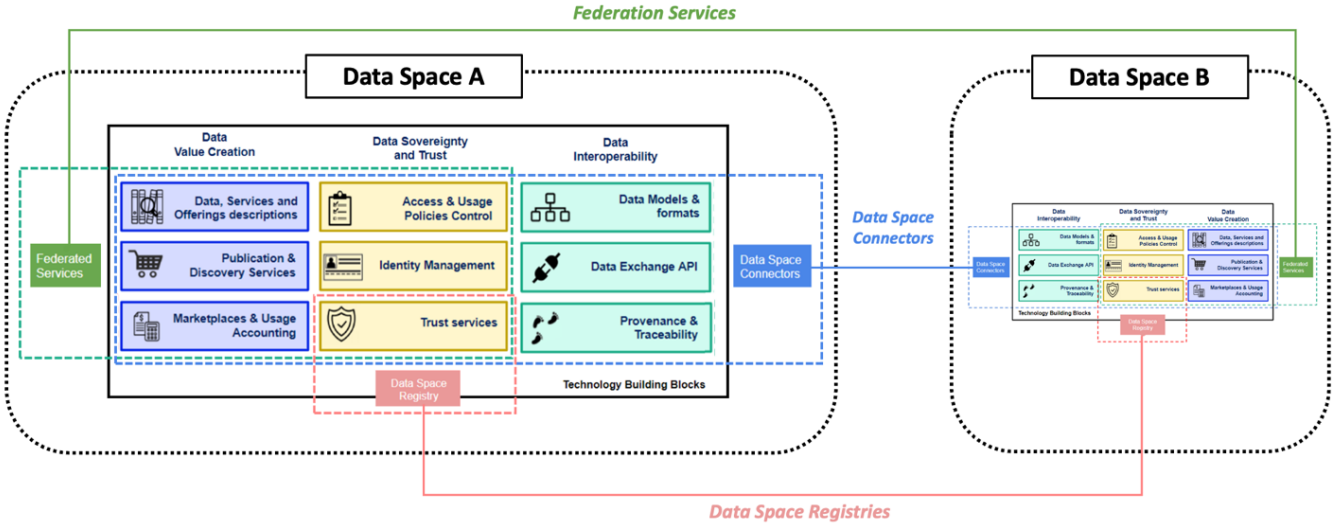

Fig. 15. Technical components enable interoperability between different data spaces (DSBA).

6.6.1. A harmonized architectural model

From the above, it is not too difficult to come up with a harmonized architectural model for the set of (potentially different) data spaces, composed of:

1. A global governance body responsible for guaranteeing interoperability between different data spaces, establishing common interoperability rules, and designating global interoperability service providers.
2. Common interoperability rules established by the global governance body, which cover the aforementioned legal, organizational, technical and semantic aspects.
3. Common services responsible for carrying out said interoperability.
4. Data spaces responsible for applying the common interoperability rules in their area, allowing them to establish their own domain interoperability framework, as long as this is not incompatible with the global one.

Additionally, the harmonized model for different data spaces decomposes each data sharing and exploitation process into 4 layers. These are:

i. *Business*. Final data space services, with their characteristics and conditions.
ii. *Trust Framework*. It is the general framework that gives confidence to data sharing, and therefore makes it possible. It is based both on regulations and regulations, as well as on the configuration of the different components (both in the individual data spaces and in relation to interoperability between different spaces), as well as on the management of said configuration. It is articulated in a practical way based on services that provide trust and sovereignty capabilities.
iii. *Federation*. Services that support the creation of ecosystems of service providers and consumers.
iv. *Data processing*. These are the specific services necessary for the technical execution of data sharing.

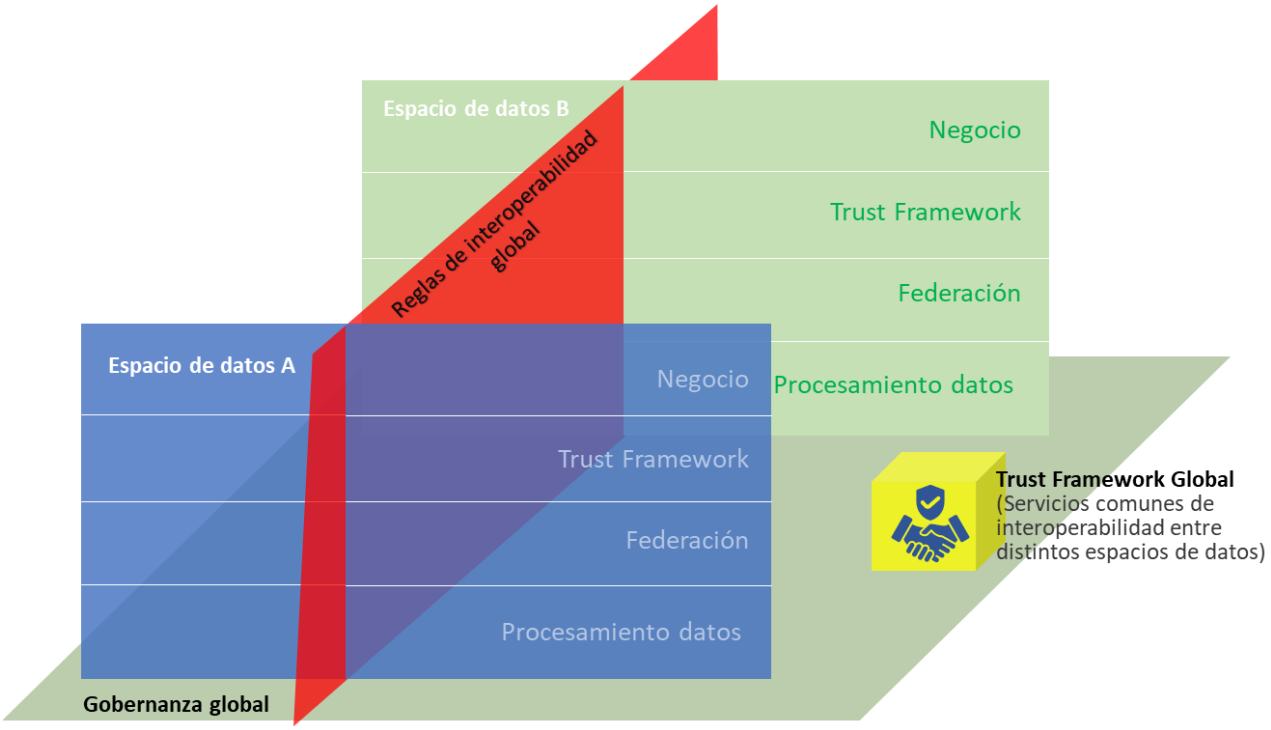

Fig. 16. A simplified harmonized architectural model for data spaces.

Legal, organizational, semantic and technical interoperability rules are applied to each of the layers in one way or another.

7. Organization and control over data spaces

7.1. Organizational models

Data ecosystems illustrate the need to harmonize internal data policies with those mechanisms to be applied beyond purely organizational boundaries. This harmonization provides clarity and supports the development of new forms of collaboration, with the governance challenges of these ecosystems being similar to the harmonized management of different service providers.

The classic literature of social and administrative sciences defines hierarchy, market, and network governance models as different ways of coordinating and organizing services, transactions, and activities among the different participants in a system. A more recent one can be added to this. Together, these four theoretical models can also be used to describe governance models in data spaces:

- The **hierarchical organizational model** exercises control through the authority of a dominant actor, who controls the decision-making process, and designs the formal coordination procedures of individual participants. In the case of data spaces, this model would be applied in platform environments where the owners of the infrastructure (both technological and soft*:* models and agreements that mediate the exchange) have hierarchical control over the participants. It is efficient when there is a strong need for control, coordination and alignment of incentives.
- The **hub (or marketplace) organization model** is characterized by strict compliance with the contractual terms that govern relatively standardized and independent transactions or activities (generally between peers), with the general regulatory framework of the hub providing the precise level of trust and coordination. which can be more or less exhaustive. The model is efficient when there are a large number of buyers and sellers, low transaction costs, and clear information. In the context of data spaces, we can assimilate these models to data market platforms ("data marketplaces"), where the relationships between consumers and data providers are based on market forces, even unfolding under the conditions of contour stipulated by the provider of the mediation platform.
- The **network organization model** enacts a system where participants demonstrate interdependent capabilities and collaboration based on consensus, reciprocity, and the search for collective goals and benefits. The model is efficient when there is a need for flexibility, learning and collaboration, being clearly applicable to a federation model or ecosystem of data federations, where evolution over time would facilitate the increase and mutual trust between participants.

  Within this model, taking into account how its governance is managed, we could consider:
  - *Leadership governance*, where a single entity acts as the central coordinator of the entire data ecosystem. This actor has the capacity to propose and define the interactions within the network, resembling its control to the hierarchical governance mode described above, although offering autonomy to the participants to adapt to said rules.
  - *Governance through a Network Administrative Organization* (NAO). This model has a neutral coordination body that governs the ecosystem, normally without directly participating in the exchange of services. This body requires the trust of the participants and the consensus on the objectives to be achieved, as critical attributes for the effectiveness of the collaboration.
  - *Shared governance*, where all organizations within the network share control and responsibility for the governance of the data ecosystem, without excessive formal structures, and where the relevance

of each decision depends on the level of involvement of its participants. This model usually requires clarity in goals, and a high level of maturity among participants.

- The **bazaar-type organization model** is a construct characterized by open licenses and commitments driven by the will to distribute information, or by the intrinsic motivation for reputational improvement; or even a tribal model that emphasizes the role of identity, reputation, and community in coordinating transactions and activities. Examples of this model are Open Data communities.

7.2. Examples of organizational models in data sharing within the digital economy

The following table shows the different organizational models, with real examples.

Organizational model	Description	Control element	Example in a data space environment	Comment
Hierarchical	Control exercised by a dominant authority	Authority	Typical case of a centralized sharing environment, for example common social networks	There is an authority that centrally controls the activity both technologically and from a business point of view.
Marketplace	Based on transactions between participants. In principle, everything agreed between them is valid	Assets and their prices	The model articulated by *AWS Data Exchange* or by the *Dawex platform* (e.g., the AgDataHub data space)	There is freedom in setting prices and availability of own resources, but compliance with transactions is strictly set based on the commercial terms established in the contracts.
network	The participants enact interdependent capacities, and there is a collaboration based on consensus, reciprocity and the search for collective goals and benefits.	Common rules	Catena-X , EONA-X	It is the typical case of a federation or data ecosystem, in which the participants submit (or agree to) common rules of operation, but the autonomy of each of them is maintained, not only in relation to their participation, but also in computer systems and technological standards
Bazaar	Based on a specific contract (open license), it promotes openness of participants and favors positive network externalities	Open licenses	*Wikipedia*, or open data initiatives	It is characterized by low levels of control, the existence of open licenses and the commitment of each participant, driven by the will to distribute information.

Table 4. Characterization of organizational models for data space environments, with examples.

7.3. Areas and decision points regarding governance

The scope of data space governance established by the *Data Spaces Support Center has previously been adopted* [40], according to which this consists of the creation of a framework made up of *"a set of standards, rules and practices that apply to the creation, administration and operation of a data space with a certain scope, and that has mechanisms to force its compliance and for conflict resolution"*.

The creation of an area of these characteristics involves making a set of decisions that largely translate into how the interoperability framework is guaranteed. Various decision points for organizational, semantic, and technical interoperability levels are described below, along with several alternatives, as examples. In this version we have chosen to ignore the legal one, since we understand this is a *sine qua non* area, as we cannot recommend the operation of data spaces that do not comply with European, national, or industrial-specific regulation. In subsequent versions we would also like to deepen the analysis of different policies chosen for specific data spaces, thereby outlining useful prescriptions in an area whose scope and heterogeneity is extensive.

7.3.1. Decisions to ensure organizational interoperability

Decision area	Decision points	Alternatives[6]
Data space participants and roles	Brokerage service provider role	The role of intermediary service provider is open to any applicant (as long as they are not a data resource provider).
		Intermediation services are provided exclusively by the platform provider
	Opening to new participants	Participation in the role of supplier and consumer open to any economic agent. Admission rules based on technological criteria described in interoperability standards
		Participation and role subject to the discretion of the promoter, membership in a sector or bilateral agreements with the governance of the data space
Business models	Business autonomy of suppliers and consumers	Complete business autonomy with peer-to-peer relationships. The only limitation for the provision of services is compliance with legal, semantic and technological interoperability standards. Providers can partner to jointly provide services
		Autonomy restricted by promoter. The promoter must approve any new service offered in the data space prior to its launch.
Interrelationship with other data spaces	Operation of the data space within a decentralized network of ecosystems	Users of a data space can interrelate with others through a network of sectorial and intersectorial ecosystems
		The data space forms a silo to which only users registered in it have access

Table 5. Decisions helping to ensure organizational interoperability.

7.3.2. Decisions to ensure semantic interoperability

Decision areas	Decision points	Alternatives [6]

[6]These alternatives are not exhaustive, but rather represent the two possible extremes.

Description of service offers	Variety of representable resources	It allows describing different resources (data, applications for data exploitation and computational resources)
		It is only allowed to define service offers based on data sets
	Flexibility in the description of services	There is a flexible information model based on vocabularies and schemas that allows defining composable services that integrate heterogeneous resources (data, applications and computational resources), even from different providers.
		The information model does not allow the composition of services
	Definition of bid resources using recognized standards	The information model allows describing the resources of the service offers through vocabularies and models based on open standards (e.g., data sets through DCAT-AP)
		The information model does not provide descriptions of resources using common standards
Data sovereignty	Definition of usage and access policies to resources	Providers express the policies for use and access to resources through a standards-based information model (e.g., ODRL). The definition of these policies is part of the service description
		The data space does not have an information model to define access and use policies for resources.
Formats	Data model formats	The modeled entities are registered using formats based on open standards for exchange (e.g., W3C VC)
		Information model entities are recorded using proprietary formats
Accordance	Existence of compliance service	There is a verification service for the objects of the information model that guarantees participants the validity of their formal composition and content.
		There is no compliance service

Table 6. Decisions helping to ensure semantic interoperability.

7.3.3. Decisions to ensure technological interoperability

Decision area	Decision points	Alternatives [6]
Data space architecture	Global architecture selection	The data space is based on an open reference architecture whose specifications and components are made available under an open-source license.
		Data space architecture is proprietary
Solution components	Reuse of building blocks	The technological solution uses only components made available in open-source format, based on de jure or de facto standards.
		The technological solution is based on proprietary developments

Table 7. Decisions helping to ensure technological interoperability.

8. Data space architectures

Keeping in mind that the ultimate objective of a data sharing scenario is the satisfaction of a business need, there are many data sharing scenarios, and for each of them there is an optimal architecture that fits its purpose. In other words, although different data space architecture models may exist for a single governance model, not all models are equally suitable.

Below, a classification of architectures is proposed that cover the vast majority of sharing scenarios. It can be considered that we are dealing with a data space only when it goes beyond a bilateral exchange of information, since this scenario does not replicate. It is also easy to establish an equivalence between the different architectures proposed for data spaces, and the classification of governance models previously stated.

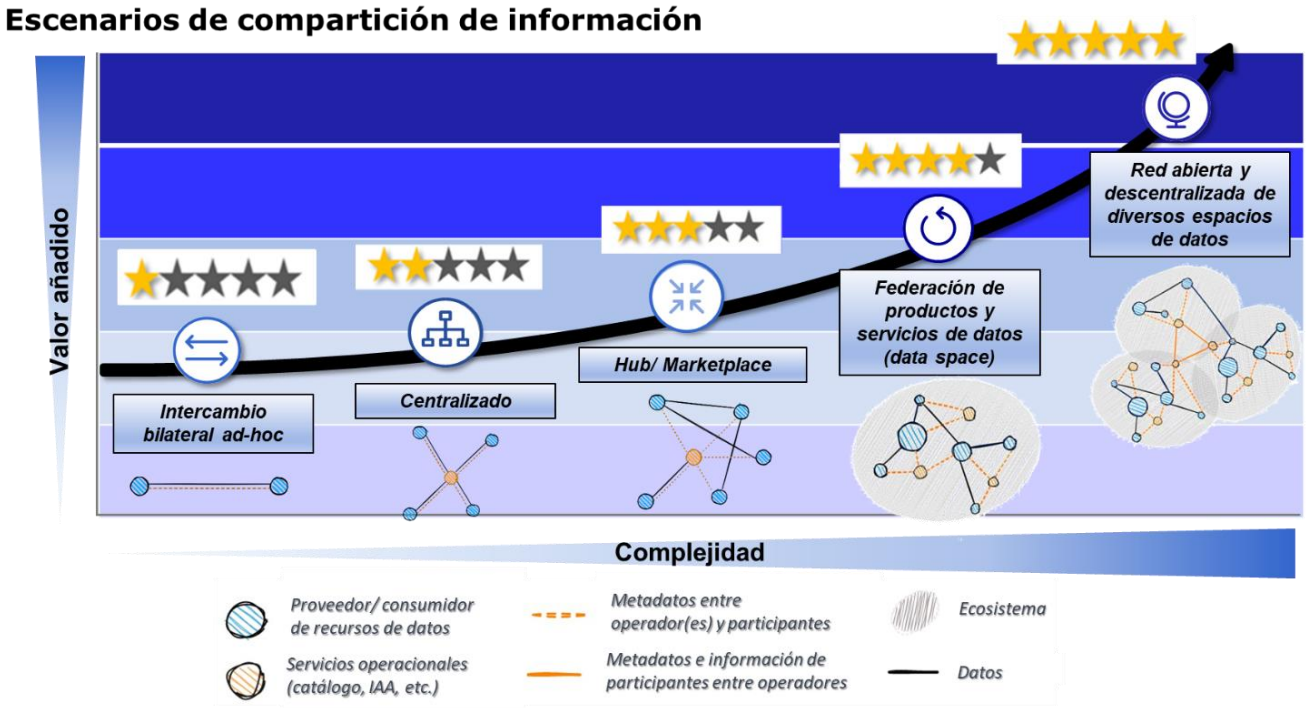

Fig 17. Information sharing scenarios, according to complexity and added value.

The ambitious European objective of establishing interoperable federated spaces for sharing data and computing resources in different industrial sectors can be considered a medium/long-term objective, given that the technologies on which it is based are still immature today. However, even without these, it is possible to advance partial solutions that allow the development of different pieces of the technological and business framework, necessary to constitute in the future those ecosystems from which to promote the deployment of the sovereign Data Economy.

