RE4DY

MANUFACTURING DATA NETWORKS

Title	D2.3: Digital 4.0 continuum reference Framework final version
Document Owners	ENG
Contributors	ENG, CHAL, KUL, ICF, FILL, GF, UNINOVA, ATL, CERTH
Dissemination	PUBLIC
Date	30/8/2024
Version	V01

Table of contents

Index of figures

Index of tables

Document Status

Leader	ENG
Internal Reviewer 1	POLIMI
Internal Reviewer 2	S21SEC
Work Package	WP2 - Digital 4.0 Continuum Reference Architecture for Active Resiliency
Deliverable	D2.3: Digital 4.0 continuum reference Framework Final version
Due Date	M24
Delivery Date	31/05/2024
Version	V01

Version History

02/02/2024	Deliverable structure, ToC, first content drafts
19/04/2024	Content updates
10/05/2024	Structure and content updates
28/06/2024	Version for peer-review
28/08/2024	Final version

Further Information

More information about the project can be found on project website: https://re4dy.eu/

Disclaimer

Project Partners

Num	Participant organisation name	Acronym
1	ASOCIACIÓN DE EMPRESAS TECNOLÓGICAS INNOVALIA	INNO
2	CHALMERS TEKNISKA HOGSKOLA AB	Chalmers
3	INTERNATIONAL DATA SPACES EV	IDSA
4	VOLKSWAGEN AUTOEUROPA, LDA	VWAE
5	ASSECO CEIT AS	CEIT
6	UNINOVA-INSTITUTO DE DESENVOLVIMENTO DE NOVAS TECNOLOGIAS-ASSOSIACAO	UNI
7	FILL GESELLSCHAFT MBH	FILL
8	AVL LIST GMBH	AVL
9	VISUAL COMPONENTS OY	VIS
10	UNIVERSIDAD MIGUEL HERNANDEZ DE ELCHE	UMH
11	ATLANTIS ENGINEERING AE	ATLANTIS
12	DATAPIXEL SL	DATA
13	CORE KENTRO KAINOTOMIAS AMKE	CORE
14	UNIVERSITETE I OSLO	UiO
15	GE AVIO	AVIO
16	ENGINEERING-INGENIERIA INFORMATICA SPA	ENG
17	POLITECNICO DI MILANO	POLIMI
18	ATOS IT SOLUTIONS AND SERVICES IBERIA SL	ATOS IT
18.1	ATOS SPAIN SA	ATOS ES
19	KATHOLIEKE UNIVERSITEIT LEUVEN	KU
20	NETCOMPANY-INTRASOFT SA	INTRA
21	NOVA ID FCT - ASSOCIACAO PARA A INOVACAO E DESENVOLVIMENTO DA FCT	NOVA
22	INDUSTRY COMMONS FOUNDATION (INSAMLINGSSTIFTELSE)	ICF
23	ETHNIKO KENTRO EREVNAS KAI TECHNOLOGIKIS ANAPTYXIS	CERTH
24	GRUPO S 21SEC GESTION SA	S21SEC
25	UNIVERSITAT POLITECNICA DE VALENCIA	UPV
26	CONSIGLIO NAZIONALE DELLE RICERCHE	CNR
27	SOCIEDAD ANDALUZA PARA EL DESARROLLO DE LAS TELECOMUNICACIONES SA	SANDETEL
28	SWITZERLAND INNOVATION PARK BIEL/BIENNE AG	SSF
29	GF MACHINING SOLUTIONS AG	GFMS ADVMAN
30	FRAISA SA	Fraisa SA
31	SIEMENS SCHWEIZ AG	SIE

List of Acronyms/Abbreviations

Acronym/Abbreviation	Description
AI	Artificial Intelligence
BDA	Big Data Application (of the ALIDA framework)
BP	Business Process
BS	Business Scenario
CAD	Computer-Aided Design
CAM	Computer-Aided Manufacturing
DaaP	Data as a Product
ERP	Enterprise Resource Planning
ETL	Extract, Transform, Load
FL	Federated Learning
FML	Federated Machine Learning
GDPR	General Data Protection Regulation
IdM	Identity Management
IoT	Internet of Things
IP	Intellectual Property
IPRO	Intellectual Property Rights Ontology
KPI	Key Performance Indicator
OML	Ontology Markup Language
PLM	Product Lifecycle Management
RA	Reference Architecture
RDF	Resource Description Framework
ROS	Robot Operating System
SHOE	Simple HTML Ontology Extension
TEF	Testing and Experimentation Facility
WP	Work Package
XAI	Explainable Artificial Intelligence

Executive Summary

This document presents the final version of the Digital 4.0 Continuum Reference Framework, comprising three main components:

- The Resiliency Model, initially described in D2.1 and now completed and consolidated.
- The Legal Framework introduced here and to be further detailed in the final deliverable, D2.4, at the conclusion of this work package.
- The Reference Architecture, detailed in D2.2 and now illustrated in terms of its various adoptions in the project pilots.

Each component plays a crucial role in the overall framework.

The Resilience Framework was developed using the IDEF0 functional modelling approach to visualize relationships between various resilience elements, with dynamic capabilities theory forming the basis for the mechanisms. The framework led to the creation of a resilience dashboard comprising a resilience compass and radar, aiding companies in assessing and developing resilience capabilities. The resilience compass allows companies to self-assess their current and future resilience capabilities across three stages (anticipation, coping, adaptation) and 11 dynamic capabilities using a Likert scale. The resilience radar helps companies detect and map risks impacting their organization and supply chain. A dashboard combines these assessments for visualizing capability development and deployment. Additionally, the study explored the feasibility of AI to enhance manufacturing resilience through data engineering, simulations, and a directed acyclic graph model. Lastly, the pilots, experimenting this framework, assessed the sustainability implications of resilience capabilities on environmental and social dimensions, revealing generally positive impacts, except for redundancy capability, which showed mixed results.

The Legal Framework outlines key aspects and implications of EU data space and data sharing legislation, focusing on various industrial sectors including aeronautical, machine tool, battery, and automotive manufacturing. It provides an overview of the relevant EU-wide legal framework pertinent to data governance and control, which is essential for the Digital 4.0 Continuum. The discussion begins with the concept of "data as a product," examining whether data can be legally classified as a product under the proposed revisions to the Product Liability Directive (PLD). While the new PLD expands the definition of products to include intangible items like digital manufacturing files, it explicitly excludes general digital data, highlighting the nuanced legal treatment of different types of digital content. Next, this framework delves into the implications of the Data Act (DA), which aims to ensure fair value distribution and foster data access and usage in the data economy. Key provisions include data access and portability rights, obligations for data holders, and the establishment of data sharing frameworks. The DA emphasizes the need for transparency and fairness in data sharing contracts, particularly concerning connected products and related services. The Data Governance Act (DGA) is also examined, focusing on its impact on data intermediation service providers and the obligations it imposes to facilitate trustworthy data sharing environments. The DGA aims to enhance the availability and management of data, ensuring secure and efficient data exchange. The intersection of data and intellectual property rights, particularly in the context of digital twins used in

various manufacturing sectors, is then analysed. It highlights the legal challenges and considerations for protecting intellectual property in a digitalized industrial landscape.

This document also reports on the setup and results of the Extended Task Force on Innovation and Standardization. This task force has facilitated workshops and collaborations resulting in various innovations, including legal ontologies, dynamic decision interfaces, and improved asset management strategies. These efforts aim to enhance efficiency and innovation within manufacturing supply chains, with alignment efforts extending to US Associate Partners.

The development of resilience and legal ontologies is also discussed, focusing on ontology engineering and semantic modelling to achieve semantic interoperability. It explores the creation of ontologies, semantic models, and their applications in data integration and knowledge representation. Various standards and languages for semantic web technologies, such as RDF, OWL, and RuleML, are also discussed. Finally, it outlines the initial development of a manufacturing resilience ontology.

The IP Rights Ontology (IPRO) aims to enhance asset management and legal certainty in digital value chains, covering six main domains: IPR recognition, criteria, data requirements, open license families, processes, and actor roles. Developed using OWL 2 Web Ontology Language, it will integrate into asset management software and be freely accessible online. Overall, it is emphasized the evolving legal landscape surrounding data as both an asset and a product, underscoring the importance of adapting to new regulations to maintain compliance and leverage data effectively in the industrial sector.

The Reference Architecture is a data-driven architecture designed for the RE4DY project that incorporates unique requirements, aspirations, and existing foundations, drawing on existing blueprints in literature and key results from previous projects. This framework consists of four layers and four vertical dimensions, enhanced by the computing networking continuum, to promote decentralization and digital continuity. Detailed in D2.2, the architecture's capabilities and methodology align with the project's goals, facilitating the convergence between Manufacturing and IT operations through comprehensive toolkits. This architecture illustrates how its building blocks align with toolkit components, providing an overview of the implementation status and the extent of coverage in implementing various building blocks. As new iterations occurred, feedback from the pilots and new implementation continued to refine and fulfil the RA building blocks. This deliverable shows practical examples of RA implementation through different architectures designed in the project pilots that align with the four specific business cases of the RA Business Layer.

1 Introduction

1.1 Context and scope of this document

The main objective of this document is to describe the final version of the Digital 4.0 Continuum Reference Framework and detail all the components that comprise it. This comprehensive framework encompasses various elements, methodologies, a reference architecture (RA), tools, and strategies designed to enhance the integration and application of digital technologies across different industrial sectors. In more detail, the document is structured as follows:

§1 Introduction:	It is the introduction to the document, meant to provide the reader a guided tour of the document sections.
§2 Resilience framework:	This section debates about the enhancements made to the Resiliency Framework, which was originally introduced in D2.1. It details the iterative improvements and refinements implemented throughout its development process.
§3 Legal Framework:	This section highlights the essential elements and implications of various EU legislative acts concerning data sharing. It also presents the establishment and outcomes of the Extended Task Force on Innovation and Standardization, which addressed legal ontologies, dynamic decision interfaces, and enhanced asset management strategies.
§4 RE4DY RA adoption:	This section focuses on the adoption of the Reference Architecture (RA), demonstrating how the pilots incorporated its building blocks and integrated them with customised implementations tailored to each pilot's specific needs.
§5 Conclusions:	This is the closing section that outlines the conclusions and next steps.
§6 References:	Section with the specific references to citations made in the document.

Within the document, particularly in "The Resilience Framework" and "RE4DY RA adoption" sections, references to the pilots are made due to their close relationship with these topics. These references are sometimes made by the pilot's progressive number, other times by their specific name or its owner, in others by the main RA business case to which they belong. To ensure clarity and convenience, a table is provided below that maps these business cases 1:1 with the corresponding pilots.

Table 1: Main RA business cases and pilots

Nr.	Main RA business case	Pilot name	Owner
1	Logistics of the future	Connected Resilient Logistics Design & Planning	VWAE
2	Megafactory & E-battery design	Electric Battery Product/Production System Engineering	AVL
3	Circular Machining	Integrated Machine Tool Performance Self - Optimisation	+GF+
4	Distributed, Green, Zero-X Manufacturing	Multi-plant Predictive ZDM Turbine Production	AVIO

1.2 Relationships among other deliverables

This deliverable consolidates and builds upon the previous two ones (D2.1 and D2.2), which laid the foundation for a framework capable of proposing not only technological solutions but also organizational, business, and legal ones. Regarding technological content, the document also references deliverables D3.1 and D3.2 for further details about the mentioned components. At the same time, it anticipates and sets the groundwork for the legal framework that will be completed in the final deliverable (D2.4) of this work package.

2 Resilience framework

The active resilience framework was developed using the IDEF0 (IEEE, 1998) functional modelling approach. The methodology for developing the framework with the various resilience elements has already been described in D2.1. The modelling approach was useful to visualize the relatedness between the different resilience elements. To summarize, the 'activity' is building manufacturing resilience, and the 'controls' or the *influencing* factors are risk management (that include risk identification, prioritization, frequency, severity and boundary of impact (Padhi, 2024)) and corresponding disruptions. The 'mechanisms' are those that *enable* the building of manufacturing resilience. In RE4DY case, the dynamic capabilities theory (Teece, 1994) was identified as an appropriate theoretical lens to develop the mechanisms for building manufacturing resilience (Figure 1).

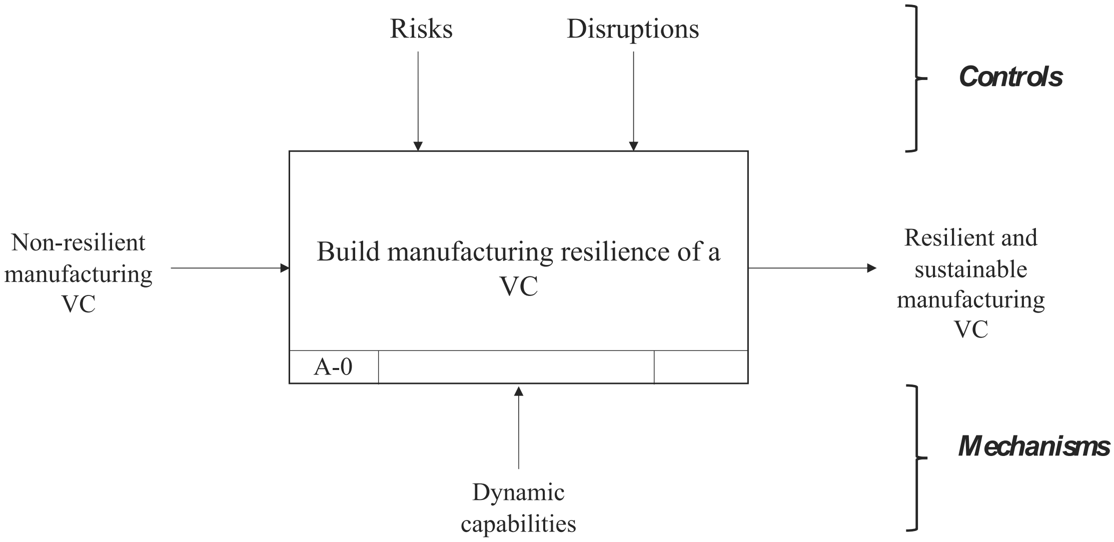

Figure 1: Active resilience framework that visualises the relationships between the different resilience elements identified (Chari, 2023)

The resilience framework formed the foundation for developing the resilience dashboard which comprised of a resilience compass (resilience capability development) and resilience radar (deployment of resilience capabilities to deal with risks and corresponding disruptions). The methodology for the dashboard design will be explained in the following section.

2.1 The resilience dashboard design

Based on the active resilience framework, the resilience dashboard was developed in two stages for measuring or assessing resilience in manufacturing organizations. In the first stage, a resilience compass was developed using a quantitative content validation index method (documented in a paper submitted to a journal) and in the second stage, a resilience radar was developed and applied in the four use cases of the project.

2.1.1 The resilience compass

The resilience compass was developed as a navigational aid to give the companies a sense of direction and assess where they are today (current state) and where they would like to be (future state) in terms of resilience capability development.

It was developed as a self-assessment tool in the form of a Microsoft excel document, which comprised of three temporal stages of resilience (anticipation, coping and adaptation), 11 dynamic capabilities (situation awareness, visibility, security, redundancy, agility, flexibility, collaboration, leadership, knowledge management, contingency planning and market position) and 54 resilience practice *Figure 2*.

Situation awareness was defined as 'the ability to sense and forecast a possible disruption through knowledge of organization/supply chain vulnerabilities, the sharing of information and corresponding activities', visibility as 'the acquisition and evaluation of information to enable transparency and awareness of the current supply chain situation, trace points of origin of entities, control disruption risks and improve decisions', security 'involves personnel security, physical security and cyber-security', redundancy 'involves maintaining excess capacity, safety stock, multiple suppliers and backup sites', agility is 'the ability to rapidly respond to unpredictable changes in demand or supply in the marketplace since customer requirements are continuously changing', flexibility is 'the ability to adapt and adjust to a disruption rather than merely withstand the damage of the disruption', collaboration is 'the exchange of information and the application of shared knowledge to decrease uncertainty', leadership is 'the execution of management in companies, which requires support from top management, engagement of employees and high-quality decision-making', knowledge management is 'the ability to learn from feedback from a disruption to develop better plans and solutions for future ones (education, training and innovation)', contingency planning 'involves supply chain reconfiguration, scenario analysis and resource reconfiguration to help organizations recover and learn from disruptions' and market position is 'related to the knowledge about financial perspectives, including financial strength, market share, cost efficiency and loss absorption'.

The resilience practices identified are listed in Table 2**.**

Table 2: Resilience practices used for the resilience capability implementation level assessment (resilience compass)

Resilience Phase	Resilience Capability	Code	Resilience Practice
Anticipation (Sense)	Situation awareness	SA1	We conduct regular risk assessments to be vigilant to risks that impact our organisation
		SA2	We conduct regular risk assessments to be vigilant to risks that impact our supply chain
		SA3	We check upcoming regulations by governmental organisations

		SA4	We check upcoming initiatives by non-governmental organisations (e.g. The UN, World Economic Forum, Ellen MacArthur Foundation, etc)
		SA5	We conduct knowledge acquisition activities to detect threats (e.g., market research, end-user surveys, use of gatekeepers, scenario planning etc)
		SA6	We conduct activities to prepare for unexpected events (e.g., emergency planning, business continuity management)
	Visibility	VI1	We are aware of 'where' disruptions occur in our organisation (disruptions are those that halt or change operations in a department, product life cycle stage, etc)
		VI2	We are aware of 'where' disruptions occur in our supply chain
		VI3	We digitally track where products are located in our operations
		VI4	We digitally track which processes have been carried out on products
		VI5	We understand what data to capture across the organisation's different functions
		VI6	We know how data is shared within the company, thus avoiding 'information silos'
	Security	SE1	We have cyber-security measures in place
		SE2	We place emphasis on the quality of the working environment for our employees (e.g. no gender discrimination, focus on mental and physical health, safety, employee benefits, etc)
		SE3	We restrict the access of data at different levels in the organisation
		SE4	We regularly conduct security audits in our organisation
	Redundancy	RE1	We have a safety stock of critical components
		RE2	We can accumulate a back-up inventory in case of emergencies
		RE3	We have a diverse supplier base (e.g. dual-sourcing, back-up suppliers, geographically dispersed suppliers, etc)
		RE4	We have a diverse customer base
		RE5	We geographically disperse our production capacity in different sites
Coping (Seize)	Agility	AG1	We spend less time to adapt to product changes (e.g. using special items from existing articles)
		AG2	We can quickly respond to disruptions without structural changes in the organisation
		AG3	We can reduce time to market
	Flexibility	FL1	We can quickly implement a wide range of changes within existing parameter configurations
		FL2	We have reshoring strategies in place to accommodate unexpected customer demands
		FL3	We can re-purpose our facilities to create alternative products in times of need
	Collaboration	CO1	The organisation works harmoniously with cross functional departments (e.g. for data sharing,

			knowledge transfer activities, collaborative tasks, etc)
		CO2	We create strategic alliances with other companies
		CO3	We have coopetition strategies in place (cooperating with a competitor to achieve a common goal)
		CO4	We share operational information externally with suppliers
		CO5	Logistics databases are integrated across the supply chain for autonomous planning
		CO6	We have technical infrastructure (digital platforms, etc) to enable collaboration between supply chain partners
		CO7	It is important for us to share knowledge with our customers
	Leadership	LE1	We have sound leadership support from motivated top-level management
		LE2	The management regularly conducts listening sessions/forums for employee feedback
		LE3	Leaders across the organization engage in scenario planning exercises
		LE4	We have sustainable logistics strategies in our organisation
		LE5	We effectively communicate within all levels of the organisation
		LE6	We have reward systems that create a safe environment to escalate and address issues
		LE7	We engage our staff in continuous improvement processes
Adaptation (Transform)	Knowledge management	KM1	We conduct knowledge empowerment training workshops for the up skilling/re-skilling of our employees
		KM2	We conduct multi-skill training of new employees to avoid quick turnover rates
		KM3	We act on previously generated knowledge (change management)
		KM4	We continuously learn after a disruption occurs
		KM5	We capture all relevant data needed to maintain our operations (e.g. system data or the transfer of tacit knowledge from operators to make it more implicit)
		KM6	We employ methods to help machines understand and process data (e.g. data mining, AI, natural language processing, etc)
	Contingency planning	CP1	We have scenario planning practices to think of different futures (e.g. order books that can be applied in different industries, etc)
		CP2	We stress test our system with disruptions to identify system configurations that result in lowest degradation and fastest recovery
		CP3	We have supply chain integration strategies (e.g. develop common infrastructure solutions, create end-end connection with suppliers for combined decision making, knowledge creation)

		CP4	We design our production so that it can cope with different unpredictable events
	Market position	MP1	We have knowledge on financial health across the end-to-end supply chain
		MP2	We have quality-based performance measures in terms of costs related to product quality (after delivery)
		MP3	We have quality-based performance measures in terms of costs related to internal failure costs (before delivery)

In the first step of the resilience assessment process where the resilience compass was implemented, workshops at each of the pilots helped in ascertaining the levels of implementation in the different capabilities under the three stages. Accordingly, these were rated on a Likert scale from 0-5, 5 being the highest level of implementation: 0-Not ready to implement, 1-Never/does not exist, 2-Sometimes/to some extent, 3-Frequently/partly exists, 4-Mostly/often exists, 5-Always/definitely exists. to assess their current level of implementation of these practices in their organization. The exercise was repeated for the future state as well.

The highest value of 5 was not required in all the capabilities to maintain resilience. This varies from company to company based on their own objectives for being resilient or the number of resources they are willing to invest in to improve their resilience level from their current to future state. The companies also gave a high/low rating for the short-term and long-term value of implementing such a practice. This was done to give companies the awareness of aligning the value of current/future implementation strategies.

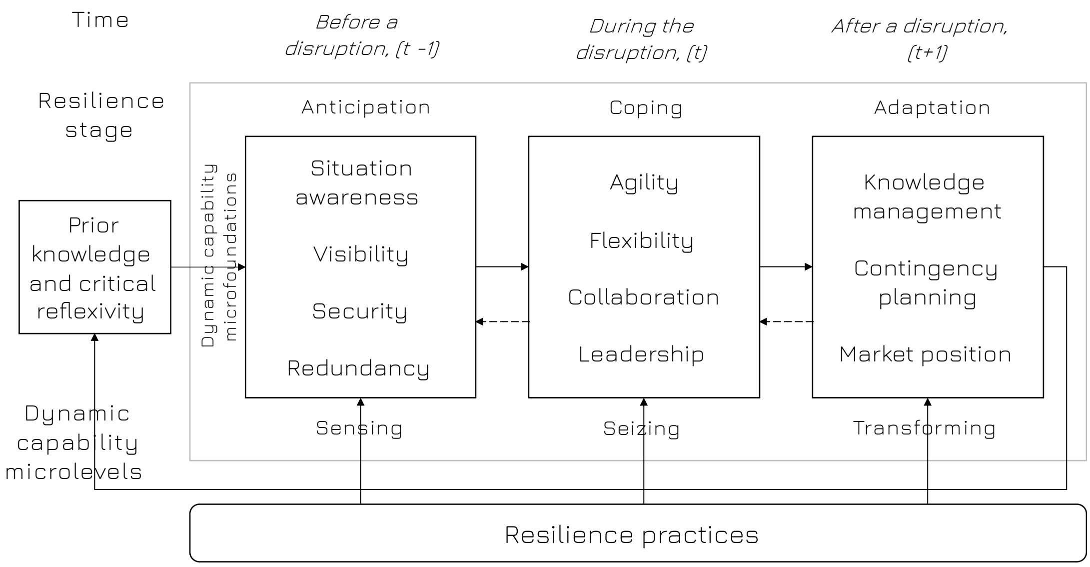

Figure 2: Relationship between the resilience stages, capabilities and practices

The compass has been shown in *Figure 3*. It consists of the three coloured resilience stages: the anticipation stage in blue, the coping stage in yellow and the adaptation stage in green. Each of the capabilities under these stages consisted of dark and shaded bands which represent the current state.

