DATA-POWERED BUSINESS ECOSYSTEM BUILDING

POSITION PAPER | SEPTEMBER 2022

Data-economy is increasingly raising due to high-value applications such as data analytics, big data or artificial intelligence. These applications require not only the soft infrastructure indispensable for data sharing and exchange (e.g., data spaces based on common design principles for a dynamic, secure and seamless flow of data between parties and domains), but also a business ecosystem in the form of digital value chains based on trust, data sovereignty and data quality, which includes EU-born digital assets and parties, to enlarge existing or build new networks of collaboration.

Under the coordination and leadership of Innovalia, in the framework of the Horizon 2020 CSA "OPEN DEI- Aligning Reference Architectures, Open Platforms, and Large-Scale Pilots in Digitalizing European Industry", Task Force 2 with a solid collaboration of more than 30 experts coming from key sectors seeks to provide common guidance for stakeholders to develop and participate in data-driven digital business ecosystems and value chains, enhancing the use of emerging digital innovation tools that are necessary for digital transformation.

Contributing organizations

innovalia ASSOCIATION	IDC	AtoS	EHTEL Collaborating for Digital Health and Care in Europe
ENGINEERING	INTERNATIONAL DATA SPACES ASSOCIATION	E.ON Energy Research Center | RWTH AACHEN UNIVERSITY	eit Digital
TECNOALIMENTI	Fondazione Politecnico di Milano		

Contributing Projects

OPENDEI			

Publisher

IDC

Editors

Oscar Lázaro, Asociación Innovalia

Carmen Polcaro, Asociación Innovalia

Authors and Contributors

Audren Layeux (CARSA)

Christian Krug (VDI)

Michela Magas (Industry Commons Foundations)

Agnieszka Radziwon (Aarhus University)

Alberto Dognini, RWTH

Mariana Faraldi, TCS

Sergio Gusmeroli, FPM

Marc Lang, EHTEL

Jesus Martinez, ATOS

Oscar Lazaro, INNOVALIA

Angelo Marguglio, ENGINEERING

Giorgio Micheletti, IDC

Luc Nicolas, EHTEL

Silvia Piai, IDC

Carmen Polcaro, INNOVALIA

Golboo Pourabdollahian, IDC

Riccardo Zanetti, ENGINEERING

Table of Contents

List of Figures

No se encuentran elementos de tabla de ilustraciones.

List of Tables

1. Introduction

2. IP and Industry Agreements

The purpose of this chapter is to contribute to the Open DEI framework with strategies for the management of IP and Industry Agreements across industrial domains, particularly from the point of view of stimulating open innovation in a data-driven economy. The paper examines emerging use cases that have been identified to require novel IP strategies or that carry potential for innovative industrial agreement solutions.

2.1. Framework conditions

2.1.1. Open DEI towards industry commons

Open DEI aims to unite several verticals including health, manufacturing, agrifood, mobility, as well as umbrella concepts such as Smart Cities (including construction and energy domains) towards industry commons. The aim of creating an industry commons is to provide several enabling layers supporting effective interaction, optimisation, sustainability and innovation for cross-domain data-driven market activities.

To stimulate effective exchange between domains, industry commons rely on a comprehensive interoperability strategy including several levels of functionality, from data reference models for cross-domain Interoperability to application protocols between cross-domain applications and, at a more granular level, interoperability between detailed functional components of cross-domain applications. From the perspective of enterprise integration, and by extension business ecosystem integration, the management and support of IP and Industry Agreements (IAs) is the key additional enabling layer for the creation of cross-domain open innovation and for the valorisation of data in the industry commons.

2.1.2. IP and Industry Agreements in the context of the EU Data Act

The proposal for EU Data Act adopted by the European Commission on 23 February 2022[1], includes several directives and recommendations that strongly support and advocate for open innovation approaches. The EU's ambition is to encourage as many actors as possible, regardless of their size, to participate in the data economy. It wishes to protect participants, especially SMEs, against unfair contractual terms imposed by parties enjoying significantly stronger market positions. It intends to create framework conditions where EU businesses, especially SMEs, have more possibilities to compete and innovate on the basis of data they generate thanks to data access and portability rights.

Open innovation approaches are also applied to proprietary data in specific scenarios. The Data Act introduces an obligation for private companies to unlock their data in exceptional situations of high public interest, such as floods or wildfires. It also intends to enable users of smart objects access to the data those objects generate so that they can share this data with third parties and spur the development of a broader range of services.

Impact studies which follow the Data Act proposal, ask for measures to ensure citizen empowerment in a human-centric data economy, development of rights strategies for co-generated data, and the establishment of metadata standards (both technical and legal) across or within sectors for data sharing. They ask for the introduction of a legislative framework for

[1] https://ec.europa.eu/commission/presscorner/detail/en/ip_22_1113

fairness controls in data sharing contracts, and a horizontal cross-domain legislative framework with access rights for data re-use.

The development of rights strategies for open innovation is thus seen to contribute to the ultimate goal, that is to unite all of the domain-specific data spaces into a single European data space and create a genuine single market for data.

2.1.3. The legal framework and associated challenges

The term "industry agreements" is not a term of art in law or in policy. A recent study commissioned by the EC focused on the concept of industry agreements and how they operate in multilateral value chains [ref]. Industry agreements are a mechanism for reaching a common understanding of functionalities or architectures and specifications. They may include a variety of tools, such as, for example, federated digital infrastructures, common vocabulary standards, and also relational contractual agreements.

Industry agreements share some common features with codes of conduct and standards but they are distinct regulatory mechanisms. For example, standards influence data and technological specifications, but they do not offer governance guidance. Codes of conduct, on the other hand, typically encompass behavioural norms and some form of a governance mechanism. However, codes of conduct have limited influence on technological specifications and they are not typically considered in discussions of new specifications.

Industry agreements address the need for coordination among economic operators in different jurisdictions and across different value chains. Their open, cross-border nature serves as a natural counterbalance to the limitations imposed by the legal principle of territoriality. The principle of territoriality dictates that national rules govern the subject matter within a particular territory. Intellectual property rights are a case in point. They are exclusive rights that subsist in subject matter and can be exercised solely on territorial basis. With the exception of some limited harmonisation in certain aspects of IP law, questions of both ownership and licensing remain heavily dependent on national interpretation of legal rules. This is a challenge for industries and value chains that extend across multiple jurisdictions.

Typically, there are two stages in the lifecycle of an IP right. The first stage concerns how rights subsist in the subject matter, e.g. works when we talk about copyright or inventions in the case of patents. This happens either by operation of law, e.g. copyright, or through an application and registration regime, e.g. patents. The second stage concerns the licensing of rights relating to the exploitation of the work, invention or other subject matter.

A first problem is, therefore, identifying the rights that subsist in the subject matter. The modalities of a legal transaction depend on the type of subject matter and the applicable law. A second problem is to identify the proprietors and the allocation of rights, specifically considering scenarios of individual and shared rights in cases of data pooling. Finally, when transacting with these rights, it must be clear what transactional tools are most effective and efficient. Many actors, particularly SMEs and start-ups, find the existing practices in data licensing confusing and part of the reason is that many of these practices draw on licensing models initially created for open source software.

Subsistence of rights. It is essential to know what is owned, who owns it and how it is licensed. These questions require a fact-intensive inquiry which may prove particularly challenging if we do not know what rights might subsist in the subject matter.

Data are a problematic subject matter from the perspective of intellectual property law. The law covering data is in some ways extremely unpredictable and confusing. While compilations of data or other material which by reason of the selection or arrangement of their contents constitute intellectual creations is protected as such, this protection does not extend to the data or material itself, and it is without prejudice to any copyright subsisting in the data or material itself [TRIPS Agreement]. Thus, unlike software, for example, data is not categorically protected by copyright. In spite of this, various data licences have been developed in recent years and the majority have been modelled after open source software licences.

Open source software licensing models are supported by copyright law. They operate on the basis that computer programmes as subject matter are protected by copyright. Most of the existing open data licensing models rely on the same assumption which, however, is erroneous. There is no legal certainty as to whether copyright subsists in data and, even if it does, how the different doctrines of copyright law apply to data. For instance, the law is unsettled on the status of machine learning models, e.g. as independent works or as adaptations of the training dataset. This means that data licensing models may rest on shaky ground. Even though they have some traits in common, the categories in data licensing and copyright licensing are conceptually distinct, e.g. the legal distinction between static and dynamic linking in software licensing which does not exist in data licensing. The law should clarify which different forms of rights might subsist in data and the threshold for subsistence of any such rights as they flow along the value chain. Overcoming this challenge would engender trust and make data transactions more predictable and reliable for stakeholders.

Tracking of rights. Another challenge is how to trace and track ownership in multilateral digital value chains. Tracking ownership is essential to improve transparency in the value chain. Transferring rights in data is much more dynamic and is likely to occur more often than, for example, the assignment of a patent application or patent. Furthermore, there are no registers of data transactions at the ecosystem level to facilitate transparency and engender trust.

As data travels across the value chain, it undergoes changes and transformations which may produce legal consequences. It is unclear, for example, whether collective contributions by different economic operators to the same data set give rise to any shared rights and whether the independent contributions merit protection by means of exclusive rights. Furthermore, if any independent contribution merits protection, a downstream user may find it impossible to get assurances that every contributor has granted a licence where the number of contributors may be in the order of thousands. These issues are well understood in the practice of open source software licensing where the doctrines of sole and joint ownership have provided some degree of clarity. While these doctrines are unlikely to apply to data transactions because of the uncertainty around the forms of exclusive rights that may subsist in data, the open source development and licensing models may inspire new solutions for collaborative data ecosystems.

Transacting with rights. Given the scale, dynamics and sheer amount of transactions taking place in the data economy, adequate technological infrastructure should be put in place to support economic operators. One possible solution is the automated licensing and clearance of rights based on distributed ledger technologies and smart contracts. This could enable transparency and improve legal certainty by increasing the predictability of data transactions. An economic operator interested in procuring a data set would not only know what rights might have subsisted in the data set but also with whom they may need to negotiate access and usage rights. Finally, automated negotiation mechanisms may facilitate the process and perhaps even relegate some low-risk transactions to entirely machine-to-machine interaction.

2.2. Opportunities

2.2.1. Open Innovation as driver of new data ecosystems

Open innovation environments present opportunities for a new strategic framework for intellectual property, based on use cases that result from collaborative and cross-domain development. The act of collaboration creates opportunities and new incentives, but also the risk of misunderstandings and disputes. Successful open innovation collaborative environments set out clear rules of engagement for a level playing field that stimulates trust among all participating stakeholders. The IP framework must therefore work together with open innovation codes of conduct and guidelines for best practice. It must particularly consider the impact of IP registration on sustainability strategies and focus on the need to track and trace the complete lifecycle of materials and products across DVCs in order to achieve the ambitions of the Green Deal. It must encourage rather than stifle innovation that will bring sustainable solutions and incentivise behaviours that contribute to the Green Deal's grand vision. Solving the Green Deal challenges cannot rely on a "winner takes all" mentality, but on systems of incentivisation that encourage collaborative co-creation that generates greater value for all stakeholders. This includes systems that, on the one hand, stimulate the larger organisations to contribute their background IP in a way that becomes beneficial to them, and on the other, allow SMEs to build additional value with this background IP to amplify its potential, and add their contribution to the value network.

The framework must also consider the shift in power dynamics that results from an entirely data-driven economy. The value in data sets may not always be contained within the data themselves but in the knowledge of how to leverage and apply the information they contain. In the example of neural networks, systems are creating vast amounts of data that often cannot be parsed in real time, so may require visualisations that highlight salient moments according to use case-driven parameters. Value of these data may only become apparent at the level of application, and IP strategies may be applied at the point of valorisation. Novel incentives, behaviours, business practice and new business models emerge out of these paradigm shifts.