8.1. Data space as a central data node

This type of data spaces is characterized by having an operator and a centralized infrastructure designed to load and exploit data from multiple sources. All data loading and exploitation operations are carried out by a centralized technical team. More than a sharing environment, it is a joint exploitation environment that can however become very sophisticated.

The platform operator schedules the storage of data in the centralized infrastructure using the appropriate ingestion tools. Once deposited, a team of engineers and data scientists designs and executes the analytical processes that add value to the available datasets, and deliver the result to domain users, who act as if they were clients of a service.

The typical case of this type of solutions is a Data Lake, understood as a warehouse in which heterogeneous (and originally dispersed) data assets are replicated as they are in their place of origin, or after the appropriate transformations. They are usually supported by a distributed file system (for example, the *Hadoop File System*) accompanied by an ecosystem of tools for ingestion, cleaning, transformation, storage and exploitation through business intelligence and Machine Learning technologies.

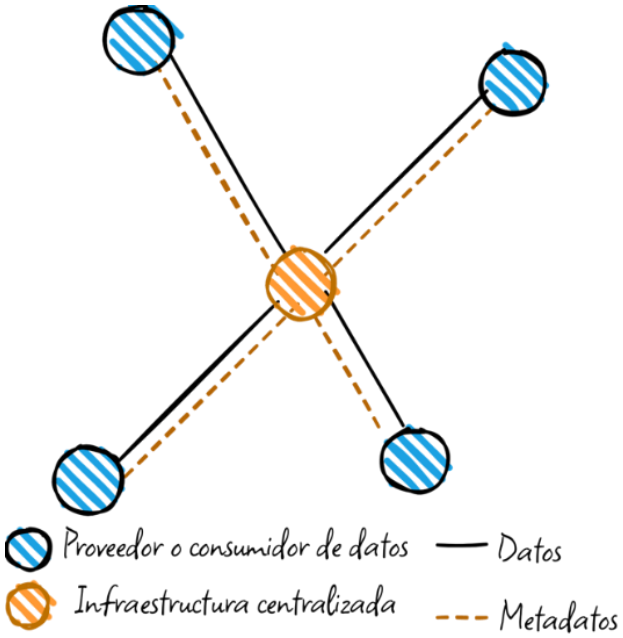

Fig 18. Centralized data-sharing scenario.

Centralized data spaces do not tend to offer tools to exercise data sovereignty in sophisticated ways, nor portability. This is a consequence of the fact that they are systems that were initially designed for internal use of organizations, so the establishment of data use policies is not a critical aspect. In return, they have robust access control systems, typically based on roles, attributes and/or auditing.

8.2. Data space as a data hub, or a data marketplace

The data space model as a data hub or "data marketplace" is characterized by having a centralized infrastructure, managed by a single operator, which offers a catalog (based on metadata) with the offer of available data resources. In this type of architecture, sharing is promoted in such a way that potential consumers go through the catalog first. Thus, after "finding each other", consumers and providers identify and authenticate themselves just before sharing. In general, the operator of this type of platform charges a fee to participants for creating and maintaining the service, and each resource offered is therefore susceptible to financial compensation.

hub service provider exercises domain governance independently, establishing the operating rules and selecting the technology to be used to support the service, which is usually its property. A priori, these platforms are not inherently intended to interoperate with other data spaces, which means that the operator

is not required to comply with large interoperability standards and is not required to offer its technology through open licenses, which It does not mean that this circumstance is not interesting to occur.

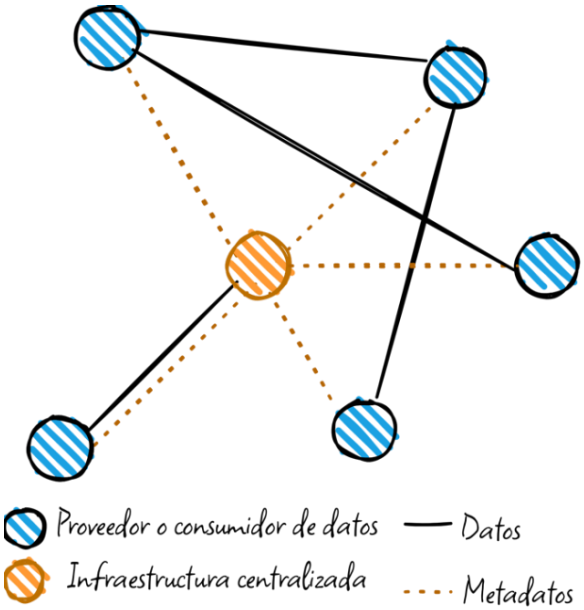

Fig 19. Diagram of a 'Hub' or marketplace-type architecture.

The hub provider is responsible for the services necessary for the operation of the data resource exchange. At a minimum, it must offer services to manage participants (registration, cancellation, suspension, etc.), guarantee security (identification, authentication, authorization), manage the metadata catalog with its associated vocabularies and ontologies, and access and access contracts. use of available resources, as well as maintaining a record of operations and payment mechanisms. These platforms are usually open and allow any interested party to join the sharing or buying and selling service. The operator is also in charge of managing requests, maintaining the software, managing incidents and any changes, among other tasks typically associated with the provision of an ICT service and business management.

In this scenario, it is possible to define data use policies (sovereignty), which can be applied by data providers who wish to establish restrictions on how their data is used, although this is not always offered and in any case it is not usually be easy to enforce (beyond a contractual safeguard). Likewise, the portability of resources is a valued functionality. In this sense, hubs can store suppliers' data in their infrastructure, or simply act as intermediaries and facilitate contact between suppliers and consumers so that they can download between them (*peer-to-peer*). Additionally, it is not uncommon for these platforms to offer value-added services, such as applications for data cleaning and processing.

8.3. Data space as a digital data federation

a federation Data is defined as a community formed by autonomous actors that consume, produce, or provide data and value-added services, or offer services (identification, sovereignty, traceability of operations, catalog of data offerings, vocabularies, etc.) for others to carry out these sharing activities. Each actor performs one or more roles, and is connected to the other participants through relationships such that their collaboration and competition promotes self-regulation, and prevents situations of dominance over the sharing environment (which are more typical in simpler scenarios, such as those mentioned above). in sections 8.1and 8.2).

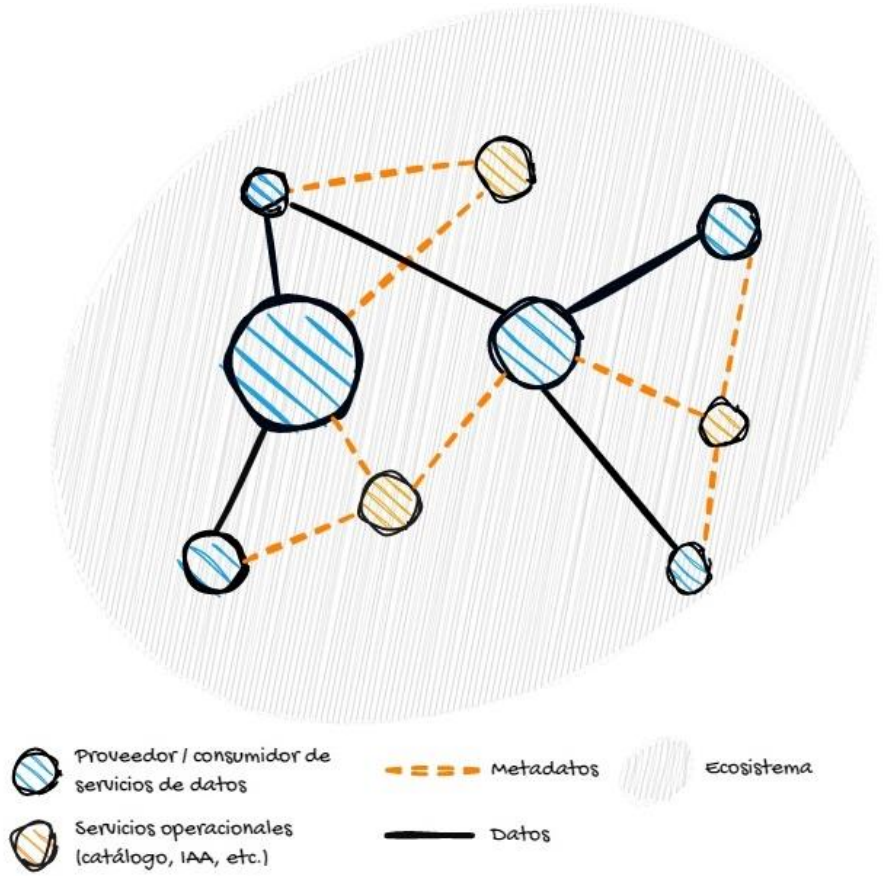

Fig 20. Diagram of a 'data ecosystem' type architecture.

A digital federation is understood to be that business environment constituted by its own participants when they relate to each other autonomously and as equals, and in accordance with a practice in which there are no dominant actors firmly determining their rules. A data federation is a digital ecosystem in which the business in question is data sharing and exploitation. These ecosystems are based on community, transparency, innovation and the ability to scale and generate shared benefits. Unlike platform models, where much of the value is retained in intermediation, here we are committed to an ecosystem model that allows participants to maintain their autonomy and, at the same time, collaborate based on peer-to-peer transactions (*peer-to-peer*).

Expressed another way, in a federation the system itself is made up of the community formed by the participants themselves, who compete and collaborate mutually in the provision and consumption of resources and services through equal-to-equal relationships, but based on the assumption that they must be the participants themselves who, based on broad autonomy, create and develop new services for their consumption. The federation paradigm is considered to be opposed to the platform paradigm, in which there is a clear distinction between the system itself and its users.

Specific features of these federated data spaces include:

- **Governance.** There must be a governance model based on mandatory interoperability rules, which guarantee the development of the data sharing business on an equal basis by the *providers and consumers of data resources in the ecosystem*. The governance rules will guarantee the reduction of barriers to entry and exit from the federation, except for those that relate to the guarantee of interoperability and cybersecurity and sustainability in the development of the business.

- There will be a *data space operator* in charge of the technical and operational tasks necessary for the operation of the system (identity, authentication and access, support, maintenance, registrations and cancellations, system supervision, etc.) who -however- will not be able to provide own data sharing services (neither data offering nor processing). The offering of such services will be carried out by the participants in the data space on the supply side: providers of data sets, or data processing services.
- There will be the figure of the *intermediary service provider*, responsible for offering value-added services that facilitate data sharing. Among them are, for example, service catalogs, activity and audit logging, or application stores for data processing and exploitation (not the applications themselves). The provision of this type of services will be open to interested parties who wish to provide them, as long as they do not meet disabling conditions of force majeure, for example due to regulatory limitations.
- Decision-making in the data space aspires to be participatory, both in technological and business matters, so that there is no dominant operator that can make unilateral decisions about its characteristics and evolution. This seeks to guarantee that the generation of innovation and value of the system is sustainable and capillary, based on promoting the participation of different types of interest groups.
- The interoperability rules will allow access to value-added services (intermediation) and the provision/consumption of data resources with cybersecurity and trust, but always autonomously by suppliers and consumers, as part of an agnostic environment that does not offers an advantage to any participant in the data space (principle of "competition on equal terms").

- **Architecture.** The technological architecture of the proposed solution will follow the distributed model based on federation, in which there is no need for centralized components in the provision and consumption of data sharing services, but rather the services are served and consumed directly in relationships of equal to equal between consumers and suppliers (*peer-to-peer*). The only exception to this rule is for identity and trust services, but even in that case decentralized mechanisms can be explored.
- **Features**. At least the following functionalities must be offered:
  - Secure data exchange between participants.
  - Data models and data formats subject to exchange.
  - Traceability and lineage of said data sets.
  - Data sovereignty, understood as the ability to define and enforce data access and use policies by the holders of their access rights.
  - Recording of data sharing activity for auditing and reporting.
  - Tools for publishing and searching data (i.e., catalogue/s).
- **Minimum construction components** (*building blocks*). The European Commission defines *building blocks* as basic digital infrastructures that can be reused for the composition of complex digital services. The current state of the art of technology makes different reference architectures available to interested parties, as well as technical components of different degrees of maturity for the creation of data spaces that support the aforementioned European strategic values.

  To boost the creation of these ecosystems, as well as mitigate dominant positions that lead to technological dependence, it is desirable that the operational and intermediation components in a data space be made generally available in an open source software regime. This is the case of the best-known European initiatives in the development of this type of systems. As an example, and not exhaustively, we

can mention those attached to the Gaia-X initiative, the International Data Spaces Association or the FIWARE community, as well as some pieces developed from programs financed by the European Commission (e.g. Connecting Europe Facility).

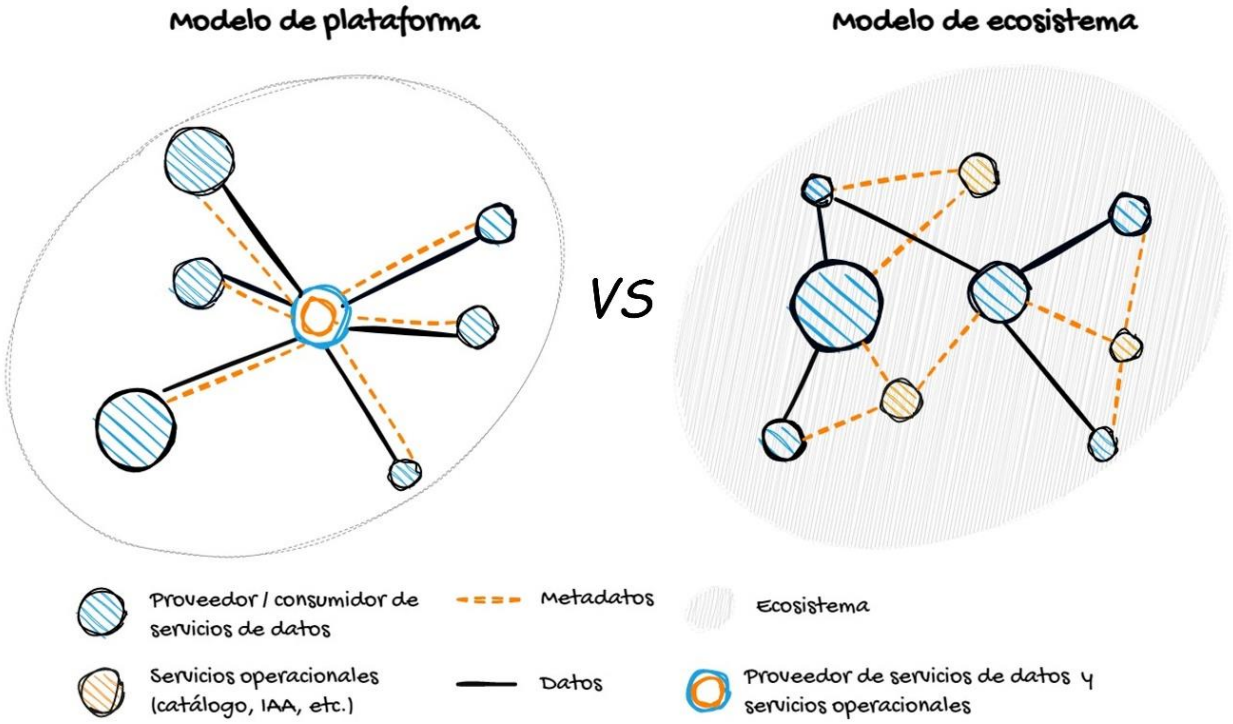

Fig 21. Platform model VS ecosystem model.

8.4. Data space as a network of data federations

The aspirational trend for data spaces is the creation of interconnected ecosystems in decentralized networks. This is the model that is driving the Gaia-X initiative, and which represents a maximum degree of interoperability and openness. In an architecture of this type, a consumer can consult the catalogs with the offer of resources in his data space, as well as the rest of the connected data spaces, because they are interoperable and - in the same way - a data provider or Other value-added services can also make their resources available to all consumers of the rest of the data spaces.

Gaia-X calls each data space or domain a "federation", which is operated by a *federator* who is responsible for offering the essential services for its operation. This federator is also responsible for connecting its domain with the rest of the data spaces (federations). Under this paradigm, Gaia-X also allows expanding the offer of resources not only to data sets, but also to applications (software resources) and computational resources.

Such an environment requires a great effort to select and define interoperability standards. The governance model is tiered: in each federation, those responsible for governance are the federators, but for the whole to work it is necessary that there be global governance that determines the interoperability rules between domains and global services such as, for example, a global identity service. This is explained in much greater detail in section 6.3.

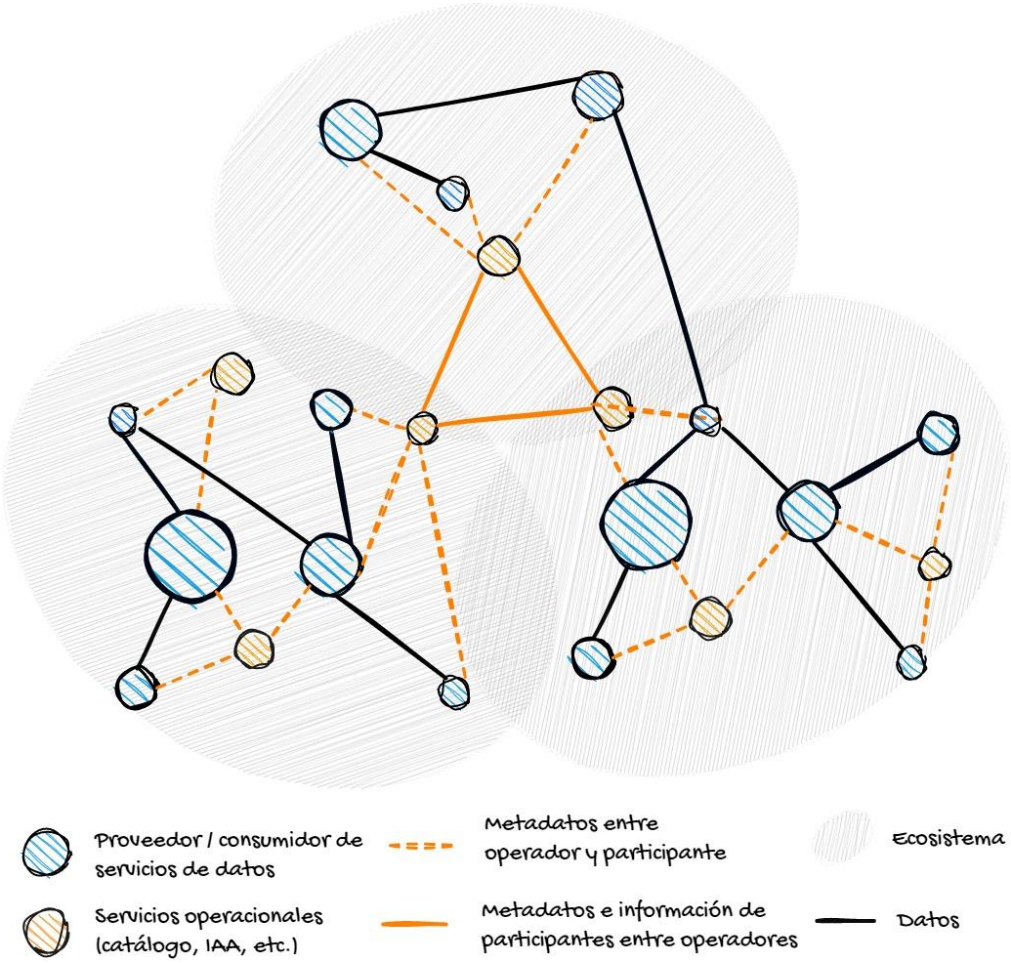

Fig 22. Diagram of an 'Ecosystem of ecosystems/data spaces' architecture.