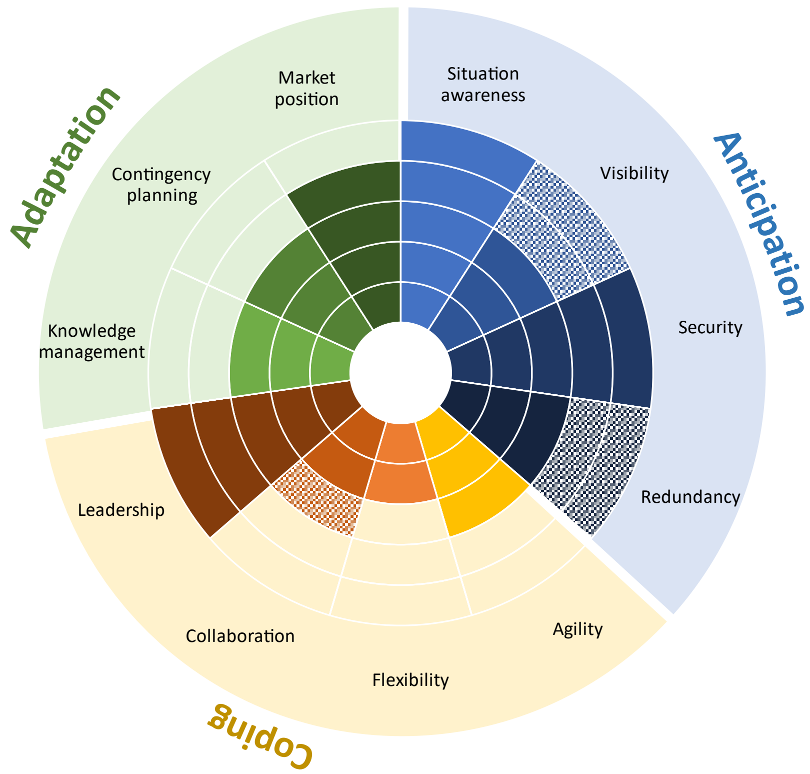

Figure 3: The resilience compass

An aggregate of the practices under each capability was represented by the dark bands and the highest value chosen for a practice was represented by the shaded bands. The dots in the figure represent the future state. If the difference in the aggregate value in the current state was >1 in relation to the aggregate value in the future state, then this was represented by a red dot. If the difference was <0.5, this was represented by a green dot and if the difference in the values was between 0.5 and 1, then this was represented by an orange dot. Donut graphs were used to create the compass.

Data was collected from each of the pilots in the form of workshops and focused interviews. During the workshop, the pilots were clearly provided with information about how the compass elements worked to do the self-assessments. Then, the interviews helped in understanding how the pilots would use the results from the assessments and the value derived from using such an assessment tool in their organizations.

2.1.2 The resilience radar

Organizations that want to be resilient should employ effective risk management processes (DuHadway S., 2017) that can help assess risks that impact them. *Figure 4* shows a time t-1, that is, a 'pre-disruptive' phase when risks are observed. These risks can then give rise to a disruption at time t, after which organizations need to recover in the 'post-disruption' phase.

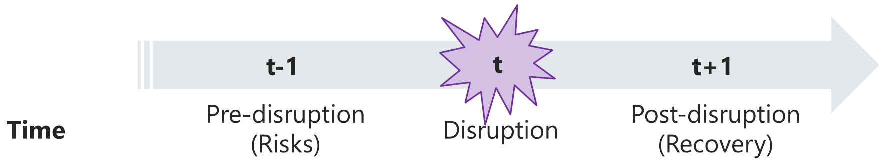

Figure 4: Risks and disruptions

The resilience radar was developed as a second step in the resilience assessment process. It can help companies detect risks that impact not only their organization but also those that impact their supply chain. A graphic visualization was created in the form of a diagram where the organization was imagined to be operating at the center of a 'radar' and which was prone to different risks (*Figure 4*). The risk or radar assessment was carried out for the current state, i.e., how the companies were currently dealing with the different risks that impact their organizations.

The pilots were then asked to list the different risks that impact their organization based on their frequency (how often they occur), severity (how critical the impacts are), the risk category (internal of within the organization, external to the organization or in the supply chain and environmental or outside the supply chain), risk type (Padhi, 2024) (operational, policy/regulation, financial, operational, demand, supply, system, behavioural and cybersecurity/safety). Then, the pilots were asked to map the risks according to the proximity from the organization (*Figure 5*).

If a risk lay close to the organisation (in the coping stage), it could mean two things: that the organisation was either firefighting such risks on a regular basis or that such risks occurred at a lower frequency and had less severe impacts to warrant the number of resources currently being used in the company to deal with such risks. Risks that were the furthest away meant that the company had enough anticipatory capabilities to proactively mitigate such a risk before it became a disruption. Risks that were mapped in the middle (adaptation) region occurred before and the organization learnt from corresponding disruptions so that they can be prepared when they occur again.

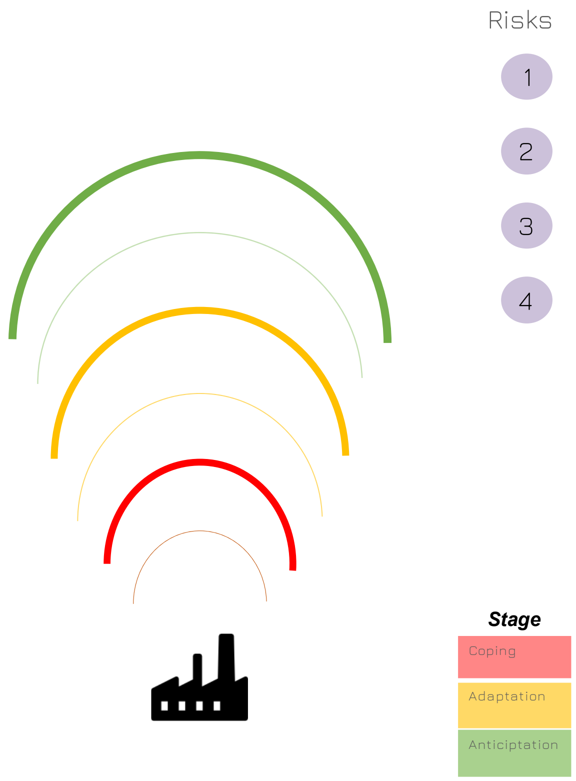

Figure 5: The resilience radar

2.2 Implementation in the "Logistics for the future" business case

This pilot consists of one company: VWAE.

2.2.1 The resilience compass assessment in Pilot 1

With respect to the resilience practices depicted in Table 2, Pilot 1 described several reasons for choosing the different implementation levels for the resilience practices. For SA1 and SA2 the level of 5 was relevant as each direct supplier to the pilot is assessed on a regular basis. SA4 was not applicable for the current local context of the pilot and was not rated. In general, the logistics scenario planning is extensively performed mainly to accommodate fluctuations in production/client demand or occasional volume fluctuations and hence a level of 3 was chosen for SA5. The same level was chosen for SA6: Emergency planning activities take place on an ad hoc basis depending on the events taking place. Some examples of the events that trigger these activities are truck driver strike, infrastructural accidents at supplier, quality issues on a part, etc.

VI1 had a level of 5 because in the automotive industry the product life cycle is defined very early where all contingencies to change/adapt to the new product are performed. VI5 had a level of 2 in the current state because the logistics systems track a whole range of data sets and for their daily activities only the most relevant data for internal processes are scrutinized. VI6 (with a current level of 2) is a topic that has mustered high levels of attention in recent times. The organization is making a significant effort to have logistics data available to every key stakeholder of the process.

In terms of security, SE2 had a current level of 3. Pilot 1 mentioned that there are internal strategies to keep the well-being of their employees at satisfactory levels. For example, "All about people" is a forum/initiative managed by high hierarchy staff to discuss and act on employees' wellbeing, benefits, etc. RE3 is a topic that the pilot has been keen to improve in recent years. In summary, there is a huge portfolio of suppliers available to each plant, however, it is not obvious that these strategies work as intended i.e.: fast reactions times without significant disruptions on the supply chain. Hence an implementation level of 2 was chosen for this practice. RE5 was not applicable to their industry. They mentioned in fact, that it is preferable to have the whole supply chain as close to the plant as possible. The same went with AG1 since the automotive industry is very rigid and any new product that requires assembly processes, must follow most of the procedures done in the previous product. FL2 was highly unlikely to happen. Usually when there are spikes in demand the main focus is to help the supplier to meet that demand by investing/adapting or optimizing its production process. Since the automotive industry is extremely rigid and complex, FL3 is highly unlikely to be carried out on a short to medium term.

The collaboration practices CO3 and CO7 were not applicable as such endeavours were planned in the past, however, for legal/compliance reasons fruitful results were never achieved. In addition, sharing knowledge with customers, at least at the plant level was not relevant (customer care is done at the central level). In general, transparency is key for the pilot, and suppliers have access to their ERP systems. Any collaboration that could potentially improve the logistics process is well received. Leadership practices were given a >4 rating. This is because the management and coordination teams promote initiatives that actively engage with staff providing a sound support for the daily business. The pilot has 'zero impact logistics as a part of their leadership initiatives, which aligned with the LE4 practice. In terms of LE6, all the staff perform compliance/anti-corruption/code of conduct assessments on a yearly basis. There is also the dissemination of the "Volkswagen Group Essentials" that encourages employees to speak up upon suspicion of foul play or upcoming bottlenecks.

For KM1, these initiatives are promoted usually in team building activities offsite. Additionally, the company takes big emphasis on training with a Training Academy (ATEC) planted right beside the premises where employees can ask or can be assigned to take courses to improve their skills. In addition, direct employees (working directly with the product) are not allowed to perform on the shopfloor area until extensive training. Indirect employees (office staff) have learning periods for handovers when assigned to new positions. The pilot mentioned that KM5 is a topic where it is reasonable to be sensible in giving the rating, as data capture might be more expensive process-wise (reading of labels for example). And capturing all relevant data is extraordinarily important when it comes to financial gains. KM6 activities are only in recent years being explored at VWAE. In general, the KM levels were rated <3 where the practices frequently existed.

CP activities were frequently implemented. For instance, preparation for new models and knowing their product life cycle means that preparation for different planning scenarios were very important to the pilot. In addition, stress tests are performed in key moments of the product life cycle, e.g., the launch phase and upon completion of major projects.

However, since the automotive industry is very rigid, it is not feasible to create options (as a back log) for unpredictable scenarios. There are, however, contingencies in place to deal with emergencies. For example, manual/emergency processes are planned with the internal supply chain of VWAE. With regards to MP2, product quality (finished product) is taken very seriously within the VW Group. There is an entire department responsible to keep quality high through reporting and various checkpoints for analysis. These quality checks are recurrent and if a part does not comply with standards it is rejected. All MP activities were rated at a level of 4.

2.2.2 The resilience radar assessment in pilot 1

The pilot described 14 risks (Table 3) that currently impact their organization and as described in Section 2.1.2, they were categorized based on their frequency, severity, boundary (where they occurred), type of risk and level of impact. The corresponding disruptions arising from such risks were also identified.

Table 3: Risks identified for pilot 1

No.	Risk	Freq.	Severity	Risk Category	Risk type	Level of impact
1	Suboptimal logistics configurations	Low	Medium	Industrial (Within SC)	Operational	Worker
2	Accident done with logistics equipment within plant premises	Low	High	Organisational (Within firm)	Operational	Machine
3	Suboptimal engineering configurations when releasing parts for production such as, technical configurations	Low	High	Organisational (Within firm)	System	Worker
4	Misinterpretation of line feeding or sequencing tasks caused by logistics service provider	Low	Low	Organisational (Within firm)	Operational	Supply Chain
5	IT issues (locally)	Low	Medium	Organisational (Within firm)	Cybersecurity and safety	Factory

6	Malfunction of automation devices	Low	Medium	Industrial (Within SC)	Operational	Machine
7	Accidents caused by human errors	Low	High	Environmental (Outside SC)	Disruption	Factory
8	Damage (supplier or plant premises) to infrastructure and/or material triggered by natural causes (supplier or plant premises).	Low	High	Environmental (Outside SC)	Disruption	Factory, supply chain
9	Raw material shortage	Low	High	Industrial (Within SC)	Demand	Supply Chain
10	Delays on external supply chain, or damages to parts/containers, specifically, with regards to incoming material moving by any means of transportation, earth, sea, or air.	High	Medium	Environmental (Outside SC)	Disruption	Supply Chain
11	Geo-Political issues between neighbouring countries or internal political policies	Medium	High	Environmental (Outside SC)	Policy/Regulation	Supply Chain
12	Technological advancements in core automotive technologies	Low	High	Industrial (Within SC)	Demand	Factory

13	Issues with parts quality invalidating bulk or big batches of parts	Medium	Medium	Industrial (Within SC)	Operational	Supply Chain
14	IT issues (headquarters)	Low	High	Organisational (Within firm)	Cybersecurity and safety	Factory

Most risks gave rise to production disturbances. Some examples given were human errors when sequencing parts to the assembly (these rarely occur), i.e., sequencing a part in the wrong order/position, or deliver the wrong container to assembly. These can cause small hourly disruptions in production cadencies. Disruptions caused by issues on systems at a local plant level IT architecture can cause local systems going offline due to data breaches or data integrity issues. It is then impossible to proceed with logistics internal supply chain because of lack of system support. Damage to infrastructure and/or material triggered by natural causes (supplier or plant premises), such as floods, wildfires, or other incidents of similar nature can cause severe disruptions which can range timewise from days to weeks. It usually means additional work to find alternative suppliers and/or means to keep production going. Generally, critical disruptions produced by central IT architecture issues can take significant effort timewise to be resolved. Legacy systems can go offline due to data breaches or data integrity issues. It can then become impossible to proceed with logistics in the internal supply chain because of lack of system support. Since these are dependent from headquarters, disruptions of this nature can take days to solve, and render the productive/logistic process inviable. Advancements in core automotive technologies can cause sudden realignment of logistics and productions strategies which require extensive rearrangements of internal and external supply chains.

2.3 Implementation in the "Megafactory & E-battery" business case

This pilot consists of two companies: FILL and AVL. Resilience assessments were performed at both companies individually.

2.3.1 The resilience compass assessment in Pilot 2

FILL implements several situation awareness resilience practices to deal with various types of risks that impacts the organization. They mentioned that regular risk assessments are a key to success for their technology leadership strategy. CE certifications are mandatory for risk assessment for products that impact the organization from the "technology push" side (technical risks), VUCA (Volatility, Uncertainty, Complexity, and Ambiguity) is from "market demand" side and covers a wide range of factors (economical risks) and with standardization certifications, e.g., ISO9001 (organizational risks). In terms of national legislation and standards to be met, the branch is conservative from a machine builder point of view and changes are long-term (rating of 2). With quotations requirements, FILL is informed about needs for standards and legislations, so they make a stage-gate decision for request offers.

Visibility practices were in general rated lower for current efforts as it was described that digital continuity seems to be mandatory for products in connected factories which currently have Computer-Aided Design (CAD) and data interoperability barriers. For the process level, software solutions in organizations are heterogenous and often not or weakly connected, as large software solutions e.g. Siemens PLM or 3DExperience are expensive. Additionally, a loss to qualifications and skills in organization for change is a barrier. Security practices scored high as the organization performs penetration tests for their systems with external organizations. Although redundancy is an important resilience enabler, stock and storage (an important Key Performance Indicator, KPI, in logistics) in general has to be reduced from an economical point of view. FILL has a diverse customer base and runs different competence centres dealing with different industry branches, reducing the risks of market changes. FILL does business in sport, plastics, composite, casting, machine tools, robotic manufacturing, woodworking and customized engineering processes areas.

Agility was rated higher values: At FILL there is a high standardization done for series products and they also seek customers with the same requirements. In addition, FILL is organized as a matrix organization: every employee can be asked to support besides their defined role, so they can react immediately to VUCA events. The company[is built and there is a failure where remanufacturing is required immediately, the company can remanufacture a steel welded structure, as well as paint and machine it in 24 hours. Those capacities for manufacturing can also reduce the time to market for some projects which can be hugely beneficial. FILL is also highly flexible and can re-purpose its facilities. For short-term needs, this is an advantage, even if in a long-term horizon continuous changes are considered risky. A vision and a mission need to be fixed for the long-term and adapt to megatrends in VUCA. They evaluate the market quarterly and when a market struggles, other products are run.

FILL's development is based on fair partnerships, and some 'collaboration' practices were set to 4. In terms of knowledge management, there are two relevant data sharing concepts: one is the interaction with the customer where data regarding requirements, project management, CAD of products are exchanged, and standards are defined. The second one is the supply chain with stock, data for manufacturing, purchasing, drawings, and layouts. Hence these practices most often exist (level of 4). Digitalization in the supply chain will be key for optimization. The most important aspect will be the balancing between stock and just-in-time, as costs in logistics is not value-adding but increases indirect costs.

Fill has continuous improvement processes and anonymous whistleblowing practices to make their employees feel safe and to create a healthy work environment. Similar practices such as these, under the leadership capability, were rated 5. Continuous learning after a disruption was deemed important as the 'culture' on dealing with failures is the most important. It does not matter who failed, but to figure out why someone failed and how to prevent to fail again. In our increasingly complex world, it is key to manage knowledge or know-how, especially because of the generation which is going to retire soon (brain drain due to retirement). In terms of market position, there is often a change in ownership of large enterprises due to capitalism. With those that FILL has a strategic partnership with, the mindset and financial health aspects are shared, however this can change rapidly due to uncertainties.

Due to AVL's high standards, risk assessments are conducted regularly, but since the company only provides engineering services, the supply chain dimension is not considered (hence the current state rating of 3 for this practice). Digitally tracking products in operations is not always possible due to high flexibility in production and frequently changing prototypes. AVL has to some extent a diverse supplier base. This is not always possible due to the innovative nature of their services and only single sources are available. Like FILL, AVL also has a matrix project management approach which helps the company in reshoring strategies. Coopetition strategies frequently exist but only with research groups and within platforms. Most prototype parts are provided by the OEM and hence they do not have logistics databases or other technical infrastructure for collaboration with their supply chain partners.

AVL has strong leadership practices in place where managers and the skill teams carry out regular meeting, employee feedback takes place across several channels and communication takes place across all levels of the organization. Most of the scenario planning activities were not a part of their main business due to the innovative nature of their products for the automotive industry. Quality-based performance measures frequently exist and are provided by their cost engineering team starting with the concept phase of the product.

2.3.2 The resilience radar assessment in pilot 2

Fourteen risks were identified at FILL and twelve at AVL. These are shown in Table 4 (where the first twelve ones refer to AVL too, as specified in brackets).

Table 4: Risks identified at FILL

No.	Risk	Freq	Severity	Risk Category	Risk type	Level of impact
1	Mismatch on customer requirements lead to higher prices e.g. over-engineering	High (Both)	Low (FILL) Medium (AVL)	Industrial (Within SC)	Demand	organisation
2	Information loss due to data transfer or wrong datasets of customer CAD, product data	Low (Both)	Low (Both)	Organisational (Within firm)	Behavioural	Worker
3	Skilled labour shortage	Medium (Both)	High (FILL) Medium (AVL)	Environmental (Outside SC)	Policy/Regulation	Organisation
4	Dependency on transport and supply chain, no access to ship or railroad (usually managed by the customer for AVL)	High (Both)	Medium (FILL) High (AVL)	Organisational (Within firm)	Operational	Supply chain
5	Mismanaged leadership	Low (Both)	Low (Both)	Organisational (Within firm)	Operational	Company
6	Whistleblowing by employees	Medium (Both)	Medium (Both)	Organisational (Within firm)	Behavioural	Employee

7	Impression of company does not necessarily match reality, e.g. image, employer branding	Medium (Both)	Medium (Both)	Organisational (Within firm)	Operational	Company
8	Cyber-security and IT risks	Medium (FILL) High (AVL)	High (FILL) Medium (AVL)	Environmental (Outside SC)	Cybersecurity and safety	Company
9	Crisis in market (e.g. combustion engine will stop 2030)	High (Both)	Medium (Both)	Environmental (Outside SC)	Disruption	Regional (local)
10	SDG Contributions foreseen for industry sector not fulfilled, e.g. Scope 3 purchased parts are lacking in CO2 footprint transparency	Medium (Both)	Medium (Both)	Industrial (Within SC)	Operational	Company
11	SDG Contribution for Scope 2- FILL/AVL is not able to consume its energy from 100% renewable sources	Low (Both)	Medium (FILL) Low (AVL)	Environmental (Outside SC)	Policy/Regulation	Regional
12	Cashflow risk, the needed amount of cash for transformation in CO2 neutral is not available	Low (Both)	Medium (Both)	Environmental (Outside SC)	Financial	Company
13	Patent risk, competitor is protecting the market (FILL)	Medium (FILL)	Medium (FILL)	Industrial (Within SC)	Operational	Company
14	Natural disaster e.g. flooding (FILL)	Medium (FILL)	Medium (FILL)	Environmental (Outside SC)	Disruption	Regional

FILL is known for their high-quality products and technical solutions; however, this can prove to be detrimental for their business as over-engineering and corresponding increase in prices means that the ROI of R&D can be difficult to reach. As the company grows, whistleblowing issues could increase due to trust and commitment issues. This further leads to loss of critical information. This loss of trust can have further ripple effects where the image or reputation of the company is lost in terms of brand name which further leads to less competitiveness in the human resource market. Cyber-security and IT risks are increasingly common concerns for being resilient and data loss due to such risks can lead to financial chaos due to customer contracts and the rebuilding of IT systems. If the correct Scope 3 emissions are not shared by the suppliers, this can lead to lack of flexibility and trust with the suppliers. Competitors acting faster with patents can cause a barrier as series automation lags behind in R&D which makes it difficult to catch up with such market trends. Regional natural disasters such as flooding is happening more often nowadays and can cause a shortage on employee labour. The company itself is in quite a geographically safe location, but employees living in areas where public transport can be affected can lead to partial shutdown of operations for days or weeks depending on severity.

For risk 3, due to the constant change of product development from battery over internal combustion engine to fuel cell the skill of each designer is needed to fulfil the customer specifications. AVL is prepared for such risks and works in skill teams, but this causes troubles in changing employees during projects. For risk 4, the prototyping parts are purchased by the customers and supplied to AVL for assembly. There is no direct contact between the part supplier and AVL. The disruption due to this risk is that the parts for assembly come very late and makes it difficult to adapt the production environment with real parts. In the future, digital twins could make it easier to adapt beforehand and only make small changes when parts arrive.

2.4 Implementation in the "Circular machining" business case

Pilot 3 consists of two companies: GF and Fraisa. Resilience assessments were performed at both companies individually.

2.4.1 The resilience compass assessment in Pilot 3

At Fraisa, most of the current resilience capability levels were 3 or 4 for the situation awareness capability. The company has plans to Install automatic internal (customer behaviour) and external (web data) data analysis to detect changes as early as possible in the future. Manual competition screening, patent observations and scenario planning are a part of their regular planning activities and hence they scored higher on these aspects. Fraisa maintains a good data exchange with its suppliers and personal contacts, to avoid unintended surprises and are able to react with sufficient resources when disruptions occur. In terms of digitally tracking products, its production papers are barcoded, and product pallets have a RFID chip. A laser marking of each blank already at the receiving department is in preparation, in order to assign all data generated in the production to the product in the future. For data and security, high values of 4 and 5 levels were chosen. This because some practices were already covered by SAP, or production figures are already linked. Fraisa is also close to obtain a Level A with respect to cyber security. In terms of personnel security, the company has stringent personnel regulations which prohibits any kind of discrimination and mobbing. Respectful treatment of each other is embedded in the corporate culture at Fraisa.