The following use cases illustrate some of the recent changes of business incentives and behaviours, and changes of technological power dynamics.

2.2.2. MTF Labs open innovation community use case

Incentivised by the potential of emerging data-driven market applications, large organisations contribute datasets, components and products into a tech transfer toolkit, to be used for novel applications by the MTF Labs[2] community of innovators. The innovators solve several challenges related to products, technical aspects, market adoption, and emerging market scenarios. Background IP that is embedded into a new product is tracked and reveals a new market possibility. The breakthrough application idea is registered as innovation IP, so that the innovator or SME who created it has a vested interest in taking it to market. Several prototypes are ready after 4 months of incubation and tested on a community of early adopters. In month 5 a patent application is filed for the most marketable solution.

[2] MTF labs emerged from a DG CNECT CSA and have been running for the past 10 years. 21 labs have been run across global locations over three continents, with over 8000 contributors to date. 74% of the CNECT Advisory Forum innovation recommendations for 2018-2020 were based directly on real use cases from MTF Labs. https://mtflabs.net/

2.2.3. Chemical industry use case

In 2019 SAP published the guideline booklet showing how they can manage the entire chemical industry production from start to finish through data driven systems[3]. Much discussion followed questioning the implications this had on power structures and on the ownership of the chemical industry's tools of production. If data-driven systems were the primary tools of production, the tangible assets of the chemical industry's production line could now be considered as content managed by the DVC. The resulting paradigm shift requires a review of IP and IA dynamics.

2.2.4. Analysis of open Innovation use cases

Use Case	ENHANCES	OBSOLESCES	RETRIEVES	REVERSES
2.2.2 MTF open innovation community	Innovation IP Layer built on top of industry source IP by diverse, creative and interdisciplinary community of experts. Rapid knowledge exchange.	In-house innovation as the sole source of incremental IP development, focus groups and satisfaction surveys as guides or success indicators.	Creative discovery at the intersection of knowledge categories, testing with experts and early adopters to feed observations and discoveries back into the development cycle.	Loss of control of the planned trajectory of product development, target market focus or feature roll-out
2.2.3 Chemical Industry	Full process management through data driven systems	Requirement for arcane sector-specific knowledge and industrial gatekeepers of materials knowledge	The domain of business as a discipline in itself, understood as a system of processes and practices independent of the specific inputs or outputs of an individual organisation	Entire industrial sector reclassified as 'content provider' to software and data companies

2.2.5. Recommendations for open innovation as a driver of new data ecosystems

- Ensure all EU data-driven platforms operate on the same ”railway tracks”.
- Create rules of engagement for a level-playing field for IP & IAs where all stakeholders can benefit.
- Ensure alignment between multilateral DVCs and the Green Deal ambitions.
- Support industrial testbeds where industries can effectively experiment within open innovation systems to create sustainable solutions.
- Democratise access to know-how and knowledge hidden in patents.

[3] https://www.sap.com/documents/2019/04/66cf30b9-497d-0010-87a3-c30de2ffd8ff.html

- Highlight pros and cons of solutions for different stakeholder groups – it's not a "one size fits all".

2.3. Boosting SME engagement and innovation with FAIR and JUST data

The essence of leveraging all intellectual outputs in open innovation environments is ensuring attribution and accountability. Aside from setting up basic codes of conduct and rules of engagement, DVCs offer the possibility of designing automatised and semi-automatised systems of attribution and accountability, including the tracking of IP in the DVC and annotations of data provenance.

Industry requires assurances about commercially available datasets[4]. “FAIRification” of data (certification that the dataset is FAIR), and JUST data annotation (a practice for the RRI researcher and data owner) have proven to be extremely useful for the valorisation of commercial datasets. This presents opportunities for small and large players alike. Aside from assurances, certifications offer additional knowledge and information about the potential quality and performance of the dataset, often assisted by innovation SME applications, that build further value for their assets with layers of additional IP.

This is one of the reasons why is it particularly interesting for industry to be involved in the spaces of the commons. Every single step that is taken with their data assets builds further knowledge and adds value to it. The IP valorisation grows and scales with its use. The interconnected system allows for notifications from every step of deployment and tracking of their IP through the value network. This allows IP owners to make informed decisions about further investments and market development.

The following use cases highlight the optimisation and valorisation of datasets brought by FAIR and ethical approaches to data management.

2.3.1. JUST data annotation use case

A computer scientist is using a FAIR dataset of energy usage patterns among city dwellers to optimise energy supply. When modelling several usage scenarios, the scientist realises that the dataset contributes to optimisation only under certain parameters. Wishing to understand better the data bias and provenance, the scientist requests from the data provider to annotate the dataset according to JUST data principles[5]. The data owner/provider contributes responsible data annotation that aims to be Judicious, Unbiased, Safe and Transparent, and provides information about the geographical location where data was collected, relevant environmental circumstances, the population demographic that was sampled, and highlights any data privacy or security issues. The computer scientist can now make an informed decision on whether further data is required in order to optimise the energy supply, and can annotate the existing dataset further by noting both successful and unsuccessful application scenarios as reference for the next user of the dataset.

---

[4] https://www.eoscsecretariat.eu/cocreating-eosc/expanding-eosc-engagement-wider-public-sector-and-private-sectors-eosc

[5] JUST data principles have been proposed to complement FAIR data by placing into focus the responsible researcher or data owner. The proposal was welcomed by the EOSC Secretariat and is part of its strategy for the expansion of EOSC to innovation and industry stakeholders (ibid) as well as the EOSC SRIA (https://eosc.eu/sria). It is also part of the recommendations for DG CNECT in a study of AI in the CCS (https://www.technopolis-group.com/report/study-on-opportunities-and-challenges-of-artificial-intelligence-ai-technologies-for-the-cultural-and-creative-sectors/).

2.3.2. Whistle rights use case

A textiles multinational is developing a data-driven system for effective tracking of product supply networks. The organisation aims to embed ethical approaches to supply chain management by tracking the working conditions of workers and prevent labour violations. In doing so, the organisation must protect the workers' data, while collecting sufficient information to be able to make informed decisions and take appropriate corrective actions. The textiles multinational partners with an innovative SME that is building software in collaboration with the Whistle project at Cambridge University[6]. Whistle's tools allow for reports to be collected through smart phones and social media via a secure platform. All identities and associated data are verified, anonymised and aggregated into an ICT SME-designed dashboard that assists multinationals in decision-making processes. The anonymised information about the supply chain working conditions can instigate systemic change inside the multinational's organisational structure.

2.3.3. Analysis of SME engagement and innovation use cases

Use Case	ENHANCES	OBSOLESCES	RETRIEVES	REVERSES
2.3.1 JUST data annotation	Decision-making and broader ethical context for application of data sets	Replication and reinforcement of inherent or unconscious bias	First-hand field experience for third party users of data sets	Scope creep: the expansion and dilution of primary data research into related but peripheral concerns
2.3.2 Whistle rights	Convenient self-reporting and unobtrusive research to enhance staff wellbeing and productivity. Autonomy.	Dehumanised workforce following standardised and mechanised processes.	Creative and collaborative individual contribution to shared goals	Micromanagement and surveillance, mismatch of reporting and actual work performed in order to conform with perceived expectation.

2.3.4. Recommendations for boosting SME engagement and innovation with FAIR and JUST data

- Design automatised and semi-automatised systems of attribution and accountability, including the tracking of IP in the DVC and annotations of data provenance.
- Include FAIRification of commercial datasets for assurance and quality purposes.
- Introduce guidelines for JUST data annotation to ensure RRI data practices.

[6] http://thewhistle.soc.srcf.net/projects/

- Consider the intellectual property value of annotations of data for use in e.g. AI software training.
- Assist students, researchers, startups & SMEs to make use of and understand IP information to boost development, avoid reinventing known concepts, handle legal risks with competitors' rights and support more efficient business strategies.

2.4. Managing IP in multilateral DVCs

The core question for IP management in multilateral DVCs is which IP will be generated and how should it be protected and shared among stakeholders.

The digitisation of all industries, where existing and new knowledge is turned into data, including representation of all tangible industry assets in digital twins, provides infinite scope for interoperability and innovation between sectors. IP assets that are part of the data commons thus generated, may include copyrights, patents, trademarks and designs that focus on different aspects of a creation, novelty or industrial applicability. Bringing them together requires alignment over a shared top-level *metadata framework*, that is FAIR and references the related open or protected documentation, smart contracts and industry agreements.

Clarity over types of innovation is essential for a successful IP strategy. *Incremental innovation* is the standard practice applied to most innovation within large organisations that optimises and improves existing technologies and business models. *Radical innovation* creates a new technology that serves the existing business model. *Disruptive innovation* creates a new business model that leverages the same technology. *Architectural innovation* changes both the technology and the business model simultaneously[7].

The value and potential economic leverage of these different categories of innovation IP vary significantly in scale across different global business ecosystems. For example, the United States tends to be more conducive to patenting new business models than the EU, where patents are mostly based on scientific breakthroughs. Innovation in the technological domain without corresponding innovation in the business model (and vice-versa) can often realise short term results but may become redundant over time.

Since process optimisation is high priority for data-driven business models, *accuracy and timeliness of the data* becomes the key driver of successful systemic solutions for enterprises. Obtaining and enforcing exclusive rights over the raw data is not a sustainable business model. Data are a highly time- and context-sensitive resources. Business models built on licensing exclusive rights over out-of-date, inaccurate or biased data sets are unlikely to remain profitable in the long run. It is not exclusive rights, but timely and convenient access that allow economic operators to tap into the value that lies in high-quality data, i.e. by focusing on providing 'exclusive rights-agnostic' access to accurate and contextually relevant data with guaranteed quality of service parameters.

The following use cases illustrate business challenges across different industrial domains, that affect the design of supporting IP and IAs.

2.4.1. Dynamic metadata in manufacturing use case

The digital twin of a tangible manufacturing asset performs localised cognitive processes. It tracks the asset performance as well as environmental conditions for its optimisation. In addition, the digital twin is linked to updates of enterprise integration data and can model quantitative or qualitative design values in response to business requirements. A change in production values triggers an update of IP values in the digital twin and the IP metadata is automatically updated to reflect the new valorisation state.

[7] Pisano, G. (2019), Creative Construction: The DNA of Sustained Innovation, Public Affairs.

2.4.2. Otter.ai use case

An EU SME proposes to record and transcribe their confidential business meetings using Otter.ai software. The AI-assisted software analyses and trains on the speech patterns and accents of each of the contributing speakers. In order to evaluate potential risks and gaps, the SME investigates whether the data analytics generated by the software can be made available to them. They discuss levels of risk associated with their rights as providers of the data, particularly in respect of making their data available to Otter.ai's partners such as the video conferencing software Zoom. They attempt to evaluate the risk of their business confidentiality being exposed to Otter.ai's VC investors. The SME concludes that it requires advice and guidance from EU regulatory bodies.

2.4.3. Medical annotations use case

A clinician is evaluating data about specific medical treatments to assist in choosing the correct treatment for a patient. Aside from data from clinical trials, the clinician is able to access the annotations database that has compiled data about applied treatment scenarios. The clinician's peers have annotated the dataset based on their experience of applying a specific medical treatment on particular use cases assisted by tools developed by innovative SMEs such as labelstud.io[8] and prodi.gy[9]. The annotations are FAIR and searchable according to tags and keywords, allowing the clinician to find the closest use cases, and ascertain which treatment produced good results. The clinician finds the annotations save time in applying the correct treatment, increase success rates and reduce risk. The pool of knowledge created through the clinicians' annotations is incentivised through attribution of IP to each contributor which is then compensated when valorised by e.g. AI software that leverages patterns in those annotations. Annotations therefore create intellectual property that is valuable in its own right.