8.5. Complementarity between scenarios

Although there is no direct scalability or portability between a centralized and a federated data platform model, there is nevertheless a significant complementarity between them. And this is because, despite the differences in their implementation and functionalities, both centralized and federated models can be adapted to the characteristics of data spaces (see chapter 3).

Their interaction can therefore be used to deploy an **evolutionary deployment strategy,** where the simplest model serves to start projects, generate best practices and valuable lessons, and increase maturity and awareness of the latent value in data and its paradigms and management schemes, aspects that will later serve to scale these scenarios towards more sophisticated models. We therefore believe that this complementarity between paradigms is a powerful tool to optimize the digitalization of industries based on data, in a realistic and capillary way.

In this sense, the resources dedicated to centralized systems also contribute to generating a greater critical mass in federated systems, as the former are considered nodes of the latter, as reflected in **Fig 23. Deployment of the platform model as a node of a federated ecosystem.**Fig 23, thus contributing to its expansion and strengthening.

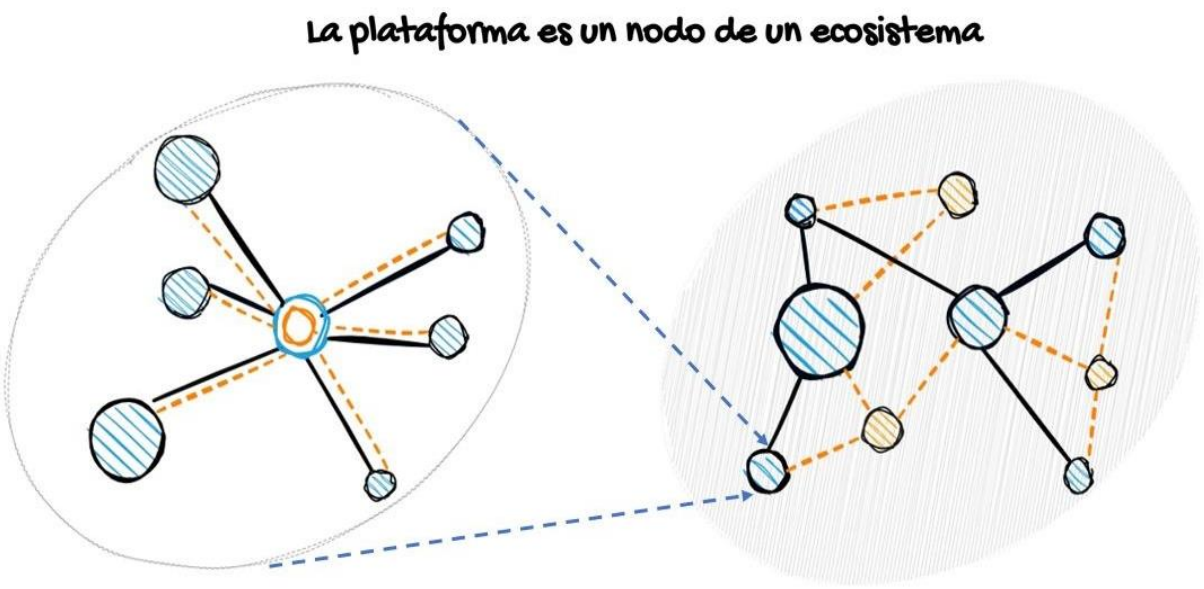

Fig 23. Deployment of the platform model as a node of a federated ecosystem.

9. Prominent actors and technological solutions within the European community

For approximately five years, there have been different initiatives within the European Union that aim to develop technology for the construction of data spaces with the aforementioned characteristics. The promoters of these initiatives work on the definition of standards, specifications and components that they make available to the community in open source format. In parallel, organizations have emerged that seek technological convergence of available solutions.

9.1. Prominent actors in the development of data spaces

Below, the most relevant actors in the development of technologies for the construction of data spaces with European values are described. Keep in mind that it is not intended to offer an exhaustive list but rather to point out those organizations with the most weight in the definition of specifications and the development of components:

- The **FIWARE Foundation** is an organization with extensive experience in the deployment of digital solutions for smart cities, and has to its credit the specification of ETSI standards for modeling and data transfer, such as the NGSI-LD . The foundation does not provide a complete architectural model for data spaces as such, but rather dedicates its efforts to the development of components (available at [41]) such as the *TRUE Connector* (connector for access to data spaces) or the *KeyRock component* that, integrated into the connector, allows identity management of participants in a data space. This has allowed them to define, together with the iSHARE Foundation, one of the most advanced models for deploying data spaces in practice [30].
- The **iSHARE Foundation** is a non-profit organization whose objective is the development of a legal and technological framework for data sharing. Its activity, however, is more focused on the sharing of records than on the data sets themselves. However, some of its components are reusable for the construction of data spaces [42]. To their credit, they have deployed a data space in the Netherlands dedicated to logistics record transactions that reportedly has over a million subscribers.

- **International Data Space Association** (IDSA). IDSA aims to develop standards for the construction, governance and operation of data spaces with European values. It offers a complete reference architecture and a core set of reference components [43]. Today, the technology they offer is the one with the most initiatives underway, although most are still in a pilot state.
- The **Gaia-X European Association for Data & Cloud AISBL** (hereinafter, Gaia-X) is a private initiative founded as a non-profit association under Belgian legislation in 2021, which currently has more than 350 associate members from all over the world. world. Their work includes the design of a data space architecture that, compared to other alternatives, expands the scope of services offered by data space infrastructures to that of cloud computing services. Particularly relevant is the framework that the association has defined to ensure interoperability between different data spaces (*Trust Framework*). In addition to an architectural model, it offers some reference components that meet its specifications [44].
- The **Big Data Value Association** (BDVA). The BDVA is a non-profit organization that aims to develop an innovation ecosystem that enables the data-driven digital transformation of the European economy and society. It has more than 230 members across Europe, including a balanced composition of large companies, SMEs, research organizations, and also users. Its fundamental lines of work revolve around supporting innovation, highlighting the collaboration with the EuroHPC joint venture and also with the industrial Big Data and AI community, to support the development of a reliable and ethical AI, which complies with the values and EU regulations. In the field of data spaces, the BDVA is part of the DSSC and collaborates with the EUHubs4Data project in the creation of a European federation of innovation spaces with data (*i-Spaces*), developing a relevant role in community creation, promoting collaboration, interoperability and adoption of data sharing projects in Europe.
- **MyData Global** is an international non-profit association that helps people and organizations benefit from personal data in a people-centric way. The central idea of MyData is that individuals should have control over the data that concerns them. With this activist-type approach, we seek to strengthen people's digital rights (a *human-centric approach*), while opening up new opportunities for companies to develop innovative services based on personal data, always built on mutual trust. In the field of data spaces, MyData plays a crucial role in promoting the creation of ecosystems in which natural persons can share and use their data in an ethical and beneficial way, while contributing to a fairer digital society. sustainable and prosperous.

9.2. Initiatives aimed at fostering convergent solutions

The existence of different organizations dedicated to the standardization and specification of data spaces has resulted in the emergence of non-interoperable technologies, which can cause the emergence of data silos, constituting a threat to the creation of a single data market.

As a consequence of the above, a series of initiatives are underway that seek the convergence of technological solutions. These initiatives are:

- **Data Space Business Alliance** (DSBA). The DSBA is an alliance formed by the associations Gaia-X, IDSA, Big Data Value Association and the FIWARE foundation. It therefore brings together the main actors in the technical development of data spaces. Its objective is to join forces around this innovative paradigm, coordinately disseminate its benefits, as well as achieve convergence between the technologies developed by each of its members. In this last sense, and to date, they have published two versions of a technological convergence document, which has served as inspiration and guidance in this document to describe the common technological framework of a data space.

- **Data Spaces Support Center** (DSSC). The DSSC is a project funded by the European Commission in which a consortium of 25 European organizations participates, including members of the DSBA, which play a prominent role in the project. Their activities partly overlap with those of the DSBA, but in this case they are officially commissioned by the EC. Its main task is to define common requirements for data spaces and accelerate their deployment. The consortium has also been tasked with setting requirements for the SIMPL project (below).
- **'Smart middleware for a European cloud federation and for the European data spaces'** (SIMPL) project is a project funded by the European Commission within the framework of the *Digital Europe Program* and endowed with up to €150M. It is currently in the award phase through open bidding, and its objective is the development of middleware for the creation of interoperable federated data spaces, whose specifications largely resemble those of Gaia-X. The project has a duration of 36 months, and it is not expected that there will be a basic version (*Minimum Valuable Product*) before the beginning of 2025. One of the requirements of the project is that specifications and components made available be reused for the construction of the middleware for the rest of the existing initiatives for the construction of data spaces, as well as using - as far as possible - the digital public infrastructures made available by the EC.

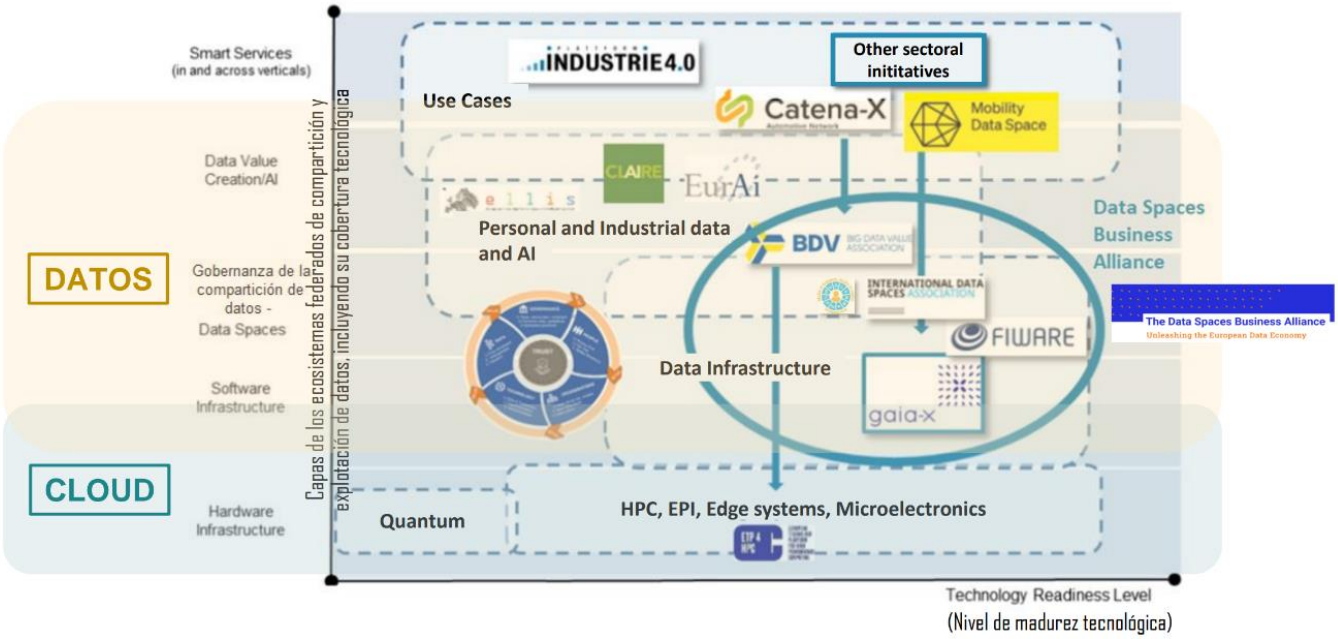

Fig 24. Data value chain across technological coverage, with several European initiatives portrayed according to technological maturity (based on BDVA, dated 2022).

9.3. Summary of the main European technological initiatives for data spaces

More generally, the following table shows a synthesis of the main technological initiatives for the construction of data spaces and for the convergence of solutions, coming from the DATES project [16].

Aspect/ Initiative	Primary Focus	Origin	Technology Base	Key Features	Use Cases	Governance	Strengths	Weaknesses
Solid	Decentralized data ownership and web applications	Developed by Sir Tim Berners-Lee and the MIT	Web standards, Linked Data	User-owned data stores (PODs), data control	Decentralized social applications, personal data management	Open-source community	User-centric data control, promotes data privacy	Limited adoption, complexity in integration, although steps ahead have been made
Gaia-X	European data infrastructure and cloud services	European Union initiative	Cloud, federated data infrastructure	Federated services, data sovereignty, cloud framework	Cloud and data services, European digital sovereignty	Consortium of European companies and organisations	Supports digital sovereignty, strong EU backing	Complex governance, implementation challenges
IDSA	Secure data sharing and data spaces	European Commission initiative	Data sovereignty, secure data sharing	Data sovereignty, secure exchange	Cross-industry data sharing, secure data ecosystems	Association of companies and research institutions	Standardised secure data sharing, cross-industry applicability	Requires broad industry collaboration, still evolving
FIWARE	Open-source platform for smart solutions	EU's Future Internet Public-Private Partnership	Open source, IoT, smart applications	IoT integration, smart solutions, APIs	Smart cities, IoT, industry 4.0	FIWARE Foundation	Versatile IoT solutions, strong EU support	May require technical expertise, IoT focus limits scope
SIMPL	Cloud-to-edge federation for major EU dataspaces	European Union initiative	Middleware platform to support data access and interoperability among European data spaces	Open-source library for data interoperability, supports various data formats and protocols	Middleware Platform	European Commission support	Enables seamless data exchange between different systems	Adoption and standardisation Not currently operational
DSBA	Collaboration for data spaces	Collaboration of IDSA, Gaia-X, FIWARE and BDVA	Data spaces, interoperability	Promote interoperability, standards	Data space collaboration, cross-sectoral	Alliance of multiple organisations	Fosters collaboration, promotes interoperability standards	Still in development, lacks widespread recognition
DSSC	Support for data space initiatives	Supports European data space initiatives	Support and guidance for data spaces	Facilitate development of data spaces	Support and coordination for data space projects	European Commission support	Provides essential support and coordination	Dependent on EU policies and funding

PART II. Strategic axes of this activity plan

Throughout this section we proceed to draw the strategic plan for the deployment of data spaces in Spain. The ultimate objective is to generate a data-oriented economy, sustainable, inclusive and generating social value, capable of taking advantage of the potential of data through the use of innovative technological means. To do this, the strategic objectives are set and the measures that allow them to be achieved are defined, grouping them into four strategic axes.

10. Strategic objectives

The final goal of promoting the Data Economy, promoting innovation and advancing the generation of shared value, in coordination with the rest of the lines of action and purposes present in the Digital Spain 2026 Agenda [23], is articulated around the achievement of the following goals:

1. **Improving the performance of strategic productive sectors**

   Improving the performance of different productive sectors will be achieved by stimulating the generation of new solutions, services and business models from the use of shared data; and encouraging creativity, research and development of data-driven technologies and applications.

2. **Guarantee of strategic digital autonomy**

   The aim is to guarantee the national capacity to provide and control those technologies and tools critical for digitalization and, therefore, for growth, competitiveness and well-being. This will be achieved, either through the generation of said technologies themselves, or by guaranteeing their supply from other territories without this implying unilateral relations of dependency.

3. **Promotion of the generation of quality data from an ethical and effective data governance**

   Maximizing the value of the data exchanged requires ensuring its quality. This is fundamental in the effective performance of data spaces, since it directly affects the trust of the participants, and requires effective governance of the data in organizations, prior to making it available.

4. **Generating value from publicly-held data**

   It seeks to maximize the value and socioeconomic impact of public data, promoting its accessibility and use by citizens, companies and organizations within data spaces. Facilitating access, reuse and sharing of public data will generate new business opportunities, innovation and allow improvements in the provision of public services.

5. **Promoting the development of a fair and sustainable Data Economy**

   The aim is to promote an equitable and sustainable environment in the use and exploitation of data, avoiding the concentration of power and the exclusion of certain actors. This involves promoting fair competition, protecting the privacy and rights of individuals, and ensuring that the economic benefits generated by data are shared equitably and sustainably for society as a whole.

11. Strategic axes for the deployment of national data spaces

Below, the strategic axes and their corresponding measures are presented, aimed at achieving the previously stated objectives.

11.1. Axis A. Enabling elements for viability and sustainability

The precise enablers for the deployment of data spaces must be considered along their different dimensions. The business one, based on analysis of economic models, the promotion of cooperation and collaborative innovation; the legal dimension, offering responses to the contractual and regulatory considerations and needs of ecosystem participants; and the functional and operational dimension, promoting the liquidity of the ecosystems to be constituted, creating a wide range of services with which to make their exploitation more flexible and stimulate, from support communities that treasure and advance shared common knowledge.

The measures set out below make it possible to generate the necessary enablers for a viable and sustainable deployment of industrial data spaces.

- **Measure A.1 Generating adequate market conditions**

  The deployment of domain-specific data spaces requires generating adequate market conditions, and work must be done on the dynamization of data sharing across said domain, the formation and dissemination of the concept of data space, generation of methodologies and templates for the discovery and design of cases. of use, development of best practices and framework for building data spaces and community formation.

  Thus, the generation of incentives for data sharing must be addressed by promoting collaboration between the different actors involved, including government agencies, companies, non-profit organizations, researchers and citizens. In this sense, as part of an environment that encourages sharing, it is important to contribute to reducing the perceived costs surrounding it. This will be achieved by supporting more efficient solutions, developing standards of practical application, and contributing to the deployment of initiatives on data management and administration. In general, the conditions for creating trust and transparency inherent to the concept of data space must be created, enabling conditions for the process of continuous innovation around data.