In terms of differentiating the supplier base, Fraisa scored itself with a high value, as carbide suppliers are system critical and the company has multiple suppliers in 4 countries and 2 continents. In addition, not all resilience practices can be applicable in all manufacturing contexts. For instance, in terms of some agile practices, its products need min. 12-14 months development time and, at the time being, fast customization of standard products is only possible in a few cases. The organization is not designed for disruptive changes; however, Fraisa is able to introduce rapid product changes in many cases. Its production is indeed designed for the manufacturing of solid carbide endmills, and to manufacture other products (e.g., when during Covid-19 pandemic), large investments are necessary. In terms of collaborations, the company has developed cross-functional workgroups, cross-company networks and working groups to provide additional services to our customers. And although only major suppliers have access to its data via SAP vendor access, the company would like to provide its customers with tool production data in the future.

The company showed also high leadership levels with systems that create a safe working environment for their employees to escalate and address issues. Fraisa systematically conducts safety walks throughout the company and also register near accidents, as well as has several information sessions per year for all employees. An information brochure is also published yearly for employees, customers and suppliers, where key financial figures are stated. A suggestion scheme for continuous improvement of processes and products has been established. The company also invests more than 1% of its revenue in employee training activities.

With the RE4DY project, Fraisa aims at systematically collecting tool and application data to better adjust future processes and to offer new service areas, so as to improve its knowledge management capability. For scenario and contingency planning, Fraisa primarily plans at the highest consolidation level in order to consider possible macroeconomic influences. Redundancies are also planned in production as well as in supply chains, to accommodate machines/supplier failures. In the case of critical suppliers, the company checks its financial reports on a cyclical basis to obtain certainty about their financial situation. Only in the case of complaints, they receive extensive feedback from our customers. Otherwise, they only receive sporadic feedback via the field service (hence the low level marked for this resilience practice). Through the systematic data collection based on the RE4DY project, this should change significantly in the future. They measure the produced scrap very accurately and develop remedial actions with root course analysis for quality-based performance measures.

GF scored itself higher on some visibility capabilities. GF uses indeed 4DX process (4 disciplines of execution) to capture the right type of data across the organization's different functions. The company has also innovation suggestion schemes and personal development policies that help create a safe working environment for their employees. For agility, flexibility and collaboration, it scored iself at lower levels in its current state and mentioned that digital platform eco-system implementations and SAP coordination activities with its key suppliers are ongoing which would help the company reaching its future targets.

2.4.2 The resilience radar assessment in Pilot 3

Fraisa identified 14 risks that impact its organization. These are shown in Table 5.

Table 5: Risks identified in Fraisa

No	Risk	Freq	Severity	Risk Category	Risk type	Level of impact

1	Disruption of supply chain (Supply of raw materials (cemented carbide rods) from overseas suppliers (China, Japan))	low, all five years	High	Industrial (Within SC)	Supply	Factory
2	Cyber-attacks; no access to data	mid, 2 times a year	High	Organisational (Within firm)	Cybersecurity and safety	Factory
3	Swiss currency gets extremely strong (0.80 / EURO)	mid, 2 times a year	High	Industrial (Within SC)	Financial	Organisation
4	Cemented Carbide gets classified as carcinogenic	low; once in 50 years	High	Industrial (Within SC)	Supply	Business
5	New material to replace Carbide get invented	low; once in 50 years	High	Industrial (Within SC)	Supply	Business
6	New business cases are popping up (less customer interference)	low, every 10 years	High	Organisational (Within firm)	Disruption	Business
7	Receiving poor data quality from customers to feed models (application data, how they use the tools)	high; weekly	Low	Organisational (Within firm)	Supply	Business
8	Inventory get stolen	mid, every five years	Low	Organisational (Within firm)	Supply	Business
9	War in Asia (Taiwan)	low; once in 50 years	High	Environmental (Outside SC)	Disruption	Business
10	Fraisa is missing market trends	low; once in 20 years	High	Organisational (Within firm)	Demand	Organisation
11	Severe Patent issues (to avoid patent infringement)	low, once in 10 years	High	Organisational (Within firm)	Operational	Business
12	Poor succession (planning of management)	low; once in 20 years	High	Organisational (Within firm)	Operational	Organisation
13	Production costs in Switzerland are going	mid; 2 times a year	High	Organisational (Within firm)	Operational	Factory

	through the roof					
14	Cannot fill open positions	high; 5 times a year	High	Organisational (Within firm)	Operational	Business

In most cases, the risks gave rise to production disruptions but in some cases, they also gave rise to reduced cash flows (risks 3, 11 and 13). Losing customers was another important disruption mentioned by Fraisa as a result of not keeping up with market trends or if cemented carbide gets classified as carcinogenic (this risk was the most prioritized). Although raw materials come from specific Asian suppliers, Fraisa has a diverse supplier base with other suppliers in Asia as well, to manage supplier dependency risks.

Six main risks impact GF at the time being. These are shown in Table 6.

Table 6: Risks observed in GF

No	Risk	Freq	Severity	Risk Category	Risk type	Level of impact
1	Supply chain interruption, regarding machine materials, tooling, services	Low	High	Industrial (Within SC)	Supply	Impacts directly revenues from machines and tooling
2	Raw materials costs for consumables and machine components	High	High	Industrial (Within SC)	Financial	Impacts revenues as sales decrease with respect to competition
3	Cybersecurity risks related to My rConnect Platform	High	High	Organisational (Within firm)	Cybersecurity	Impacts data protection obligations and sales of digital products
4	CO2 footprint in excess with respect to new regulations	Low	Medium	Environmental (Outside SC)	Operational	Machine and consumables sales revenues reduced due to non-compliance of new norms

5	Data sharing and exchange limitations with customers	High	Medium	Industrial (Within SC)	Operational	Implementation of RE4DY innovations limited due to missing data access from/to customer site
6	Accuracy of tool wear and residual lifetime estimations	Medium	Low	Organisational (Within firm)	Operational	Benefits from tool wear monitoring and attractivity of solution are reduced

Cybersecurity risks are most prioritized, and GF has dedicated resources for dealing with it (the subsidiary company supplying the platform and connectivity has specific features and certifications related to this risk). Although most risks directly interrupt production and reduce sales for machines and consumables, GF mentioned that the risks could directly impact services delivered by the RE4DY project, i.e., market deployment could be discontinued or incomplete due to such risks.

2.5 Implementation in the "Distributed, Green, Zero-X Manufacturing" business case

2.5.1 The resilience compass assessment in Pilot 4

In general, Pilot 4 was satisfied with its resilience capability implementation levels (marked with green dots in their resilience dashboard (Section 2.6). Its cyber-security measures were compliant to regulations from both civil and military aviation and hence, these capabilities were marked as high. Knowledge management activities were rated highly as well, as this pilot intended to use AI and ML techniques to train their junior operators and prevent dependencies on skilled senior staff.

2.5.2 The resilience radar assessment in Pilot 4

Seven risks were identified in pilot 4 and are shown in Table 7

Table 7: Risks identified in pilot 4

No	Risk	Freq	Severity	Risk Category	Risk type	Level of impact
1	Lack of material due to incorrect planning in production processes or delays in	Medium	High	Organisational (Within firm)	Supply	Supply chain

	material procurement from the supplier					
2	Extraordinary maintenance due to unexpected failures on production machines	Medium	High	Organisational (Within firm)	System	Supply chain, quality
3	Expertise gap due to human resources turnover	Medium	High	Organisational (Within firm)	Behavioural	Factory, quality
4	Geopolitical factors (wars, pandemic events, macro economical dynamics)	Low	High	Environmental (Outside SC)	Disruption	Factory, supply chain
5	Air traffic market changes	Low	High	Environmental (Outside SC)	Disruption	Factory, supply chain
6	Redefinition of priorities in deliveries due to market dynamics, customers' decisions or relevant events from the fleet	High	High	Industrial (Within SC)	Demand	Factory, supply chain, quality, finance
7	Cybersecurity events impacting the availability of shop floor assets or company systems	Low	High	Organisational (Within firm)	Cybersecurity and safety	Factory, supply chain, quality, finance, safety

Most risks gave rise to a delay in customer deliveries. Some risks also impacted customer demand (reducing it) which in turn could give rise to reduced revenues. The most relevant risk for this pilot was the expertise gap due to increased employee turnover. Tacit knowledge is a key part of the production processes. Lack of expertise can decrease the shop efficiency in terms of production and proper management of the assets. It's a clear issue for the inspection processes that relies on qualified operators: it could lead to slowing down the inspection activities because of the lack of certified personnel. The pilot also mentioned that they already have novel solutions in place to manage risks as part of their Business-As-Usual (BAU) activities. For instance, risks are currently handled at three levels: organisational level (risk management register), local level (handled by the CTO and team for risk-related policies, managing vulnerabilities, product and asset lifecycle management) and an external level where service provider monitors risk-related activities 24/7.

2.6 The resilience dashboards

Based on the assessments through the resilience compass and radar, a dashboard was created that combined these efforts for effective visualization of capability development and deployment by the companies. Companies can use the dashboard to effectively navigate the various uncertainties that face them and develop corresponding resilience capabilities in the three temporal stages of resilience. This can be seen in *Figure 6*, *Figure 7*, *Figure 8*, *Figure 9*, *Figure 10* and *Figure 11* for the six companies in pilots 1-4 respectively. The risk bubbles shown in purple were the most prioritized in the companies. An interesting finding was that most of the prioritized risks are well-anticipated for. That is, the pilots have invested in such capabilities to be aware of and mitigate such risks before they become disruptions in their organisations.

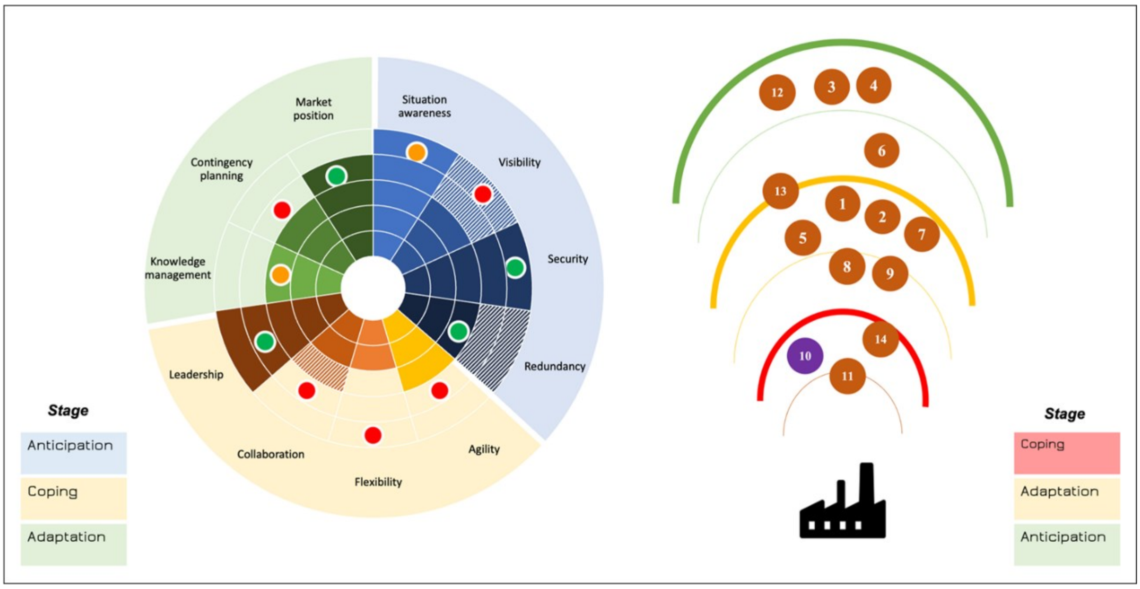

Figure 6: Resilience dashboard of VWAE

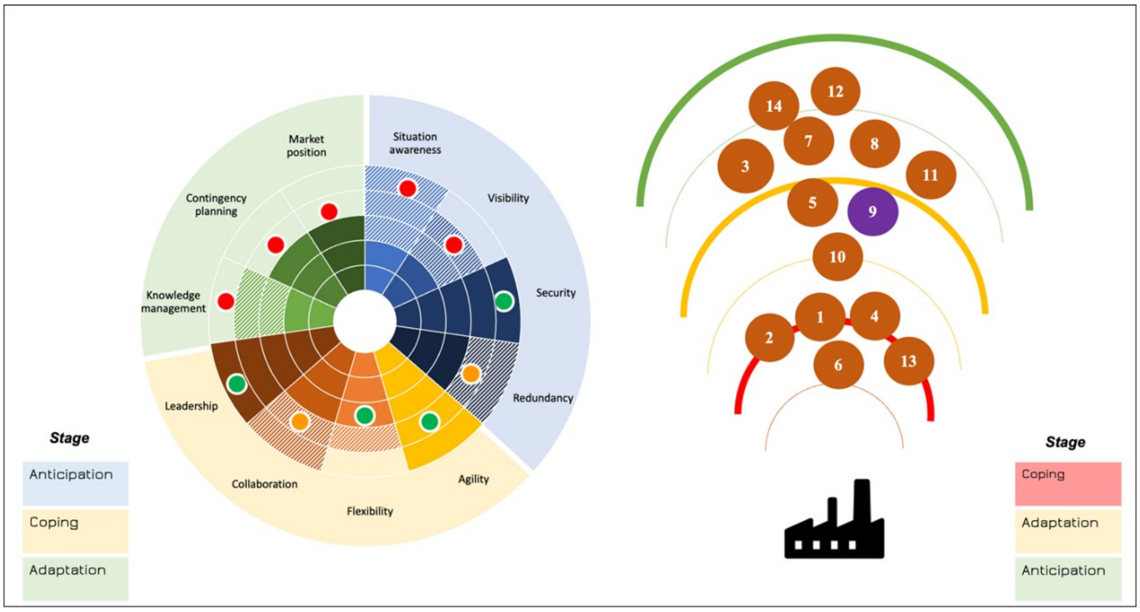

Figure 7: Resilience dashboard of FILL

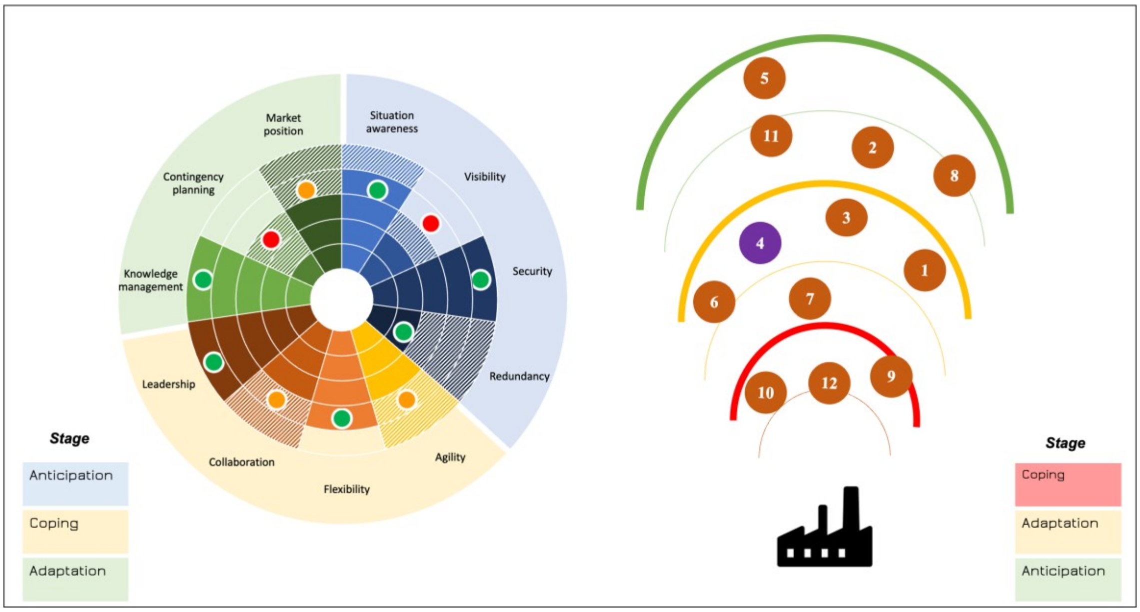

Figure 8: Resilience dashboard of AVL

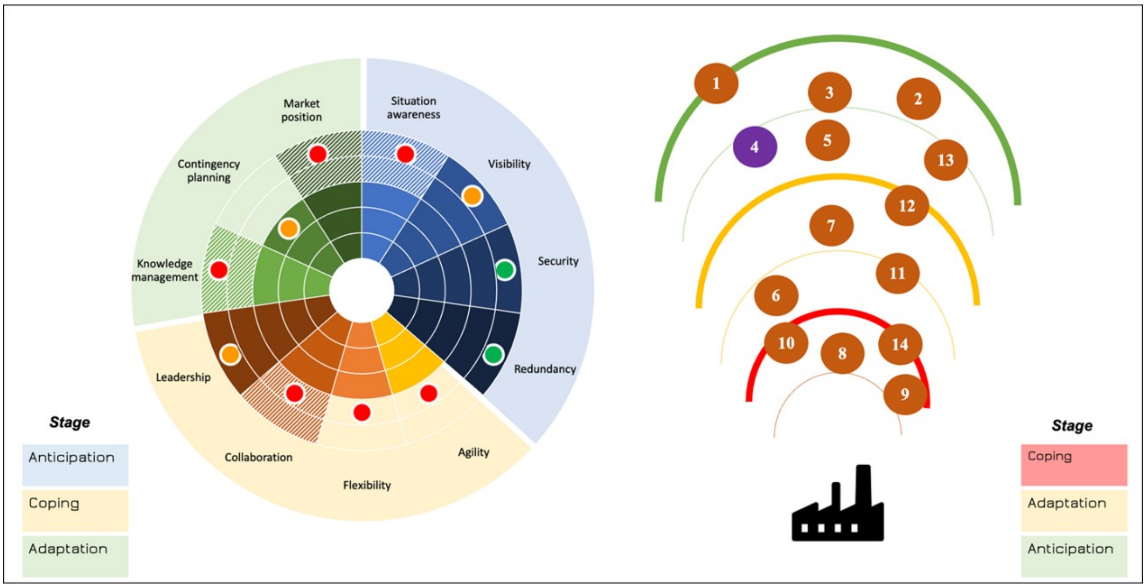

Figure 9: Resilience dashboard of Fraisa

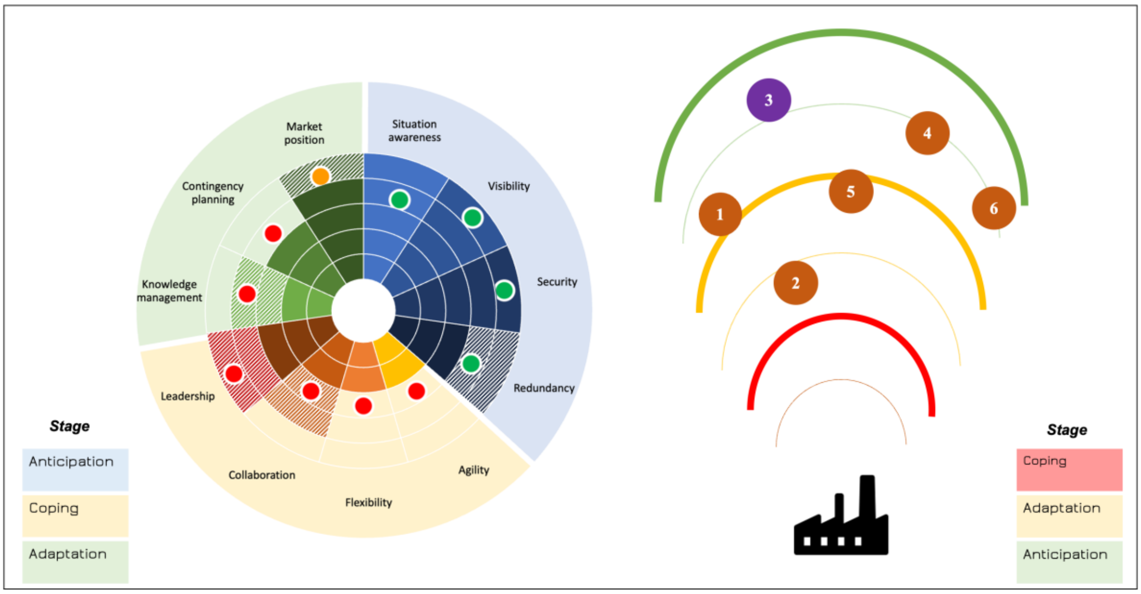

Figure 10: Resilience dashboard of GF

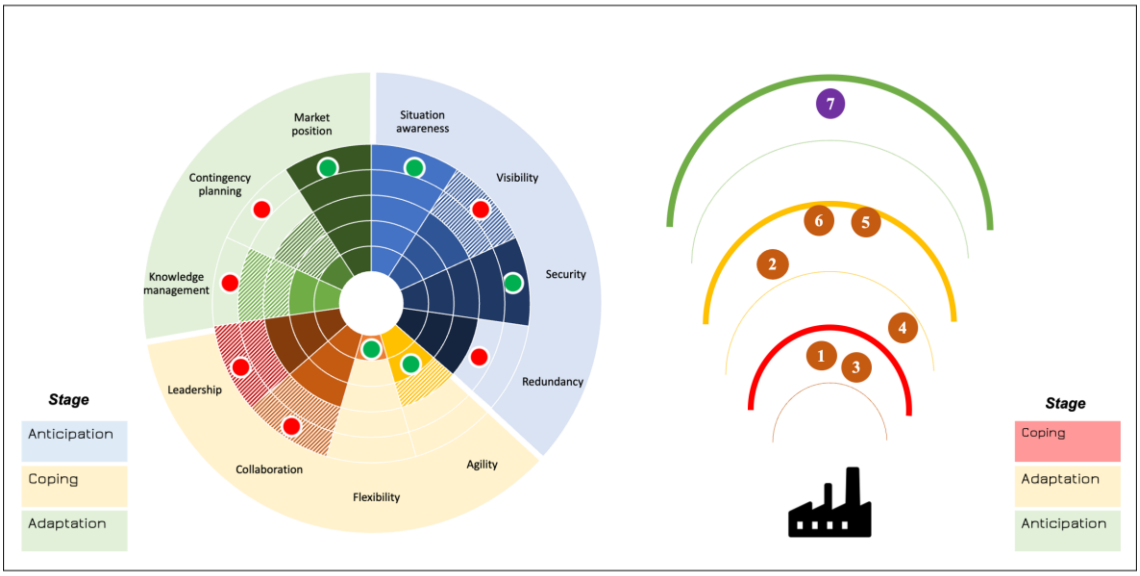

Figure 11: Resilience dashboard of Avio Aero

2.7 Data-driven approach for Artificial Intelligence implementation

To explore the feasibility of artificial intelligence methods and to improve resilience in manufacturing operations, the consortium conducted a detailed study on data engineering by simulating the intra-dependencies in a discrete manufacturing system, which are generally non-deterministic due to uncertain events. Synthetic data generation techniques were used to overcome a variety of challenges frequently encountered in artificial intelligence development:

- High quality benchmark data availability: addressing the need for accurate, representative data sets to train and validate AI models including labels.
- Complex data integration: managing data from heterogeneous sources and formats, which is critical in a fragmented manufacturing environment.
- Scalability of simulations: ensuring our models can handle increased complexity and data volume without performance degradation.
- Real-time data processing: developing capabilities for processing data in real time to provide timely insights and responses to dynamic manufacturing conditions.

Using this approach can effectively overcome these challenges, enabling further exploration and refinement of these models.