2.4.4. Analysis of managing IP in multilateral DVCs use cases

Use Case	ENHANCES	OBSOLESCES	RETRIEVES	REVERSES
2.4.1 Dynamic metadata in manufacturing	Mass customisation and adaptation. Complete lifecycle circular production. Context-sensitive alteration of IP production and generation.	One size fits all manufacturing. Excess production and waste. Inability to adjust according to environmental factors or changing conditions.	Bespoke customisation reintroduced in mass production contexts. Creation of subtlety in manufacturing iterations	Adaptive products as standard. Constant shifting of manufacturing output threatening consistency of brand identity

[8] https://labelstud.io/
[9] https://prodi.gy/

2.4.2 Otter.ai	Creates a text searchable record of internal intelligence for the organisation	Excessive reporting and administration that obstructs agile decision making	Organisational memory and the ability to return to closed-door meetings. Grouped information and internally-shared private intellectual property	Increased exposure to industrial espionage and foreign leverage of AI training through contribution to big data sets in convenience trade-offs that are difficult to calculate or mitigate.
2.4.3 Medical annotations	Shared specialist knowledge and analysis amongst peer practitioners.	Compound errors from access only to one's own subjective analysis. Reliance on peer-review process for knowledge acquisition	Clinical practice conference networking to share observations and anonymised diagnosis metadata	Potential to over-rely on consensus diagnosis rather than trust one's own expertise.

2.4.5. Recommendations for managing IP in multilateral DVCs

- Create a top-level IP metadata structure that is FAIR and applicable across all IP categories.
- Clarify the legal status of data, data products and data annotations under IP law.
- Discuss the application of sole and joint ownership doctrines in multilateral digital value chains.

Identify solutions for tracking

2.5. Incentivising rich data value networks

The transition from current data value chains into richer multilateral ones with the right incentives, requires the building of an industry commons *ecosystem of trust*[10]. Innovation is enabled by cross-domain interoperability supported by data reference models such the EU's OntoCommons EcoSystem (OCES[11]). However, interoperability requires the creation of a supporting system of systems that stimulate data exchanges in optimal, sustainable and ethical ways.

The industry commons model therefore anticipates a series of enabling systems connected and supporting the interoperability and innovation layers. *Systems of agreements* embed regulation, IAs, peer-to-peer contracts and IP registration and tracking. *Systems of resilience* include environmental sustainability and resilience strategies to mitigate against Black Swan events. *Systems of responsibility* introduce responsible AI, ethics, and Corporate Social Responsibility. At the foundational layer, *systems of beliefs* set parameters that reflect societal values, including making sure that inclusivity is foundational to the whole system.

Entering an open innovation ecosystem of trust is accompanied by multiple business and optimisation incentives[12]. Releasing industry background IP into the space of the industry commons is supported by feedback loops and notifications that highlight salient advancements in research, innovation, technology transfer, product development, marketing and HR recruitment. Participation provides additional value of FAIR data from academic research and TTOs. As innovators model new use case scenarios based on industry background IP, or build upon identified gaps across domains, they provide significant emerging market data, onboard early adopters, and identify themselves as potential partners for incubation and market acceleration.

In open innovation environments therefore, building on the shoulders of one's peers provides greater value for all, and therefore requires updates to IP strategies. As it is currently configured, patent registration actively ring-fences individual ideas. In open innovation environments, showing provenance and context reinforces, rather than detracts from an argument for novelty or patentability. It makes clear that which is new and how it advances the knowledge. Showcasing a value network of ideas can clearly distinguish one IP registration from a similar idea that has been arrived at through another DVC with different background IP. This incentivises greater participation and stimulates multiple opportunities for valorisation for involved stakeholders.

The following use cases illustrate how DVCs allow for agile integration of innovators and entrepreneurs in innovation processes.

2.5.1. Industry commons use case

An innovation SME is modelling a novel use case within the industry commons. The SME uses a range of proprietary data from materials companies to test material properties in combination with a novel use case scenario and related environmental data from public sources. All data is FAIR and reusable. The resulting dataset creates

[10] Michela Magas & Dimitris Kiritsis, 2021. Industry Commons: an ecosystem approach to horizontal enablers for sustainable cross-domain industrial innovation (a positioning paper), International Journal of Production Research, DOI: 10.1080/00207543.2021.1989514 [https://www.tandfonline.com/doi/abs/10.1080/00207543.2021.1989514]

[11] https://ontocommons.eu/

[12] The EU project RE4DY explores the business potential of data by uniting multiple platform and software stakeholders into a multilateral DVC valorisation system. https://re4dy.eu/

a layer of innovation IP that indicates optimal conditions for certain materials in the specific innovation use case. The proprietor of the best performing material dataset receives a notification of the material's optimal performance in the modelled innovation use case that reveals a potential new market application. Both the SME and the background IP holder can make an informed decision about a potential emerging market opportunity.

2.5.2. IP Stack use case

An innovation SME is an avid tester of breakthrough technologies and has good knowledge of emerging markets. It sources background IP from TTOs in the DVCs and designs technology transfer IP (hardware and software) that it combines with background IP to lower the entry barrier to emerging markets. It uses modelling software designed by a partner SME to model emerging market use cases and gather data about market deployment. The resulting dataset creates a layer of innovation IP, that combines with background IP and transfer IP to create an emerging market IP Stack. All partners in the value network are notified of the emerging market opportunity.

2.5.3. MARLs use case

An SME is producing a low-risk application that does not require a large investment upfront and long development times. They are faced with having to decide what to patent and how long to keep the development confidential. They assess patenting times as too long for timely release of the product to the market. They prefer to be the first in the market and benefit from a competitive advantage. For a timely release, the SME finds the Technology Readiness Levels alone inadequate for monitoring rapid development progress, and instead wishes to involve adopters at early TRL stages. The SME opts for Market Adoption Readiness Levels (MARLs)[13] to help finetune the application's usability and encourage early adoption. They consist of evaluating the level of risk of an application, incentivising early adoption and estimating the adoption potential, analysing the data from early adoption, and assessing technology readiness.

[13] Market Adoption Readiness Levels (MARLs) were introduced by the CNECT Advisory Forum in 2014 in order to speed up market deployment of low-risk data-driven applications.

2.5.4. Analyisis of incentivising rich data value networks use cases

Use Case	ENHANCES	OBSOLESCES	RETRIEVES	REVERSES
2.5.1 Industry commons	Pool of shared knowledge from which new products and data applications can be generated	Lockdown of IP for fear of loss of control or that others may gain competitive edge	Collaboration across industrial sectors and between large and small platers to create mutually beneficial connections through layers of tracked and attributed IP	Outsourcing of innovation to independent or smaller organisations in order to leverage new IP created from unused assets
2.5.2 IP Stack	Ability to model new use case scenarios through digital twins and interoperability of data between TTOs, industry and innovation communities	High risk and costly experimentation within real-world manufacturing and production processes	Innovation process as a collaboration between trades	Disruptive and architectural challenges to organisations as a result of combinative IP innovation
2.5.3 MARLs	Usefulness of assessment of product readiness or market fit in an age of digital twins and rapid prototyping	Measures based on NASA safety protocols for high risk, high cost ventures	Experimentation and iterative product development	Ongoing releases of minimum viable product rather than fully realised and tested applications

2.5.5. Recommendations for incentivising rich data value networks

- Build an industry commons ecosystem of trust.
- Ensure that IP & IAs are one layer of a necessary industry commons system of systems.
- Ensure that inclusivity is foundational to the DVCs.
- Ensure feedback loops and notifications for background IP owners that highlight salient advancements in research, innovation, technology transfer, product development, marketing and HR recruitment.
- Consider SMEs and innovators as providers of emerging market data, early adoption, and potential partners for incubation and market acceleration.
- Update IP registration strategies to incorporate background IP and incentivise greater participation in DVCs.
- Analyse data transactions from the perspective of exclusivity, on one hand, and the values of timeliness and accuracy, on the other.
- Develop an online ICT platform to facilitate streamlined innovation workflows, easier team collaboration and smarter follow-up on findings and learnings.
- Build upon existing open innovation frameworks that rely on derivative and appropriation mechanisms, to ongoing collaborative innovation practices that include background IP provided in DVCs.

2.6. Industry Agreements for replicability and emerging business models

Fostering participation in multilateral DVCs relies on ensuring replicability and a business continuum, and therefore requires more than cross-domain interoperability and FAIRification of data. The multilateral DVC combination of technological and industry assets with business strategies and horizon scanning for market opportunities, goes beyond enterprise integration to a new dimension of cross-domain technological and business *ecosystem integration*, that brings novel business incentives, business optimisation, new business models and access to emerging markets.

Within the context of multilateral DVCs, SMEs are under pressure to build fast, but may not always be sufficiently agile or have sufficient knowledge of both technology and business to be able to identify correctly the innovation opportunity in the market.

The following use cases illustrate models for reducing risk and increasing business opportunities for innovation SMEs.

2.6.1. IP Screener use case

A startup SME website proposes to reduce CO2 and methane emissions by developing additives for cattle fodder. The summary of their business model is copied from their website and fed into the IP Screener AI-assisted software[14]*. IP Screener analyses the text and maps out the global activity in this domain, ranking research and innovation results by relevance and visualising them on a world map. Most economic activities in this domain appear in the US, followed by Canada. In Europe the dominant market activity is in the UK, followed by Belgium, Germany and Finland. Australia, Japan, and New Zealand show some relevant market activities, that could also be targeted by the SME's business plan. A notable presence in the European market is shown to be Royal DSM, known for nutrition technologies, though less known for fodder. As part of their business plan, the startup could consider partnering with such a big player or include them in their exit strategy. The software reveals that the topic is trending, and that the trend has been increasing steadily over the past 20 years. For the benefit of investors this demonstrates that the startup idea is not hype. The software reveals a particularly prolific researcher/inventor on the topic, whom the SME could consider hiring. A classification code reveals four further species of cattle who can benefit from the invention. In summary, the software has enabled the SME to perform a risk and business assessment in a very short time. They have been working on this product for two years, but in around 10 minutes they have found a new selling proposition, relevant companies on the market, relevant updates for their business plan and even technology which could be used to improve their product. The patenting heat map shows relevant patent registration in the US, but an available market in Europe.*

2.6.2. Berkeley startuppers program use case

Berkeley offers an eight-week program with coaching and training of SME business "startuppers". This includes an obligation for SMEs to conduct 100 interviews. An AI-assisted software that scans the global research and innovation landscape, such as

[14] IP Screener provides AI-assisted software for rapid scanning of the global patenting landscape, including the entire WIPO and EPO database: https://ipscreener.com/

the one shown in 2.5.1, can help them to target the right stakeholders, investigate where to pivot, and correctly position their business proposition early in business development to reduce risk.

2.6.3. Cross-domain applications use case

An SME has identified a gap between two industry verticals, and created a hybrid solution, compatible with both data domains. The SME uses the MARLs model (see 2.4.3) to onboard early adopters from both domains and identify the uptake rate. Engagement with the breakthrough solution yields a great deal of data and results in useful business insights for product improvement. Despite a low technology readiness, data from early adoption is shown to background IP proprietors and to potential investors to illustrate market potential and help them make an informed decision on further investments.