  The generation of optimal market conditions must encourage the participation of the entire productive fabric, especially small and medium-sized companies, paying special attention to ensuring that anti-competitive practices of market concentration do not occur.

- **Measure A.2 Searching for social and territorial cohesion**

  The principle of territorial and social cohesion must be preserved both in the conception of the different data space initiatives and during their ordinary production.

  The conception and implementation of these spaces must be distributed throughout the different territories and productive sectors, taking advantage of the different innovation hubs deployed there. Thus, the role of autonomous communities and local entities is considered fundamental in the harmonious deployment of data spaces, due to their proximity and link to the territory, and their domain knowledge. Thus, it is key that they participate in said deployment by defining sectorial strategies and adopting public policies with specific territorial and social criteria.

  Part of the intended social cohesion also lies in the role of small and medium-sized businesses. Therefore, their active participation in the data spaces that are created must be encouraged, so that the advantages of their creation permeate the basis of the economic fabric. The aim is to contribute to the structural change that involves moving towards a more entrepreneurial economic model in the territory based on supporting the development of its tools and capabilities.

  Social cohesion ultimately involves involving people from different sociodemographic, economic and cultural contexts in the governance of their data. As generators of the data, your participation can help

ensure that the knowledge and value generated from it best represent your interests, avoiding promoting inequalities and guaranteeing equity in the assignment of value.

- **Measure A.3 Creation of a community**

  The concept of community is key to generating the necessary trust framework for the deployment of different data spaces. In any sector there are common and differential elements, unique to its field, the result of its specific business processes and a progressive and iterative generation of consensus and common knowledge. Community across a domain, understood as a meeting point for data, semantics, artifacts, standards, best practices, codes of conduct, constructive components and alternatives for data spaces in production.

  Creating and strengthening community requires developing dissemination and training actions that allow understanding the data space paradigm and knowing how to successfully apply use cases. Likewise, community formation can be reinforced from the public sector by supporting the effective deployment of the different national hubs of European data sharing initiatives, since these hubs are the natural place for knowledge sharing.

- **Measure A.4 Guaranteeing interoperability**

  The interoperability of the data space is one of its key elements and requires specific attention in its different dimensions (legal, organizational, semantic and technical) so that it is fully achieved. Beyond a given data space, it is also necessary to take into consideration the interconnection of different spaces. All of this implies that, adjusting to the maturity of the existing solutions, the different initiatives and standards for their adoption are proposed, assuming the minimization of the business risk that it entails.

  Consequently, guaranteeing interoperability involves facilitating the exchange of information based on the development and use of commonly used semantics and domain vocabularies. Semantics accessible in open access repositories that can be used in different use cases, even in different sectors.

- **Measure A.5 Independence of the underlying technological solution**

  Data spaces must seek to be independent of the underlying technological solution, allowing their portability and deployment in different physical infrastructures. This encourages the generation of an innovative ecosystem around solutions and services, where the key is the generation of business opportunities, of shared benefit among the participants, ideally beyond the short term, seeking their sustainability in the future.

  Given the lack of maturity of the technological developments on which data spaces are based, the priority must be the generation of business cases that value data sharing and exploitation scenarios. Simple solutions must be sought that do not raise entry barriers for small participants and guarantee ease of adoption and use.

  The business ecosystems thus generated are not closed to possible technological evolutions, which can allow taking advantage of the different European initiatives and standards once they are considered sufficiently mature to mitigate the business risk of their adoption.

- **Measure A.6 Deployment of laboratories**

  Overcoming the current lack of technological maturity requires actions such as the development of demonstrator environments, authentic spaces for technology and business experimentation, capable, in turn, of exerting driving capacity in the dynamization of the concept of data space.

  The experimentation laboratories will allow, without the requirement of long-term sustainability and market results, to test the deployment of the building blocks of the different alternatives for constructing

data spaces. Neutral, reliable platforms, upholding the European principles of data spaces, will be provided for the design, construction and experimentation of technological solutions for data spaces. This can cover both solutions close to development and research, as well as solutions closer to the market.

Knowledge and mastery of the technologies involved will allow Spanish companies and technology industries to position themselves in the *European Digital Single Market*, while continuing to advance Spanish technological sovereignty within the general European framework.

11.2. Axis B. Reliable data governance

The deployment of data spaces in the various sectors requires adequate data governance both in the data space itself and internally in the participating organizations.

- **Measure B.1 Establishing governance frameworks**

  The establishment of governance framework criteria for data spaces must be addressed throughout its different dimensions, guaranteeing transparency in data management, with special consideration of aspects related to privacy and security. Clear and public policies must be generated that explain how data is managed; as well as implement control and supervision measures to guarantee compliance with said policies.

  Designing new ways to address the risks inherent to data exchange in open ecosystems, and the resulting regulatory compliance, are key factors and must be prevented from becoming disincentives. Due to the large volume of data and algorithmic complexity, the supervision and regulation of data spaces, digital services and artificial intelligence systems surpass traditional human capabilities, making regulatory efficiency a challenge. Technological solutions, known as “Regulatory Technology” (*RegTech*), can be key to overcoming these challenges, since they improve regulatory compliance, reduce costs, improve risk management, promote innovation and encourage collaboration between various stakeholders. Trust is fundamental in the Data Economy and can be supported by RegTech through a clear assignment of rights, responsibilities and roles between various actors, being able to help small and medium-sized companies to deploy innovative digital services and products responsibly by reducing the compliance burden.

- **Measure B.2 Cybersecurity and data confidentiality**

  Data space cybersecurity must be present at all times. Data spaces will be combined, added and recomposed, deployed on common software infrastructures. If these data spaces do not provide the same level of cybersecurity from the beginning, the combined data will be, in terms of security, at the lowest common denominator and this will weaken the trust of its participants. Therefore, when building data spaces, it is necessary to create and manage security conditions with that perspective.

  Data sovereignty and data privacy are also key determinants of data spaces and, in their deployment, a fundamental part is the technologies necessary to guarantee them. Thus, as its deployment progresses, we will see the development of technologies aimed at guaranteeing decentralized identity management, probably supported by blockchain infrastructures; and, following the philosophy present in the deployment of Web 3.0, privacy guarantee technologies (PET) such as data computing, synthetic data generation, differential privacy or secure multi-party computing; as well as technologies aimed at satisfying regulatory compliance without attention.

- **Measure B.3 Assessing organizational maturity**

  The effective sharing of data requires internal data governance by the different intervening actors, governance developed considering the standards in this regard (UNE specifications 0077, 0078, 0079

[45]). With a view to promoting the evaluation of the different actors involved in a data sharing scenario, data processing maturity evaluations will be facilitated in accordance with the UNE 0080 and ISO 33000 specifications.

- **Measure B.4 Improving the quality of exchanged data**

  Data of proven quality does not have the same value as data that lacks it. Data quality directly affects the effectiveness and usefulness of data spaces, as well as decision-making based on that data or its uses by artificial intelligence tools. Low-quality data has a negative impact on the confidence of participants, questioning the viability of the system.

  A program for continuous improvement of data quality must be established (UNE specifications 0079, 0081 [45]), understood as the degree to which a set of data meets a series of characteristics such as completeness, validity, precision, uniqueness and consistency. consistent with the intended use of them. In this sense, the improvement program must establish uniform criteria with which to demonstrate the acquired level of quality.

  The quality of the data ends up having an impact on the artificial intelligence systems that make use of it in the context of a data space. Having adequate knowledge about the origin, scope, collection, preparation, correction, representativeness and privacy of a data set can allow choosing the appropriate techniques and tools, allowing the deployment of bias mitigation strategies.

11.3. Axis C. Driving efforts in key sectors

The deployment of data spaces in the various economic sectors requires selecting those considered strategic and focusing the development of driving projects on them. Its promotion requires the generation of public-private collaboration actions involving administrations, companies, technology centers, innovation centers, specialized centers to support digitalization, as well as the academic field, among others.

- **Measure C.1 Targeted support for industrial sectors**

  In the process of reindustrialization of the national economy, the aim is to attract or maintain in the country those phases of the value or production chains with the highest added value, seeking to promote shorter and more resilient value chains, as well as being more environmentally sustainable. The intensive use of data favors the digitalization of the value chains of the different economic sectors, promoting innovation and value generation, allowing better strategic decisions to be made and the creation of new value-added products and services.

  Data spaces must reinforce the role of the Spanish industry in those sectors where it currently holds leadership positions. Data spaces must support the digital performance of the key sectors of the Spanish economy, including at least its development in the agri-food sector, the sustainable mobility sector, the health sector, the tourism sector and the commerce sector. Promoting collaboration between companies is key to promoting international leadership.

  Thus, the development of data spaces must support the deployment of the connected vehicle, reinforcing the reindustrialization and electrification of mobility. In the health sector, it will promote cutting-edge solutions and precision medicine will be promoted, through the analysis and exploitation of data. For the agri-food sector, data spaces can improve its competitiveness, sustainability, traceability and food safety. In tourism, extracting value from data can lead to the development and sustainability of this essential industry. Finally, we must take advantage of the enormous potential of the language economy sector, of which Spain is a benchmark, seeking to lead the development of artificial intelligence in Spanish.

- **Measure C.2 Response to social challenges**

  Data spaces are a key tool in the response to contemporary social challenges, making a significant contribution to the achievement of the Sustainable Development Goals and being considered free from causing significant harm to the environment (DNSH principle, "*Do No Significant Harm*").

  Thus, on the one hand, we face an aging population, with greater needs for medical care and social assistance that put pressure on health and social services systems. In this context, data can be crucial to optimize the care economy and address the demographic challenge, through the collection, analysis and sharing of relevant data.

  On the other hand, the depopulation of rural areas in Spain is a problem that could be mitigated through decisive support for the rural economy. In this scenario, data and data spaces can play a vital role in promoting more effective policies and strategies for rural economic development, facilitating job creation, improving quality of life and promoting sustainability. in these regions.

  Finally, we are facing the ecological transition, a crucial challenge to address the impacts of climate change and ensure a sustainable future for future generations. Data spaces can be an important catalyst for this transition towards a more sustainable and environmentally friendly economic model. It is crucial that all actions to deploy data spaces are consistent with the principle of not causing significant damage to the environment, pursuing environmental sustainability and energy efficiency.

11.4. Axis D. Key role for public administrations

Administrations must become a fundamental part of the common European data spaces, adopting, depending on the case, a regulatory, operator or data provider/consumer role, giving rise to new public-private collaboration environments based on data. In this way, a radically transformative effect can be induced at a social level, enabling the sustainable growth of the regions, the diversification of the economy, and the development of new markets, lines of activity and business models based on data ("data-driven").

In any case, the regulatory role of public administrations is essential, as a formulator of public policies from which to establish the regulations (laws, policies, standards, etc.) with which to deploy, operate and participate in data spaces, or from where set the requirements for resilience and digital security, data protection, antitrust and intellectual property protection, or how to address the certification of data providers and data space operators, including the design of standards for the fair incorporation and user participation, and all this without neglecting its environmental sustainability.

- **Measure D.1 Involvement in industrial deployments**

  In specific sectors where public intervention solves a market failure, or there is a decided public vocation or legal mandate in this regard, or the criticality and sensitivity of the data so advises, the corresponding body may adopt, within the framework of its powers, a more active role in the design and operation of data spaces. The objective would not be so much to create a single large sectorial space, but rather to develop specific actions in specific situations, coordinated with the rest of the actions deployed in the sector.

  In any case, the sectorial knowledge of the different organizations involved is essential and must be present in the different public-private collaboration mechanisms that are addressed. This collaboration must be aimed at achieving the collectivization of the value generated, the long-term sustainability and self-sufficiency of the solutions addressed.

- **Measure D.2 Mobilizing the value of public data**

Actions must be promoted that make large volumes of data of proven value available to the productive fabric held by public administrations. These — although not necessarily or only them — are the source of data generation, and it will be necessary to encourage their publication.

The *AGE Data Platform* (Measure 6 "Transparent data management and exchange" within the 'Public Administration Digitization Plan') can be used to make information published as open data accessible, both internally and to other administrations. This release must be made under service conditions that enable linking of systems in production.

The high value data sets (HVDS) provided by the AGE would also be accessible from the Data Platform (via API, or bulk download), making them accessible with the appropriate service levels, internally and externally, paying attention whenever possible to its geospatial component. The ubiquity and interdisciplinary function of geospatial data makes them particularly valuable as base data on which to base other information, and their publication should be encouraged.

- **Measure D.3. Value-capture by administrations**

  Public administrations can obtain a clear return from their participation in domain and industrial data spaces not only by providing information, but also by accessing information provided by other participants. A first facet of this value capture derives from the promotion of transparency involved in participating in a data space, enabling an environment of responsibility and openness, which leads to greater trust in public institutions and a better understanding of government policies.

  The potential of valuable information present in the private sector must also be captured. Through their participation in data spaces, administrations have the opportunity to obtain crucial information from different industries and companies, which can be of great value for decision-making and policy formulation, allowing administrations to adjust their strategies to better reflect market conditions and needs.

  Finally, data altruism actions should also be encouraged, in the context of corporate social responsibility, encouraging organizations to share their information for the collective benefit within an environment of sovereignty and trust. Thus, if cases are identified in which the economic or social benefits of data sharing are manifest and this does not occur, such action could be motivatedly urged with the aim of responding to significant social problems, or seeking socially acceptable conditions. beneficial in terms of competition, regulation and public policies.

- **Measure D.4 Public administrations as an example**

  The exemplary role of public administrations, endorsing the design principles of data spaces, is fundamental. Strengthening collaboration within the public sector is embodied in public sector data spaces. The objective is, assuming a much more interdisciplinary and interdepartmental approach and taking advantage of the latest technologies, to project on a larger scale the current methodologies, specifications and practices related to information processing, achieving a fluid and continuous exchange of data between administrations, industrial sectors. and citizens, that generates advantages and opportunities for the different actors involved, and always keeping in mind the necessary privacy and security considerations.

  Public sector data spaces would be built around the *AGE Data Platform*, provided by the General Secretariat of Digital Administration as a common service, taking advantage of its storage capabilities, analytical possibilities and data governance tools.

12. Objectives, axes and measures

The table below reflects how the measures present in the different strategic axes (horizontal) directly or indirectly support the achievement of the strategic objectives (vertical).

	Support for strategic productive sectors	Guarantee of strategic digital autonomy	Promotion of the generation of quality data	Valuation of public data	Fair and sustainable Data Economy
A. Enabling elements of viability and sustainability					
A.1 Generating adequate market conditions	X	X			X
A.2 Searching for social and territorial cohesion	X				X
A.3 Creation of a community	X			X	X
A.4 Guaranteeing Interoperability	X	X	X		X
A.5 Independence of the technological solution		X			X
A.6 Deployment of laboratories	X	X			
B. Trustworthy data governance					
B.1 Establishing governance frameworks	X		X		X
B.2 Cybersecurity and data confidentiality		X			X
B.3 Assessing organizational maturity	X		X		
B.4 Improving the quality of exchanged data	X		X		
C. Driving efforts in key sectors					
C.1 Targeted support for industrial sectors	X			X	
C.2 Response to social challenges					X
D. Determinant role of public administrations					
D.1 Involvement in industrial deployments	X				X
D.2 Mobilizing the value of public data			X	X	
D.3. Value-capture by administrations			X	X	X
D.4 Public administrations as an example		X		X	X

Table 8. Measures present in strategic axes and objectives.

PART III. Heuristical considerations when deploying data spaces

Beyond the situation analysis carried out in the previous sections, it is also interesting to prescribe specific actions that facilitate the achievement of the stated objectives and the correct performance of the strategic lines and measures described above, while allowing us to define various considerations. The initial conception of data spaces will be carried out around the satisfaction of specific business needs, articulated around use cases. These use cases are a practical application of data processes and technologies to achieve a shared business objective, within the data space environment.

The latter is therefore a facilitator of satisfying a business need. The process of launching a data space must respond to the different dimensions of data spaces described above, satisfy the characteristics listed, establish the appropriate governance models, and rely as much as possible on preconceived technological solutions.

In any case, the conception of the data space and its implementation must always take into account its scalability (both of use cases, data sets and services, and of participants), as well as its interoperability with other data spaces. In the near future, once different spaces have been deployed, faced with an opportunity in the data market or an unsatisfied need for information, or simply faced with a business opportunity in data intermediation, an organization would proceed to study the different spaces of existing data, choosing the one that - covering your business requirements - offers better conditions of service, access, quality and/or price.

13. Generation of the right conditions for conceiving & deploying data spaces

Each sector and thematic domain has its own specificity in terms of data types, data flows, business models and stakeholder needs. Therefore, the best way to advance the conception of data spaces is a community-based approach, through the co-design and co-creation of data spaces that consider the specific context of each domain. Thus, the deployment of industrial data spaces requires active listening to the voice of the market, which leads to the compression of the data dynamics in each sector, and the generation of the necessary market conditions.

13.1. Listening the voice of the market

In building solid and effective strategies for the development of data spaces, it is essential to prioritize and value the voice of the market. Actively listening to this promotes greater alignment between technological capabilities and the functional demands of the industry, encourages innovation and ensures that the efforts invested in the creation of data spaces are genuinely beneficial and profitable for all actors involved. . Ultimately, giving prominence to the voice of the market in this context reinforces the relevance, sustainability and positive impact of domain data spaces in the digital economy.

Thus, we believe in promoting the representation of the industry in the definition and implementation of industrial data spaces, through a general methodology and a process of continuous improvement, which we can summarize around the following steps:

- **Industry experience and innovation:** Use cases enacted from the market play a fundamental role, allowing experimentation, learning and innovation in different sectors. These use cases can be carried out both internally in organizations and in external environments, whether open or closed, public, private or hybrid. These experiences help identify the specific needs and challenges of each sector, as well as explore

new business and development opportunities. It is the basis on which to support organizational incentives to explore and capitalize on the benefits of these new data sharing and exploitation paradigms.

- **Consolidation of experiences and knowledge:** The aforementioned sectorial projects (use cases) accumulate valuable experiences and results, which can be consolidated at the national level from the different *data-driven* innovation communities. These communities drive and coordinate the respective sector ecosystems, shaping new projects in (and updates to) already operational data spaces. The consolidation of these experiences between communities also leads to greater interrelation between different ecosystems, generating interrelated domain communities that develop and increase shared knowledge. This national cooperation encourages the emergence of new projects and growth opportunities.
- **Convergence with existing European models:** As - from the design - national data spaces must be interconnected with European ones, learning at a national level will also seek convergence with advances at a general level. The EC is working on the compilation and synthesis of different blueprints and models for data spaces from the DSSC project, with the idea of eventually shaping a unified one that serves to deploy the common European spaces, guarantors of that interoperability. In this sense, we believe that Spanish experiences should also be reflected in the construction of said European *blueprint*, thus facilitating the desired interconnection, and enabling our spaces to access data resources from a much more extensive ecosystem.