The discrete manufacturing system has been represented as a Directed Acyclic Graph (DAG), where each node symbolised a distinct operation or process, and the edges represented the flow of materials or information. This graph structure eased the mapping of the manufacturing process in a clear and structured manner. System's dynamics were modelled by incorporating a variety of time-stamped but uncertain events that could affect each node. The current model was validated on an existing simulation model representing an experimental production line at Chalmers University of Technology.

In thismodel, the adjacency matrix describing the graph serves as an input parameter, allowing to study diverse systems ranging from small to large and simulate complex intra-dependencies. This configurability enables tailoring the simulations to different scales and complexities, easing the study of AI model robustness across various settings. Additionally, the analysis maintened the flexibility to define which data within the simulation are observable and which are not, depending on the existing data infrastructure.

Moreover, the consortium can specifically designate which elements of the observable data serve as input variables and which are used as labels, forming essential data-label pairs for supervised AI model evaluation. This distinction is vital for applying our AI models to tasks such as forecasting, causal inference, or anomaly detection. By clearly defining and manipulating input and label data, the consortium can rigorously test and validate our AI models, ensuring they are capable of making accurate predictions and identifying patterns based on the simulated data provided.

With this configuration, the ways AI models perform in small settings can be examined, such as those typically encountered in Small and Medium Enterprises (SMEs), where resources might be limited, and operational dynamics differ significantly from larger corporations. Simultaneously, the analysis can be scaled up to larger models, which are characteristic of bigger companies with extensive operations and more complex data ecosystems. This diverse capability enhances the ability to adapt AI technologies to a broad range of industrial contexts, providing valuable insights into the scalability and applicability of AI solutions in diverse manufacturing environments.

2.8 The sustainability implications of building resilience

The pilots were also asked to assess if the resilience capability had a positive or negative implication on the triple bottom line of sustainability (Elkington, 1998) (environmental, social and economic pillars). The economic pillar was not considered as no business would operate without a profitable objective. In addition, previous literature on resilience has also primarily considered the economic dimension (Ivanov, 2020) (Rashid A.H.M., 2014) without focusing on the other two pillars. Hence, the assessment focused on the environmental and social dimensions. These were mapped as shown in *Figure 12*. This additional step was carried out in the study, since it has been previously observed that there could be relation between resilience and sustainability, and empirical evidence from the companies could corroborate these theoretical findings. Some pilots described the positive and negative sustainability impacts of developing resilience capabilities, while some other ones described the implications in terms of Sustainable Development Goals (SDGs) (UN, 2015).

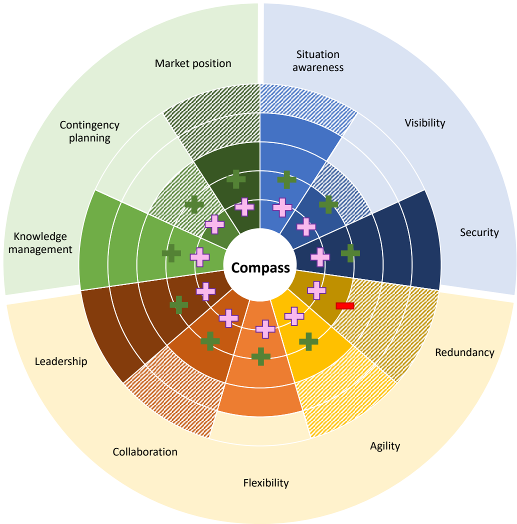

Figure 12: Sustainability implications of building resilience capabilities

On the whole, all the pilots described positive implications of building resilience capabilities on environmental and social sustainability, except for the redundancy capability which had conflicting impacts. For instance, VWAE and GF described a negative relationship of redundancy to the environmental dimension with comments such as – 'With higher levels of redundancy, these options often come at costs of burning through more resources' and 'Stocks may negatively influence the carbon footprint of production sites. Secondary suppliers are a more environmentally friendly strategy to implement', respectively. Whereas Fraisa described 'Without redundancies, we would have to react extremely quickly and with little optimization. This would lead to significantly more transport and less environmentally friendly transport, thus greatly increasing the CO2 inefficiency', implying a positive relationship.

In terms of the environmental impact of implementing the visibility capability, Fraisa said '[..] we have switched to CO2 neutral electricity, we have installed photovoltaics on the roof, we have switched the heating to wood pallets, and we are working with our suppliers to ensure that the raw materials used can be produced as CO2 neutral as possible. We are installing RFID technology to track products and product pallets' showing a tremendous awareness to contribute to ongoing sustainability challenges and the use of technologies to promote such practices.

3 Legal, governance, and ontological frameworks

3.1 EU Data space and data sharing legislation key aspects and implications

Each RE4DY pilot is active in a different industrial sector, including aeronautical manufacturing (Avio Aero), machine tool manufacturing (GF/FRAISA), battery manufacturing (AVL), and automotive manufacturing (VWAE). This results in the applicability of a variety of sectoral, national, and international regulations regarding, e.g., airworthiness, machinery safety, and cybersecurity.

Rather than providing industry-specific guidance on these issues, this section instead supports and enables the broader Digital 4.0 Continuum by outlining the EU-wide, cross-sectoral legal framework that specifically relates to data and the sharing, governance, and control thereof. The focus is not on creating a comprehensive account of these laws and the surrounding legal questions, but on bringing to light select issues that are particularly relevant to the Digital 4.0 Continuum. Subsection 3.1.1 considers the concept of "data as a product" from a regulatory perspective, clarifying whether data may be considered a 'product' in the legal sense with reference to instruments such as the new proposed Product Liability Directive (PLD). Subsection 3.1.2 introduces Regulation (EU) 2023/2854 (the Data Act or 'DA') and analyses its implications for RE4DY use cases, particularly with regards to data control, accessibility, and the conclusion of data sharing contracts. Subsection 3.1.3 introduces Regulation (EU) 2022/868 (the Data Governance Act or 'DGA') and analyses its implications for RE4DY data spaces, especially with regard to the concept of data intermediation service providers and the latter's obligations. Lastly, Subsection 3.1.4¡Error! No se encuentra el origen de la referencia. analyses the interface between d ata and intellectual property rights, as well as the application of intellectual property rights to digital twins.

3.1.1 Legal Perspectives on Data as a Product

It is important to understand whether or not data is a product in the legal sense, as this classification may carry with itself additional requirements with regard to, e.g., the product's attributes, manufacturing process, or performance, as well as implications with regard to, e.g., liability in the case of harm caused by the product. It is also important to preface any analysis with the disclaimer that the definition of a 'product' is not universal and varies between legal fields and EU instruments,

At present, Council Directive 85/374/EEC (Product Liability Directive or 'PLD') defines the term product as encompassing "all movables", as well as "electricity" (Art. 2 PLD). Data or information in its digital form would therefore fall outside of the scope of the term 'product', unless integrated in a tangible object (Buiten, 2021).

This reflects the general understanding of products as physical objects that has only recently begun to change following the Commission's 2018 evaluation of the PLD[1] and the broader 2020 European strategy for data.[2]

However, the PLD is currently undergoing revision considering new technological developments. Its new version, having been approved by the European Parliament on 12 March 2024, is now pending approval from the Council of the European Union. The revised definition of product now explicitly includes intangible items, as it reads: "'product' means all movables, even if integrated into, or inter-connected with, another movable or an immovable; it includes electricity, digital manufacturing files, raw materials and software" (Art. 4.1 Revised PLD proposal). A digital manufacturing file, in turn, is defined under Article 4.2 revised PLD as: "a digital version of, or digital template for, a movable which contains the functional information necessary to produce a tangible item by enabling the automated control of machinery or tools." Beyond digital manufacturing files in particular, earlier versions of the proposed new PLD left it generally ambiguous as to whether the definition of products included not just software as an intangible product, but data as well.

Yet, the newest version from 12 March 2024 states in Recital 13 that "Information is not, however, to be considered a product, and product liability rules should therefore not apply to the content of digital files, such as media files or e-books or the source code of software." If these provisions survive unaltered and the revised PLD is adopted, it would therefore be the case that data cannot constitute a product insofar as liability is concerned. At first glance, it appears contradictory that digital files and their contents are not products, yet digital manufacturing files are products.

This confusion is addressed in Recital 16 of the revised PLD proposal, which clarifies that "Whereas digital files as such are not products within the scope of this Directive, digital manufacturing files, which contain the functional information necessary to produce a tangible item by enabling the automated control of machinery or tools, such as drills, lathes, mills and 3D printers, should be considered to be products, in order to ensure the protection of natural persons in cases where such files are defective." In other words, specific packages of functional data may be considered a product, insofar as they can enable the automated control of machinery to produce a tangible item and cause harm through that production or its resultant item. This represents a specific and limited extension of the definition of products that is relevant for the context of RE4DY use cases.

RE4DY digital twins leverage real-time data to control manufacturing machinery across the automotive, aeronautical, battery, and machine tool sectors. The digital twins themselves are composite systems that consist of a computer program, algorithm(s), and an underlying model of a system. From the outset, the software component of RE4DY digital twins will straightforwardly fall under the scope of a revised PLD 'product'. But, with regard to data productization, whenever digital twins have as their input or output digital manufacturing files (e.g., CAD/CAM files as in some components of the AVL pilot), it is important to consider that these inputs and outputs will likely qualify as (functional data)

[1] {COM(2018) 246 final} - {SWD(2018) 158 final}, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018SC0157

[2] COM(2020) 66 final, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0066

products and be subject to the liability regime of the revised PLD (the specific provisions of which are beyond the scope of this section).

Moreover, the PLD's regulatory logic does not explicitly require that harm be caused only by the faulty item created from a defective digital manufacturing file. Recital 16 of the revised PLD gives an example of a defective CAD file 'product' being used to create a 3D printed good which then causes harm, though it is not stipulated that in all cases the harm must come from the produced good rather than the action of the machinery or tools automatically controlled by the defective digital manufacturing file. In the context of RE4DY cognitive digital twins, this raises the question of whether the revised PLD's regulatory logic may be extended on a sectoral level to classify as 'products' any files modelling a physical manufacturing system and guiding that system's actions in the manufacturing process, despite not *per se* including the information to produce novel tangible items in most cases. At present, this question is hypothetical – the current wording of the revised PLD proposal would not classify such digital twin data components as 'products', due to their lack of instructions for producing tangible items.

In any case, with the exception of digital manufacturing files, it is possible to note that data does not constitute a product viz-a-viz liability rules, while software does. One could also look at Regulation (EU) 2023/988 (General Product Safety Regulation, or 'GPSR'), which, although geared towards the protection of consumers (i.e., natural persons), likewise modernizes the notion of 'product'. In Article 3.1, the GPSR states: "'product' means any item, whether or not it is interconnected to other items, supplied or made available, whether for consideration or not, […]". It is ambiguous from this definition whether intangible objects such as software or data are to be considered 'items' and therefore 'products', or whether they are merely constitutive elements of products. During the legislative process, the EU decided not to adopt amendments that sought to explicitly state that products could be "any item, tangible or intangible" and that tried to clarify that embedded software and stand-alone software would fall under the scope of the GPSR.[3] In that respect, novel product safety provisions offer no interpretative guidance on the legal dimensions of RE4DY DaaP.

On the other hand, legislation such as the proposed "Regulation on the transparency and integrity of Environmental, Social and Governance (ESG) rating activities" (proposed ESG Ratings Regulation)[4] explicitly recognise the existence of 'data products'. Whereas the initial proposal of this regulation never mentioned the term 'data product'[5], the text adopted by the European Parliament on 24 April 2024 explicitly makes numerous references to the same term, although it never defines a data product. Data products are likewise frequently mentioned but not defined under the International Organization of Securities Commissions' 2021 report on ESG ratings and data product providers[6], which the ESG Regulation is informed by and makes reference to. Nevertheless, these texts

---

[3] Procedure 2021/0170(COD), Document A9-0191/2022, available at: https://www.europarl.europa.eu/doceo/document/A-9-2022-0191_EN.html

[4] Document P9_TA(2024)0347, EP compromise text, available at: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0347_EN.pdf

[5] {SEC(2023) 241 final} - {SWD(2023) 204 final} - {SWD(2023) 207 final}, available at: https://www.europarl.europa.eu/RegData/docs_autres_institutions/commission_europeenne/com/2023/0314/COM_COM(2023)0314_EN.pdf

[6] Document FR09/21, available at: https://www.iosco.org/library/pubdocs/pdf/IOSCOPD690.pdf

represent an acceptance of data productization on a sectoral level, signaling that data products and data as a product may be more readily regulated in specific industry contexts, rather than in general, broadly applicable laws.

The concept of data products also plays a role in the regulatory regime of Implementing Regulation (EU) 2020/469 on air traffic management and air navigation services[7]. Therein, Annex III defines a "data product specification" as "a detailed description of a data set or a collection of data sets together with additional information that will enable it to be created, supplied to and used by another party". This term reflects the regulator's recognition that structured datasets accompanied by sufficient metadata constitute complete products in their own right. Yet, contrasting the Implementing Regulation on air traffic management and air navigation services and the proposed ESG Ratings Regulation on one hand, and the PLD and GPSR on the other hand, it also becomes possible to differentiate two distinct incarnations of the broad concept of data as a product. As a term of law exemplified under the PLD and the GPSR, 'product' is a strictly defined category and generally excludes from its scope pure information. Separately, however, "data products" are evident in the ESG Ratings Regulation and the Implementing Regulation on air traffic management and air navigation services as a term of industry that the law is evolving to regulate on a sectoral basis, rather than a fundamentally legal concept. In this context, data products are usually broadly defined, if at all explicitly defined, and include within their scope data and metadata. The classification of a dataset as a "data product", being a rather descriptive term under either of the examined legal instruments, carries significantly less legal obligations relating to the dataset, as opposed to the classification of data or software as a "product" under liability or safety law, which carries significant legal ramifications. Across both "data products" and "data as a product", it is also possible to note a common trend, in that productization under the law in generally corresponds with the risk inherent in low-quality data. By regulating data products on a sectoral basis or by recognizing data itself as a product, productization allows the regulator to assign responsibility over data in instances where data quality is of paramount importance.

Overall, the understanding of data as a product in EU legislation is currently evolving and marked by conceptual differences between regulatory fields. For the most part, neither raw nor processed data is a "product" in the legal sense, underlining the need for RE4DY pilots to distinguish between technical and legal terminology. Yet, exceptions to the norm exist, including "ESG data products" and digital manufacturing files in the field of liability. The latter especially are data products with legal implications for the assignment of responsibilities in RE4DY cognitive digital threads.

3.1.2 The Data Act

The Data Act[8] entered into force on 11 January 2024 and will be applicable from 12 September 2025. Its stated purpose is ensuring fairness in the allocation of value from data among actors in the data economy and fostering access to and use of data.[9]To this end, the DA establishes frameworks for data sharing (Chapters III and IV), for switching data processing services (Chapter VI), as well as for participating in data spaces and for

[7] Implementing Regulation (EU) 2020/469, available at: https://eur-lex.europa.eu/eli/reg_impl/2020/469/oj

[8] Regulation (EU) 2023/2854, available at: https://eur-lex.europa.eu/eli/reg/2023/2854

[9] Document SEC(2022) 81 final} - {SWD(2022) 34 final} - {SWD(2022) 35 final}, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022PC0068

employing smart contracts in data sharing agreements (Chapter VIII). Furthermore, the DA creates novel data control rights, namely new rights to data access and portability (Chapters II and VI).

Key concepts under the DA include connected products and related services, which are respectively defined in Articles 2.5-2.6 as items that obtain, generate or collect data concerning their use or environment and that are able to communicate product data on one hand, and digital services vital for the functioning of the connected product on the other hand. Key actors under the DA include data holders, which Article 2.13 defines as natural or legal persons that have the right or obligation to use and make available product and related services data, as well as users, which Article 2.12 defines as natural or legal persons that own or have had rights transferred over connected products, or that receive related services.

Importantly, Article 3.1 DA endows users of connected products and related services with the right to access readily available product and service data (including relevant metadata). Where relevant and technically feasible, this accessibility should be direct, rather than requiring the cooperation of the data holder or product manufacturer. By default, the user's access to product and related service data must also be easy, secure, free of charge, and in a comprehensive, structured, commonly used, and machine-readable format (Art. 3.1 DA). The user of a connected product or related service is also entitled to certain pre-contractual information under Article 3.1 DA, including (i) the characteristics of the data generated by the IoT device, (ii) whether and to what purpose the prospective data holder intends to use generated data itself and whether it intends to allow third parties to use the data, and (iii) how the user may exercise control over the data, including how they may exercise their rights of data access and portability.

For RE4DY ecosystem actors, these provisions provide further impetus toward an open data ecosystem, where device manufacturers and service providers are obliged to furnish users with valuable data from machine operations. The providers of tools, sensors, and related service in all pilot use cases are expected to abide by such data access obligations by adopting "data accessibility by design". However, these obligations are not absolute – two exceptions relevant for RE4DY use cases warrant mentioning.

The first exception concerns information that is inferred or derived from raw data, which Recital 15 of the DA clarifies as "the outcome of additional investments into assigning values or insights from the data, in particular by means of proprietary, complex algorithms, including those that are a part of proprietary software". One given example of such data is information derived by means of sensor fusion, which is a core feature in digital twins in predictive maintenance applications (Liu Z., 2018). Therefore, the obligation of RE4DY product and service providers to freely share 'first-line' data covers measurements by homogenous sensors, but not complex insights of the kind that underlie many of the pilots. Using the GF/FRAISA pilot as an example, GF as a user of smart tools made by FRAISA may be entitled to freely access information on torque or cutting forces, but not complex insights as to whether a tool is worn out and requires maintenance.

The second exception in data access rights concerns data that contains trade secrets. According to Articles 4.6-4.8, data sharing may be refused pursuant to a substantiated and communicated finding by the data holder that the trade secrets in their data have not been guaranteed appropriate protection via technical or organizational measures. There exists uncertainty as to the dynamic that these rules will create in practice, as the DA seems to expect 'ex-ante' identification of trade secrets in data, whereas trade secrets are usually confirmed 'ex-post' in courts and dispute settlement bodies. This results in the possibility that trade secrets protection will be overclaimed in data, leading to unjustified restrictions in data sharing (Mylly, 2024). RE4DY actors are therefore advised to keep in mind the project goal of open and transparent data ecosystems and avoid overreliance on trade secrets protections within the DA if better contractual or IPR alternatives exist.

Beyond the data access right, the DA also furnishes users of connected products or related services with an enhanced right to data portability under Article 5, which compels data holders to make available readily available data to third parties at a user's request. Like the data access right, the portability right is contingent upon appropriate measures being implemented to preserve trade secrets in data where relevant. Article 8 DA obliges data holders who have been obliged to share data under Article 5 DA to do so under fair, reasonable, and non-discriminatory (FRAND) terms and conditions. Even so, Article 9 DA allows for data holders to ask for reasonable compensation in such cases and it is worth noting that this compensation may even include a margin above the inherent costs incurred in making the data available and the investments in the collection and production of data. Until the European Commission publishes guidance on FRAND terms and conditions and on the calculation of reasonable compensation pursuant to Article 42 DA, it is necessary for RE4DY actors to draw inspiration from existing market practices. Inspiration may come in particular from the field of Standard Essential Patents (SEPs), in whose licensing FRAND terms play a central role (Drexl, 2017).

The DA's novel rights over data are of high relevance to RE4DY actors, as they extend previously existing rights from the realm of personal data protection to also cover legal persons and non-personal data, in addition to natural persons and personal data as was the case under the GDPR.

Of note, too, is the DA's regime for data sharing contracts between enterprises. This regime does not concern only data generated by connected products or related services, but more broadly any form of data that underpins a data sharing contract between enterprises. With this in mind, Article 13 DA regulates data sharing contracts by addressing the issue of unfair contractual terms. First, Article 13.1 DA stipulates that a term in a data sharing contract concerning access to and the use of data or liability and remedies for the breach or the termination of data related obligations shall not be binding if it is simultaneously unilaterally imposed and unfair. Article 13.4 DA identifies as always unfair those contractual terms that (i) exclude or limit the imposer's liability for intentional acts or gross negligence, (ii) exclude the imposed party's available remedies in the case of non-performance of the contract or limit the imposer's liability for breach of contractual obligations, and (iii) give the imposer the exclusive right to interpret contractual terms or determine the conformity of supplied data with the contract. Consequently, Article 13.5 DA establishes a longer list of contractual terms that are presumed to be unfair but may in theory be proven otherwise. Examples of such terms include terms that limit the imposed party from using the data it provides or generates during the contract in an adequate manner (Art. 13.5.c DA), and terms that allow the imposer to access and use the other party's data in a manner that is significantly detrimental to latter's legitimate interests, which might be the case if data is commercially sensitive, or protected by trade secrets or intellectual property rights (Art. 13.5.b DA). Finally, Article 13.3 DA broadly defines unfair contractual terms as terms "of such a nature that [their] use grossly deviates from good commercial practice in data access and use, contrary to good faith and fair dealing." This general definition may be relied upon in cases where a term that is suspected of being unfair does not figure in the lists of *per se* and presumed unfair contractual terms. With the DA's contractual regime in mind, it is important for RE4DY contracting parties to carefully consider the fairness of provisions that grant them exclusive rights over data or that limit another party's use of its own data. In addition, it is to be kept in mind that the DA forbids anticompetitive practices that see a user, data holder, or third-party data recipient using shared or generated data to develop competing products or derive economic insights with respect to another entity in the data sharing arrangement (Arts. 4.10, 4.13, 5.6, and 6.2.e DA).

Finally, the broader RE4DY industrial data space and its pilot use cases which foresee the creation of data spaces should carefully note the interoperability requirements for data space participants in DA Chapter VIII. These requirements are listed under Article 33.1 DA and include: (i) describing, where possible in a machine-readable format, dataset contents, use restrictions, licenses, data collection methodologies, data quality, and uncertainty; (ii) describing the data structures, data formats, vocabularies, classification schemes, taxonomies, and code lists; (iii) describing the technical means to access data (e.g., APIs) and their terms of use and quality of service, so as to enable automatic access and transmission of data between parties; (iv) where applicable, providing the means to enable interoperability of tools for automated execution of data sharing agreements, such as smart contracts. It is therefore advisable that these requirements are taken into account and responsibilities for them are established in a data space's cooperation agreement or its individual service-level agreements. RE4DY actors may also look forward to the eventual adoption of harmonised standards by the Commission, since Article 33.3 DA establishes that entities complying with such harmonised standards are assumed to automatically be in conformity with Article 33.1 DA's essential requirements for data space participants.

3.1.3 The Data Governance Act

The Data Governance Act[10] came into force on 23 June 2022 and became applicable on 24 September 2023. The DGA seeks to build trust in data sharing among individuals and undertakings by establishing three separate legal regimes. The first of these is a legal regime governing the conditions, fees, and procedures for re-using data held by public sector bodies that is protected on the grounds of confidentiality, intellectual property rights, or personal data protection (Chapter II). The DGA's second regime governs data intermediation service providers (DISPs) (Chapter III). The third regime of the DGA governs data altruism, defined as the voluntary sharing of data for no reward beyond compensation of costs and in pursuit of objectives of general interest (Chapter IV). Of these regimes, the provisions on DISPs are particularly relevant for the Digital 4.0 Continuum.

The DGA considers data intermediation services to be vital enablers of common European data spaces due to their non-discriminatory orchestration of data-driven ecosystems. Correspondingly, data intermediation service providers face demanding requirements concerning their form and manner of operation. It is therefore important to consider which RE4DY use cases may involve the provision of data intermediation services and what consequences intermediation services might have for the actors involved.