2.6.4. Analysis of IAs for replicability and emerging business models use cases

Use Case	ENHANCES	OBSOLESCES	RETRIEVES	REVERSES
2.6.1 IP Screener	Patent search speed and ability to locate high quality work happening in the field that may connect with proposed developments	Expense and limitation of patent search focusing only on elimination of possible competing ideas	IP registration as frontier knowledge creation rather than territory claiming	Reverse engineering of patent application through elimination of searchable terms
2.6.2 Berkeley startuppers programme	Rapid AI feedback on market fit and early pivot recommendations for new SMEs and entrepreneurs.	Tenacity and luck as the key defining features of successful market innovations. Fixed 'grand visions' and missions	Flexible Intellectual property as the central valuable asset of an entrepreneurial team.	Market-led rather than market-driving or market-creating innovation. Regression to the mean rather than outliers.
2.6.3 Open Source	Talent pool of collaboration for the creation of useful services and applications. Creative forks and unforeseen innovation leaps.	In-house development teams working to a single coordinated plan. Proprietary software as the sole economic model.	Mission-driven innovation, moonshots and alignment of multitudes around a single vision	Corporate exploitation of community contribution.

2.6.5. Recommendations for Industry Agreements for replicability and emerging business models

- Move beyond enterprise integration to a multilateral DVC cross-domain technological and business ecosystem integration.
- Clarify the legal effect of automated data transactions from the perspective of IP law
- Identify key elements in data-driven business models that determine the choice of open, closed or mixed IP strategies.
- Raise IP awareness as a business tool and educate users to be able to have more mature and valuable discussions around it.
- Implement AI to support SMEs in reviewing and benefiting from the mapping of their innovation landscape, to make use of extracted business intelligence and other IP activities on the market.
- Identify where IAs are needed and where more informal agreements can yield timely results.

2.7. DIH as key players in multilateral DVCs

The European Digital Innovation Hubs were created as ecosystems where various SME stakeholders could be integrated in larger value networks, contribute to knowledge transfer and to the building of innovative solutions. Much of DIH activities are already aligned with open innovation strategies and therefore well placed to capitalise on the new business and IP strategies coming from the DVCs.

Tracking IP in value networks, making SME IP FAIR through a standardised top-level metadata, ensuring registration of SME *innovation IP* built on top of background IP provided by large organisations, layers of attribution, accountability, as well as confidentiality and protection of SME trade secrets, are all models that can support SME participation in future multilateral DVCs.

The following use cases demonstrate the value that open innovation rules of engagement and multilateral DVC strategies can bring to DIH SMEs.

2.7.1. Open Source use case

In a connected DVC that tracks IP valorisation, an organisation uses open source code in a software product that is to be licenced under a proprietary licence. 95% of the final product is based on public domain or open source material and 5% is novel. Since the application of the open source code is trackable in the value network and attributed to the SME that originated it, the SME receives a notification of the code reuse and the system triggers a minimum level of compensation for the SME.

2.7.2. Patent for primary industry use case

At seed stage, an SME uses a microboard supplied by a large semiconductor supplier that offers novel affordances for embedded ware using data from proprietary sources. It creates a prototype that allows the use of bodily gesture to change menus hands-free while on the move. The system recognises the difference between regular movement (e.g. jogging) and deliberate gesture that gives instructions to a control centre. The SME registers the invention as an innovation IP layer in the DVC. The system is then tested in conjunction with operators of primary industry vehicles who ordinarily have to stop operations in order to answer phone calls or react to an emergency. Prior to testing the novel application, the potential partner from primary industry had been testing ways to improve the operator's controls using expensive screen and camera equipment. This was raising the price of the vehicle too much for the corporate buyer. The novel system designed by the SME proves to be very effective at optimising communications and adding security features while keeping the costs of implementation low. The SME receives an offer for deployment of the product in primary industry vehicles. The SME then negotiates a percentage profit with the semiconductor supplier that owns the proprietary IP embedded in the solution, and seeks competitive offers for similar embedded ware. With clarity over the registration of innovation IP in the DVC, the SME proceeds to file a patent in the domain of primary industry.

2.7.3. Analysis of DIH as key players in multilateral DVCs use cases

Use Case	ENHANCES	OBSOLESCES	RETRIEVES	REVERSES
2.7.1 Cross-domain applications	Shared data formats and standardisation enable hybrid innovation. New market categories and novel solutions to societal challenges	Industrial silos and locking down of specialist IP within specific industry sectors	The company as organisation of collaborators with complementary skills and knowledge bases. Age of Discovery.	Shallow understanding and misapplication of complex and deep sector-specific knowledge
2.7.2 Patent for primary industry	Identification of innovation in one domain as useful, applicable or even essential in other, seemingly unrelated verticals	Expensive investment in large scale incremental innovation	Ideas and innovation as cross-boundary IP applicable across industry sectors. Potential to 'invent the wheel' with a simple concept applicable and revolutionary across a wide range of domains.	Inventors and innovators disincentivised by working in sectors other than those that inspired them to create the original IP.

2.7.4. Recommendations for DIH as key players in multilateral DVCs

- Support SME participation and awareness of business and IP opportunities in multilateral DVCs in the next EU frameworks.
- Ensure that participation in DVCs is stimulated by valorisation strategies that are sufficiently affordable and sustainable for SMEs.
- Ensure DIH SMEs benefit from attribution, accountability, layers of confidentiality and protection, the tracking of IP in value networks, FAIR data strategies for a standardised top-level metadata, and registration of innovation IP built on top of background IP provided by large organisations.

Create mechanisms for support of arbitration and dispute resolution to protect EU DIH SMEs IP

2.8. Conclusions on IP and Industry Agreements

3. Pillar 4: Multilateral Digital Value Chain (DVC) Business Continuity Index

The main objective of this fourth pillar is to produce a methodology for the development of a framework to assess the digital maturity of modern Digital Value Chains (DVC) — with applicability across all industrial sectors. The framework is conceived to provide orientation and guidance to decision-makers willing to engage in cross-DVC collaboration schemes.

Such an ambitious target requires to reduce the complexity by dividing our methodology into four steps: (1) definition of scenarios for multilateral data exchanges, (2) assessment of interconnection of existing assets, (3) gap analysis (4) improvement of existing pathways.

1. **Definition of scenarios for multilateral data exchanges.**

This paper provides a series of **3 scenarios** (value-chain data exchange, multilateral data exchange, data spaces) to model how data exchange takes place in contemporary industrial sectors. The goal of these scenarios is to help stakeholders to understand in what situations and for what purposes the framework can be applied. Decision-makers should start by identifying **in which scenarios their targeted business cases take place** and what it involves in terms of strategy, costs and efforts.

2. **Assessment of existing assets** and development of the **Business Continuity Index (BCI)**.

The assessment of digital maturity for business continuity **should start by investigating at the DVC level for at least two reasons**. Firstly, because this is easier and most cost/time-effective. Secondly, because many key factors to embark in the data economy are not controlled by private players. Instead, they take place at the value chain level — a single player does not necessarily have control (e.g., data sharing legislation or multilateral data exchange architecture).

The framework is therefore developed as a series of 8 criteria to measure digital maturity of DVCs. This assessment is designed to provide the analyst with a **Business Continuity Index (BCI)** to summarise the level of digital maturity of the analysed value chain(s). This simple approach provides a powerful tool to orient efforts and strategies of industrial players. For example, a DVC with a very low BCI is unlikely to offer an interesting target to undertake the efforts necessary to engage in a business partnership.

3. **Gap analysis and definition of scenarios**

The gap analysis acts as a bridge between the assessment of existing assets and improvement of pathways. The gap analysis is composed of 8 diagnostics (one for each criterion of the maturity assessment) to assess what is missing, based on the DVC assessment, to carry the targeted business case (is it feasible at all, what would be the cost, risks, etc). Then doing the gap analysis, the view point should move from a DVC perspective to **more specific analysis based on the business case** and type of organisations targeted.

4. **Improvement of existing pathways**.

The improvement of existing pathways requires to identify remediation measures to address all the identified gaps, assess feasibility, costs and adjust the business case.

3.1. Scenarios for multilateral data exchanges

Transparency and interoperability are the key enablers for resilient supply-chains and production systems to ensure business continuity, either at the company level or for transversal cooperations and exchanges.

The aim of the following implementation approaches is to show, how data exchanges can be realised, while preserving the autonomy of those involved and at the same time exploit the values of data exchanges. For this purpose, three possible topologies were developed:

1. Value-chain specific Data Exchange
2. Multilateral Data Exchange – Using a common Data Exchange Framework / Data Space
3. Cross Sectorial Interaction – Data Exchange between companies in two different Data Exchange Frameworks / Data Spaces

Basis for all three implementation approaches is **the Asset Administration Shell (AAS).** The AAS is the data image of an object (all kind of entities like machine, software, humans etc.) and contains all the information that characterizes the object and its behaviour. The life cycle of products, devices and documents can be documented with these digital images[15]. This approach allows to exchange asset information in a structured, interoperable way across companies in the engineering phase as well as the operation phase.

3.1.1. Value-chain specific Data Exchange

Nowadays, productive value creation in industrial applications takes place in globally distributed value networks. Such a value network includes a variety of companies, such as OEMs (original equipment manufacturers), suppliers at various tier levels, machine suppliers, system integrators and IT system solution providers. These companies are accompanied by providers of additional services for the organisation and execution of global logistics and service processes.

The value-chain's specific data exchange typology concerns the exchange of data along a specific value-chain. The participants are familiar with each other at each respective level. This means that the OEM knows the tier 1 supplier and the tier 1 supplier knows the tier 2 supplier.

The basic use case to demonstrate this kind of topology is the **Collaborative Condition Monitoring**[16] [17] (CCM), that represents the mechanisms needed to exchange data across company borders considering autonomy aspects.

CCM is based on cross-company cooperation within the framework of the "three-point fractal" consisting of a component supplier, a machine supplier (integrator) and a factory operator, with the aim of generating economic added value for all parties involved (Figure 1). Figure 1 represents machine supplier integrating several components from many component suppliers. He delivers this machine to a factory operator, who integrates the machine in his production and operates it afterwards.

[15] Platform Industrie 4.0, 03/02/2022, *Asset Administration Shell – Reading guide*, https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/AAS-ReadingGuide_202201.html
[16] Platform Industrie 4.0, 30/05/2022, *Details of the Asset Administration Shell*, *https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/Details_of_the_Asset_Administration_Shell_Part1_V3.html*
[17] Platform Industrie 4.0, 29/09/2020, *Collaborative data-drive business models*, https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/collaborative-data-driven-business-models.html

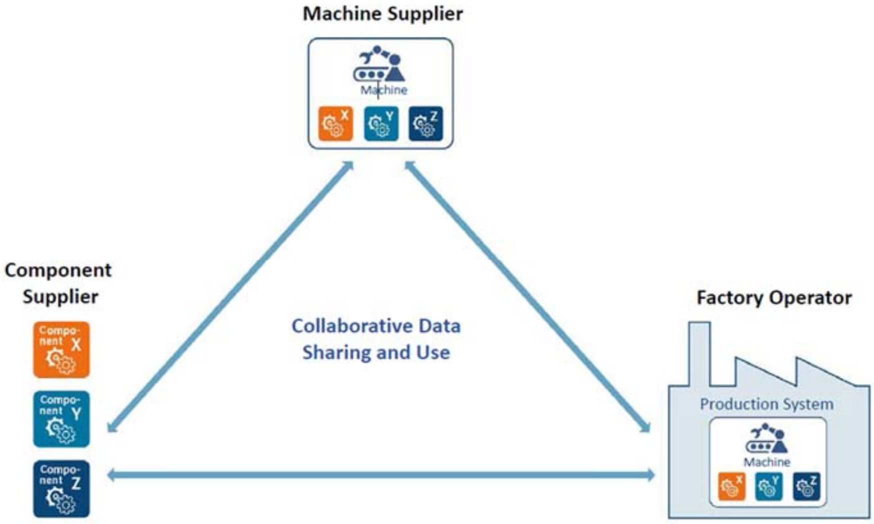


Figure 1 „Three Point Fractal“ of Cross-Company Data Sharing in industries

The specific aim of the CCM use case is to identify the need for maintenance measures in good time by monitoring the condition data to avoid a possible failure of the production facilities. Therefore, the condition data of the machines and of their integrated components operated by the factory operator are transmitted from the factory operator's automation network via his intranet to the factory operator's AAS server and are stored there in sub models of the asset's AAS (digital twinning).