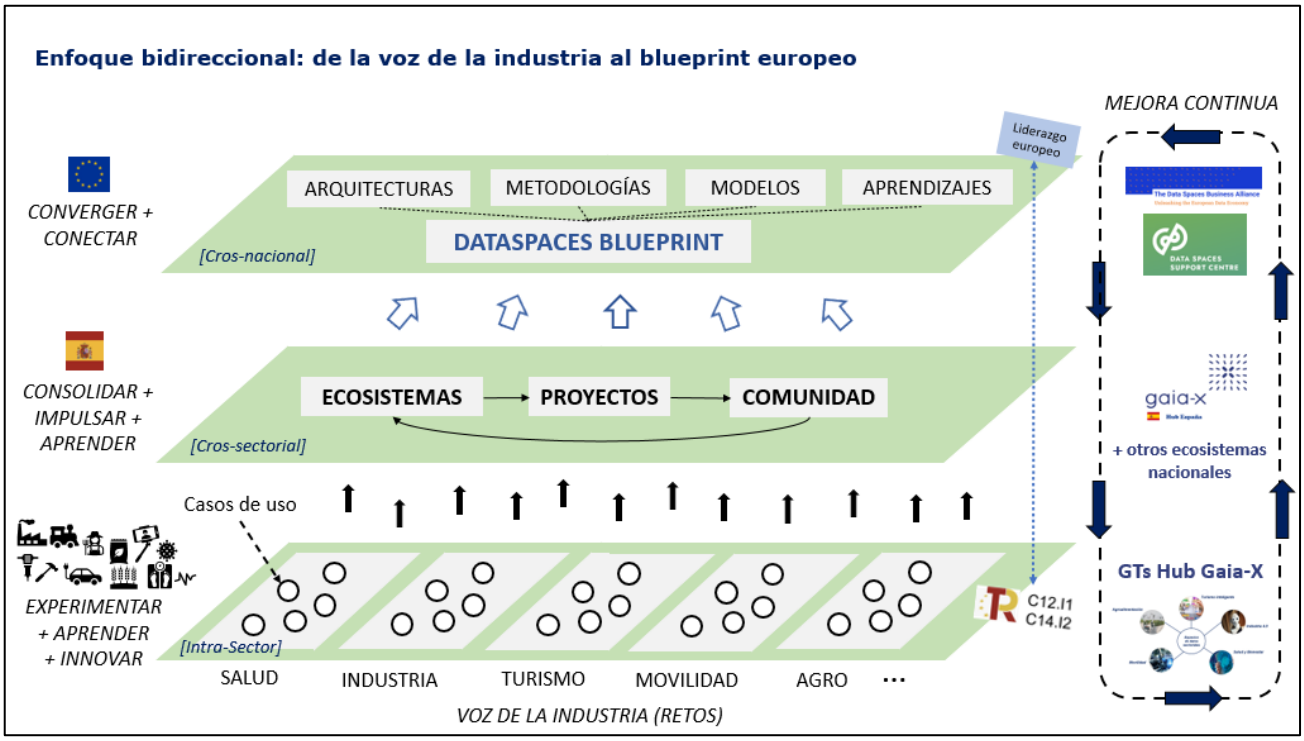

Fig 25. Bidirectional approach: From the voice of the Spanish industry to the European blueprint.

Progress is being made in the generation of boundary conditions that enable the viability and sustainability of data spaces. These boundary conditions, understood as generating the capabilities of the ecosystem of participants around the concepts of community and promoting their capacity to act, in line with European specifications, will enable the generation of use cases, and their subsequent deployment in spaces of data. Through experimentation, the consolidation of experiences and convergence with European specifications, innovation is promoted, domain communities are generated and collaboration is encouraged at national and European level.

The communication and publication of the progress made, both internally and outside the organization, is essential to ensure progress and promote transparency of actions.

13.2. Analyzing the data potential across a given industry/ domain

The actions mentioned above should lead to a deep knowledge of the potential of data in the sector under study. This knowledge is essential to thoroughly understand the value that data can provide in said sector, providing us with a comprehensive and enriching vision of how data spaces can be implemented and used in the specific sector, promoting transformation and innovation in said domain.

This domain knowledge should cover various key aspects that help visualize the strategic importance of data in said sector and its evolution over time. Thus, it should be analyzed how data management has been changing in the sector, examining the main trends and evolutions. These actions will allow us to understand the economic and social impact that data can have, highlighting the importance of effective data management. In addition, current and future sector initiatives that drive the development of data spaces will be identified and analyzed, exploring the interrelationships between data spaces and the digitization of the sector.

By studying the market analysis and value chain of the sector, it will be examined how data spaces can add value along the value chain. The main actors and their relationship with data spaces will be identified, and emerging trends and developments that may influence these spaces are highlighted. Likewise, the importance of data ecosystems and the creation of communities around data in the sector under study will be considered. These concepts are key to promoting collaboration and the exchange of knowledge between the different actors in the sector.

Legal and regulatory analysis is also essential, and the laws and regulations that impact data management in the sector must be listed. It examines how these regulations specifically affect the sector and are considered when designing and managing data spaces. In addition, strategies and initiatives that can drive the development of future data spaces in the sector must be identified. These strategies seek to promote the adoption and effective use of data spaces, taking into account the needs and particularities of the sector in question.

Finally, the future of data spaces in the sector will be projected in the medium and long term. Current predictions and trends, as well as technological innovations that may influence them, should be analyzed. Strategic areas of interest are also identified and potential future risks and challenges that could arise in relation to data spaces in the sector are considered.

14. Designing data spaces

The conception of a data space must be guided by the satisfaction of an unresolved business need. Its conception process involves an evolution from the use case to a minimum data space, from which to proceed to its scalability, always keeping in mind its interoperability with other data spaces (from the same or different sectors).

The first step would be to identify a specific use case within a certain sector. This use case must address a specific problem and have the potential to generate value through the exchange of data, ideally solving the problem in a way that would not be possible working individually. Once decided on its viability, it will be necessary to carry out the initial design of the use case, defining the actors involved, the necessary data, the required processes and the expected results.

With the feasibility and design of the use case already defined and validated, a minimum data space (MVDS) is developed. The MVDS is an initial, functional version of the data space that focuses on meeting the minimum needs of the identified use case. Once this has been developed, it undergoes a validation process, and feedback from users and stakeholders is collected. This allows you to identify possible improvements, adjustments and additional functionality that may be necessary to optimize the use of data space and meet the needs of the use case. Based on the feedback received, the aforementioned MVDS can be scaled and generalized to become part of the data spaces present in a certain sector, having to resolve at that time the issues related to its scalability and interoperability.

14.1. Design and feasibility analysis for data sharing use-cases

Any data sharing action and eventual conception or participation in a data space must be preceded by a feasibility study to determine its scope, evaluate its potential and limit its complexity. This process is not trivial; It is not immediate to discover unmet business needs, exceeding technological constraints, covering regulatory, organizational and governance aspects, and seeking to satisfy a certain business opportunity or challenge of the different actors in the industries involved.

Its deployment should be understood as an open innovation process, subject to a cycle of results analysis and continuous improvement. Challenges in developing use cases typically include:

- Identify challenges or business opportunities achievable through data exchange
- Clearly determine the scope of the use case to ensure the focus and timelines for its development
- Evaluate the potential value created by the use case
- Evaluate how the scope of the use case satisfies interoperability and trust requirements
- Create a use case design that is scalable and reusable.
- Meet economic and emotional expectations
- Study how to visualize results. Set key effectiveness indicators.

These actions can be included in three stages present in the development of any use case: opportunity detection, feasibility evaluation and use case design. In the first stage, only the opportunity part will be analyzed, defining aspects such as the customer experience, the value chain or the economic-emotional business case at a high level. It is possible that once you move on to the next stage, and begin to fill out the feasibility evaluation templates, you will have to think again in more detail about the opportunity and, in the same way, when you advance to the stage design with respect to the previous two. This iterative process will facilitate the completion of the templates, moving through the different sections as the use case matures.

The Spanish *National* Data Office, exercising its function of energizing the governance, management, sharing and use of data, has published two guides to facilitate the implementation of data sharing use cases: the 'Use Case Design Guide' [46] and the 'Use Case Feasibility Assessment Guide' [47].

14.2. Conceiving data spaces from the use-case perspective

The process of capturing a use case within a minimal data space can be described as a journey through different stages where different aspects across the business, legal, operational, functional and technological dimensions are responded.

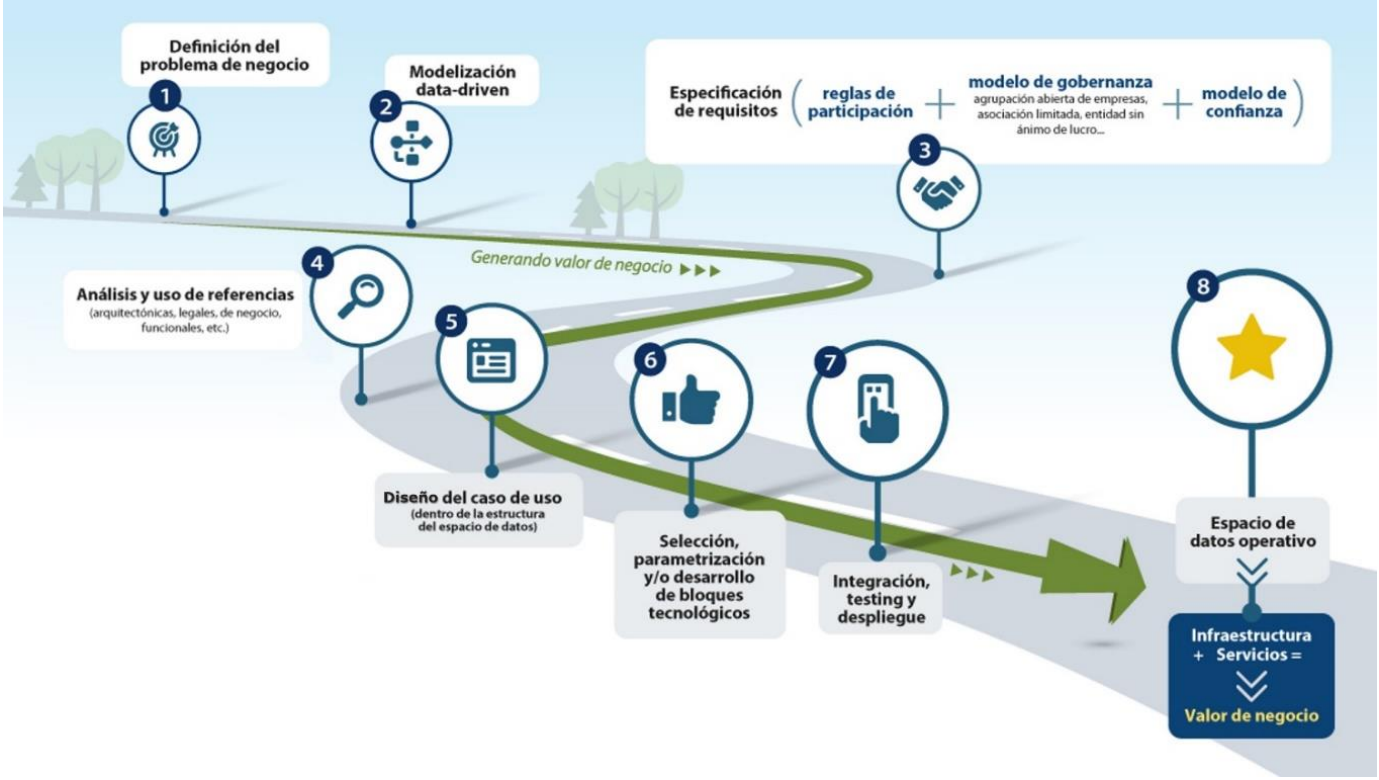

Fig 26. Developmental model for use cases, giving birth and populating data spaces.

The stages considered within this initial conception model of the data space are the following:

- **Stage 1: Definition of the business problem**. In this phase, a group of potential participants detects an opportunity around the sharing of their data —- until then in silos —- and its corresponding exploitation. This opportunity may refer to new products or services, efficiency improvements, or the resolution of a business problem. That is, there is a business objective that the group is capable of solving together, sharing data. The 'Guide for evaluating the viability of use cases' will help the correct development of this stage, also serving as support in the following ones.

- ***Data-driven* modeling**. In this phase, those elements that serve to structure and organize the data for making strategic decisions based on its exploitation will be identified. It involves defining a model that possibly uses multidisciplinary tools to achieve business results. It is the part that is traditionally associated with data science tasks.

- **Stage 3: Consensus on requirements specification**. Here, the actors who sponsor the use case must establish the relationship model that they will have during this collaborative project around the data. This model must define and establish the rules of participation, establish a common set of policies and governance model, and create a model of trust that acts as the root of said relationship.

- **Stage 4: Analysis and use of references.** Before carrying out Phase 5 “Design of the use case”, and in order not to create it from scratch, the multiple existing references related to the objective of being able to reuse resources and techniques already developed will be identified, analyzed and synthesized. For example, models, methodologies, artifacts, templates, technological components or solutions as a

service. Thus, just as an architect designing a building can reuse recognized standards, in the world of data spaces there are also models on which to paint the components and processes of a use case.

- **Stage 5: Use case design**. The design is the means of expression of the ideas of those who have defined and agreed on the use case, and must explicitly include the solutions proposed for each of the parts of its development. This design is unique for each use case, undertaking the development of a functional technical document that includes all the characteristics of the case to be developed, based on the references identified in the previous stage. The 'Use Case Design Guide' will help the correct development of this stage. To carry out this stage, it will also be necessary to carry out project management activities related to the definition of scope and planning (times, costs, risks, resources, etc.).

- **Stage 6: Selection, parameterization and/or technological development**. Technology enables the deployment of data transformation and exploitation, favoring the entire life cycle, from its collection to its value. In this phase, the infrastructure that supports the use case is implemented, understood as the collection of tools, platforms, applications and/or pieces of software necessary for the operation of the application. To carry out this phase, it will also be necessary to carry out project management activities related to the definition, planning and execution of the development of the selected technology, generating the information or documentation necessary for its correct achievement.

- **Stage 7: Integration, testing and deployment**. Like any technological construction process, the use case will go through the integration, testing and deployment phases. Integration work and tests on functional, usability, exploratory nature, acceptance, etc., will help us achieve the desired configuration for the operational deployment of the use case. In the case of wishing to incorporate a use case into a pre-existing data space, the integration would seek to fit within its structure, which means modeling the requirements of said use case within the processes and building blocks of the data space. To carry out this phase, it will also be necessary to carry out project management activities related to the execution and control of the integration, testing and deployment of the use case as an operational data space, generating the information or documentation necessary for its correct achievement. and evaluation.

- **Stage 8: Operational data space**. The end point of this journey is the live use case, which will employ digital services deployed on top of the data space fabric. This implies that the life cycle of value creation based on shared data would have been efficiently articulated, and business revenue would be obtained according to the original approach. In this stage, service management activities related to the monitoring and control of the data space already in place will be carried out, generating the information or documentation necessary for its correct operational management.

  However, this does not prevent the data space from continuing to evolve after the fact, since its vocation is to grow either with the entry of new challenges, or with the entry of new actors to existing use cases. In fact, the scalability of the model is one of its unique benefits, being able to consider a federation from the design that can support different resources and functionalities in the future.

This oversimplified model can also be equally seen from a slightly different perspective, perhaps less dynamic, that emphasizes the building phases mapped to the different layers making up data spaces. In particular, the following diagram shows a pyramid-like structure where we have sliced layers to indicate that they represent relevant technical domains (not just technological though, as Data Science acumen can also be approached from a mathematical and data modeling perspective).

The diagram shows how domain layers lie on top of each other, and this is done with a purpose. It also contains sample graphical representations of some of the deliverables, frameworks or tools within each domain layer, color-coded so they can be more easily associated.

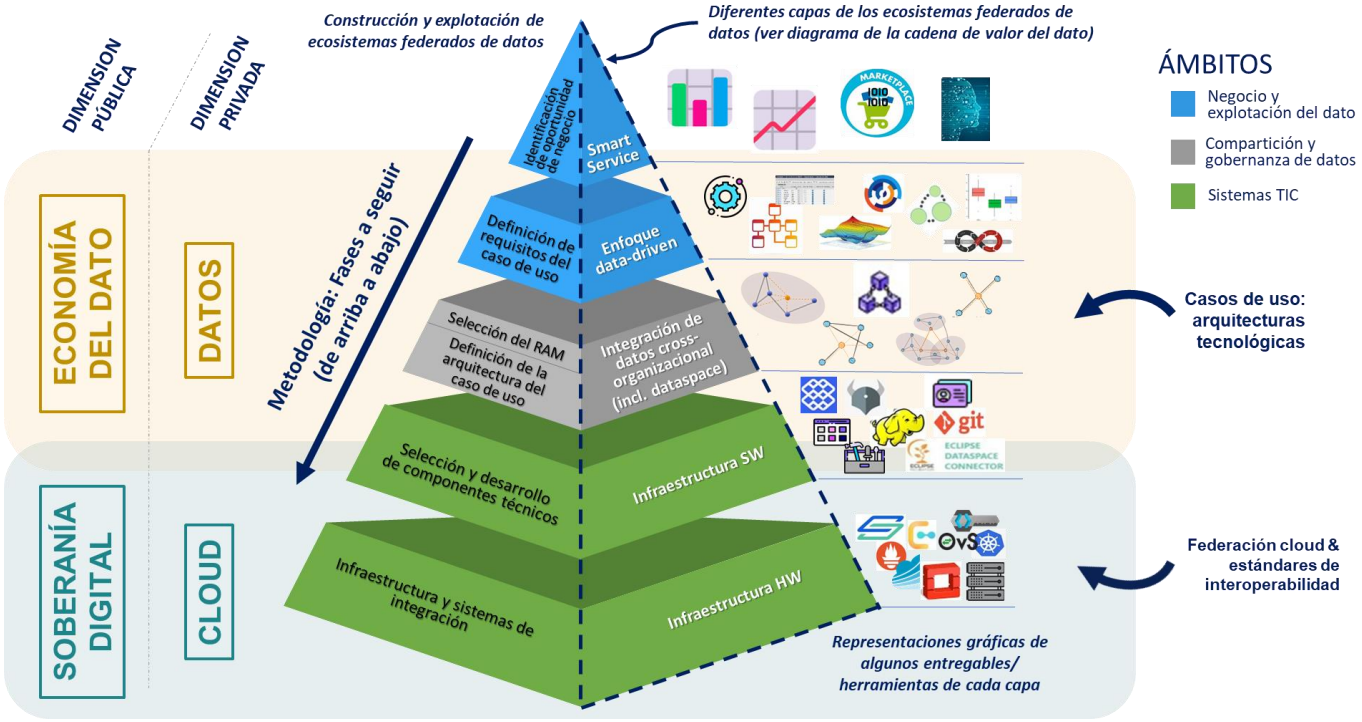

Fig 27. Technical layers within a methodology for building data spaces.