First, it must be noted that some of the DGA's basic definitions differ from similarly named terms in the DA. For example, the DGA defines a 'data holder' as "a legal person or a natural

---

[10] Regulation (EU) 2022/868, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022R0868

person who is not a data subject with respect to the specific data in question, which has the right to grant access to or to share certain personal data or non-personal data" (Art. 2.8 DGA). Compared to the previously-mentioned definition of data holder under the DA, the DGA omits the requirement that data holders must have a right to 'use' the data that they hold – it is sufficient for an entity to be able to grant access or to share certain data. Furthermore, the DGA's concept of data holder applies to any sort of data, whereas data holders under the DA are only defined with respect to data from connected products or related services.

A 'data user' in the DGA refers "a natural or legal person who has lawful access to certain personal or non-personal data and has the right [. . .] to use that data for commercial or non-commercial purposes" (Art. 2.9 DGA. This is distinct from the concept of 'user' under the DA, which referred to a person who uses connected products and/or related services, rather than data as is the case under the DGA.

Consequently, it should be kept in mind that DGA and DA terminology is similar but not equivalent, and that DGA concepts are usually broader in scope than their DA relatives. Roles in a data ecosystem should be assigned separately with reference to each law's specificities.

A DIS is defined under Article 2.11 DGA as "a service which aims to establish commercial relationships for the purposes of data sharing between an undetermined number of data subjects and data holders on the one hand and data users on the other, through technical, legal or other means, including for the purpose of exercising the rights of data subjects in relation to personal data". This definition excludes from its scope services that add substantial value to data via aggregation, enrichment, or transformation and which then directly license the newly resulting data to data users (Art. 2.11.a DGA). Also excluded are data sharing services offered by public sector bodies that do not aim to establish commercial relationships (Art. 2.11.d DGA) and services that focus on the intermediation of copyright-protected content (Art. 2.11.b DGA). Most notably, the definition also excludes services that are used by either a single data holder or by multiple legal persons in a closed group (Art. 2.11.c DGA).

The research on DISPs undertaken by KUL has resulted in the publication of an in-depth White Paper on the Definition of Data Intermediation Services (Bobev T., 2023). This subsection will highlight some of the main challenges in delineating DISPs that the White Paper identified. The first of these challenges comes with the use of the phrase "commercial relationships" in the DIS definition, which is a new and less clearly defined concept in EU law than formulations such as "business" and "economic activity". Consultations with the Commission reveal that this divergence in terminology was unintentional and "commercial relationships" should be seen as synonymous with "economic activities".

Another, more persistent challenge stems from the notion that DISPs intermediate between "an undetermined number" of entities rather than within closed groups. There is yet no clear threshold or test to distinguish between an open and a closed group in practice. Article 2.11.c DGA gives the following examples of closed groups: "supplier or customer relationships or collaborations established by contract, in particular those that have as a main objective to ensure the functionalities of objects and devices connected to the Internet of Things". Yet, these examples are illustrative rather than exhaustive – they do not reflect the full gamut of closed groups that may be found in data sharing ecosystems. Furthermore, not all contractual collaborations result in closed groups, leading to the question of what sort of contractual mechanism may distinguish closed groups from open ones (e.g., contractual accession following approval by current members, strict contractual novation, or another threshold entirely). On this matter, consultations with the Commission reveal that a helpful indicator of closed groups may indeed be the existence of a treaty-like mechanism that allows all current members of a data sharing ecosystem to decide on new users who can use the service, though this indicator may be of questionable utility in the case of ecosystems with a great number of service users. Conversely, a service provider who is free to decide who can use the intermediation service usually indicates the existence of an open group and therefore of a DISP, regardless of the contractual terms surrounding the service's use.

The obligations of a DISP are detailed under DGA Articles 11 and 12. Some obligations are relatively routine, such as Article 11 DGA's requirement to notify a competent authority of the planned intermediation service's parameters, including details of the DISP entity's identity and the specie of data intermediation being planned according to Article 10. Most of the remaining obligations are significantly more demanding, beginning with the requirement that a DISP must be a separate legal entity that does not offer services beyond those specifically facilitating the exchange of data (Art. 12.a DGA).

The data intermediation service must be offered under fair, transparent, and non-discriminatory terms that are unbundled from the use of any other service (Art. 12.f DGA), and the DISP may not use the intermediated data for any purpose other than to put it at the disposal of the data user (Art. 12.a DGA). The DISP is also expected to take appropriate measures to ensure the security (Art. 12.l DGA), confidentiality (Art. 12.j DGA), and interoperability (Arts. 12.d and 12.i DGA) of its intermediated data, as well as to ensure a reasonable continuity of its service in the event of insolvency (Art. 12.h DGA). Finally, DISPs intermediating between data subjects and legal persons are endowed with a fiduciary duty to act in the data subject's best interest (Art. 12.m DGA).

RE4DY pilots would be well advised to carefully consider whether a data intermediation service is appropriate for their use case. Considering that the pilots are close collaborations between closed groups or supplier/customer relationships, it would be easy to avoid falling under the scope of the DGA and having to spin off a DISP entity at first. However, as the data spaces mature and extend to include new actors along the digital value chain (as envisioned, for example, by the Avio Aero pilot), the issue of DISPs becomes more relevant. It is recommended that any actors who are unprepared or unwilling to oversee DISP entities consider avenues to contractually circumscribe their membership, pursuant to DGA's framework.

3.1.4 Intellectual Property Rights in Digital Twins and Data

Intellectual property rights are a key component of the Digital 4.0 legal framework, particularly in light of the fact that pilots have repeatedly pointed to data ownership as an issue that creates barriers to data sharing and innovation. To ensure that RE4DY project partners may operate in a trusted environment that stimulates data sharing, it is relevant to clarify how (intellectual) property rights apply to data and digital twins.

There is no generally accepted exclusive property right over information (Geiregat, 2022), (World Bank, 2021) (Thouvenin, 2021). In other words, it is not possible to claim 'ownership' over data in a strict sense of the word. Rather, it is possible to own the various intellectual property rights that attach to data, and therefore to exercise control over data via those rights. Yet, data and datasets are only part of RE4DY's true subject matter – digital twins.

As mentioned in 3.1.1, digital twins are composite subject matter that consist of multiple discrete components: input and output data, software, hardware, and AI models that underpin an ongoing simulation. Each of these components may individually qualify for protection via one or several IP rights. A variety of IP rights exist, though the rights most relevant for the digital twins' context include copyright, patent, the *sui generis* database right, and trade secrets.

The input and output data of a digital twin may be effectively protected either through the *sui generis* database right or through trade secrets. The *sui generis* database right is established under Directive 96/9/EC, which grants that the maker of a database the right to prevent extraction or re-utilization of the whole database or of substantial parts thereof (Art. 7.1 Database Directive). This right is contingent upon the database maker demonstrating a substantial investment in obtaining, verifying, or presenting the database's contents. Crucially, this right does not extend to databases whose data the maker has generated/created themselves.[11] Furthermore, this right has been restricted under Chapter X of the Data Act, where Article 43 establishes that the *sui generis* database shall not apply when data is obtained *or* generated by a connected product or related service. Beyond reiterating the lack of protection for databases comprised of generated data, this provision is important in that it renders even the data *obtained* by most IoT or 'smart' devices ineligible for protection by the *sui generis* database right. As a result, the utility of the *sui generis* database right is limited in RE4DY's use cases, where smart devices equipped with sensors are expected to generate novel datasets.

Another recourse is to protect input or output data via trade secrets. Trade secrets, as defined under Article 2.1 of Directive (EU) 2016/943 (Trade Secrets Directive or 'TSD') constitute information that is (i) secret, i.e. "not, as a body or in the precise configuration and assembly of its components, generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question"; (ii) of commercial value due to being secret; (iii) subject to reasonable steps, given the circumstances, to be kept secret. Problematically, current scholarship favours the position that individual datums and unprocessed ('raw') machine-generated data cannot be protected by trade secrets (Aplin T., 2023). Datasets composed of derived or inferred data, however, may warrant trade secret protection and remain valid objects of trade secret protection. Interestingly, the Data Act, as discussed previously in this document, anticipates and mandates appropriate measures to protect trade secrets inherent in data falling under scope, despite its scope only encompassing raw data generated by connected products, which, as stated above, is typically considered ineligible for trade secrets protection. Between the ambiguities introduced by the Data Act, the fact that trade secrets are only truly enforceable *ex ante*, and the practical difficulties involved in controlling and tracing information, as well as in detecting and proving misappropriations, trade secrets are a relatively unpopular form of legal protection for data (Aplin T., 2023). It seems more practical and reliable for RE4DY project partners seeking to assert control over digital twin data to do so via contract law. The forthcoming industrial agreements

[11] CJEU Case C-203/02, *The British Horseracing Board v William Hill Organization Ltd.*, paras 31-33. ECLI:EU:C:2004:695

proposed within the RE4DY project are intended to provide a tool that helps address the need for dependable and clear assignments of rights over data.

Copyright is a possible avenue for RE4DY project partners to assert exclusive control over the software components of digital twins. Pursuant to Articles 1.1-1.2 of Directive 2009/24/EC (Computer Programs Directive), the expression of an original computer program is protected as a literary work within the meaning of the Berne Convention. As such, the expression of a computer program's source code, object code, architecture, preparatory materials and design work, as well of as its various audio-visual elements and interfaces are all protected under copyright. Copyright does not, however, protect ideas and principles (including algorithms) that underlie computer programs, per Article 1.2 and Recital 11 of the Computer Programs Directive.

Finally, patents represent a promising protection modality for 'inventions' that are at once novel, inventive, and industrially applicable. A patent therefore allows a digital twin's particular combination of hardware, simulation, models, and area of application to benefit from an exclusive right. The European Patent Office has confirmed via case law (EPO EBA, G 0001/19) and guidelines (EPO Guidelines for Examination, Part G-II, 3.3.2) that computer-implemented simulations and machine learning models are patentable inventions in principle. Predictive maintenance is, furthermore, an EPO-recognized field of AI technical application, which contributes to the technical character and, therefore, to the inventiveness of a digital twin patent claim.[12]

In the RE4DY digital twin context, patents are therefore expected to constitute a key mode of intellectual property protection. However, unlike copyright protection, which automatically attaches to its subject matter upon the creation of a work of authorship, patents are granted following a successful application. Patent applications are considered on a case-by-case basis following a detailed consideration of their claims and documentation, and it is therefore not possible to speak broadly about the patentability of RE4DY pilot digital twins, aside from the fact that both their field of application and their basic nature as patentable subject matter offer good prospects for a successful patent claim. Patents also carry an added advantage for 4IR projects such as RE4DY in that they combine exclusive legal protections with public visibility, transparency, and auditability, all of which are vital for the wider uptake of Industry 4.0 paradigms and the data spaces and data sharing ecosystems that underpin 4IR.

Whatever IP rights may attach to RE4DY digital twins and their data, tracking of IP rights over multilateral digital value chains remains a challenge. To reliably track IP rights across complex data transactions and transformations, semantic interoperability must extend to the legal dimension. A machine-readable legal ontology of IP rights would be the first step toward such interoperability and is therefore under development, as will be discussed in Section 3.2.

3.2 Task force setup and results

An Extended Task Force on Innovation and Standardisation that includes eight partners and two associate US partners (IOF/ASU and NIST) has been established. The Task Force has been following closely legal and regulatory developments, such as the Data Act and Data Governance Act. It is also developing solutions for the ontology data layer to be

[12] See, e.g. https://www.epo.org/en/news-events/in-focus/ict/artificial-intelligence

included in the DaaP marketplace. Moreover, it has designed a prototype of dynamic interface visualisations for MVN knowledge management and improved decision making, including sustainability management including recycling and reuse, vertical and horizontal alignment of knowledge management using ontologies, and enterprise integration asset management.

Currently in development by the Extended Task Force are the following innovations:

- Development of a VW Knowledge Graph for internal logistics (lead: IOF/UiO)
- Alignment with Data Act, Data Governance Act and continuous development of a Legal Ontology of IP Rights (lead: KU Leuven)
- Development and testing of dynamic interfaces to assist decision-making – MVN sustainability, knowledge and asset management (lead: ICF)
- LLM-driven supply chain Risk Management and Resilience Ontology (lead: Chalmers)
- AI-assisted Asset Management with IP and business landscaping (lead: ICF)
- Asset-centred & Integrating Pilot Platform (lead: ICF)

The objective of the work on a VW Logistics Knowledge Graph (KG) is to develop and document implementation of an internal Knowledge Graph, and an ontology system to address manufacturing internal logistics bottlenecks. This stream of our work relies on shopfloor components and relationships/dependencies – a subsection of production steps where bottlenecks occur to evaluate KG implementation. The expected added value is an improved understanding, security and analytical capabilities as integration of ontologies and Knowledge Graphs increases the accuracy of analysis of big data instances and improves the quality of decisions. Currently, the efforts have been focused on VW requirements specifications, alignment of the TEF setup at SSF with VW use case scenario, development of the Information Modelling Framework by UiO, alignment of IMF with existing ontologies (IOF Core, SCRO), and building extensions of existing ontologies to reflect use case requirements. A draft version of the KG is currently in progress. The ultimate objective is to create an internal logistical ontology combined with documented experiences and recommendations.

The objective of the work on the Legal Ontology of IP rights is to use it in application scenarios to assist tracking of IP and licensing in MVNs. This stream of research rests on the assumption that a novel Legal Ontology of IP Rights dataset would improve understanding, traceability, and legal certainty with regard to digital asset rights. The added value of this ontology would result in increased data-sharing results from greater trust and transparency in the data-sharing ecosystem. We aim to have one successful application of the legal ontology to track IP across assets was already achieved in Month 9 of the project. In the second half of the project the objective is to test this ontology in the context of the RE4DY demonstrators.

The objective of the Dynamic Interface & Visualisation work stream is to develop and implement innovative dynamic interfaces integrating sustainability, knowledge and asset management for improved decision-making. This facilitates more effective decision-making with efficient management of complex, heterogeneous data and assets with ontology-driven management systems, as well as interoperability harmonisation, quality and consistency provided by KGs, cost/risk reduction with standards for data checks and secure data transfer. Our interim KPIs included the delivery of two ontologies referenced +

one interactive environment prototyped and were reached already at Month 9 of the project and resulted in an accepted peer-reviewed paper (Calero C., 2006). The ultimate objective is to build several ontologies referenced by a MVN Knowledge Graph - logistics, legal assets and sustainability, accompanied by a visualisation of internal logistical flows, documented evaluations and recommendations.

The objective of the Risk Management and Resilience Ontology work stream is to test and evaluate innovative search mechanisms using LLMs and AI for replacement materials during manufacturing supply chain failures. This work relies on internal and external time series data contributing more effective and accurate search for replacement products. The added value lies in the adaptation of search mechanisms for sustainability, resilience, and circularity, and providing support for the design and development of products and production processes to be more resilient. The first iteration of a model has been provided, with an improved performance model expected to be ready for testing in the final project TEF.

The objective of the Asset Management with IP and Business Landscaping work stream is to enable access to the IP and business landscape, identifying potential innovation markets, gaps in current patenting and replacement suppliers. This work relies on WIPO, EPO and worldwide patenting databases (normalised) accessed via NLP query and API to provide a faster evaluation of the IP and patent business space within supply chains. The added value of this work lies in optimising search & AI supported workflows with Knowledge Graphs and faster localisation of potential collaboration partners and IP gaps. The first iteration of AI-assisted decision support in the interim reporting period is being followed by automatic support to decision-making workflows, optimised search with KG for knowledge-based interaction and narrowing down of options for the final project TEF.

The Asset-centred and Integrating Pilot Platform work stream aims to build an integrated GUI system and examples of interconnectivity with different data and service sources, including ontology resource access. This work relies on environment integrating data from supply chains, web-based tools and ontologies, JSON data exchange and direct database access, and access to different external multimedia and other data and information. Its added value lies in the combined meta-data and portfolio manager for resources in production and supply chains, and the enhanced ability to dynamically add interoperability to future services and data. The first pilot of user interface and core selection of information types/interoperability with external services and data sources is being followed by the development of ontology management tools access and supporting resources, interoperability with visualisation tools and processes.

Three of these innovations spearheaded by the Extended Task Force have been aligned with the Associate Partners from the USA (IOF, ASU and NIST):

- VW Knowledge Graph: Provides IOF Supply Chain WG with a real-life use case supported by data that can be used for extending and validating the Supply Chain Reference Ontology.
- Resilience Ontology: Aligned with an objective of the NSF Proto-OKN (Open Knowledge Graph) project where the resiliency of manufacturing supply chains will be assessed through semantic reasoning.
- Legal Ontology of IP rights: Aligned with an objective of the NSF Proto-OKN project where the patents will be analysed to identify and predict the trends and

trajectories in manufacturing technologies and assess the readiness of Small and Medium-Sized Manufacturers (SMMs) in adopting those technologies.

3.3 RE4DY Resilience and Legal ontologies

Data from different platforms and sources might be heterogeneous in syntax, schema, or semantics, which make data integration and data interoperability difficult. Ontology engineering and semantic modelling provide solutions to achieve semantic interoperability in a heterogeneous information system. The following sections introduce the concepts of ontology engineering and semantic modelling, as well as some relevant standards and languages. Several existing ontologies are reviewed in the end of this chapter.

3.3.1 Ontology Engineering

Ontology engineering is the general term of methodologies and methods for building ontologies. Ontology engineering refers to "The set of activities that concern the ontology development and the ontology lifecycle, the methods and methodologies for building ontologies and the tool suites and languages that support them." (Ameri, 2022) The results of ontology engineering provide domain knowledge representation to be reused efficiently and prevent waste of time and money which are usually caused by non-shared knowledge. It helps Information Technology (IT) to operate with interoperability and standardization.

3.3.2 Semantic Modelling

Ontology represents the nature of being, becoming, existence, and so on in the way of philosophy. One of the most well-known is: "ontology is an explicit, formal specification of a shared conceptualization of a domain of interest" (Gruber, 1993).

Semantic modelling can help defining the data and the relationships between entities (Calero C., 2006). An information model provides the ability to abstract different kind of data and provides an understanding of how the data elements are related. A semantic model is a type of information model that supports the modelling of entities and their relationships. The total set of entities in a semantic model comprises the taxonomy of classes that can be used to represent the real world.

The main objective of semantic modelling techniques is to define the meaning of data within the context of its correlation, and to model the domain world in the abstract level. The benefits of exploiting semantic data models for business applications are mainly as follows:

- Avoiding misunderstanding: by providing a clear, accessible, agreed set of terms, relations as a trusted source and discussions, misunderstandings can easily be resolved.
- Conduct reasoning: by being machine understandable and through the usage of logic statements (rules), ontologies enable automatic reasoning and inference which leads to automatic generation of new and implicit knowledge.
- Leverage resources: by extending and relating an application ontology to external ontological resources, via manual or automatic mapping and merging processes, the need for repetition of entire design process for every application domain is eliminated.

- Improve interoperability: semantic models can serve as a basis for schema matching to support systems' interoperability in close environments where systems, tools and data sources have no common recognition of data type and relationships.

Ontologies provide formal models of domain knowledge exploited in different ways. Therefore, ontology plays a significant role for various knowledge-intensive applications. Depending on corresponding languages, several different knowledge representation formalisms exist. However, they share a common set of components such as classes, relations, formal axioms and instances.

- Classes represent concepts, which are taken in a broad sense. For instance, in the Product Lifecycle domain, concepts are Life Cycle phase, Product, Activity, Resources, Event, and so on. Classes in ontology are usually organized in taxonomies through which inheritance mechanisms can be applied.
- Relations represent a type of association between concepts of the domain. They are formally defined as any subset of a product of n sets, that is: $R \subset C1xC2x...xCn$. Ontologies usually contain binary relations. The first argument is known as the domain of the relation, and the second argument is the range.
- Formal axioms serve to model sentences that are always true. They are normally used to represent knowledge that cannot be formally defined by the other components. In addition, formal axioms are used to verify the consistency of the ontology itself or the consistency of the knowledge stored in a knowledge base. Formal axioms are very useful to infer new knowledge.
- Instances are used to represent elements or individuals in an ontology.

As a Design Rationale (DR), ontology can be used as follows (Mizoguchi R., 1998):

- Level 1: Used as a common vocabulary for communication among distributed agents.
- Level 2: Used as a conceptual schema of a relational database. Structural information of concepts and relations among them is used. Conceptualization in a database is nothing other than conceptual schema. Data retrieval from a database is easily done when there is an agreement on its conceptual schema.
- Level 3: Used as the backbone information for a user of a certain knowledge base. Levels higher than this plays role of the ontology, which has something to do with "content".
- Level 4: Used for answering competence questions.
- Level 5: Standardization
  - Standardization of terminology (at the same level of Level 1)
  - Standardization of meaning of concepts
  - Standardization of components of target objects (domain ontology).
  - Standardization of components of tasks (task ontology)
- Level 6: Used for transformation of databases considering the differences of the meaning of conceptual schema. This requires not only the structural transformation but also semantic transformation.
- Level 7: Used for reusing knowledge of a knowledge base using DR information.
- Level 8: Used for reorganizing a knowledge base based on DR information.

3.3.3 Standards and Languages for Semantic Web

Ontology Markup Language (OML)

OML was developed at the University of Washington, is partially based on Simple HTML Ontology Extension (SHOE). In fact, it was first considered an XML serialization of SHOE. Hence, OML and SHOE share many features. Four different levels of OML exist: OML Core is related to logical aspects of the language and is included by the rest of the layers; Simple OML maps directly to Resource Description Framework (RDF); Abbreviated OML includes conceptual graphs features; and Standard OML is the most expressive version of OML. We selected Simple OML, because the higher layers don't provide more components than the ones identified in our framework. These higher layers are tightly related to the representation of conceptual graphs. There are no other tools for authoring OML ontologies other than existing general-purpose XML edition tools.

XML-based Ontology Exchange Language (XOL)

The US bioinformatics community designed XOL for the exchange of ontology definitions among a heterogeneous set of software systems in their domain. Researchers developed it after studying the representational needs of experts in bioinformatics. They selected Ontolingua (a Tool for Collaborative Ontology Construction) and OML as the basis for creating XOL, merging the high expressiveness of OKBC-Lite, a subset of the Open Knowledge Based Connectivity protocol, and the syntax of OML, based on XML. There are no tools that allow the development of ontologies using XOL. However, since XOL files use XML syntax, we can use an XML editor to author XOL files.

Simple HTML Ontology Extension (SHOE)

SHOE is a small extension to HTML which allows web page authors to annotate their web documents with machine-readable knowledge. SHOE makes real intelligent agent software on the web possible. HTML was never meant for computer consumption; its function is for displaying data for humans to read. The "knowledge" on a web page is in a human-readable language (usually English), laid out with tables and graphics and frames in ways that we as humans comprehend visually. Unfortunately, intelligent agents aren't human. Even with state-of-the-art natural language technology, getting a computer to read and understand web documents is very difficult. This makes it very difficult to create an intelligent agent that can wander the web on its own, reading and comprehending web pages as it goes. SHOE eliminates this problem by making it possible for web pages to include knowledge that intelligent agents can actually read.

Ontology Interchange Language (OIL)

OIL was developed in the OntoKnowledge project (www.ontoknowledge.org/OIL), permits semantic interoperability between Web resources. Its syntax and semantics are based on existing proposals (OKBC, XOL, and RDF(S)), providing modelling primitives commonly used in frame-based approaches to ontological engineering (concepts, taxonomies of concepts, relations, and so on), and formal semantics and reasoning support found in description logic approaches (a subset of first order logic that maintains a high expressive power, together with decidability and an efficient inference mechanism). OIL, built on top of RDF(S), has the following layers: Core OIL groups the OIL primitives that have a direct mapping to RDF(S) primitives; Standard OIL is the complete OIL model, using more primitives than the ones defined in RDF(S); Instance OIL adds instances of concepts and roles to the previous model; and Heavy OIL is the layer for future extensions of OIL. OILEd, Protégé2000, and WebODE can be used to author OIL ontologies. OIL's syntax is not only expressed in XML but can also be presented in ASCII. We use ASCII for our examples.