From this starting point two use cases can be differentiated: ”Use Case A - Value-chain Specific Data Exchange - Direct Between Several Companies” and “Use Case B - Value-chain Specific Data Exchange – With Data Trustee“

Use Case A - Value-chain Specific Data Exchange - Direct Between Several Companies

The factory operator transmits the sub model data[18] to request value-added partners. In order to do this, he has set up a specific data policy[19] for each data set and must enforce it on every request (Figure 2).

Since it can be assumed that the manufacturers of the components are most capable to assess their components, the component supplier (tier 2 supplier) can request the condition data of his components from the factory operator or from the machine supplier as well. The factory operator and the machine supplier only grant access to the component supplier if they comply with their individual data policies.

Through this process the machine supplier and the component suppliers become data consumers and the factory operator becomes the data provider.

[18] Operating data indicated by the thermometer in the graphic
[19] Indicated by the check list beside the thermometer

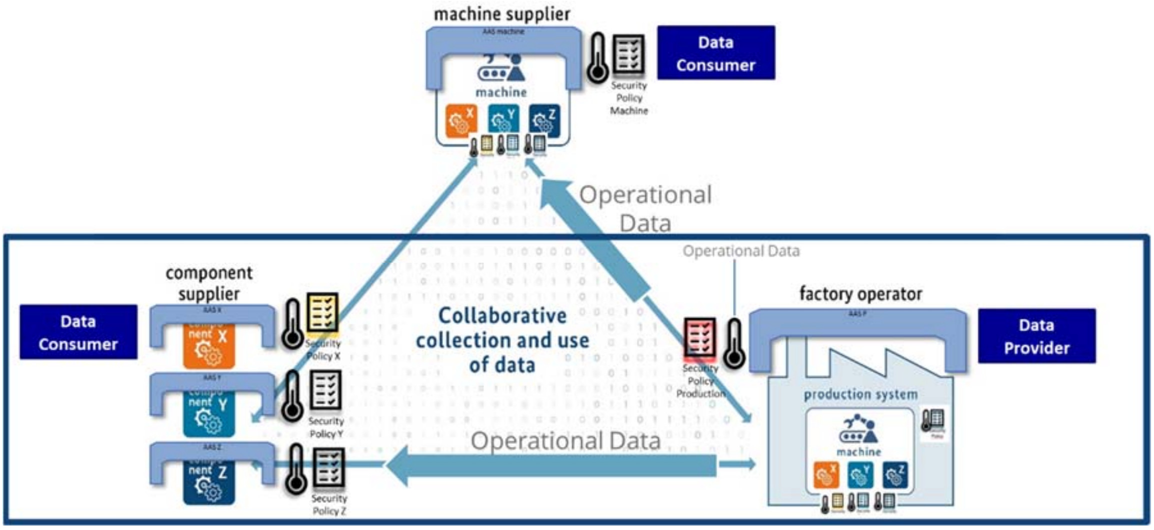


Figure 2 Value-chain Specific Data Exchange - Direct between participating companies

Use Case B - Value-chain Specific Data Exchange – With Data Trustee

The factory operator transfers the sub model data of the AAS, including a suitable data policy to a service provider on the Internet, who manages the data as a trustee (Figure 3). Since it can be assumed that the manufacturers of the components are most capable to assess their components, the tier 2 supplier (component supplier) can query the data of the components he has manufactured from the data trustee. However, the data trustee only grants access if the component manufacturer has been granted appropriate rights by the factory operator.

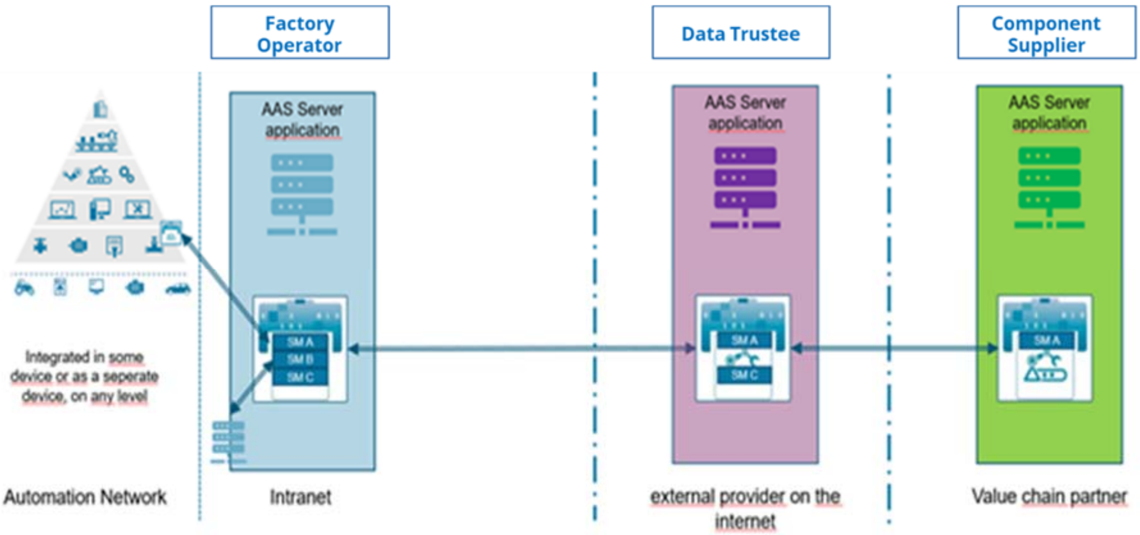


Figure 3 Value-chain Specific Data Exchange – With Data Trustee

By doing so, business continuity of production can be ensured and, if the supply-chain is considered, potential failures of supply can be avoided or rescheduling of supply-chains can be conducted in case of necessary maintenances at the supplier's production. CCM is thereby just an example. Any other data-based service can be executed using this approach.

3.1.2. Multilateral Data Exchange – Using a common Data Exchange Framework / Data Space

All entities in the global value production network have a variety of different bilateral relationships with their partners in the value-chain. The complexity of these value networks can be modelled by scaling the three-point fractal[20] (Figure 4). This "simple" multiplication does not lead to transparency as it maintains a bilateral data exchange with no transparency up- and down-stream of the value-chain beyond the adjacent delivery level of the companies.

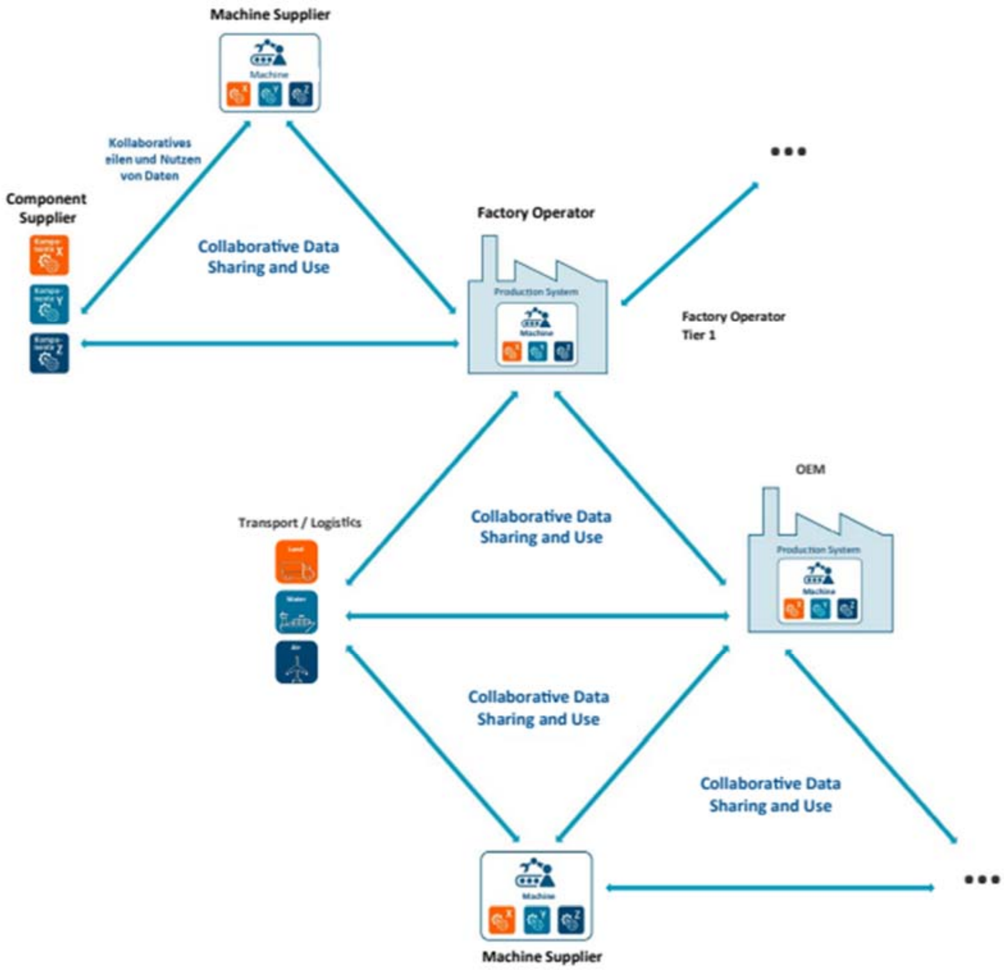


Figure 4 Supply-chain represented by the minimal „three-point fractal" (Source: [Plattform Industrie 4.0 - Multilateral data sharing in industry (plattform-i40.de)](plattform-i40.de))

At present, business models or activities designed to improve efficiency in many manufacturing companies involve bilateral data exchange between two participating companies (as sketched in Figure 2). However, for additional efficiency or to leverage market potential in the manufacturing industry, a change in mindset is needed. From purely bilateral data exchange to holistic, standardised and multilateral sharing of data by multiple stakeholders (Figure 5). It is key to business continuity to be able to monitor the whole supply-chain and be thereby in the position to adopt the supply-chain if a disruption occurs down- or up-stream including the connected domains like logistics.

Thereby new suppliers, whose products meet the requirements, can be found faster and can be easily integrated into the digital supply-chain.