14.3. Community, trust and simplicity

The first steps in the conception of a data space can be favored by taking into account measures to enhance the generation of community, the creation of trust and the simplicity of the system [48].

- **Motivation for the common good:** It is interesting that when starting a new data space, the parties involved share interests, and that the main focus is not exclusively economic. Data sharing based on shared interests becomes a catalyst for mutual growth and achievement of common goals. This collaborative approach encourages the creation of sustainable solutions, promotes transparency, and serves to strengthen long-term relationships.
- **Involvement of owners of trusted and highly influential data sources:** It is a recommended practice to begin with data sharing based on a series of well-established and highly reliable sources, it being important that the owners of these are few, but influential, of in such a way that these pioneers serve as a driving element of the ecosystem, attracting new actors until the necessary critical mass is reached.
- **Simplicity in the service:** Simple solutions must be sought, which is not necessarily equivalent to simple. Thus, it is important to clearly explain the operating conditions of the data space. In this sense, before developing a regulatory and governance framework from scratch, it is good practice to reuse existing frameworks, which have proven useful, and adapt it to the operation of one's data space, as necessary. The initial exchange of data and services should be focused on simplicity, avoiding unnecessary complexities initially, relying as much as possible on available standard solutions and tools. Starting with data sharing whose uses do not include information subject to privacy restrictions, perhaps with an initial objective of consumption and internal improvement, initially far from an intense commercial return, helps

to generate a system of comfort between owners and suppliers that will be able to establish the collaboration bases in a future scaling of the data space.

14.4. Data space scalability

The data space conceived in this way must pay special attention, from its design, to its scalability. Once the data space is up and running, a continuous improvement process is established to optimize its performance and adapt it to new needs and trends in the sector. User feedback is collected and analyzed, updates and improvements are made, and new opportunities for growth and expansion of the data space in the sector are explored.

Scalability should be understood as the possibility of hosting new use cases, new data sets subject to exchange, new participants, and new value-added services, adapting to future needs. This involves developing a long-term vision from the start and considering how the solution can evolve as its scope expands and new participants are added.

In addition, simplicity must be sought in the proposed solutions. It is essential to avoid entry barriers for small participants, promoting ease of adoption and use of the solution. This allows a wide range of actors, regardless of size or resources, to benefit from and contribute to the use case. In technical terms, an agile and simple deployment of the solution is sought, but at the same time, robust and scalable. This implies using technologies and architectures that allow efficient development and implementation without major difficulties.

14.5. Best practices in governing data spaces

Establishing and maintaining an effective data space governance system can be a significant challenge. As organizational boundaries are crossed and data is shared, questions arise about the reliability, security, privacy, and appropriate use of the data. To address these challenges and ensure effective data space governance, it is not possible to provide universal recipes, although it is possible to list best practices that can be taken into consideration:

- **Collaboration and communication.** Promoting a culture of collaboration and open communication among all stakeholders is essential for the effectiveness of the governance proposed for the data space. This involves fostering an environment where ideas, knowledge and perspectives are shared, and where feedback can be given and received constructively. Organizations must work together to solve problems, identify opportunities and develop strategies to enhance data value. Effective communication can also help ensure that all participants understand and follow established governance policies and protocols.
- **Monitoring and continuous improvement.** The effectiveness of data space governance must be continually monitored and improved to respond to the changing needs and challenges of organizations. This may involve carrying out system audits, reviews of established policies, as well as performance evaluations of the services provided and the quality of the data exchanged, with implications for its correct governance and management. Organizations must be willing to adapt their approaches and strategies based on the results of these assessments. Continuous improvement may also involve ongoing training and education of staff to ensure they have the necessary knowledge and skills to effectively handle data.
- **Transparency.** Transparency is a key component of data space governance, helping with regulatory compliance and building and maintaining trust. This means being open and honest about how data is collected, used and shared, establishing its provenance and responsibility for any data exchanged. The governance policies and procedures of the data space must be clearly communicated to all its

participants. Transparency can also help build trust between organizations and ensure that they all comply with established policies and protocols.

- **Adaptability.** Data space governance must be adaptable to respond to the changing needs, challenges and opportunities of organizations and their environments. This may require periodic review and updating of your policies and procedures, including the adoption of new approaches and technologies for managing services and data. Adaptability also involves being flexible in terms of collaboration with other organizations, and being able to adapt to their needs and requirements. This is in fact what underlies the EIF interoperability framework that we saw in PART I. Concept of a data space.

15. Practical considerations for useful and far-reaching shared data

15.1. The importance of the quality of data exchanged

In a constantly changing environment and in which data generation is growing exponentially, it is necessary to establish common processes throughout the life cycle of the data assets of any organization. Having well-governed, managed data with adequate levels of quality becomes a necessity for all types of institutions, and a common evaluation methodology is necessary that can help with continuous improvement of said processes and allows evaluating the maturity of a data organization. standardized form.

In any case, the governance of data spaces should not be confused with the governance and management of data deployed by the different participants within their organizational contour. Clearly, without the existence of data governance and management in an organization, it will be difficult to have quality data to offer to the data space, nor will it be possible to take full advantage of the data consumed within a data space.

The *National Data Office* has sponsored, promoted and participated in the generation of the national specification UNE 0080:2023 [49], which aims to respond to the need for a process evaluation guide based on international standards. A document with which organizations, both public and private, can systematically evaluate the current state or degree of achievement of their work processes in terms of governance, management and data quality management based on the definition given for those processes. in the UNE specifications of 'Data governance' UNE 0077:2023 [50], 'Data management' UNE 0078:2023 [51], and 'Data quality management' UNE 0079:2023 [52].

The development of a homogeneous maturity assessment framework for an organization, in relation to its data processing, provides the opportunity to analyze in detail the internal procedures it carries out, identifying weaknesses associated with these procedures and allowing the establishment of action plans to its optimization and improvement. Strengthening these internal processes ensures greater reliability and quality of data, which also results in greater and easier participation in new data markets.

Additionally, when an organization wants to integrate its data into a data space for future intermediation, it would be advisable to perform a quality assessment, labeling the data set appropriately. The UNE 0081:2023 specification [53] provides a quality model and a quality assessment methodology from which to approach the aforementioned accreditation. Data of proven quality has a usefulness and value different from that which lacks it, positioning the former in a preferential place within the competitive market. Beyond the evaluation of the quality of the data product, its management and continuous improvement can be addressed following the quality management instructions of UNE 0079:2023 [52].

Within a specific data space, or sectorially, different quality scales can be thought of that prioritize some of the characteristics or properties; or even in the existence of predefined common business rules that facilitate the comparison of results. All these decisions must be made within the framework of the aforementioned governance of the data space. In any case, the quality of the data must be part of the metadata of the dataset

or the service of the data product offered. Ideally, it should be done in such a way that its verification can be carried out in an automated manner, along with the rest of the technical compliance ("*technical enforcement*") aimed at guaranteeing data sovereignty.

Thus, the quality of a data set, or some of its characteristics or properties, could be accredited through a verifiable credential, ultimately verifiable against the supplying entity using some type of distributed registration technology. This automated validation does not exclude the opportunity to demonstrate quality within a data marketplace *through* some type of visual badge according to the level achieved.

15.2. Perspectives on privacy

Data spaces must guarantee adequate treatment of privacy, generating trust and legal security, in line with the regulatory framework established by the General Data Protection Regulation (GDPR, Regulation (EU) 2016/679) [6]ensuring a high level and protective uniform. Data Spaces represent significant opportunities, but require critical and objective analysis to ensure that their implementation is compatible with fundamental rights and freedoms.

As mentioned, data spaces present complexity in terms of organization, laws, technology, and scale, covering various subjects, data, regions and times. Industrial sector data spaces will not always handle personal data, but -if they do- they must consider privacy from the design, considering the appropriate risk analysis. The Spanish Data Protection Agency (AEPD) addresses these issues in the document '*Approach to data spaces from the perspective of the RGP* ' [54], aimed at data controllers and processors involved in data spaces , as well as DPDs, data protection advisors, and all parties involved in data processing within the framework of a data sharing model.

Privacy Enhanced Technologies (PETs) enable the sharing and exploitation of sensitive data. Some techniques protect the input data, others focus on preserving the privacy of the output data, and a third group protects the algorithms involved using artificial intelligence techniques and tools. Beyond classic anonymization solutions, the most prominent technologies, at different levels of maturity, include *Secure Multiparty Computing*, *Differential Privacy*, *Homomorphic Encryption*, *Federated Learning*, *Zero-Knowledge Proofs*, *Trusted Execution Environments* (TEEs), and *Pseudonymization*. In any case, the opportunity and maturity of the selected technology must be analyzed, considering its associated costs within the appropriate risk analysis.

15.3. Licensing of open data across data spaces

Open data, both those provided by the public sector and the private sector, are a fundamental part of data spaces, and those actions that make them available to the productive fabric with clear and homogeneous access conditions must be promoted.

Facilitating the reuse of public sector data requires analyzing the role that licensing formulas can play in this area. Directive 2019/1024 [9] establishes that, in any case, licenses must be proportionate, non-discriminatory, justified by public interest objectives, and that they do not unnecessarily hinder the possibilities of reuse or restrict competition. This obligation has led our legislator to opt for the use of open licenses (art. 9 Law 37/2007), whether a government open license or a standardized open license.

Open licenses are those that are available online; allow anyone, for any purpose, to access, use, modify and share data and content free of charge; and that are based on open formats. These licenses indicate the specific conditions applicable at each moment over time, including their validity period and the conditions for each type of information. The best-known free open licenses are *Creative Commons* and *Open Data Commons*.

The use of CC-BY 4.0 licenses allows us to offer a response that is highly compatible with the ultimate intention of homogeneous access to various data sources. In fact, the Commission's implementing regulation 2023/138, which develops the provisions of *Directive (EU) 2019/1024* [9] in aspects related to specific high-value data sets, it establishes in its article 4 that these sets "will be made available for reuse under the conditions of dedication to the public domain of *Creative Commons* (CC0) or, alternatively, of the *Creative Commons BY 4.0* license or any equivalent or less restrictive open license".

The purpose of the recommendation of these licenses is, on the one hand, to guarantee that public data is made available with a minimum legal restriction and, on the other hand, that the conditions of reuse are harmonized within the European Union to facilitate the development of cross-border applications and services. Therefore, it seems reasonable to think that, if this recommendation is followed, extending it to all open data sets would be logical to ensure that legal interoperability is promoted, and because it would greatly facilitate the reuse of all open datasets, both for national and foreign agents.

16. Interconnection across data spaces

As has been expressed, under the data space paradigm there is room for both centralized environments for the agglutination of information and generation of value-added services, with or without economic compensation, as well as innovative federated environments for data sharing and federated exploitation of computational resources. or algorithmic.

Thus, the creation of data spaces on different technologies and under different data sharing paradigms is enabled, making their interconnection a critical aspect, understood as interoperability between different data spaces. The different data spaces built must keep in mind that in the near future they must be able to dialogue. And they must be, as stated in section 6.1.2, through the legal, organizational, semantic and technical levels, for which the appropriate technological and semantic standards, as well as policies, norms and procedures, must be defined.

At the level of the different European common spaces to be established, it is possible that -for example- the European mobility data space should interact and be linked with the tourism or health data space. Therefore, it will be necessary to accommodate different datasets and services, and different layers of data cleaning, enrichment and analysis, which could be separated by legal, institutional, infrastructure or even administrative silos. All this, with a view to taking advantage of the synergies between the different initiatives and with the ultimate goal of paving the way towards the interoperable data federation foreseen in the European Data Strategy.

Even with the current situation of lack of complete and mature technical solutions, it is advisable to establish criteria to be met by the different data spaces created today. These can be conceived on the basis of different maturity scenarios, and assessing the innovative appetite of each sector, as well as the national strategic positioning. In the absence of a complete family of standards, and starting as a minimum possible criterion of guaranteeing interoperability between different data spaces, and with a view to covering at least semantic interoperability, a minimum of services available in each data space is proposed, as well as a unique catalog of the data resources present in each of them:

- The service catalog would be prepared in accordance with Gaia-X specifications, supported by verifiable credentials, and with its conformity potentially validable via the Gaia-X Digital Clearing House. At least one (1) data service should exist for each data space.
- The data catalog may also reference its corresponding service, which will have precise information on the different resources available in said data space, for access and use. The way to prepare the catalog

will be around the DCAT-AP standard. This would allow the federation of the data catalogs thus formed to be carried out in an analogous and synergistic manner to how data held by the public sector is being federated today.

And this is because at the European level it is planned to federate high-value data sets (in compliance with the *Open Data and Reuse of Public Sector Information Directive*), as well as data sets accessible for research purposes (*Data Governance Act*) through a single federated catalogue, which in Spain will take shape around the *data.gob.es* portal.

For this reason, and based on the use of a compatible mechanism throughout the different industrial data spaces, the range of action of the national portal is expanded (which in this way reflects the catalogs in said data spaces), and also allows the latter to interoperate with public resources in a much more native way.

In the medium term, beyond these minimum criteria and as the maturity of existing solutions progresses, progress must be made in the use of "technical compliance" standards for identification, authentication, authorization, as well as communication and modeling. of data and metadata. For this, it is essential to monitor the maturity in the deployment of Gaia-X and the rest of the European technological initiatives for data spaces, in a journey towards a convergent technological framework that is already underway within the *Data Spaces Business Alliance* and the *Data Spaces Support Centre* .

17. Characteristics to consider in developing data spaces

There is no single way to address the different dimensions of a data space and answer the questions raised there. But it is possible, from the principles intrinsic to data spaces in general, and federated ones in particular, to state a set of business and technical characteristics that the different data spaces must follow.

17.1. Data spaces' business characteristics

ID	Characteristic	Justification	Importance
N01	Existence of a data space governance model.	Its presence favors transparency in management and trust in the administration of the system.	Fundamental
N02	Definition of a mechanism of financial and non-financial incentives for data exchange, as well as a mechanism of agreements between the parties.	Ensure that the necessary conditions are created to ensure the sustainability of the data space.	Fundamental
N03	Definition of an ontology and semantics that connects with the specific characteristics of each sector (business language).	Guarantee that the data space is a tool aligned with the business needs of the actors involved.	Fundamental
N04	Define an ROI matrix for each participant in the data space.	Ensure the maximum level of involvement of data space participants, connecting with incentive mechanisms, value proposition.	Recommendable
N05	Define a matrix of OKRs and KPIs that allow measuring the success of the use of the data space in different dimensions:	Ensure that the effectiveness of the data space is monitored and that primary and secondary objectives are	Recommendable

	use, shared data, participants, services offered...). Creation of a monitoring and control dashboard.	achieved. Guarantee control of its evolution and facilitate a strategy of continuous improvement and capacity for action to correct possible deviations.	
N06	Define the personalized value proposal appropriate to the needs of each group of participants in its different dimensions, and establish a communication strategy with the objective of attracting new members and improving the NPS of the members.	Convert the data space into a service aligned with the needs and expectations of its users and *stakeholders*, and provide it with a layer of visibility that guarantees its growth and scalability, as well as its sustainability.	Fundamental
N07	Define a successful Change Management strategy that incorporates a team structure, sponsorship model, special tactics (for anticipated resistance) and risk assessment.	Optimize the adoption cycle of the data space, especially in its initial "*awareness*" and "*interest*" phases.	Recommendable
N08	Define a project *roadmap following* project management *frameworks (such as PMI, agile...) that allow correct control and possible scaling of the project.*	Ensure the optimal execution of the program associated with a *journey* (phases) and its control, monitoring and possible escalation by a data space program management team.	Recommendable
N09	Definition of the Project team: roles, responsibilities, etc.	Ensure the optimal execution of the program associated with a *journey* (phases) and its control, monitoring and possible escalation by a data space program management team.	Recommendable
N10	Definition of a data space knowledge management strategy that guarantees access to appropriate knowledge for all users.	Ensure learning within the data space. *Learning Company* Approach (continuous learning and improvement).	Recommendable
N11	Definition of a learning strategy throughout the *journey* that guarantees the replicability of micro projects and/or use cases and that facilitates the creation of "*best practices*" to share within the data space and with other data spaces.	Ensure the scalability of the data space and its ability to accelerate the adoption of sector data spaces in other industries. Ensure public-private learning and collaboration (data space models).	Recommendable
N12	Definition of an initial portfolio of use cases associated with business problems (following typical methodologies, see e.g., [46], [47]).	Guarantee that the data space is connected to business creation in the industry and that the Data Economy model is disseminated within the sector.	Recommendable
N13	Definition of a system for measuring the impact on the users' business derived from the data space.	Determine the ROI of data spaces in sectors and their contribution to	Recommendable

		growth and creation of wealth and employment	
N14	Definition of a *stakeholder*-services-knowledge map for the data space.	Ensure an adequate balance of actors in the data space and that the capabilities and services are adequate to obtain results.	Optional
N15	Existence of a code of conduct considering ethical aspects of data processing.	Adherence to codes of conduct promotes trust. In particular, the monitoring of the principles set forth in the Charter of Digital Rights.	Fundamental

Table 9. Table of business requirements in data spaces.