DARPA Agent Markup Language + OIL (DAML+OIL)

DAML+OIL has been developed by a joint committee from the US and the European Union (IST) in the context of DAML, a DARPA project for allowing semantic interoperability in XML. Hence, DAML+OIL shares the same objective as OIL. DAML+OIL is built on RDF(S). Its name implicitly suggests that there is a tight relationship with OIL. It replaces the initial specification, which was called DAML-ONT, and was also based on the OIL language. OILEd, OntoEdit, Protégé2000, and WebODE are tools that can author DAML+OIL ontologies.

Web Ontology Language (OWL)

OWL is the result of the work of the W3C Web Ontology Working Group. This language derived from DAML+OIL and, as the previous languages, is intended for publishing and sharing ontologies in the Web. OWL is built upon RDF(S), has a layered structure and is divided into three sublanguages: OWL Lite, OWL DL and OWL Full. OWL is grounded on Description Logics and its semantics are described in two different ways: as an extension of the RDF(S) model theory and as a direct model-theoretic semantics of OWL. Both have the same semantic consequences on OWL ontologies.

- OWL 2: OWL 2 is an extension and revision of OWL that adds new functionality with respect to OWL; some of the new features are syntactic sugar (e.g., disjoint union of classes) while others offer new expressivity. OWL 2 includes three different profiles (i.e., sublanguages) that offer important advantages in particular application scenarios, each trading off different aspects of OWL's expressive power in return for different computational and/or implementation benefits. These profiles are:
- OWL 2 EL: It is particularly suitable for applications where very large ontologies are needed, and where expressive power can be traded for performance guarantees.
- OWL 2 QL: It is particularly suitable for applications where relatively lightweight ontologies are used to organize large numbers of individuals and where it is useful or necessary to access the data directly via relational queries (e.g., SQL).
- OWL 2 RL: It is particularly suitable for applications where relatively lightweight ontologies are used to organize large numbers of individuals and where it is useful or necessary to operate directly on data in the form of RDF triples. OWL 2 ontologies: The Direct Semantics that assigns meaning directly to ontology structures and the RDF-Based Semantics that assigns meaning directly to RDF graphs.

Resource Description Framework (RDF)

RDF, developed by the W3C for describing Web resources, allows the specification of the semantics of data based on XML in a standardized, interoperable manner. It also provides mechanisms to explicitly represent services, processes, and business models, while allowing recognition of nonexplicit information. The RDF data model is equivalent to the semantic networks formalism. It consists of three object types:

- Resources are described by RDF expressions and are always named by URIs plus optional anchor IDs
- Properties define specific aspects, characteristics, attributes, or relations used to describe a resource

- Statements assign a value for a property in a specific resource (this value might be another RDF statement)

The RDF data model does not provide mechanisms for defining the relationships between properties (attributes) and resources—this is the role of RDFS. RDFS offers primitives for defining knowledge models that are closer to frame-based approaches. RDF(S) is widely used as a representation format in many tools and projects, such as Amaya, Protégé, Mozilla, SilRI, and so on.

According to W3C, RDF model has advantages as follows:

- The RDF model is made up of triples: as such, it can be efficiently implemented and stored; other models requiring variable-length fields would require a more cumbersome implementation.
- The RDF model is essentially the canonicalization of a (directed) graph and has all the advantages (and generality) of structuring information using graphs.
- The basic RDF model can be processed even in absence of detailed information (an "RDF schema") on the semantics: it already allows basic inferences to take place, since it can be logically seen as a fact basis.
- The RDF model has the important property of being modular.

The union of knowledge (directed graphs) is mapped into the union of the corresponding RDF structures. Since RDF is a standard model for data interchange and is a W3C recommendation designed to standardize the definition and use of metadata-descriptions of Web-based resources, it is well suited to representing data. As knowledge representation, when it comes to semantic interoperability, RDF has significant advantages (Noy NF, 2001): The object-attribute structure provides natural semantic units because all objects are independent entities. A domain model—defining objects and relationships—can be represented naturally in RDF. To find mappings between two RDF descriptions, techniques from research in knowledge representation are directly applicable.

Rules are widely recognized to be a major part of the frontier of the Semantic Web, and critical to the early adoption and applications of knowledge-based techniques in e-business, especially enterprise integration and B2B e-commerce. This includes Knowledge Representation (KR) theory and algorithms; mark-up languages based on such KR; engines, translators, and other tools; relationships to standardization efforts; and, not least, applications. Interest and activity in the area of Rules for the Semantic Web has grown rapidly over the last years.

Known rule systems fall into three broad categories: first-order, logic-programming, and action rules. These paradigms share little in the way of syntax and semantics. Moreover, there are large differences between systems even within the same paradigm.

Rule Interchange Format (RIF)

RIF is a W3C supported standard for exchanging rules among rule systems and in particular, among Web Rule Engines. RIF is focused on exchange rather than on trying to develop a single one-fits-all rule language because, in contrast to other semantic web standards, such as RDF and OWL, it is clear by the involved working groups that a single language would not satisfy the needs of many popular paradigms for using rules in knowledge representation and business modelling. But even rule exchange alone is recognized as a daunting task.

Regarding RIF, the approach taken by the working group was to design a family of languages, called dialects, with rigorously specified syntax and semantics. The family of RIF dialects is intended to be uniform and extensible. RIF uniformity means that dialects are expected to share as much as possible of the existing syntactic and semantic apparatus. Extensibility here means that it should be possible for motivated experts to define a new RIF dialect as a syntactic extension to an existing RIF dialect, with new elements corresponding to desired additional functionality. These new RIF dialects would be non-standard when defined but might eventually become standards. Because of the emphasis on rigor, the word format in the name of RIF is somewhat of an understatement. RIF in fact provides more than just a format. However, the concept of format is essential to the way RIF is intended to be used. Ultimately, the medium of exchange between different rule systems is XML, a format for data exchange. Central to the idea behind rule exchange through RIF is that different systems will provide syntactic mappings from their native languages to RIF dialects and back. These mappings are required to be semantics-preserving, and thus rule sets can be communicated from one system to another provided that the systems can talk through a suitable dialect, which they both support. The RIF Working Group has focused on two kinds of dialects: logic-based dialects and dialects for rules with actions. Generally, logic-based dialects include languages that employ different types of logic, such as first-order logic (often restricted to Horn logic) or non-first-order logics underlying the various logic programming languages (e.g., logic programming under the well-founded or stable semantics). The rules-with-actions dialects include production rule systems, such as Jess, Drools and JRules, as well as reactive (or event-condition action) rules, such as Reaction RuleML. Due to the limited resources of the RIF Working Group, it defined only two logic dialects, the Basic Logic Dialect (RIF-BLD) and a subset, the RIF Core Dialect, shared with RIF-PRD; the Production Rule Dialect (RIFPRD) is the only rules-with-actions dialect defined by the group. Other dialects are expected to be defined by the various user communities.

Rule Markup Language (RuleML)

RuleML constitutes a family of Web rule languages which contains derivation (deduction) rule languages, which themselves have a web-based Datalog language as their inner core. Datalog RuleML's atomic formulas can be (un)keyed and (un)ordered. Inheriting the Datalog features, Hornlog RuleML adds functional expressions as terms. In Hornlog with equality, such misinterpreted (constructor-like) functions are complemented by interpreted (equationdefined) functions. These are described by further orthogonal dimensions "single- vs. setvalued" and "first- vs. higher-order". Combined modal logics apply special relations as operators to atoms with a misinterpreted relation, complementing the usual interpreted ones (Boley H, 2010).

RuleML is a markup language developed to express both forward (bottom-up) and backward (top-down) rules in XML for deduction, rewriting, and further inferential-transformational tasks. A number of markup languages that are defined as part of RuleML are the following:

- Mathematical Markup Language17 (MathML)
- DARPA Agent Markup Language18 (DAML)
- Predictive Model Markup Language19 (PMML)
- Attribute Grammars in XML20 (AG-markup)

- Extensible Stylesheet Language Transformations21 (XSLT)

Semantic Web Rule Language (SWRL)

The Semantic Web Rule Language (SWRL) is based on a combination of the OWL DL and OWL Lite sublanguages of the OWL Web Ontology Language with the Unary/Binary Datalog RuleML sublanguages of the Rule Markup Language (RuleML). The proposal extends the set of OWL axioms to include Horn-like rules. It thus enables Horn-like rules to be combined with an OWL knowledge base. A high-level abstract syntax is provided that extends the OWL abstract syntax described in the OWL Syntaxes22 document. An extension of the OWL model-theoretic semantics is also given to provide a formal meaning for OWL ontologies including rules written in this abstract syntax.

The proposed rules are of the form of an implication between an antecedent (body) and consequent (head). The intended meaning can be read as: whenever the conditions specified in the antecedent hold, then the conditions specified in the consequent must also hold. Both the antecedent (body) and consequent (head) consist of zero or more atoms. An empty antecedent is treated as trivially true (i.e. satisfied by every interpretation), so the consequent must also be satisfied by every interpretation; an empty consequent is treated as trivially false (i.e., not satisfied by any interpretation), so the antecedent must also not be satisfied by any interpretation. Multiple atoms are treated as a conjunction. Note that rules with conjunctive consequents could easily be transformed (via the Lloyd-Topor transformations) (Lloyd, 2012) into multiple rules each with an atomic consequent.

Atoms in these rules can be of the form C(x), P(x,y), sameAs(x,y) or differentFrom(x,y), where C is an OWL description, P is an OWL property, and x,y are either variables, OWL individuals or OWL data values. An XML syntax is also given for these rules based on RuleML and the OWL XML presentation syntax. Furthermore, an RDF concrete syntax based on the OWL RDF/XML exchange syntax is presented. The rule syntaxes are illustrated with several running examples.

3.3.4 Initial outline of the manufacturing resilience ontology

The resilience ontology is under development and will only be briefly defined here.

Figure 13-14: Main classes in the manufacturing resilience ontology

The resilience ontology is built using the open source Protégé ontology editor and based on the basic formal ontology (BFO). This is currently being developed in collaboration the RE4DY Extended Task Force on Innovation and Standardization. The nomenclature of the classes was defined according to the Information Modelling Framework (IMF) Ontology 2023).

The resilience ontology defines the different classes that are important to consider for building manufacturing resilience *Figure 13*.

Specifically, all stakeholders in a supply chain need to jointly build resilience as the resilience of the supply chain is as strong as its weakest link. Here, the stakeholder was defined as an 'actor' where the production facility or plant, along with the distribution centre and market customer were defined according to (Calero C., 2006), (Mizoguchi R., 1998) with the addition of 'suppliers' as a key stakeholder. The size of the business was also considered with a distinction between large MNCs and smaller SMEs. Next, the type of manufacturing domain is also important to consider as resilience can vary and have different objectives there. For the context of the RE4DY project, we have four manufacturing domains, and these are highlighted in *Figure 13*. We have only considered discrete manufacturing in this project and cases in the process industry will also need to be evaluated as resilience could differ there.

To address risk management aspects for building manufacturing resilience, four classes were defined: 'ResilienceMeans', 'Risk', 'RiskAttributes' and 'Disruption'. These are shown in *Figure 15* along with the expanded sub-classes under each.

Figure 15-16: Classes related to risk management

Resilience can be brought about (the means) by dynamic capabilities (Chari A., 2024) in three temporal stages of anticipation, coping and adaptation and corresponding capabilities under each of these stages have been defined in the resilience ontology. These capabilities provide a starting point to build resilience and can belong to more than one stage based on the type of resilience practice implemented at the company. More details are documented in an unpublished journal paper (Chari A., 2024). Risks are unintended events that can pose a threat to a company and in some cases create a disruption if the risk is not handled appropriately. These disruptions can impact the organisation at various levels and have different source categories and types (Padhi, 2024) as indicated in the ontology. Risks can also have varying levels of frequency (how often they occur) and severity (impact of disruption).

For the next stage of the resilience ontology development, object properties and relationships between the classes will be defined.

3.3.5A Legal Ontology of IP Rights definition

The goal of the intellectual property rights ontology (IPRO) is to assist in asset management and facilitate trust and legal certainty in multilateral digital value chains. The IPRO is designed to model six main subject matter domains: (i) the intellectual property rights recognized within the European Union's legal framework, (ii) the substantive criteria that delineate each IPR, (iii) the data requirements that must be fulfilled to register certain IPR, (iv) the most common families of open software and data licenses, (v) the processes and requirements that together form software and data license frameworks of permissions and prohibitions, (vi) the roles and classes of actors implicated in the exercise of IPR. The IPRO is designed with a view toward its eventual integration into asset management software and automated compliance solutions.

The IPRO is built in the OWL 2 Web Ontology Language using the open source Protégé 5 ontology editor developed by the Stanford Center for Biomedical Informatics Research. With the assistance of the State University of New York at Buffalo, the IPRO has been integrated into the top-level Basic Formal Ontology (BFO), so as to allow for interoperability with other domain ontologies such as Chalmers' Resilience Ontology.

The IPRO is being developed offline and shared between members of the RE4DY Extended Task Force on Innovation and Standardization. Once the ontology is completed, it will be published online with the support of ICF and INNO, where it will be freely accessible for the general public. The publication is currently planned to take the form of a GitHub Page (as opposed to a repository).

Intellectual Property Rights

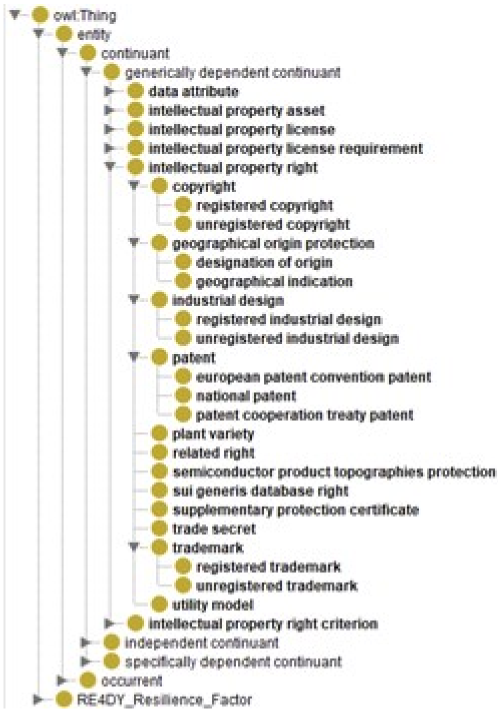

Figure 17: IPR in IPRO

The IPRO is grounded on the identification of the various IPR in the European legal framework. It distinguishes subtypes of IPR based on both substantive legal criteria (e.g., designation of origin vs geographical indication) and based on procedural differences (e.g., registered copyright vs unregistered copyright). Insofar as there exist multiple pathways to being granted a single right (e.g., a patent), each of these pathways has also been identified so that their criteria may be described.

Intellectual Property Criteria

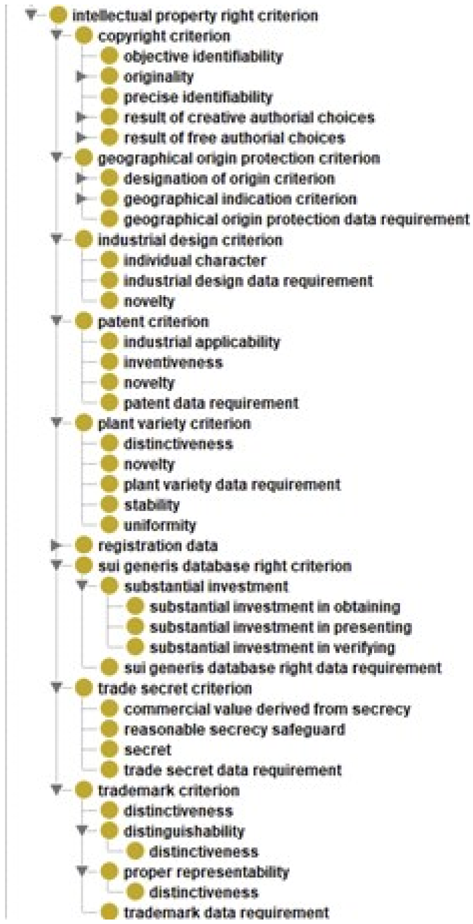

Figure 18: IPR Criteria in IPRO

IPRO defines the criteria that must be present for each IPR to be realized. Different criteria have different relationships with their corresponding rights. For example, whereas a *sui generis* database right requires substantial investment in *at least one* of obtaining, presenting, or verifying the data, a patent requires novelty *and* inventiveness *and* industrial applicability.

Minimum IPR Data Requirements

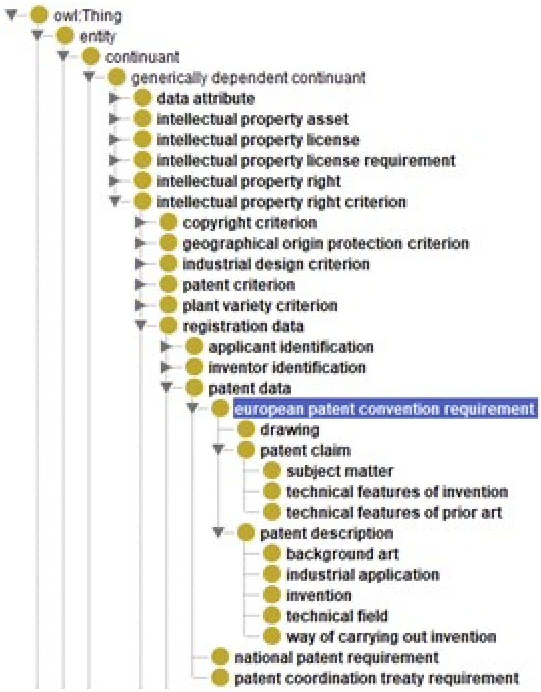

Figure 19: Minimum Data Requirements for IPR

The IPRO identifies the minimum data elements that must accompany each IPR's registration. The IPRO does not exhaustively define all the optional data that *may* accompany IPR registration. Not all IPR must be registered, so not all IPR have minimum data requirements described.

Common Open License Families

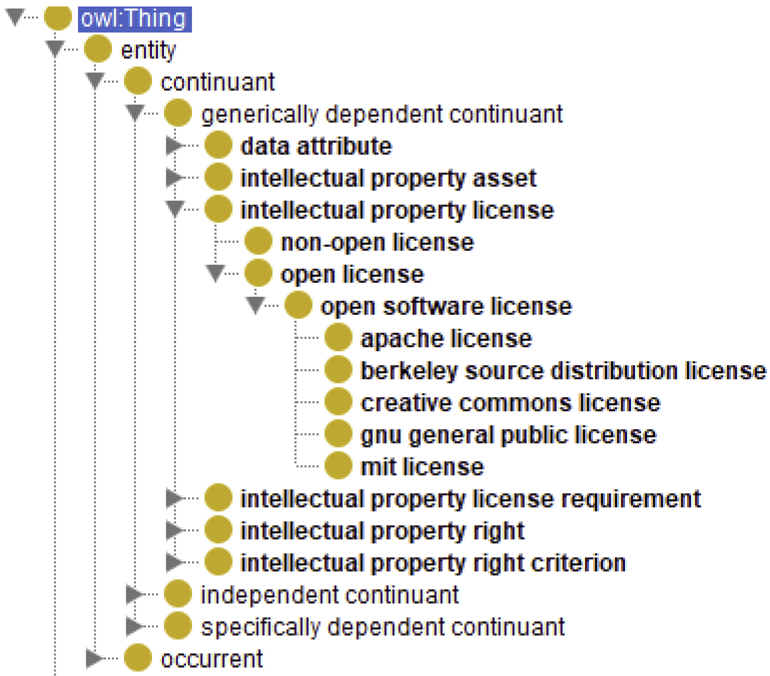

Figure 20: WIP list of Open License Families

The IPRO contains classes for the most common open software licenses, whose terms frequently carry over into open data licenses. Each class entity constitutes a license family, whereas specific licenses (e.g., GNU GPL 3.0 or Creative Commons Attribution 4.0 International) are individuals within the ontology.

IPR Processes and License Requirements

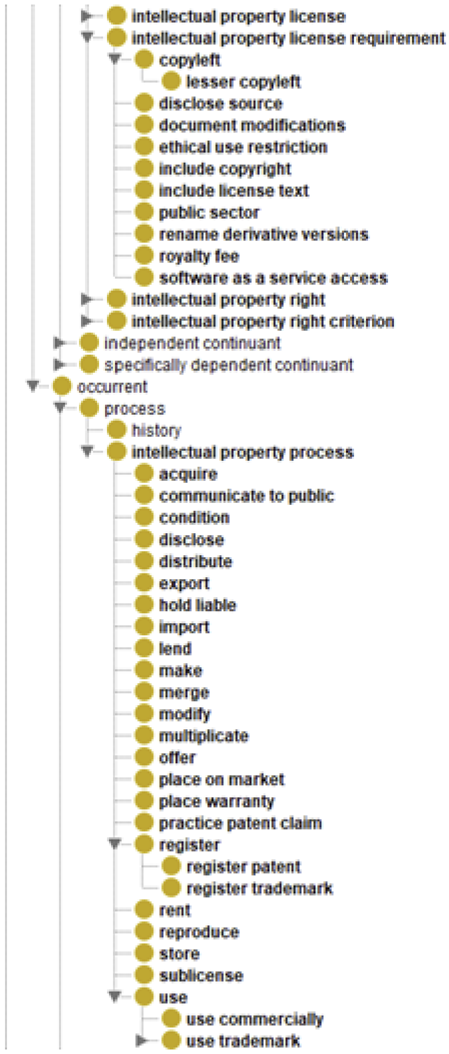

Figure 21: IP Processes and (Software) License Requirements

The IPRO broadly defines IP-relevant processes implicated in IP law. These processes are not limited to the field of software or data licenses.

On the other hand, the IPRO also delineates the specific requirements imposed by open software licenses, as formulated under the European Commission's Join up Licensing Assistant.

IP Roles and Actors

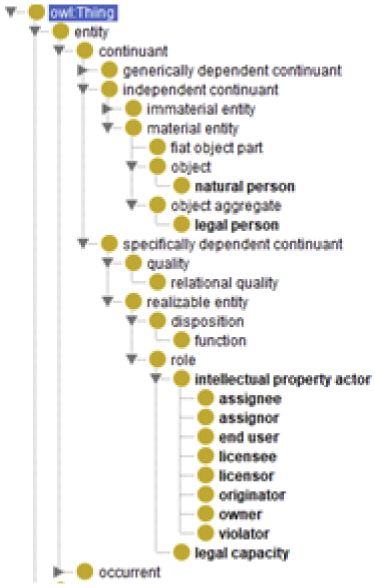

Figure 22: WIP Roles and Actors in Intellectual Property

The IPRO contains a basic schematic of possible roles and types of persons that may be assigned to entities that interface with the field of intellectual property.