[20] Basis model see Figure 1

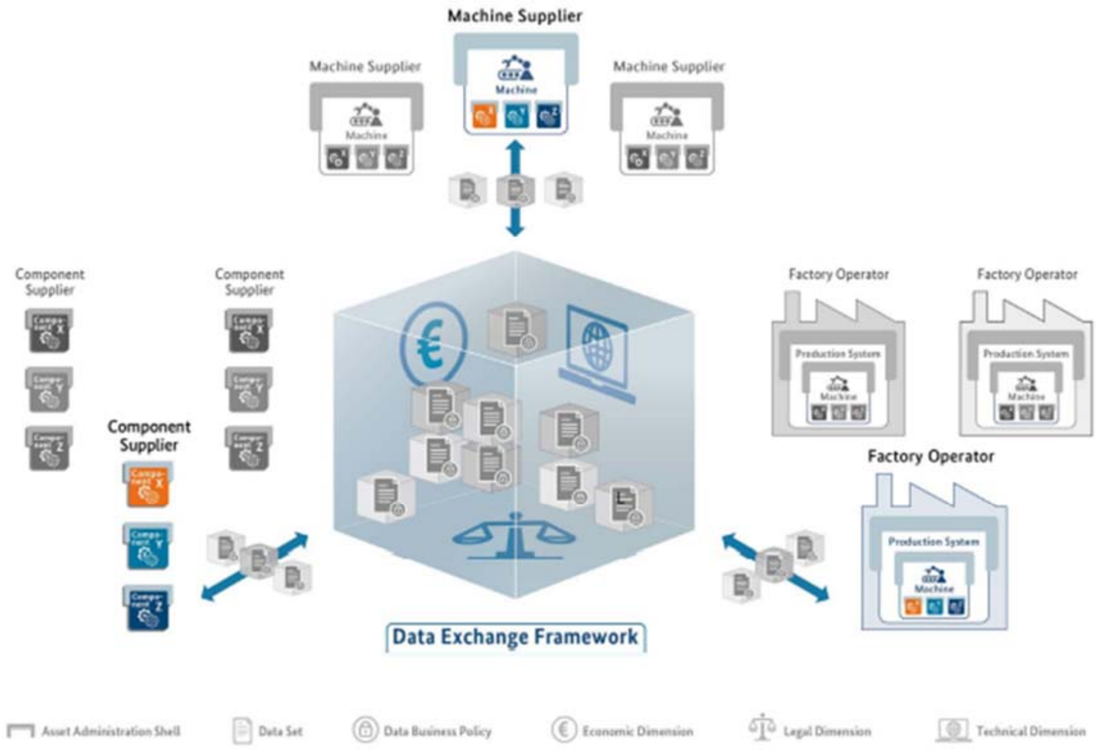


Figure 5 Data Exchange Framework Industry 4.0 (Source: [Plattform Industrie 4.0 - Multilateral data sharing in industry (plattform-i40.de)](#))

This **Data Exchange Framework** plays an important role **as a Data Space**. It is used to establish the framework conditions (technical, legal, and economic) for cross-company data exchange (see Figure 5). In accordance to a road network, the infrastructural components of the data exchange framework are to be built, based on an collaborative/cooperative approach among all stake holders and users, with appropriate fees paid for use.

Within multilateral sharing of data, the autonomy of the participants must be continuously preserved. Therefore, the opportunity must be provided in order to connect the generated data sets with data business policies that determine the type of permitted usage or application, covering both technical and economic aspects of data usage.

While the need to design the technical and legal dimensions of the data exchange framework, needed to implement the use case seems obvious, the economic dimension plays a particular role. Since the economic dimension concerns the share of added value resulting from the multilateral sharing of data between the participants. Hence, making it the driving force behind implementation.

One of the biggest challenges building sector-specific data frameworks/ data spaces is the development of a uniform syntax and semantic uniformity of attributes. Therefore, the theoretically safest way is to develop a basic ontology, based on which the sector-specific ontologies are created. The sector-specific supplements developed in the process flow into the basic ontology and are then available again for the development of further sector-specific ontologies.

In practice, however, one often encounters historically grown structures that are difficult to replace by a new ontology. Consequently, a widespread approach is to develop only the ontology for the exchange format for application-specific ontologies (local ontologies) based on sector-specific standards. Participants in the data room are then required to set up connectors to their internal, possibly proprietary data structures. This procedure often leads to the participants adapting their internal data structures to those of the data space as to reduce translation efforts. As a result, a data space-specific ontology can slowly develop, for which frictional losses must be expected because of the adjustment.

After each potential change of one of the connected participants (e.g., by update of some features), the connector of the participant must be checked for its correct functional reliability. This is to the effect of ruling out errors caused by incorrectly interpreted data and the resulting possible production downtimes respectively to ensure business continuity.

3.1.3. Cross-Sectorial Interaction – Data Exchange between companies in two different Data Exchange Frameworks / Data Spaces

Companies are often participants in several supply-chains and work together with companies in other sectors (Figure 6). While collaboration in one sector can be managed, using the previously presented Data Exchange Framework/ Data Spaces, cross-sector collaboration faces additional challenges.

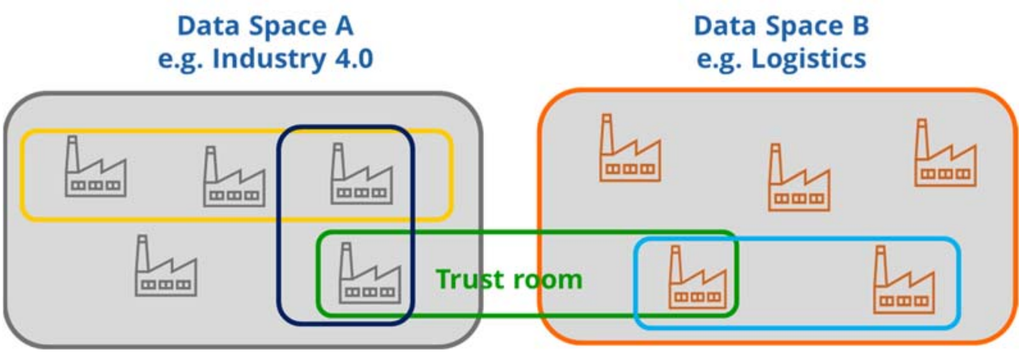


Figure 6 Cross-sector cooperation

These additional challenges are reflected in the necessary cross-sectorial sematic consistency of the attributes used to describe both the economic, legal and technical framework, as well as the content of the data sets and data policies (**Errore. L'origine riferimento non è stata trovata.**). A correct transfer of the data records and a valid examination of the data policies can only be conducted through the cross-sector semantic uniformity and a uniform syntax. That becomes very clear when considering physical dimensions like weight (kg, lbs etc.), length (cm, inch) or volumes (l, gallon).

The above-described sector-specific procedure to achieve a uniform semantic and synctatic can be transferred to cooperation across sector boundaries. However, the period for the standardisation of the ontologies is likely to be incalculably long. In most cases it will therefore be appropriate to compare the sector-specific ontologies of the exchange formats and to provide connectors for cross-sector data exchange, taking syntax and semantics into account.

After each change in the ontology of one of the connected data spaces, these connectors must be checked for their correct functional reliability to prevent errors caused by incorrectly interpreted data and the resulting possible production downtimes respectively to ensure business continuity.

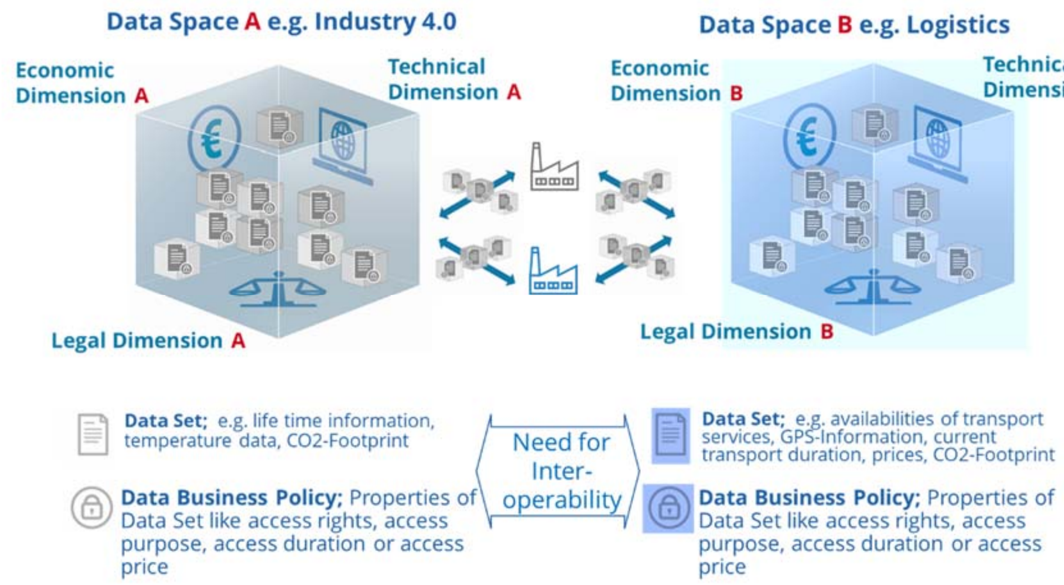


Figure 7 Cross-Sectorial Interaction – Data Exchange between companies in two different Data Exchange Frameworks / Data Spaces

3.1.4. Collaboration of Ecosystems

The European economy is characterised by diversity and plurality. Therefore, requirements for data ecosystems will differ. At the same time, there is a special added value in the cross-sector and cross-industry data exchange to optimize supply-chains, because this is the only way that supply-chains with their interactions in different domains can be monitored and, if necessary, adjusted to ensure business continuity.

Hence, it is essential that the specific ecosystems are interoperable with each other. At the same time, specific data ecosystems with sector- and industry-specific ontologies (syntax, semantics, relationships) keep entry barriers low for participants, e.g. for SMEs.

To bring these two pillars together, it is vital to speak an externally consistent language (as introduced in the AAS approach above) or to perform translations between the data ecosystems while interacting. Which approach fits best to the specific situation can be figured out by the interconnectivity/interoperability-check

.

3.2. Interconnection of existing assets (maturity assessment)

The framework is built as a ranking system of 5 qualitative levels organised in a logical and successive order:[21]

1) **Siloed level**: internal activities remain compartmentalised.
2) **Foundational level**: companies develop efforts for internal integration. Most efforts are done internally with proprietary formats, standards and systems. However, these efforts already contribute to make cross-company exchanges possible.
3) **Supply-chain integration**: efforts and results for interoperability are scaling-up but achievements are confined to small alliances, usually involved in direct business relationships as introduced in the Use Case A further below.[22].
4) **Foundational multilateral-DVC (M-DVC)**: some efforts to expand supply-chain initiatives into truly multilateral initiatives developed in the section 4.2.1.2 further below.[23].
5) **Mature m-DVC**: coexisting to bilateral exchanges, multilateral practices are common practice supported by mature infrastructure and enabled by a digital ecosystem as introduced in section 4.2.1.3.[24]

It is important to highlight that the framework is conceived for application at the **sectorial level**. Applying this scale at the company level would lead to inconsistencies as any value chain perspective starts from the idea that a company is already integrated into a supply-chain. To some extent, this framework could still be applied at the company level at the condition that both the siloed-level and the foundational-level are removed from the framework.