17.2. Data spaces' technical characteristics

ID	Characteristic	Justification	Importance
T01	Solution based on an available reference implementation sponsored by the Data Space Business Alliance (Gaia-X, IDSA, FIWARE, Simpl).	Need to promote data spaces aligned with the strategy and developments sponsored by the European Union. Ideally certified by an authorized third party.	Recommendable
T02	Technological solution independent of the underlying technological infrastructure.	The system must be portable to different environments.	Recommendable
T03	Technological solution available to the community under open source license.	Entry barriers are lowered, helping to launch similar initiatives.	Recommendable
T04	Guarantee of cybersecurity and privacy of data space.	It is necessary to ensure compliance with security management standards in the system.	Fundamental
T05	Monitoring interoperability within the data space	Legal, organizational, semantic and technical interoperability	Fundamental
T06	Reliable interconnection (in practice) with other data space initiatives.	The formation of silos is avoided and the coherent deployment of the Data Economy is favored.	Recommendable
T07	Existence of a future evolution plan to adapt to the Gaia-X Trust Framework and use of the Gaia-X Digital Clearing House.	The Gaia-X Trust Framework is the leading initiative to formalize data space interoperability.	Optional
T08	Data space as an open and heterogeneous ecosystem without dominant actors and without disproportionate entry and exit barriers.	Ecosystems favor a fair market, innovation and prevent the emergence of actors that exercise dominant positions.	Fundamental

T09	Existence of mechanisms to guarantee trust (identification, authentication and authorization).	It is necessary that participants identify themselves and be able to be authenticated at all times, stating the authorizations granted.	Fundamental
T10	Trust (identification, authentication and authorization) will be based on decentralized identifiers (*Self-Sovereign Identity*) and the use of DLT (*Distributed Ledger Technology*).	SSI-based trust is the convergence solution of choice for leading technology providers. Its use facilitates future migration towards an interoperable system.	Optional
T11	Existence of tools for recording the activity of the participants.	Necessary to audit the use of data space and to calculate compensation between participants, if applicable.	Fundamental
T12	Existence of data sovereignty mechanisms, understood as the ability of the owners of the rights of access and use of data to define policies for their exercise.	Sovereignty is one of the pillars of data spaces with European values.	Fundamental
T13	Definition of data sovereignty so that it is susceptible to automated verification (*enforcement*).	Formalizes the implementation of data sovereignty. Formation of *data share agreements* in the form of *smart contracts*.	Optional
T14	Use of ontologies, semantics and recognized vocabularies and standard formats.	Need to promote the use of ontologies and semantics and standard formats to achieve interoperability. If the creation of a model is considered necessary, it will be published on a specialized portal.	Recommendable
T15	Existence of mechanisms to ensure traceability and lineage of data.	It is necessary that the origin of the data be transparently described.	Fundamental
T16	Possibility of specifying mechanisms to verify and ensure data quality.	The quality of the data exchanged helps generate value.	Recommendable
T17	Existence of resource catalog publishing tools.	Essential functionality for publishing and discovering resources available in the data space. Both data sets and applications and, if applicable, computational resources.	Fundamental
T18	Existence of secure data exchange mechanisms between peers (*peer to peer*) via APIs.	A centralized model in which there is a central data repository is not desirable, but rather the promotion of ecosystems.	Recommendable

T19	Provision of *marketplace-type tools* for the offer and consumption of available resources.	The available intermediation tools and services must have as one of their purposes the creation of a resource market.	Recommendable
T20	Offering intermediation and value-added services, including the deployment of solutions based on innovative technologies (AI).	A data space must offer, at a minimum, the possibility of sharing not only data, but processing applications that add value to the data.	Recommendable
T21	Use of PET technologies (Privacy Enhancing Technologies)	PET technologies allow the exploitation of data while guaranteeing privacy, making it possible to use data sets that could not otherwise be shared.	Recommendable
T22	Offering computing resources as shared resources.	The supply of computational resources increases the value of data space.	Optional

Table 10. Table of technical requirements in data spaces.

Annex I. Resources to meet the criteria of governmental grants for data spaces (Ministerial Order TDF/1461/2023)

CRITERION	SUPPORT RESOURCES[7]	CORRESPONDENCE CHARACTERISTICS OF DATA SPACES
A. Technical solution		
1. Features and functionalities:		
– Quality, sustainability, simplicity and coherence of Architecture.	Chap. 6.6, Chap. 8, [55], [40], [56], Ch. 1 of [36], [57], [58],[59]	T18
– Reuse of components from European initiatives.	Chap. 0, [42], [44], [41],[43]	-
– Infrastructure deployment suitability. Use of European public infrastructures.	Chap. 9.2, [60], [61], [62],[63]	-
– Portability, independence of technological infrastructure.	Chap. 11.1,[64]	T02
– Legal, organizational, semantic and technical interoperability in the data space.	Chap. 6.1.2, [65],[66]	T05
– Use of innovative technologies (Artificial Intelligence, Privacy Enhancement Technologies –PET–, Distributed Registry Technologies –DLT–, Distributed Learning, synthetic data…).	Chap. 3.4, [67], [68], [69], [70], [71], [72], [73],[74]	T21
2. Alignment with international standards:		
– Alignment with reference architectures of Data Spaces.	Chap. 9, [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], Part II + Part IV of[85]	T01
– Use or generation of widely used semantics.	Chap. 7.3.2, [86],[87]	S14 , N03
3. Information security:		
– Guarantee of security and privacy of information.	Chap. 11.2, Chap. 15.2, [54], [88], [89], [90], [91],[92]	T04
– Mechanisms to guarantee data sovereignty.	Chap. 5.2.4.2, [93], Chap. 8 of [85],[94]	T12
– Identity, authentication and access systems.	Chap. 5.2.4.2, [95], [96], [97],[98]	T09
B. Data sharing ecosystem		
4. Alignment with national strategy:		
– Alignment with the domain and industrial priorities defined in the calls.	Chap. 11.3, [99],[100]	N13
– Alignment with the 'Activity plan for the deployment of data spaces' of the National Data Office.	Chap. 16, [101],[102]	T17
5. Data Space business plan:		
– Existence and viability of the business model.	Chap. 1.2, Chap. 5.2.1, [103], Chap. 4.4 of	N02 , N04

[7] Referenced chapters without a bibliographic note follow correspond to sections of this same document. For example "Chap 6.6" corresponds to ' Architectural and software components'.

	[36], [104], [105], [106], [107]	
– Potential for domain innovation.	Chap. 13 , Part III of [85], [108], [109], [110]	T08
6. Participants: quality and completeness:		
– Participation of private companies.	[111], [112], [24], [113]	N14
– Domain representativeness of participants.	Chap. 13.2	N14
– Participant escalation plan.	Chap. 14.4, Chap. 14.5, [114]	T17
7. Data Space governance system:		
– Depth and quality of the governance system.	Chap. 6.4, Chap. 6.5, [115], Chap. 3 + Ch. 4 of [116]	T08
– Relationship model, decision-making process, operation model, service level agreements.	Chap. 5.2.3, [117]	N01
– Guarantee of regulatory compliance.	Chap. 5.2.2, [118], [119], [120], Chap. 3 of [116]	N01
– Ethical considerations.	[121]	N15
8. Dissemination of results and recruitment of new participants:		
– Communication plan and recruitment of new participants.	Chap. 14.4, Chap. 14.5, [122]	S17 , N06
– Plan for disseminating the acquired knowledge. Making code and results available in open repositories. Community formation.	Chap. 14.3, [123], [124]	S03 , N10
C. Definition of Subsidized Project		
9. Data sharing model:		
– Definition of the data to be exchanged.	Chap. 14.1, [125], [126]	T15
– Definition of providers and consumers of each data, as well as intermediaries.	Chap. 6.3, [48], [127]	T11
– Data quality guarantee.	Chap. 15.1, [128], [129], [130]	T16
10. Use cases:		
– Relevance of use cases.	Chap. 13, Ch. 3 of [36], [131], [132], [133], [134]	N12
– Detailed definition of the use case.	Chap. 14.1, Chap. 14.2, [46], [47]	N03 , N08
– Contribution of value to the data economy.	Chap. 2, [135], [136], [137]	T20 , N02 , N04
– Expansion plan to other use cases.	Chap. 14.4	T19 , N06 , N11
11. Interconnection with other data spaces:		
– Interconnection between data spaces.	Chap. 6.6.1, [138], [139], [140], [141]	S06 , N11
12. Work Plan:		
– Quality work plan.	-	-

– Work plan feasibility.	Chap. 14.1, [129], [49], [142], [50], [51]	N05 , N07 , N09
13. Budget:		
– Budget appropriate to objectives.	-	N04 , N08

Annex II. Open initiatives with experience in data spaces

The conception, governance and management of data spaces requires a solid understanding of the concepts, best practices, challenges and opportunities that these spaces represent. In this regard, there are several sources of information and support resources that can provide additional guidance.

- The data.gob.es portal, beyond its main source of information related to the world of open data, also provides news, articles and documents about current data spaces in Spain and Europe. The *data spaces* label makes it easy to find these sources.
- The Data Spaces Support Center (DSSC) is the European initiative to support the development of common European data spaces in different sectors and economic domains. The DSSC establishes a network of stakeholders that includes all relevant organizations and initiatives involved in the development of data spaces, offering advice, resources, training and support.

  One of the first resources developed by the DSSC has been a data space starter kit that provides a basic initial guide to understand what the basic elements of a data space are and how to address the different challenges that arise when it comes to data space. to build them. This document includes an annex with a complete glossary of references.
- The Spanish Gaia-X Hub is a national meeting point for the community promoting the Data Economy. The hub develops activities leading to the dissemination, promotion and training in the field of the Data Economy, generating complete functional and technical studies, sectorial analyses and various proofs of concept.

  Within the association, five thematic working groups have been promoted aimed at different sectors (industry, mobility, agri-food, health and tourism) and a group of innovative technologies for transversal technical support. These groups are working on defining the characteristic challenges of each sector, articulating them around use cases, seeking to demonstrate the value of shared data within a data space.

Annex III. Bibliographic references

[1] European Commission, «Com(2020) 66 "A European Data Strategy",» Feb 2020. [Online]. Available: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020DC0066&from=ES.

[2] European Parliament, "The ubiquity of the digital single market (Technical Sheets 2023)," [Online]. Available: https://www.europarl.europa.eu/factsheets/es/sheet/43/la-ubicuidad-del-mercado-unico-digital . [Last access: 2023].

[3] European Commission, "European Data Market study 2021-2023," May 2023. [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/results-new-european-data-market-study-2021-2023.

[4] Official Journal of the European Union, "European Declaration on Digital Rights and Principles for the Digital Decade," 2023. [Online]. Available: https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32023C0123(01).

[5] Government of Spain, "Bill of Digital Rights," 2021. [Online]. Available: https://www.lamoncloa.gob.es/presidente/activityes/Documents/2021/140721-Carta_Derechos_Digitales_RedEs.pdf.

[6] Official Journal of the European Union, "Regulation (EU) 2016/679 relating to the protection of natural persons with regard to the processing of personal data and the free circulation of such data," Apr 2016. [Online]. Available: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32016R0679.

[7] Official Journal of the European Union, "Regulation (EU) 2022/868 on European data governance," Jun 2022. [Online]. Available: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32022R0868.

[8] European Parliament and Council, "Regulation (EU) 20233/2854 of the European Parliament and of the Council of 13 December 2023 on harmonized rules for fair access to and use of data (Data Regulation)," 13 Dec 2023 . [Online]. Available: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202302854.

[9] Official Journal of the European Union, "Directive (EU) 2019/1024 on open data and the reuse of public sector information," Jun 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32019L1024&from=ES.

[10] Official Journal of the European Union, "Regulation (EU) 2022/1925 on contestable and equitable markets in the digital sector," 14 Sep 2022. [Online]. Available: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32022R1925.

[11] Official Journal of the European Union, "Regulation (EU) 2022/2065 relating to a single market for digital services," 19 Oct 2022. [Online]. Available: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32022R2065.

[12] Official Journal of the European Union, "Regulation (EU) 2018/1807 on a framework for the free circulation of non-personal data in the EU," Nov 2018. [Online]. Available: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32018R1807&from=ES.

[13] EU Council, "Artificial Intelligence Regulation: Council and Parliament reach agreement on world's first rules on artificial intelligence," Dec 2023. [Online]. Available: https://www.consilium.europa.eu/es/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the- first-worldwide-rules-for-ai/.

[14] European Commission, "Proposed Regulation on 'managed security services' amendment.," Apr 2023. [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/proposed-regulation-managed-security-services-amendment.

[15] European Commission, “The European Interoperability Framework (EIF) in detail,” [Online]. Available: https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-

observatory/european-interoperability-framework-detail.

[16] «CSA - Dates European Data Space for Tourism,» [Online]. Available: https://www.tourismdataspace-csa.eu/.

[17] «CSA - DSFT Data Space For Tourism,» [Online]. Available: https://dsft.modul.ac.at/.

[18] «CSA - Data Space for Skills (DS4Skills),» [Online]. Available: https://www.skillsdataspace.eu/.

[19] «CSA - PrepDSpace4Mobility,» [Online]. Available: https://mobilitydataspace-csa.eu/.

[20] «CSA - AGRIDATASPACE,» [Online]. Available: https://agridataspace-csa.eu/.

[21] «CSA - Green Deal Data Space GREAT,» [Online]. Available: https://www.greatproject.eu/.

[22] «CSA - Data Space 4.0 Manufacturing,» [Online]. Available: https://manufacturingdataspace-csa.eu/.

[23] Ministries of Economic Affairs and Digital Transformation. Government of Spain., «Spain Digital 2026,» 2022. [Online]. Available: https://espanadigital.gob.es/.

[24] data.gob, "The European Commission launches a support center for data exchange," Sep 2019. [Online]. Available: https://datos.gob.es/es/noticia/la-comision-europea-pone-en-marcha-un-centro-de-soporte-para-el-intercambio-de-datos.

[25] EU Support Center for Data Sharing, “Data Sharing Practice Examples: A digest of practice examples published on the SCDS,” [Online]. Available: https://eudatasharing.eu/data-sharing-practice-examples. [Last access: Mar 2022].

[26] European Commission, Joint Research Centre, «Farrell, E., Minghini, M., Kotsev, A., et al., European data spaces : Scientific insights into data sharing and utilization at scale, Publications Office of the European Union Luxembourg, 2023 , ISBN 978-92-68-03,” 2023. [Online]. Available: https://data.europa.eu/doi/10.2760/400188.

[27] Data Sharing Coalition, “Data Sharing Canvas: A stepping stone towards cross-domain data sharing at scale,” Apr 2021. [Online]. Available: https://datasharingcoalition.eu/app/uploads/2021/04/data-sharing-canvas-30-04-2021.pdf.

[28] International Data Spaces Association (IDSA), "International Data Spaces Association.," [Online]. Available: https://internationaldataspaces.org/publications/most-important-documents/.

[29] Big Data Value Association (BDVA), “Towards a European-governed data sharing space,” Nov 2020. [Online]. Available: https://bdva.eu/download/80/data-spaces/3238/towards-a-european-governed-data-sharing-space-bdva-position-paper-2.pdf.

[30] i4Trust. Fiware, iShare & FICodes, «i4Trust. Data Spaces for effective and trusted data sharing.,” Sep 2021. [Online]. Available: https://cdn.on.expo-x.com/storage/37519/i4trust-buildingblocks-202109-1.pdf.

[31] Gaia-X, “Gaia-X: Vision & Strategy,” Dec 2021. [Online]. Available: https://gaia-x.eu/wp-content/uploads/2021/12/Vision-Strategy.pdf.

[32] European Commission, “Staff working document on data spaces,” Feb 2022. [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces.

[33] European Commission, “Second staff working document on data spaces,” Jan 2024. [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/second-staff-working-document-data-spaces.

[34] Directorate-General for Research and Innovation, “Turning FAIR into reality : final report and action plan from the European Commission expert group on FAIR data,” 2018. [Online]. Available: https://data.europa.eu/doi/10.2777/1524.

[35] iSHARE. [Online]. Available: https://ishare.eu/home/about-ishare/the-foundation/history/.

[36] L. Nagel and D. Lycklama, “OpenDEI: Design Principles for Data Spaces - Position Paper (v1.0),” Jul 2021. [Online]. Available: https://www.opendei.eu/wp-content/uploads/2022/03/Position-Paper-Design-Principles-for-Data-Spaces.pdf.

[37] Geonovum, “Exploring data space initiatives (v0.52),” 2023. [Online]. Available: https://www.geonovum.nl/uploads/documents/Exploring%20data%20space%20initiatives%20v0.52_EN%20publication%20version%20Geonovum.pdf.

[38] Data Spaces Support Centre, “Data Spaces Blueprint (v0.5),” Sep 2023. [Online]. Available: https://dssc.eu/space/BPE/179175433/Data+Spaces+Blueprint+%7C+Version+0.5+%7C+September+2023.

[39] Data Spaces Support Centre, “Collection of Standards and Technologies Landscape (v1.0),” Oct 2023. [Online]. Available: https://dssc.eu/space/SE1/185794561/Collection+of+Standards+and+Technologies+landscape+%7C+Version+1.0+%7C+October+2023.

[40] Data Spaces Support Center, “DSSC Glossary,” Mar 2023. [Online]. Available: https://dssc.eu/space/Glossary/55443460/DSSC+Glossary+%7C+Version+1.0+%7C+March+2023.

[41] The FIWARE Foundation, “GitHub FIWARE,” [Online]. Available: https://github.com/Fiware. [Last access: Feb 2024].

[42] The iSHARE Foundation, “GitHub iSHARE,” [Online]. Available: https://github.com/iSHAREScheme. [Last access: Feb 2024].

[43] IDSA (International Data Spaces Association), “GitHub IDSA,” [Online]. Available: https://github.com/International-Data-Spaces-Association. [Last access: Feb 2024].

[44] Gaia-X AISBL, “GitLab Gaia-X AISBL,” [Online]. Available: https://gitlab.com/gaia-x. [Last access: Feb 2024].

[45] data.gob, «UNE Specifications – Government, management and data quality.», Mar 2023. [Online]. Available: https://datos.gob.es/es/blog/specifications-une-gobierno-gestion-y-calidad-del-dato.

[46] Data Office, “Use case design guide,” 2023. [Online]. Available: https://datos.gob.es/sites/default/files/blog/file/guia_de_diseno_de_casos_de_uso_0.pdf.

[47] Data Office, «Use case viability evaluation guide.», 2023. [Online]. Available: https://datos.gob.es/sites/default/files/blog/file/esap_technical_recommendations_for_member_states_harvesting_guidelines_1_0.pdf.

[48] iSHARE, «CookBook for Data Spaces (Quick Start Guide for Data Spaces),» Aug 2023. [Online]. Available: https://ishare.eu/wp-content/uploads/2023/09/iShare-Data-Spaces-Kookboek_UK-v1.pdf.

[49] UNE, "UNE 0080 Government Evaluation Guide, Management and Data Quality Management," Jun 2023. [Online]. Available: https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0071383.