4 RE4DY RA Adoption

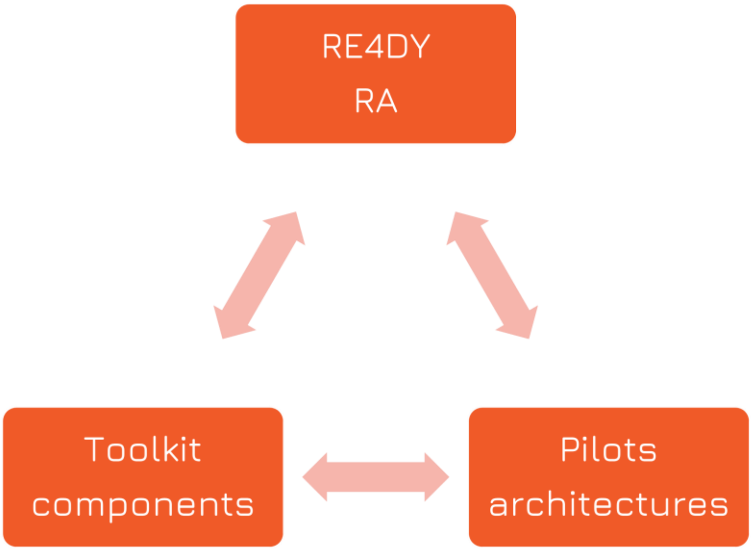

Figure 23: RA and mutual influences

The initial version of this framework, described in D2.2, introduced an iterative approach to incorporate valuable feedback and insights gathered during the development lifecycle. Illustrated in the diagram above are the clear connections between the Reference Architecture (RA), Toolkit components, and Pilots' architectures. The Reference Architecture served as the foundation for shaping the toolkit components, which are essential elements in the various architectures designed for the pilots. The double-headed arrows signify mutual influences, originating from the pilot needs impacting the RA design. The RA, with its building blocks, then defined the components necessary to fulfil the requested capabilities, leveraging a selected base of components to avoid starting from scratch or reinventing the wheel. Similarly, the toolkit components were influenced by the pilot's requirements, completing the loop within testbed environments. This iterative process ensures that the framework continuously evolves to meet the dynamic needs of the users and stakeholders, fostering a more robust and adaptable system.

This approach ensured that each pilot could design their own architecture tailored to their expectations, starting from an adequate level of abstraction, like the reference architecture (RA) of RE4DY. Such RA provides a common basis for all pilots to address the project challenges. It can be identified in each pilot's design, for example, through the components implementing its building blocks, as illustrated in the following examples.

4.1 RA in the Toolkit definition

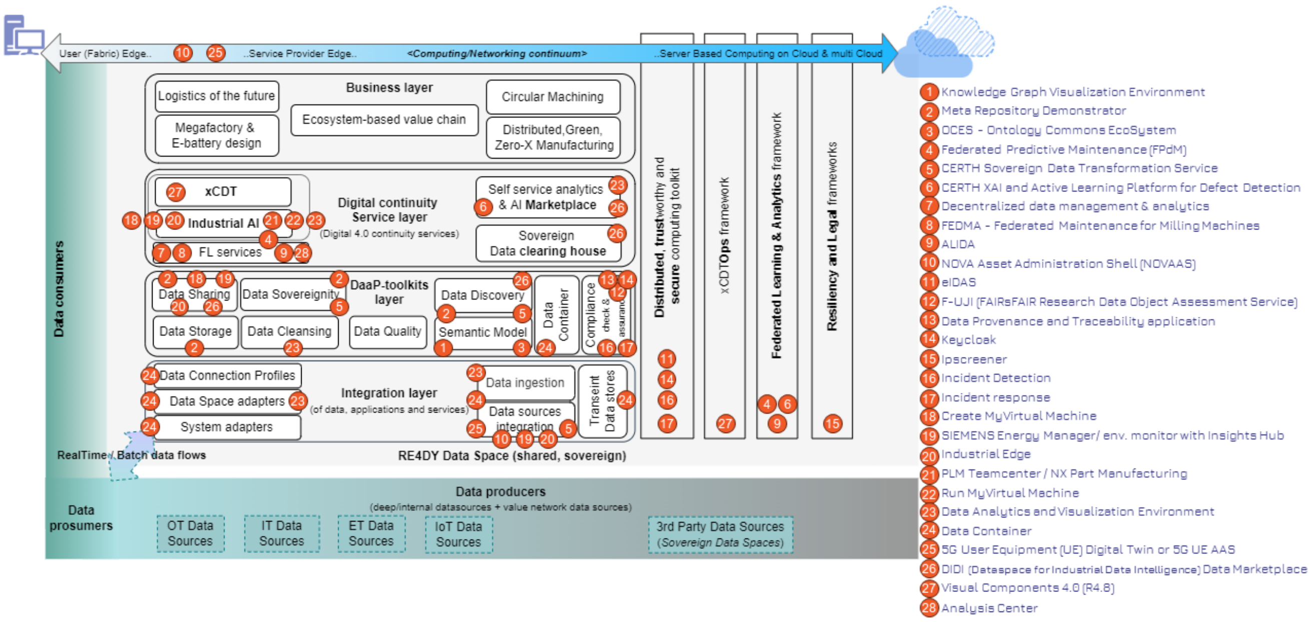

Figure 24: RA & RE4DY Toolkit components

The image above illustrates the thoughtfully constructed data-driven architecture designed for the project, taking into account its unique requirements, aspirations, and existing foundations (based on existing blueprints in literature as well as on key results from previous projects). This architectural framework consists of four layers and four vertical dimensions, complemented by the computing networking continuum, aimed at enhancing decentralization and digital continuity. The capabilities provided and the methodology employed in developing this blueprint were detailed in D2.2, along with the motivations that led to its creation. These motivations, summarised below, include:

1. There's a need to strengthen decentralisation across infrastructures, services, and data planes.
2. Implement Digital Continuity across computing, networking, and deployment to ensure seamless utilization of the digital thread, regardless of the location of data and applications.
3. Embrace not only a data-driven approach but also fully leverage the concept of DaaP, offering a marketable digitization of the value chain (data space).
4. Ease natural convergence between Manufacturing and IT operations by integrating toolkits that cover the entire lifecycle of an Industrial Data Platform, from realization (design and development) to commissioning (integration and validation), and, finally, to operation (management and maintenance) of all software artifacts.

In this image, the architecture illustrates how its building blocks align with toolkit components, listed on the right side. This offers an implementation overview, effectively depicting the current status of the toolkit and its extent of coverage in implementing the various building blocks. As such, it doesn't have to be considered exhaustive or comprehensive. In fact, as long as iterations occur based on the schema shown earlier, new feedback naturally becomes available from the pilots, and new implementations become feasible to fulfil the RA building blocks.

4.2 RA instances in the Pilots' architectures

In the following subsections, practical examples of RA implementation through the different architectures designed in the pilots are described. These pilots are aligned with the four specific business cases of the RA Business Layer, which are of highest priority in the EU for the development of a common manufacturing data space. Table 1 in the Introduction section provides a map of these business cases with the corresponding pilots for clarity and convenience.

Each pilot aims to assess and to showcase the technologies provided by the RE4DY toolkit, developed as part of the WP3 tasks. Furthermore, all pilots explore custom-built vertical solutions tailored to their project's ambitions, utilizing these technologies, and addressed within WP5.

To illustrate how each pilot has implemented its own architecture, a standardized approach has been adopted across all pilots. Each pilot showcases its solution for specific Business Processes (BP) using two main types of images.

The first image depicts the pilot's connection with the RE4DY Reference Architecture (RA) and its building blocks. Implementations are listed in the legend on the right and mapped within the architecture design. Components of the toolkit are identified by numbers enclosed in orange circles by default, with green circles indicating components that have been adopted in the pilot.

The second type of image presents the pilot's implemented architecture, with the RE4DY toolkit list in the legend on the right facilitating the mapping of the implemented architecture to the toolkit components and then the RE4DY RA. In this type of image, only the numbers of the mapped components are enclosed in orange circles.

4.2.1 Connected Resilient Logistics Design & Planning

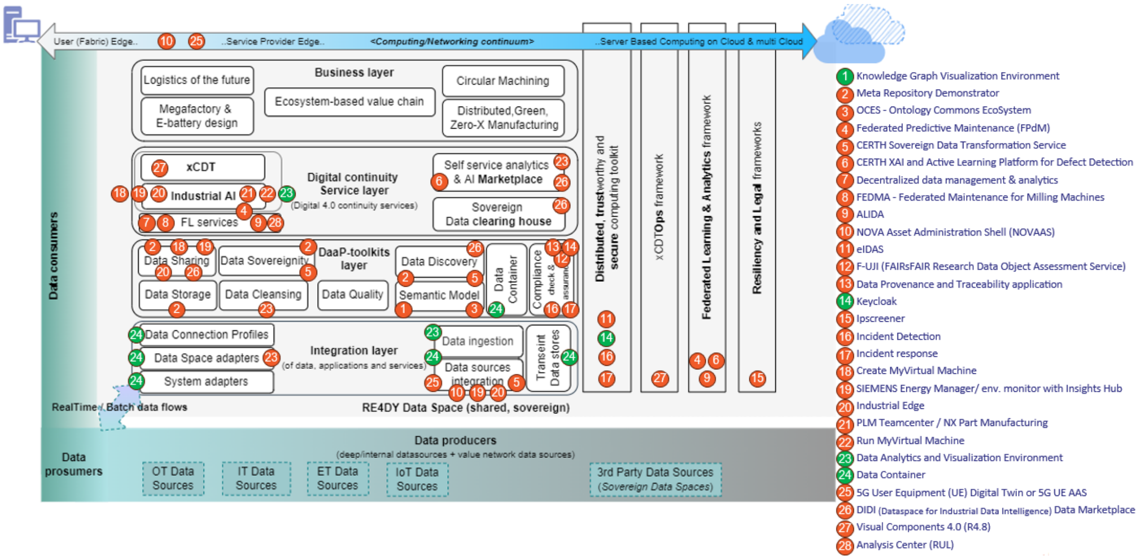

Figure 25: RA & toolkit components in the VWAE's pilot

The Connected Resilient Logistics Design & Planning will focus on the internal logistics processes. As documented in previous deliverables, it addresses the following main objectives:

BP1 - Autonomous Planning – Objective is to automate the analysis leading to the optimum scenarios.

BP2 – Shopfloor Implementation – Objective is to ease operation in the shopfloor through digitalization.

BP3 – Resource Optimization – Objective is to process equipment data to optimize the line feeding equipment efficiency.

Figure 26 depicts the architectural solutions designed for the three different business scenarios in this pilot.

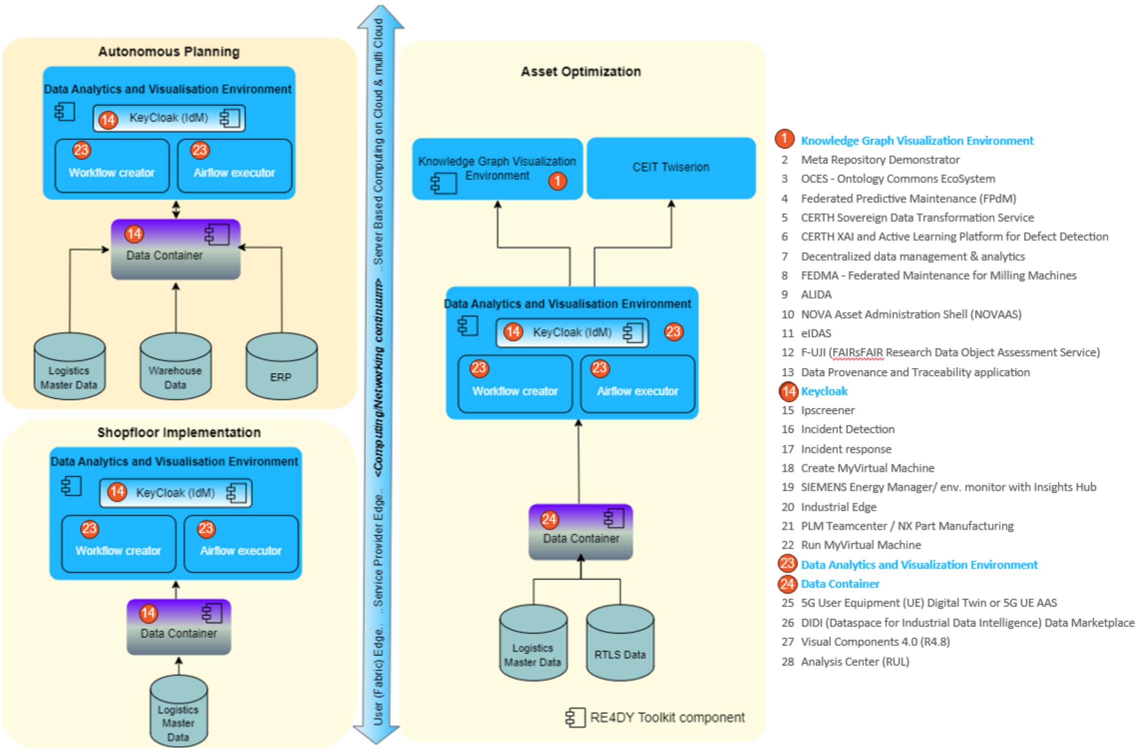

Figure 26: VWAE pilot architecture

In the picture, there are three main frames displaying the different solution implementations for each business process. The standard toolkit components are marked with an icon, as indicated in the legend.

On the first frame "Autonomous planning" the integration of 3 different datasets into the data container and the Data analytics and Visualisation environment is crucial to get analytics and run algorithms on the data to get the desired optimizations. To achieve this, it leverages the following components from the RE4DY toolkit:

KeyCloak (IdM): This toolkit component secures access to the pilot applications. It is embedded within the Data Analytics and Visualisation Environment.

Data Analytics and Visualisation Environment: This component enables the automation of data tasks such as ETL or the use of algorithms. This provides the environment to create workflows (workflow creator) and the infrastructure for them to run (Airflow executor). In this particular BP, it is used to load and merge the data automatically and apply algorithms to detect inefficiencies and suggest optimizations.

DataContainer: This component ensures access to data and metadata by incorporating controls, filters, and automatic conversion in the context of the DaaP concept experimentation.

The second frame "Shopfloor implementation" describes the connection of data to the shopfloor. This is achieved by communication from the Data analytics and visualisation Environment directly with the hardware on the shopfloor (E-Papers) in near-real-time.

KeyCloak (IdM): It serves the same functions as in the BP described before.

Data Analytics and Visualisation Environment: It serves the same functions as in the BP described before.

DataContainer: It serves the same functions as in the BP described before.

Finally, the third frame "Resource Optimization" describes the solution for BP3. This business process focuses on analysing data from the line-feeding equipment to understand what is happening in the shopfloor and allow for better planning and resource optimization.

Knowledge Graph Visualization Environment: This component will be used to query the knowledge graph created for this logistics processes.

KeyCloak (IdM): It serves the same functions as in the BP described before.

Data Analytics and Visualisation Environment: It serves the same functions as in the BP described before.

DataContainer: It serves the same functions as in the BP described before.

CEIT Twiserion: This component is used to create a digital twin of the logistics processes and line-feeding equipment to get insights and to allow "what-if" analysis.

Table 8: Main RA Components adoption in the pilot

Id	Component name	RA implemented BB (Building Block)	RA Layer
1	Knowledge Graph Visualization Environment		
14	KeyCloak	compliance check and assurance	DaaP-toolkits layer
23	Data Analytics and Visualisation Environment	Industrial AI and Self-service Analytics	Digital continuity service layer
24	Data Container	Data container	DaaP-toolkits layer
		ALL	Integration layer

4.2.2 Electric Battery Product/Production System Engineering

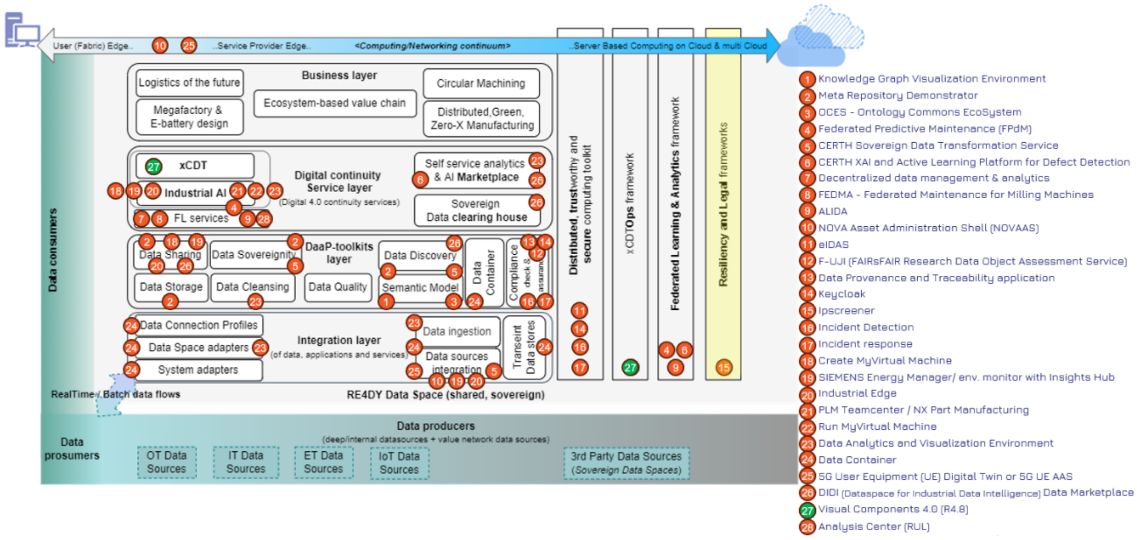

Figure 27: RA & toolkit components in the Fill/AVL pilot

The Electric Battery Product/Production System Engineering pilot operates in the context of manufacturing of traction battery modules and packs for e-mobility solutions. As documented in previous deliverables, it addresses the following three main challenges:

- The goal in RE4DY is the data-driven digital value network to decrease timely and cost-effectively e-battery package component complexity.
- It will concurrently accelerate product design workflows and highly customized flexible serial production of battery packages for OEM and TIER-1.
- The RE4DY engineering industrial data fabric will be capable of building data-driven active resiliency strategies over multi-vendor platforms and ecosystems of partners and suppliers.
- It will improve the industrial ability to adapt and respond to disruptions and unplanned events through “resilient-by-design” products and processes.
- The digital thread-data management methods to build resilient manufacturing networks for new e-battery technologies demand integration of engineering data space with engineering data fabric.

These challenges correspond to distinct business scenarios, each with its own objectives and expected benefits, which are:

BP1 - Agility

- Agility – Time to adapt to product change reduction (-30%)
- Speed to market – Lot-size-1 engineering (-15%)
- Less time designing new solutions (-10-15%)

BP2 - Sustainability

- Sustainability – Energy and resource efficiency increase (15%)
- Connect smart factory
  - Efficient connection of machines line

- Safety of workers
    - Zero emissions, zero overwork, zero injuries

BP3 - Customisaton

- Customisation – Battery package optimizations (+10-15%)

BP4 – Productivity OEE

- Availability – Unplanned downtime reduction (-20%)
- Service cost reduction (-15%)

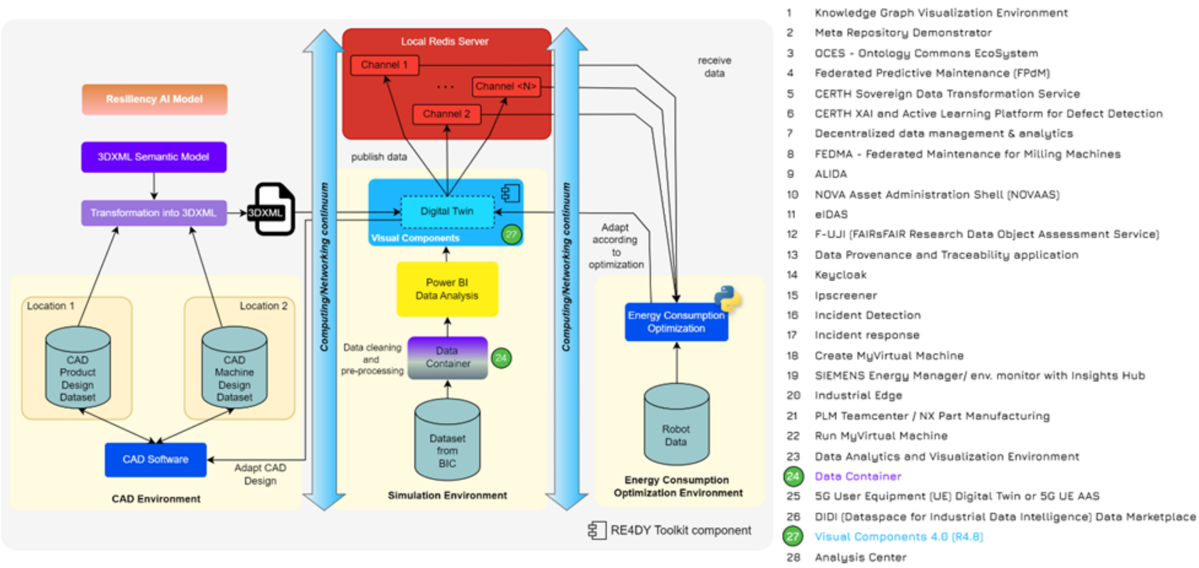

Figure 28: FILL/AVL pilot architecture

The figure above depicts the architecture used in this pilot. The central component in this architecture is Visual Components. It is used to process any data coming from three different environments, which can roughly be categorised in

- CAD environment
- Machine Data environment
- Robotic Energy Consumption Monitoring environment

The architecture should be further explained in the following paragraphs.

CAD Architecture

To create a digital twin in Visual Components, the user needs to provide CAD data from the factory, including information about machines and workplaces, as well as product CAD data. While machinery data is typically sourced from suppliers, product data is generated during the design process. Caused by the different sources and the fact that different customers use different CAD-software, a standardised file format was in demand.

Therefore, 3DXML was introduced. The XML based format was developed by Dassault Systèmes can easily be exported from CATIA and SOLIDWORKS, as well as its direct compatibility with Visual Components. The data within 3DXML files can be read and written using freeware tools such as the Windows Editor or Visual Studio. These files comply with a predefined structure with a set of tags that are commonly readable, while also allowing for extensions to provide additional product information. Furthermore, 3DXML files have a small file size, resulting in quicker loading of parts. All these properties are crucial for quick customisations on the product.

Machine Database Architecture

Machine data are generated during the production process and saved on a server which is managed with Microsoft SQL system. This dataset contains information from sensors, as well as predefined product information such as the production program and process data which is necessary for tracking the components throughout the entire production process.

All the data collected is then cleaned, processed and analysed with a system created specifically for the use in the Battery Innovation Center with Power BI. The results of the analysis, including the battery cell voltage values, which are crucial for the use in the product, and of the cycle time, as a database, are then transferred in Visual Components to get realistic values for failure quotes and varieties in cycle times. This enables a more accurate representation of the production process through the digital twin and, out of it, better forecasts for new production processes, leading to an earlier market placement.

Robotic Energy Consumption Monitoring Architecture

To compute the energy consumption with the use of third-party python frameworks a stand-alone application was necessary. In order to be able to calculate the energy consumption of robots, the path movements, or more precisely the axis angles, must first be known. To transfer these data from Visual Components (VC) to the stand-alone application, communication must be established between the python application and VC. There were initially two approaches to establishing this communication. Firstly, via a local client/server connection using the socket module and secondly via a pub/sub connection using Redis. Since a large number of robots sending data either require several servers or the data can be mixed up, the pub/sub connection was chosen in the architecture.

With this architectural information, it is now appropriate to delve into greater detail about some of the components.

Machine Data Analysis with PowerBI

As mentioned above, machine data is analysed with Power BI. This process should now be further explained. Five different datapoints saved at the server are most valuable for later usage in the digital twin. These are:

- Battery Cell Voltage
- Stacking Force
- Process Start Time
- Process End Time

The first two datapoints are crucial since these values are used to identify rejects. The battery cell voltage is measured before the stacking of the cells and must fall in a range defined by a minimum and a maximum tolerated value. These values are calculated individually for every cell depending on their last measurement before shipment. The stacking force serves as an indicator for the manufacturing tolerance of the cell carrier. If the stacking forces are too high, the cell diameter does not fit the cell carrier diameter properly.

With this knowledge where failures occur in the process and how they appear, along with the machine data, the digital twin can use a very accurate model for the calculation of probabilities to describe the production process.