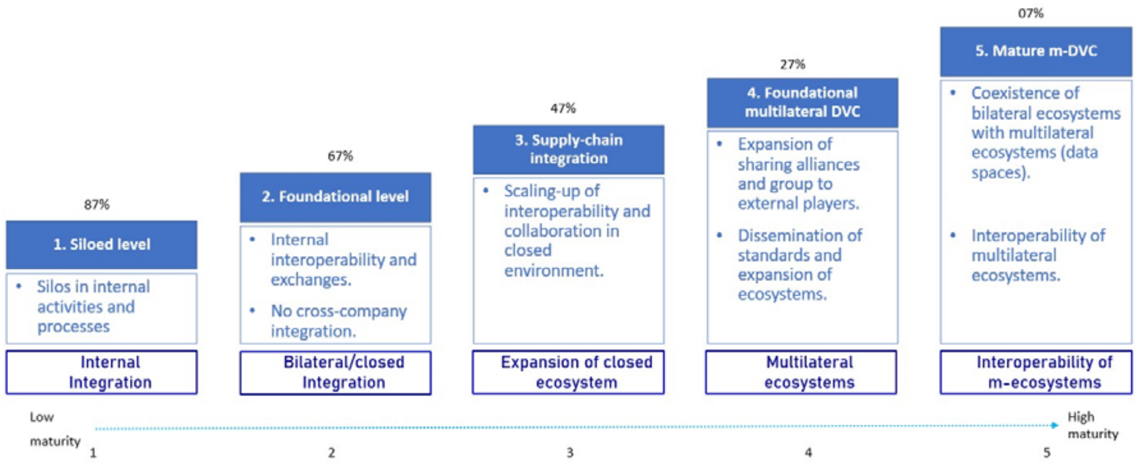
The succession of stages was organised in a logical manner. Indeed, it is highly unlikely to exchange in exchanges with companies active in a DVC where activities remain compartmentalised (level 1).

However, companies are not working in a closed environment. The State, for example, is a decisive stakeholder that can independently undertake initiatives to foster digital maturity at different stages of the scale. For instance, new legislations (e.g., the data act) open new possibilities to all European businesses to engage in multilateral data exchanges. Similarly, the EU currently undertakes important efforts for the development of Industrial Data Spaces across all economic sectors. A data space is an advanced architecture for the data economy (level 4 or 5 of the maturity scale) that can already be developed in low-maturity DVC with a small group of digitally mature players — especially when public authorities are supporting the initiative.

In other words, simultaneous efforts can be undertaken by different players at different level of the maturity scale. Consequently, progression on the maturity scale is not purely linear but should be seen as a "balancing act" as illustrated on **Errore. L'origine riferimento non è stata trovata.**.

[21] Turning these steps into quantitative grades can be explored at a later stage. This would require to identify quantitative indicators that can be normalised and turned into an overall grade.
[22] Use Case A - Value-chain Specific Data Exchange - Direct Between Several Companies
[23] Multilateral Data Exchange – Using a common Data Exchange Framework / Data Space
[24] Cross-Sectorial Interaction – Data Exchange between companies in two different Data Exchange Frameworks / Data Spaces

Figure 8: Summary of the different steps and components of the digital continuity framework (source: author's elaborations)



Each stage of the maturity scale is briefly defined on the above figure. In addition, a **score expressed in percentages** is given above each level. The idea of this grade is for the analyst to provide a summary of its analysis: what level of digital maturity is achieved at each stage of the maturity scale. The percentage figure can be given based on two sorts of inputs that can be either combined or exclusive based on whether the analyst wants to focus at the VC level or at the sectorial level:

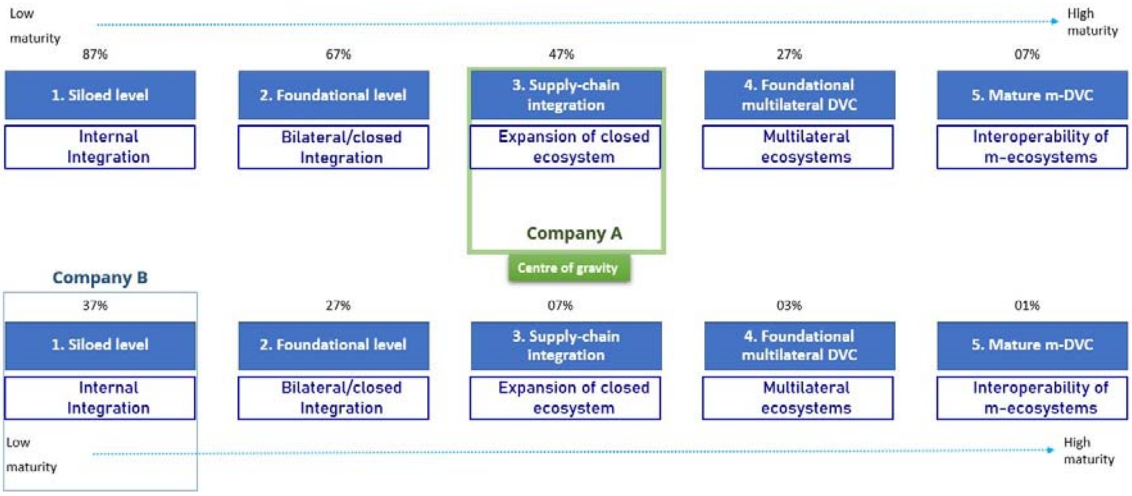
1. What percentage of companies in a sector have reached a level (between 1 and 5) of maturity;
2. That a certain percentage of companies in a specific VC have reached level 1 to 5

As the analysis will be carried at the DVC level, the score is mostly a qualitative assessment based on the indicators that are provided in **Table 1** further below.

As already stated, efforts can be engaged at the different stages of the maturity scale simultaneously. This is why it is possible to have positive scores at the most advance stages of the scale even though the first level did not reach 100% yet. This is why this paper recommends the adoption of the concept of "**Centre of Gravity" (**as illustrated on Figure 9**)**.

The centre of gravity clarifies the situation of the DVC by identifying where are the main barriers that remain to be addressed for the DVC to reach a higher level of digital maturity.

Figure 9: Mock example of the results from applying the proximity framework, with the addition of the "centre of gravity" concept (source: author's elaborations)



The **Business Continuity Index** introduced at the beginning of this paper thus involves two elements:

1. A **digital maturity score** given for each level of the maturity scale;
2. The definition of a **Centre of Gravity** for each industrial value chains.

Now that the overall logic and structure of the maturity scale has been clarified and detailed, **indicators must be defined** (a preliminary list of these indicators is provided in **Table 1**.) to clarify how digital maturity will be measured to both provide the digital maturity score and locate the centre of gravity.

The structure of the framework is designed to provide a comprehensive view of the level of digital maturity achieved by a DVC. Indicators are organised in **3 key dimensions/categories**, each divided in **2 sub-criteria.** [25]

1) **Common Vocabulary Standards (CVS**) — standards for exploitability of data:
    a. semantic (meaning of data)[26]
    b. syntactic (format of data)[27].
2) **Relational Contractual Agreement (RCA) —** legal and business framework for the exchangeability of data:
    a. legal and governmental framework (laws, rules, legislations);
    b. business and ecosystem (presence of strategies, roadmap, informal rules and even customary rights for data sharing practice.)
3) **Federated digital infrastructure (FDI)** — infrastructure for data exchange:
    a. Interfaces and platforms (interoperability of platforms and systems, e.g., cross-platforms exchangeability);

[25] This overall framework was first designed in CARSA et all, 02/12/2021, *Study on technological and economic analysis of industry agreements in current and future digital value chains*, https://op.europa.eu/en/publication-detail/-/publication/8c021023-53ee-11ec-91ac-01aa75ed71a1/language-en
[26] Semantic stands for the meaning of any symbol, in particular sentences, parts of sentences, words or parts of words. In the case of DVC maturity, this involves defining common format, names, ranges, etc.
[27] Syntactic is the rule system for combining elementary symbols into composite symbols in natural or artificial symbol systems. The assembly rules of the syntactic are opposed to the interpretation rules of the semantics. In the case of DVC, this involves data format etc.

b. Infrastructure for multilateral data exchange (e.g., availability of data centres to warehouse information on European soil as legally required).

The 8 different factors for the maturity assessment are put together on **Figure 10** to provide a simplified view of the results that the framework could provide to the analyst.

Figure 10: mock-up illustration of the final result provided by the framework (source: author's elaborations)

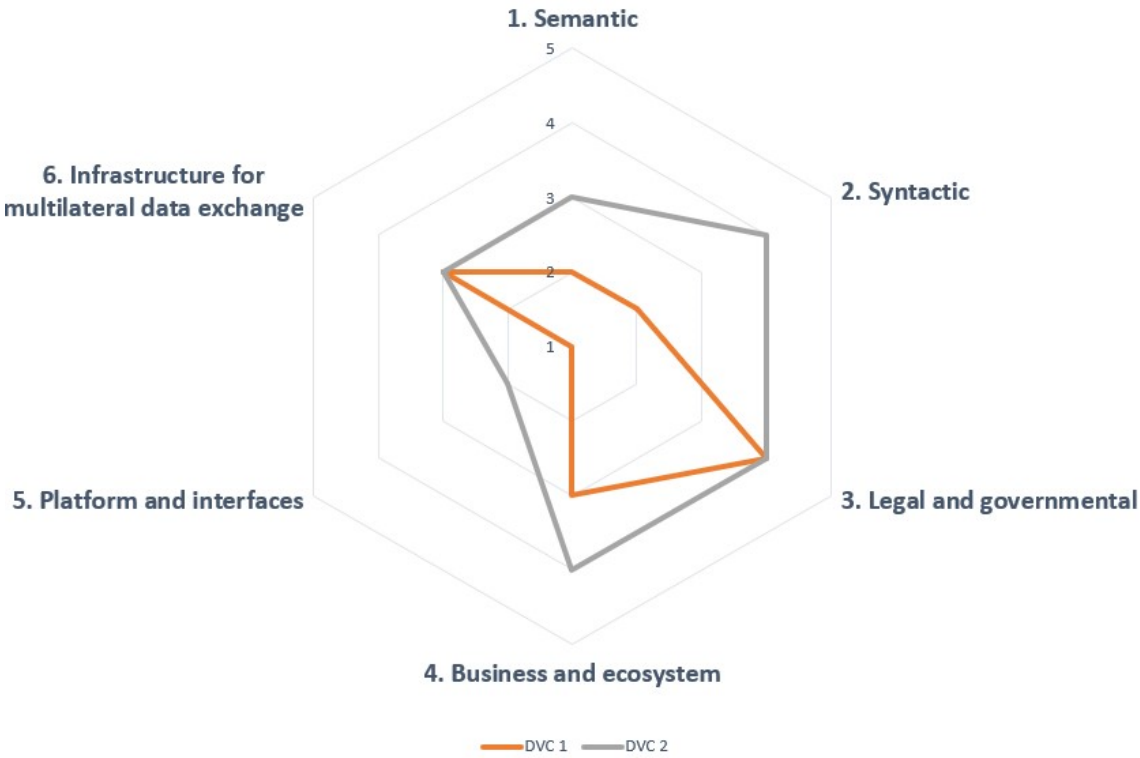


Table 1: Detailed indicators and variables for the interoperability assessment framework

Type	Criteria	1) Siloed Level	2) Foundational Level	3) Supply-chain integration	4) Foundational m-DVC	5) Mature m-DVC
CVS	A. Semantic	• Siloed industrial and business activities. • Lack of internal data exchangeability. • Lack of standardisation (need for new standard and lack of uptake).	• Creation of new standards leading to new form of complexity and incompatibility between standards. • Development of proprietary standards for internal exchanges.	• Semantic standardisation of basic function with moderate adoption rate. •Lack of standard for "niche function".	• Large level of adoption of semantic standards. • Remaining gaps for niche languages and advanced functions.	• Extension of semantic standardisation to most devices and systems. • Semantic standardisation with large uptake and implementation.
	B. Syntactic	• No sectorial data quality standards for exchange and sharing. • IoT and connected devices show an inconsistent and limited level of data quality. • No solutions for legacy systems.	• Development of proprietary solutions and standards to improve data quality. • Fragmentation of the data quality environment (lack of certification, monitoring, …). • In-house solutions for integration of legacy-system.	• Emergence of quality framework assessment in localised part of the value-chain and industrial activities. • Low adoption rate. • Advanced solution for legacy systems.	• Moderate adoption rate. • Data quality standardisation covers most functions and is no longer localised. • Inconsistencies and incompatibility of standards between devices and functions.	• High adoption rate. • Uptake of DVC-wide data quality assessment framework. • Accepted certifications schemes. • Cross-devices and systems compatibilities of standards.
RCA	A. Legal and governmental	• Only general legal principles, either at national or EU level. • No sector-specific rules for sectorial data exchanges.	• Sectorial initiatives for self-regulation are still missing. • Emergence of code	• Mature and digitalised B2G data exchanges.	• Legal framework to surround B2B data exchange. • Customary law	• Full harmonisation of European legal sectorial frameworks.