[50] UNE, "UNE 0077 Data Government," Mar 2023. [Online]. Available: https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0071116.

[51] UNE, "UNE 0078 Data Management," Mar 2023. [Online]. Available: https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0071117.

[52] UNE, "UNE 0079 Data Quality Management," Mar 2023. [Online]. Available: https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0071118.

[53] UNE, "UNE 0081 Guide to evaluating the Quality of a Data Set," Sep 2023. [Online]. Available: https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma/?c=N0071807.

[54] Spanish Data Protection Agency., «Approach to data spaces from the perspective of the RGPD.», May 2023. [Online]. Available: https://www.aepd.es/es/documento/aproximacion-espacios-datos-rgpd.pdf.

[55] Data Office, "Technological reference framework for the construction of federated data spaces," Jun 2023. [Online]. Available: data.gob/.....

[56] DSSC, “Conceptual Model of Data Spaces (v0.5),” Sep 2023. [Online]. Available: https://dssc.eu/space/CME/176554182/Conceptual+Model+of+Data+Spaces+%7C+Version+0.5+%7C+September+2023.

[57] Otto, Boris, et al., “Data Ecosystems. Conceptual Foundations, Constituents and Recommendations for Action,” Oct 2019. [Online]. Available: https://www.researchgate.net/publication/339528681_Data_Ecosystems_Conceptual_Foundations_Constituents_and_Recommendations_for_Action.

[58] OpenDEI, “Reference Architectures and Interoperability in Digital Platforms,” Sep 2022. [Online]. Available: https://www.opendei.eu/wp-content/uploads/2022/10/REFERENCE-ARCHITECTURES-AND-INTEROPERABILITY-IN-DIGITAL-PLATFORMS.pdf.

[59] Data Office, "Data Mesh: Introduction, motivation, components and tools," Feb 2024. [Online]. Available: data.gob.es/.....

[60] data.gob.es, “Cloud federations and data spaces: Learn about the Simpl project,” Mar 2023. [Online]. Available: https://datos.gob.es/es/blog/federaciones-en-la-nube-y-espacios-de-datos-conoce-el-proyecto-simpl.

[61] European Commission, “SIMPL: Streamlining cloud-to-edge federations for major EU data spaces,” Dec 2023. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/simpl.

[62] CEF Digital, “European Blockchain Services Infrastructure,” Dec 2020. [Online]. Available: https://ec.europa.eu/digital-building-blocks/sites/display/DIGITAL/European+Blockchain+Services+Infrastructure+%28EBSI%29+dashboard.

[63] European Commission, “Connecting Europe Facility program (CEF Digital) - Building blocks,” [Online]. Available: https://ec.europa.eu/digital-building-blocks/sites/display/DIGITAL/About+us. [Last access: Feb 2024].

[64] European Commission, “Data Act, Q&A: Portability,” Jun 2023. [Online]. Available: https://ec.europa.eu/commission/presscorner/detail/en/qanda_22_1114#:~:text=move%20data%20and%20applications.

[65] European Commission, “The New European Interoperability Framework,” [Online]. Available: https://ec.europa.eu/isa2/eif_en/.

[66] IDSA, “Governance for Data Space Instances Aspects and Roles for IDS Stakeholders (v0.1),” Aug 2021. [Online]. Available: https://internationaldataspaces.org/wp-content/uploads/IDSA-Position-Paper-Governance-for-Data-Space-Instances-Aspects-and-Roles-for-IDS-Stakeholders.pdf.

[67] OECD, “Emerging privacy-enhancing technologies: Current regulatory and policy approaches,” Mar 2023. [Online]. Available: https://doi.org/10.1787/bf121be4-en.

[68] Data Office, «Introduction to PET technologies,» Dec 2023. [Online]. Available: data.gob.es/.....

[69] W3C, «W3C Verifiable Credentials (VCs). Recommendation 1.1,” Mar 2022. [Online]. Available: https://www.w3.org/TR/vc-data-model/.

[70] Ocean Protocol Foundation, “Ocean protocol: User guides,” [Online]. Available: https://docs.oceanprotocol.com/user-guides. [Last access: 2023].

[71] Data Office, "Data Lakehouse: Introduction, concepts, components and examples," Feb 2024. [Online]. Available: data.gob.....

[72] Data Office, "Data Fabric: Introduction, motivation, components and tools," Feb 2024. [Online]. Available: data.gob....

[73] Data Office, "OBDM and data spaces," Dec 2023. [Online]. Available: data.gob.....

[74] Data Office, “Confidential Computing,” Mar 2023. [Online]. Available: data.gob.es/....

[75] DSSC, “Summary of assets,” 2024. [Online]. Available: https://dssc.eu/space/DDP/117211137/DSSC+Delivery+Plan+-+Summary+of+assets+publication.

[76] Ministry of Transport and Communications - Finland, “State of data spaces 2021,” Oct 2021. [Online]. Available: https://www.opendei.eu/wp-content/uploads/2022/03/LVM-state-of-data-spaces-2021-FINAL.pdf.

[77] The iSHARE Foundation, “iSHARE,” [Online]. Available: https://ishare.eu/home/about-ishare/.

[78] Data Office, «i4TRUST Description,» Jun 2022. [Online]. Available: data.gob.es/.....

[79] SITRA, “IHAN Blueprint 2.5,” Jan 2020. [Online]. Available: https://www.sitra.fi/app/uploads/2018/12/ihan-blueprint-2-5.pdf.

[80] Data Office, “IDS architectural reference model summary,” Feb 2022. [Online]. Available: data.gob.es/....

[81] IDSA, “IDS-RAM 4.0 Architectural Model,” 2023. [Online]. Available: https://docs.internationaldataspaces.org/ids-knowledgebase/v/ids-ram-4/.

[82] Data Office, "Gaia-X for Dummies," Dec 2022. [Online]. Available: data.gob.es/.....

[83] Gaia-X AISBL, “Gaia-X Architecture Document (v23.10),” Oct 2023. [Online]. Available: https://docs.gaia-x.eu/technical-committee/architecture-document/23.10/.

[84] DSBA, “Data Spaces Business Alliance Technical Convergence V2,” Apr 2023. [Online]. Available: https://data-spaces-business-alliance.eu/wp-content/uploads/dlm_uploads/Data-Spaces-Business-Alliance-Technical-Convergence-V2.pdf.

[85] B. Otto, M. ten Hompel, and S. Wrobel (Eds.), “Designing Data Spaces,” 2022. [Online]. Available: https://library.oapen.org/bitstream/handle/20.500.12657/57901/978-3-030-93975-5.pdf.

[86] SEGITTUR, «Tourism Ontology Specification,» Sep 2023. [Online]. Available: https://www.segittur.es/wp-content/uploads/2023/09/19-de-septiembre-de-2023-Especificacion-Ontologia-de-Turismo.pdf.

[87] UNE CTN 178/SC 5, «UNE 178504:2022 Digital, intelligent and connected hotel (HDIC) to smart tourist destination/smart city platforms. Requirements and recommendations,” Jun 2022. [Online]. Available: https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma/?c=norma-une-178504-2022-n0069880.

[88] AEPD, "Guides, recommendations, etc. on data protection in Data Spaces, Cloud Computing and Big Data," [Online]. Available: https://www.aepd.es/areas-de-actuacion/innovacion-y-tecnologia#Nube.

[89] Ametic, “Shared Data Spaces: Security Challenges,” Jan 2024. [Online]. Available: https://ametic.es/wp-content/uploads/2024/01/Documento-Espacios-de-Datos-Compartidos.pdf.

[90] CCN, "CCN-STIC-823 Use of cloud services," [Online]. Available: https://www.ccn-cert.cni.es/es/series-ccn-stic/800-guia-esquema-nacional-de-seguro/541-ccn-stic-823-seguro- en-entornos- cloud/file?format=html.

[91] CCN, «Data traceability in the context of the National Security Scheme (ENS),» [Online]. Available: https://www.ccn-cert.cni.es/es/series-ccn-stic/800-guia-esquema-nacional-de-seguro/541-ccn-stic-823-seguro-en-entornos- cloud/file?format=html.

[92] EU Support Center for Data Sharing, “B.7 - Technical guidance report on data security,” 2018. [Online]. Available: https://eudatasharing.eu/secure-data-sharing. [Last access: Mar 2022].

[93] Data Office, "Access and use control: XACML, Use control in IDS, OPA and other technologies," Feb 2024. [Online]. Available: data.gob.es/.....

[94] BAIDATA, "SQS, first accredited laboratory for the validation of IDS components," Feb 2023. [Online]. Available: https://baidata.eu/sqs-primer-laboratorio-acreditado-para-la-validacion-de-componentes-ids.

[95] Data Office, "Trust mechanisms in data spaces: Alternatives and comparison between solutions," Feb 2024. [Online]. Available: data.gob.es/....

[96] Data Office, "Decentralized Identity: What it is, characteristics and current situation," Feb 2024. [Online]. Available: data.gob.es/....

[97] World Wide Web Consortium, “W3C Workshop on Strong Authentication & Identity,” [Online]. Available: https://www.w3.org/Security/strong-authentication-and-identity-workshop/report.html.

[98] The European Union Agency for Cybersecurity, ENISA, “Digital Identity: Leveraging the Self-Sovereign Identity (SSI) Concept to Build Trust,” Jan 2022. [Online]. Available: https://www.enisa.europa.eu/publications/digital-identity-leveraging-the-ssi-concept-to-build-trust/@@download/fullReport.

[99] Government of Spain, "Component 12 (PRTR): Industrial Policy Spain 2030," [Online]. Available: https://planderecuperacion.gob.es/politicas-y-componentes/componente-12-politica-industrial-espana-2030.

[100] Government of Spain, «Agenda España Digital - Axis 8,» [Online]. Available: https://espanadigital.gob.es/medidas-y-avances.

[101] Gaia-X AISBL, «Self-Descriptions in Gaia-X (Hackathon #3),» [Online]. Available: https://gitlab.com/gaia-x/gaia-x-community/gx-hackathon/gx-hackathon-3/-/wikis/GX-Hackathon-3#presentations-used-during-the-hackathon.

[102] Data Office, «Semantic Web. Vocabularies. DCAT and RDF,” Mar 2023. [Online]. Available: data.gob.es/.....

[103] European Union, “New Business Models for Data-Driven Services,” Nov 2023. [Online]. Available: https://data.europa.eu/en/doc/new-business-models-data-driven-services.

[104] Hub Gaia-X Germany, “Gaia-X and Business Models: Types and Examples,” [Online]. Available: https://gaia-x-hub.de/wp-content/uploads/2023/02/Whitepaper-Gaia-X-Business-Models.pdf.

[105] Hub Gaia-X Germany, “Gaia-X for SMEs,” Nov 2022. [Online]. Available: https://gaia-x-hub.de/en/gaia-x-for-smes/.

[106] Gaia-X AISBL DSBC, “Finance Handbook (v1.0),” 2023.

[107] Gaia-X Institute, “The Economics of Data-Sharing,” 2024.

[108] Data Office, «Radiography of the Tourism Data Space in Spain,» Dec 2022. [Online]. Available: https://portal.mineco.gob.es/es-

es/comunicacion/Paginas/201222_Radiografia_Espacio_Datos_Turismo_Espana.aspx.

[109] Ricardo Valle Institute of Innovation Foundation, “Data Ecosystems,” Jun 2023. [Online]. Available: https://innovairv.com/wp-content/uploads/2024/01/Ecosistemas-del-Dato_Web-1-2.pdf.

[110] data.gob.es, “Data Spaces Business Alliance (DSBA): How to take advantage of synergies to boost data spaces,” Sep 2022. [Online]. Available: https://datos.gob.es/es/noticia/data-spaces-business-alliance-dsba-como-aprovechar-sinergias-para-impulsar-los-espacios-de.

[111] European Commission, “Guidance on private sector data sharing,” [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/private-sector-data-sharing.

[112] Capgemini, “Data Sharing Masters,” Jul 2021. [Online]. Available: https://www.capgemini.com/wp-content/uploads/2021/07/Final-Web-Version-of-Report-Data-Ecosystems.pdf.

[113] European Round Table for Industry, “B2B Data Sharing,” Jun 2021. [Online]. Available: https://ert.eu/wp-content/uploads/2021/06/ERT-Expert-Paper-B2B-Data-Sharing-FINAL.pdf.

[114] data.gob, “Fostering a flexible and capillary data economy: the importance of effective public-private collaboration,” Nov 2023. [Online]. Available: https://datos.gob.es/es/blog/fomentando-una-economia-del-dato-flexible-y-capilar-la-importancia-de-una-colaboracion-publico.

[115] IDSA, “IDSA Rulebook,” 2023. [Online]. Available: https://docs.internationaldataspaces.org/ids-knowledgebase/v/idsa-rulebook/front-matter/readme.

[116] M. Lindner, S. Straub and B. Kühn, «How to Share Data? Data-Sharing-Plattformen für Unternehmen,” Apr 2021. [Online]. Available: https://www.iit-berlin.de/wp-content/uploads/2021/04/SDW_Studie_DataSharing_ES-1.pdf.

[117] The Center of Excellence for Data Sharing and Cloud (CoE-DSC), “Market Research: How do service providers respond to planned EU data sharing initiatives and emerging Data Spaces?,” Mar 2023. [Online]. Available: https://coe-dsc.nl/market-research-how-do-service-providers-respond-to-planned-eu-data-sharing-initiatives-and-emerging-data-spaces/.

[118] data.gob, "European data regulation facing the challenge of a harmonized application that promotes data sharing," Dec 2023. [Online]. Available: https://datos.gob.es/es/blog/la-regulation-europea-en-materia-de-datos-frente-al-reto-de-una-aplicacion-armonizada-que.

[119] EU Support Center for Data Sharing, “B.1 Report on collected model contract terms,” Jul 2019. [Online]. Available: https://eudatasharing.eu/legal-aspects/report-collected-model-contract-terms. [Last access: Mar 2022].

[120] SITRA, «Rulebook for a fair data economy - Part 2 (v2.0),» Jun 2022. [Online]. Available: https://www.sitra.fi/app/uploads/2022/11/rulebook-for-a-fair-data-economy-part-2_2022-11-30.pdf.

[121] Data Office, "Ethical principles of data processing," Jan 2024. [Online]. Available: data.gob.es/....

[122] DATES European Data Space for Tourism, “Deliverable D5.1 Project Communication and Engagement Plan,” [Online]. Available: https://www.tourismdataspace-csa.eu/wp-

content/uploads/2023/06/D5.1-Communication-and-engagement-action-plan-v7.pdf.

[123] IDSA/ DSSC, “Data Spaces Radar,” [Online]. Available: https://www.dataspaces-radar.org/radar/.

[124] BAIDATA, "Developing expertise: BAIDATA and IDSA boost talent in data spaces with their professional certification program," Feb 2024. [Online]. Available: https://baidata.eu/desarrollando-expertise-baidata-e-idsa-impulsan-el-talento-en-espacios-de-datos-con-su-programa-de-certificacion-profesional.

[125] data.gob, “Datasheets for datasets,” Oct 2022. [Online]. Available: https://datos.gob.es/es/blog/documentacion-de-datos-datasheets-datasets.

[126] Data Office, “Data documentation: Datasheets for datasets,” Jul 2023. [Online]. Available: data.gob.es/....

[127] IDSA, “New Business Models for Data Spaces Grounded in Data Sovereignty,” [Online]. Available: https://internationaldataspaces.org/wp-content/uploads/IDSA-Position-Paper-New-Business-Models-sneak-preview-version.pdf.

[128] data.gob.es, "Application of the UNE 0079:2023 Quality Management Specification to open data," Sep 2023. [Online]. Available: https://datos.gob.es/es/blog/aplicacion-de-la-especificacion-une-00792023-de-gestion-de-calidad-los-datos-abiertos.

[129] data.gob.es, «UNE 0080 Specification: UNE Specifications – Government Evaluation Guide, Management and Data Quality Management,» Jun 2023. [Online]. Available: https://datos.gob.es/es/blog/specifications-une-guia-de-evaluacion-del-gobierno-gestion-y-calidad-del-dato-0.

[130] data.gob.es, “UNE 0081 Specification – Data quality evaluation guide,” Sep 2023. [Online]. Available: https://datos.gob.es/es/blog/especificacion-une-0081-guia-de-evaluacion-de-la-calidad-de-datos.

[131] IDSA, “Data spaces overview,” [Online]. Available: https://internationaldataspaces.org/wp-content/uploads/IDSA-data-spaces-overview-2022.pdf.

[132] Data Office, “Compilation of use cases for data sharing and exploitation in the tourism sector,” Jan 2024. [Online]. Available: data.gob.es/....

[133] data.gob, "European support and coordination actions for the deployment of Data Spaces," Feb 2024. [Online]. Available: https://datos.gob.es/es/blog/acciones-europeas-de-apoyo-y-coordinacion-para-el-despliegue-de-espacios-de-datos.

[134] D. Ordoñez, J. Seguí-Pons and M. Ruíz, “Defining the Balearic Islands' Tourism Data Space: An Approach to Functional and Data Requirements,” Feb 2024. [Online]. Available: https://doi.org/10.3390/data9030041.

[135] European Commission, “Data Stories,” [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/results-new-european-data-market-study-2021-2023#:~:text=Data%20stories. [Last access: Feb 2024].

[136] European Commission, "D4.2 - EU Data Landscape," Dec 2023. [Online]. Available:

https://ec.europa.eu/newsroom/dae/redirection/document/100864.

[137] Data Office, “Data Economy,” Jan 2024. [Online]. Available: data.gob.es/....

[138] IDSA, “Gaia-X and IDS,” Jan 2021. [Online]. Available: https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-Position-Paper-GAIA-X-and-IDS.pdf.

[139] DOME Project, “A Distributed Open Marketplace for Europe Cloud and Edge Services,” [Online]. Available: https://dome-marketplace.eu/.

[140] Gaia-X AISBL, “Gaia-X and Catalogues,” Jan 2024. [Online]. Available: https://gaia-x.eu/news-press/gaia-x-and-catalogues/.

[141] data.gob, "The importance of the Telco sector in the deployment of a Digital Europe," Jan 2024. [Online]. Available: https://datos.gob.es/es/blog/la-importancia-del-sector-telco-en-el-despliegue-de-una-europa-digital.

[142] UNE Magazine, “UNE specifications for government, management and data quality,” [Online]. Available: https://revista.une.org/57/specifications-une-para-gobierno-gestion-y-calidad-del-dat.html.