		• No national initiatives or solutions for data warehousing and storage. • European data space initiative in the sector is at its inception stage or did not lead to fruition.	of conducts and general principles for the sector. • Regulations for B2G data exchange. • Infrastructure, solutions and legal agreements for data warehousing. • Specific bodies are created by EU and/or national institutions to regulate the ecosystem.	• Lack of framework for B2B exchanges. • Code of conducts and general principles specific for the sector are widespread and applied (birth of customary law). • Some effort of sectorial self-regulation appears for dedicated initiatives.	starts to turn into self-regulation and common practices. • Legal framework for cross-borders data exchange is well clarified and enabled at the European level. • Sectorial legal challenges are well mapped but are not all addressed.	• Most legal barriers to data exchange are removed. • Self-regulation becomes the norm with trust anchor and some degree of monitoring (e.g., monitoring bodies in data space initiatives).
	B. Ecosystem and business culture	• Lack of exchange culture. • Most companies do not have a data strategy nor roadmap for data-sharing. • Lack of best practices and monitoring at the internal level for data exchange.	• Cultural shift starts to be observed in the sector. • Leading companies have developed roadmaps and strategies. • Best practices start to emerge.	• A sectorial culture of data exchange starts to appear but is limited in number and outreach. • SMEs starts to develop strategies and roadmaps for data sharing. • Replication of best practices starts to take place.	• Institutionalisation of sectorial culture and sharing customs. Sharing culture is widespread. Forums and arenas for discussions. • A large number of SMEs (between 30 and 50%) are embarked in the sharing ecosystem. • Best practices are widespread with cross-fertilisation.	• Structured culture with arenas for discussion, exchanges and exclusion. • More than 50% of SMEs are involved in the sharing ecosystem. • The sector shines as best practice for other industries.

FDI	A. Platform and interfaces	• Incompatible interfaces and platforms within a same company. • Need to use different platforms and interfaces for different functions or use similar equipment. • Low uptake of analytics.	• Internal data fusion environment begins to take place. • Internal and proprietary solutions for interoperable platforms and interfaces. • Internal analytical solutions start to be implemented for specific use (but remain limited by data silos and lack of data fusion).	• Internal standardisation and compatibility of platforms. • Almost full integration of internal ecosystem in data fusion ecosystem. • Competing small alliances for uptake and development of standardised platforms and interfaces. • Efforts to address data continuity for advanced analytics.	• Low level of alliance’s platforms interoperability. • New analytics IT and software companies mainly work in close bilateral relationships. • Moderate level of digital continuity for advanced analytics.	• Cross-platform interoperability. • Digitalisation of sectorial activities and digital continuity in industrial processes. • Advanced analytics leveraging internal and external data.
	B. Infrastructure for multilateral data	• Lack of solution for internal connectivity. Internal data silos remain the norm in the sector. • Pen & Papers remains the norm even though digitisation has started. • Lack of exchange protocols for internal industrial processes. • No sectorial exchange infrastructure. No protocols for sectorial data exchanges.	• Existing but incompatible standardised protocols for data exchanges between machines and functions. • Moderate digitisation of internal processes. • Single-access points for specific data sets at the sectorial level.	• Fragmented standard landscape for exchange protocols but most needs are addressed. • Several single access points have appeared (no interconnectivity or repository of access points). • High level of digitisation,	• Data spaces gain in membership and impact. • New multilateral exchange infrastructure and mechanisms are common practice. • High-level of digitalisation. • B2G data exchange is almost fully digitalised. •	• Interconnectivity of single-access points. • Increased scope of data spaces, several data spaces in the sector are interconnected. • B2G and B2B processes are mostly digitised and digitalised. • Easy access to

			• Bilateral exchanges remain the norm for most data exchanges. • Data spaces with limited membership. • No inclusiveness for SMEs.	digitalisation of processes moderately advanced. • Birth of multilateral exchange infrastructure. • SMEs involved in asymmetric relationships for data exchange (e.g., supplier-customer). • Digital infrastructure for B2G data exchange.	B2G data exchange is almost fully digitalised. • Higher involvement of SMEs through openness of multilateral alliances and data spaces.	exchange infrastructure for SMEs. • Almost full digitalisation of business activities.

3.3. Gap analysis

The application of the DVC framework and the identification of the BCI is the first step to decide whether to engage or not in multilateral data exchanges. Especially, some key factors are decisive in deciding whether or not collaboration is possible (e.g., legal barriers, lack of infrastructure for multilateral data exchange). However, in many cases, the framework will identify more mitigated situations where more explorations are necessary to identify next steps and to decide whether or not more efforts should be invested.

To do this, the **second step of the suggested methodology** is to apply a **gap analysis**. Building on the results from the maturity assessment, the gap analysis aims at circumventing the areas where more clarity/effort is needed. To say it more clearly, if a DVC highlights a lack of standardisation of data formats, it must be clarified if this barrier is a definitive one, if it can be addressed and if yes, at what cost.

At this stage, the company will need to leave the DVC approach to engage in more specific analysis of both its own resources and the exact type of organisations it wants to engage with. The company should run internal consultation of key stakeholders to identify key pain points and clarify the business case/opportunities.[28] The gap analysis should be driven by clearly defined business objectives to identify the gap between the current situation and the expected results.[29] Furthermore, the gap analysis is necessary as an intermediary stage to enable the identification and development of **remediation measures** (see section 4.4).

Table 2 below provides an overview of the different elements that have to be covered in a gap analysis.

Table 2: Improvement of existing pathways, general layout and methodology

Layer	Sub-criteria	Description
CVS	**1) Semantic**	**Exchangeability diagnostic**. Assessing if the data are exchangeable between players in the DVC.
	2) Syntactic	**Readability and meaningfulness diagnostic**. Assess data quality requirements, data flaws and assessment frameworks. This involves lack of data standards uptake internally or in the sector; lack of data quality
RCA	**1) Legal & governmental;**	**Legal diagnostic.** Non-technical elements barriers related to legislation and governmental actions.
	2) Business and ecosystem	**Cultural diagnostic.** Identification, localisation and definition of cultural resistance to sharing culture
FDI	**1) Platforms and interfaces**	**IT systems diagnostic**. Identifications of the internal and external limitations to engage in data sharing.; lack of sectorial or internal interoperability of systems and interfaces; integration of legacy systems and their data outputs.
	2) Exchange Architecture for	**Architecture diagnostic**. Is the infrastructure to engage in multilateral data exchanges available or needs to be developed? E.g., data warehousing service providers located in Europe to host data.

[28] *"During a gap analysis you should consult end users to understand the pain points and issues they have with the current data and applications. For example, some data may not conform to existing business rules and, may have some quality issues " Stakeholders that should be involved in such an approach includes project sponsor, business and data analysis, data/information architect, data consumers, security specialist, subject-matter experts, IT developers, data specialist, information governance manager, data champion, chief information governance office and data manager/officer*." National Archives of Australia, January 2019, *Interoperability development phases resources*, https://www.naa.gov.au/sites/default/files/2019-09/interoperability-development-phases-resource.pdf

[29] National Archives of Australia, January 2019, *Interoperability development phases resources*, https://www.naa.gov.au/sites/default/files/2019-09/interoperability-development-phases-resource.pdf

	multilateral exchanges	

3.4. Improvement of the existing pathways (improvement framework)

The gap analysis has only been briefly introduced in the previous section. This is because the identification of remediation measures to **improve existing pathways** is closely intertwined with the gap analysis. While both stages require to be analytically separated, they should be jointly conducted during the analysis to have a clear view on business case, opportunities and costs.

Overall, 4 steps are necessary to improve systems to engage in multilateral exchanges:

1. **Definition of the business case.** What are the objectives or use cases driving the efforts to engage in multilateral data exchange/sharing.
1. **Diagnostics and integration needs**. What has to be changed to achieve the business case? Stakeholders need to identify what are the needs to achieve business proximity (how to integrate digital and business functions, what solution exists and what the fitness of existing solutions and systems to achieve the business case);
2. **Costs-benefits assessment**. Based in the integration needs and diagnostic, what are the expected costs and benefits? Can existing assets be reused, do they require upgrade or to be eliminated?
3. **Refinement and finalisation of the business case and feasibility assessment**. Can the original business plan be adapted to address the barriers and difficulties identified in the diagnostics and lower costs? For instance, can the business case be adapted to meet the legal requirements as defined in the data act and other important European legislations?

The use case/business case will widely vary based on the industry, the sector and the objective of the stakeholders involved in the sharing initiative. Moreover, cost-benefits assessment is a well-known step for any type of business initiative. Consequently, both these stages can be left aside for the purpose of this paper, even-though they are full-fledged and crucial stages in the improvement of pathways and assets.

Table 3 below provides more details, guidelines and recommendations on how to run the **different types of diagnostics needed** to assess business proximity and engage in multilateral exchanges. Building on the framework to assess current level of maturity, these diagnostics are aimed at providing guidance to companies willing to address the identified gaps and challenges.

Table 3: Improvement of existing pathways, general layout and methodology

Diagnostic	Remediation measures
1) Exchangeability diagnostic.	• Alignment with business partners: definition of common standards to be adopted based on the business case. • Engaging in the development and implementation of semantic transformation models (interfaces between internal proprietary and common standards). • Implementation and engagement in sectorial standardisation efforts for long-term objectives (not directly related to the business case).
2) Readability and meaningfulness diagnostic.	• Engaging in the development of syntactic transformation model. • Adoption of common data quality frameworks to ensure adequate level of data quality based on the business case.
1) Legal diagnostic.	• In the short and medium term, adaptation of the business case to meet legal requirements. • In case the business case is of strategic importance in the long-term, engagement in advocacy groups and lobbying activities.
2) Cultural diagnostic.	• For internal barriers (e.g., defiance toward the data economy or lack of adapted skills), upskilling, training and awareness-raising activities. • Definition of governance structure and framework to engage in data sharing communities. • For external barriers (culture of the industry, etc), engagement in long-term strategies (joining advocacy groups, trade associations, European alliances, etc.). Results are however less certain and only likely to materialise in the long-term.
1) IT systems diagnostic	• Increasing internal data continuity across functions, departments and activities (e.g., AAS, digital threading, etc). • Development of integrated internal data-hubs for data fusion. Silo integration of legacy machines and systems to facilitate internal use and external sharing of data. • Depending of the industry, development of alliances for standardisation of platforms and IT systems (e.g., health institutions need to standardise digital process, digitalise and standardise electronic health records and harmonise digital platforms as enabling factors to engage in the data economy). • Development and implementation of data quality framework.
2) Architecture diagnostic.	• Leveraging existing European initiatives (e.g., French blue initiative, GAIA-X or IDSA) to develop a common exchange architecture. • Adoption of the IDSA marketplace architecture can reduce the reliance on external data servers.