The process starts and end times are needed to depict precise cycle times for each process and get a range for the variety of the cycle times of every single process. This information is essential for an accurate simulation. A high variety in the cycle times can lead to a need of buffers and therefore needs to be considered in terms of resiliency.

Visual Components

The Visual Components Toolkit is the main component of our architecture, this powerful simulation and offline programming software brings everything together. The Visual Components Toolkit allows to create a virtual model of the pilot to simulate different scenarios. CAD design decisions can be verified, and possible bottlenecks in the design and manufacturing process of the production line can be identified. Once the bottlenecks have been identified, they can be removed, and the optimized production line verified.

The software also improves the uptime for changes in product design, for example, different types of battery pouches can be tested and the necessary changes to the production line can be identified before the changes are implemented in the real world. Another useful aspect of the software is off-line programming for robots, which allows to check that the program does not cause collisions between the robot and its surrounding environment.

In the pilot, it also helps generating the kinematic motion data of the robot, on which the calculation of the energy consumption in the application is based.

Robotic Energy Consumption Monitoring

The custom-developed Python application allows the energy consumption of one or more robots to be monitored. This monitoring alone will allow engineers to get a feel for energy consuming movements and can help to reduce power consumption at an early stage. In the next stage of development, appropriate algorithms will automatically reduce power consumption.

The Application will include the following options:

- Optimization of the robot path motion.
- Optimization of energy consumption during braking.
- Reduction of energy consumption peaks.

The energy consumption will cause a certain trade-off with the cycle time, so this must be considered to find the optimal operating range of the robot.

The calculation of the energy consumption is based on kinematic movements that the Python application receives from other software. The advantage of the application being independent is that it can interact with software other than Visual Components. The robot parameters (e.g., mass, centre of gravity, inertia, friction coefficient, ...) are the other data needed for the calculation. As these data are not readily available, they must be determined by measurements, estimates based on similar robots or smaller models. In order to be able to use this data in other applications, a standardized file format was sought in which the robot parameters could be stored. After research, the *.urdf format, an *.xml format supported and developed by ROS, was chosen to describe robot models with a predefined structure. Some CAD systems, such as SolidWorks, also offer export of CAD data to this file format. These files can be edited and read using freeware tools in the same way as 3DXML files.

4.2.3 Integrated Machine Tool Performance Self Optimisation

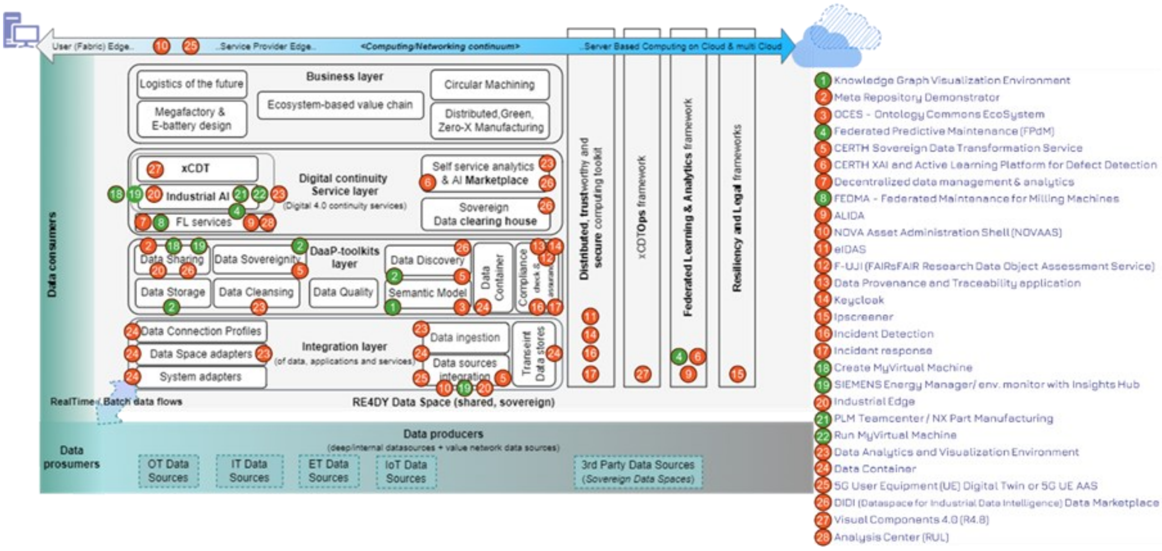

Figure 29: RA & toolkit components in the GF-FRAISA pilot

The GF Fraisa pilot implements services for tool and machine maintenance management for the case of milling technologies. As the scenario is deployed across the tool and machine lifecycle for high productivity and high precision applications, business process related to virtual planning and adaptive manufacturing and quality control are included. The challenges addressed are the following:

1. Selection of best tools for a given part manufacturing, with virtual simulation of manufacturing KPIs
2. Individual tool lifecycle management with AI prediction of tool wear for optimized tool recycling
3. Predictive maintenance of key machine components for guaranteeing high precision and maximize uptimes
4. On machine quality control of manufactured parts for adaptive manufacturing

Those challenges are associated with the corresponding business processes:

BP1 – Process Planning and Preparation

- Objective: Tool information available with CAM and machine conditions for process planning & simulation.
- Benefit: Selection of best tools and strategies for optimized machining processes.

BP2 – Tool Management and recycling

- Objective: Tool data integration for machine operation and Monitoring of tool status and timed recovery and refurbishing of tools with predictive solutions.
- Benefit: Tool recycling and refurbishing.

BP3 – Machine Maintenance

- Objective: Maintenance of critical machine components.
- Benefit: Monitoring of component status and timed warning, repairing or refurbishing process with predictive solutions.

BP4 – Adaptive Digital Manufacturing

- Objective: Machine Verification using metrology and advanced part alignment.
- Benefit: Automated in-machine metrology and feedback.

The following picture represents the architecture of the pilot for all the business processes, requiring specific modules related to federated learning FEDMA and FPdM as specific application for predictive maintenance.

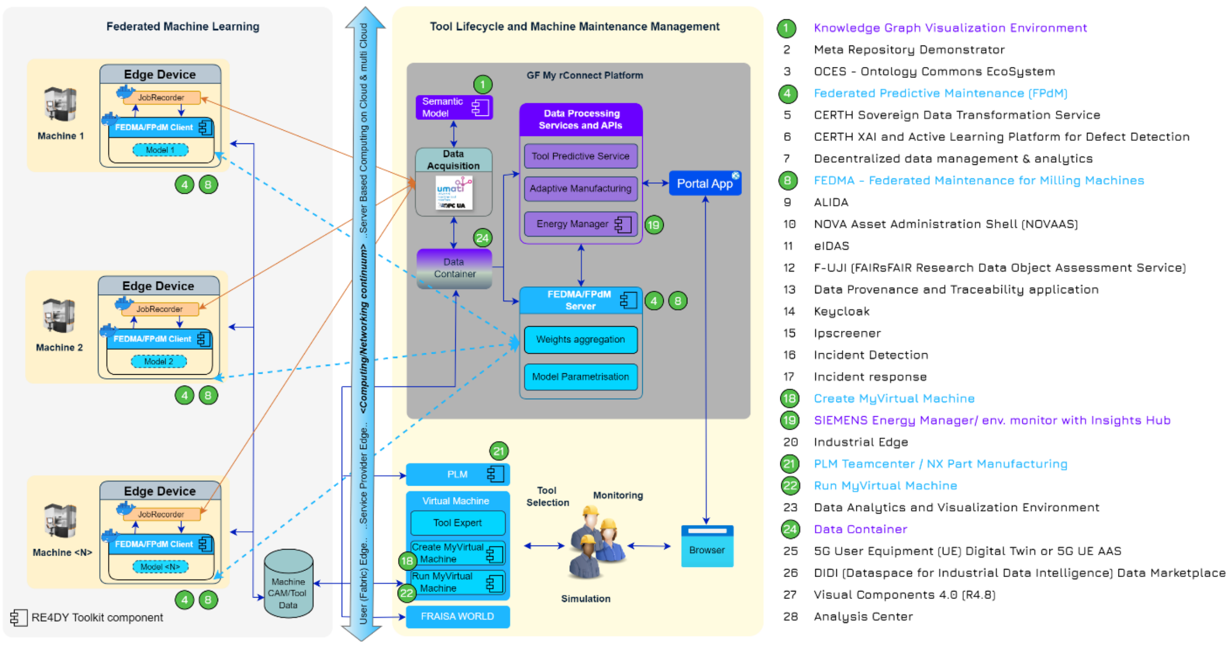

Figure 30: GF-FRAISA pilot architecture

Architecture Components and Description

Virtual Machine (Siemens/Fraisa)

Different components for setting up a virtual machine will be implemented in the Siemens Virtual environment to simulate the process based on CAD and CAM data integrated in the Siemens PLM with the selected machine. This will allow the virtual verification of the manufacturing process before proceeding with the execution. A specific add in will be included in the CAM for taking into account detailed tool data, available from Fraisa world, as well as recommendations for the optimum strategies for a given part manufacturing from the Fraisa Tool Expert module.

Data Container

The various data sources will be synchronised in a Data Container; starting from the Tool ID number and article specifications, which will be associated with the Tool holder and the selected machine. Simulation data and machine references will be integrated in this thread, which will be completed with the machine data from my rConnect for the particular process (Job recorder). This Digital thread will be the base for setting up the different applications for monitoring the tool and machine lifecycle and allow the optimised management of the full process with the federated learning solutions.

Data Semantics (UiO)

A dedicated ontology will be implemented for providing the semantics to the data related to the machine. In particular, critical components like the Drive Train test will be included as well as associated testing processes in production. The collected data will be categorised using such ontologies so to enable effective diagnosis services.

FEDMA (Core)

The FEDMA component that is deployed on the edge, will consume the data from the JobRecorder service following a predefined data model and local edge data transmission system at machine level, and after synchronisation of this data with the machine and tool related data in the Data Container.

In the GF's cloud deployment, only the Federated Learning (FL) Server will be hosted. This server will serve as the central federated learning server responsible for aggregating the weights from various FEDMA clients.

Each client will have a backend with a dedicated endpoint where requests can be made to obtain the model's output (inference). In the current diagram, the Data Acquisition service will consume the model output from each FEDMA Client. OPC UA UMATI standardised protocols will be used but not exclusively for this data collection.

FPdM (ATLAS)

For the GF use case, a federated solution for Remaining Useful Life estimation will be implemented and deployed. This will be achieved using the Federated Predictive Maintenance (FPdM) framework, which will ensure the safety of the datasets through a federated connection with the pilot.

The FPdM component is a software solution equipped with predictive functionalities, enabled by a set of microservices or sub-components. Throughout the project, a range of predictive and monitoring capabilities will be utilized to address the pilots' business cases. Specifically, for this case, the focus will be on remaining useful life estimation techniques.

The FPdM consists of two main parts, each following a standard microservices architecture, where all functionalities are implemented in distinct microservices. Communication between the services is facilitated through a central bridge:

- FL Client: This service contains the model for Remaining Useful Life estimation, which will be gradually trained and evaluated.
- FL Server: Responsible for aggregating the weights of the various models in the most optimal and suitable way, the FL server supports a variety of aggregation techniques.

Communication between the two main parts is handled via Flower Framework[13]. Visualization and notification capabilities will be offered and explored according to the scenario's needs and requirements. Both FPdM sub-components will be dockerized and deployed at locations specified by the pilots, either in the cloud or at the machine edge.

Adaptive manufacturing (Innovalia)

A module for the on-machine verification of the quality of the part will be implemented using the Innovalia platform, which will collect data from touch probes and compare the results with CAD KPIs, this in order to allow continuous improvement and adaptive manufacturing.

4.2.4 Multi-plant Predictive ZDM Turbine Production

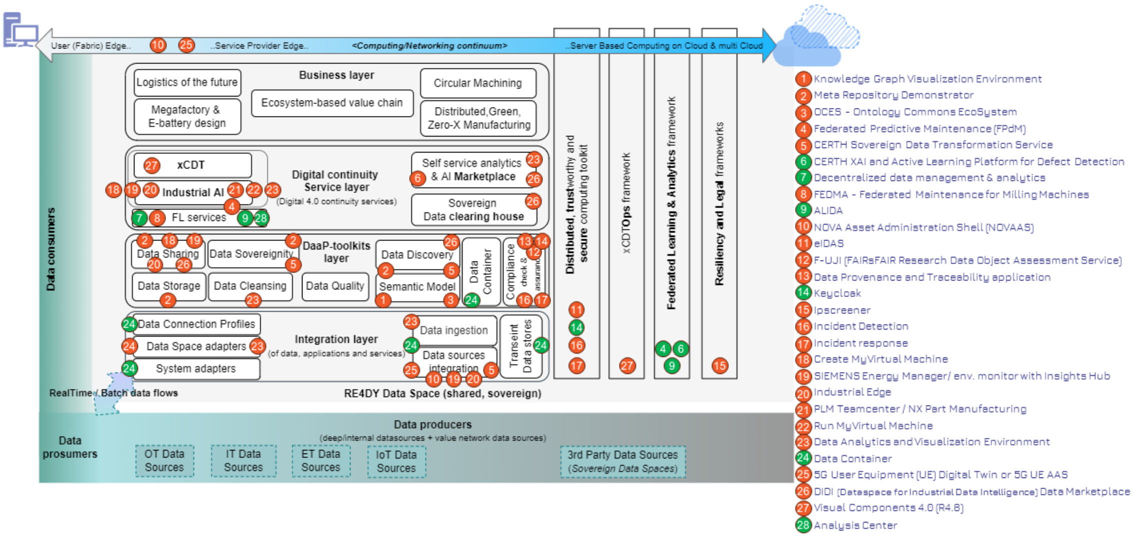

Figure 31: RA & toolkit components in the AVIO-Aero pilot

This pilot operates in the context of manufacturing and in-service maintenance of engine modules and systems for both civil and military aviation. As documented in previous deliverables, it addresses the following three main challenges:

1. Defect detection tools based on AI/ML to support quality inspection operations, highlighting to the operator possible areas to investigate.
2. A learning platform for the operators to improve the resiliency of the inspection processes, leveraging datasets and AI results.
3. Predictive quality algorithms developed for a family of products.

These challenges correspond to distinct business scenarios, each with its own objectives and expected benefits, which are:

[13] https://flower.ai/docs/framework/index.html

BS1 - AI-enabled Visual Inspection of parts

- Objective: automate Visual Inspection with AI tools, supporting the inspectors
- Benefit: shorten the inspection lead time

BS2 - Training & Certification of Visual Inspectors

- Objective: leverage the AI tool for visual inspection (BS1) to train junior inspectors
- Benefit: improve process resiliency, decoupling training from the availability of senior inspectors.

BS3 - Predictive Quality

- Objective: move from a reactive to a proactive management of quality issues
- Benefit: next gen analytical and predictive capabilities across the production process

The picture below succinctly depicts the architectural solutions designed for the three different business scenarios in this pilot.

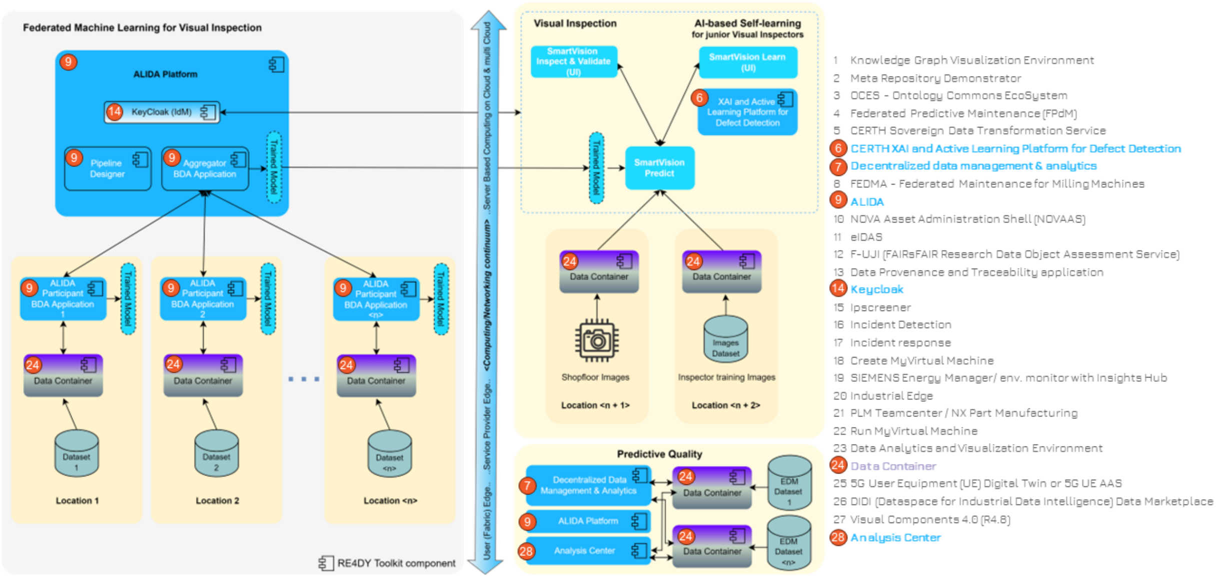

Figure 32: AVIO-Aero pilot architecture

In the picture, there are three main frames displaying various components. To distinguish between the standard toolkit components and the custom-built vertical solutions for the pilot, the former are marked with an icon, as indicated in the legend. These components are also listed on the right side along with other RE4DY Toolkit elements. In the list, components involved in this pilot are additionally identified by being numbered with an orange circle.BS

The <u>first frame</u> represents a stage that is crucial for both BS1 and BS2, as well as for BS3, making the appropriate considerations in terms of datasets and therefore algorithms, which differ for the latter compared to the first two. It features algorithms that utilize FML techniques to develop and train a model, which is then provided as a service to both the Visual Inspection and Self Learning components in the second frame. To achieve this, it leverages the following components from the RE4DY toolkit:

ALIDA and its sub-components:

ALIDA Pipeline designer: The designer component allows for the design of the Big Data pipelines and the definition and integration of the participant and aggregator BDA applications. More details about its functionality are available in sections D3.1 and D3.2.

ALIDA BDA Participant applications: an application is defined and developed for each dataset, which may originate e.g. from different shop floors or production lines. It is designed to run close to the data source to minimize data flows and thus reduce the chances of data leaks, although this proximity is not a mandatory constraint.

ALIDA BDA Aggregator application: the aggregator integrates output results produced by the participant BDA applications and builds the final trained model.

KeyCloak (IdM): this toolkit component secures access to the pilot applications. It is embedded within the ALIDA framework as well as integrated into the solutions for BS1 and BS2.

DataContainer: This component ensures access to data and metadata by incorporating controls, filters, and automatic conversion in the context of the DaaP concept experimentation. Although it is not directly related to the challenges addressed by the pilot, it is included as an additional testbed to facilitate synergy with the stand-alone experimentations carried out in WP3. *Figure 33* presents an example of its integration with the ALIDA framework.

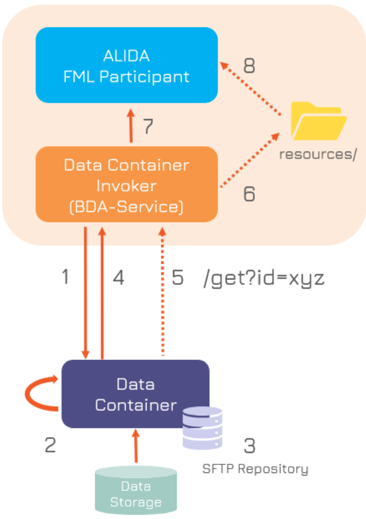

Figure 33: Hypothesis of Data Container - ALIDA Batch pipeline integration

Specifically, it is foreseen the use of an auxiliary BDA service called *Data Container Invoker,* whose task is to simplify the interaction between FML Participant - embedding the core local training logic - and the Data Container - exposing a callback-based interface to the data. Basically, the auxiliary BDA-service relieves the Data Scientist from dealing with the technicalities of the Data Container interface.

The second frame displays the UI applications specifically designed for the scenarios of BS1 and BS2: SmartVision Inspect & Validate and SmartVision Learn, respectively. Both applications interact with SmartVision Predict, a service that leverages the model previously trained and built on top of the toolkit components depicted in the first frame. Additionally, this frame includes other toolkit components, such as:

DataContainer: It holds the same value as already mentioned previously for the same component. In this particular context, it is functional in mediating access for the different types of applications involved.

CERTH XAI Interfaces:

Using explainable AI (XAI) approaches like Grad-CAM (Selvaraju R. R., 2017) and D-RISE (Petsiuk, 2020) (Randomized Input Sampling for Explanation) offers a visual explanation of the predictions made by AI models, thereby enhancing transparency and trust in these models. These techniques highlight critical areas within an image that significantly impact the model's decisions, allowing users to identify the specific pixels that contributed to the predictions. This capability is crucial, especially when dealing with ambiguous predictions, as it enables users to pinpoint and understand the model's reasoning process (Adadi, 2018), (Guidotti R., 2018).

When ambiguous predictions are identified, the most informative samples can be selected for further examination. By providing more accurate labels to these selected samples, the model can be retrained to improve its performance. This iterative process of refinement ensures that the AI system becomes progressively better at making accurate predictions over time. The retraining process is triggered when the number of newly labelled images reaches or overcomes a predefined threshold set within the active learning platform. This threshold is crucial for managing the balance between the frequency of retraining sessions and the computational resources required. To this end, the XAI interfaces provide:

1. Retraining Threshold Initialization: set the threshold for the number of new labelled images required to trigger the retraining process. This ensures that retraining occurs at optimal intervals, maximizing efficiency and effectiveness.
2. Model Selection: choose the appropriate AI model that will be used for making predictions.
3. Image Selection: select the images that need to be analysed for potential defects.
4. Model Prediction and Heat Map Representation: use the selected model to predict defects in the chosen images. The predictions are then visualized using heat maps generated by techniques like Grad-CAM and D-RISE, which highlight the areas of the image that are most influential in the model's decision-making process.
5. Identification of Selected Pixels Used for Prediction that are Out of ROI (Region of Interest): analyse the heat maps to identify any pixels that the model uses for its predictions that fall outside the predefined region of interest. This step ensures that the model focuses on the relevant parts of the image, thereby improving the accuracy of its predictions.
6. Retraining If Threshold is Reached or Surpassed: once the number of newly labelled images reaches or exceeds the predefined threshold, initiate the

retraining process. This ensures that the model continuously improves its performances basingd on the latest data and annotations.

The <u>third frame</u> defines the boundaries of BS3, which focuses on Predictive Quality by experimenting with two different ALIDA based solutions, one the Decentralised data management analytics and the other on the Analysis Center.

ALIDA and its sub-components: in BS3, as in BS1 and BS2, ALIDA is used to federally train a predictive quality machine learning model for anomalies detection.

Decentralised data management analytics:

This component develops an unsupervised classification framework that works in a distributed fashion on different machines, and integrates partial knowledge acquired at each location through the federated learning functionality provided by the ALIDA framework. The main function developed allows automatic classification of parts worked on different machines. Data available at each machine is locally analysed, and a decentralised, unsupervised clustering algorithm is implemented, allowing to cluster together similar components operated at different machines. Clustered data are provided to the process owner, for further expert assessment about the quality of the different parts. As a special case, this service can be used for unsupervised anomaly detection. Essentially, the service allows to benefit from data available at multiple machines without moving data across them, nor centralised data at any single location. Therefore, it allows for combining possibly rare data (e.g., about anomalous behaviours) across various machines, providing global knowledge without moving data from the machines where they are generated.

Analysis Center:

Based on the updated version of the architecture, in the Avio case, the focus lies solely on the analysis, specifically in developing and implementing a service that can handle Failure Detection and Identification in a federated manner. The federated aspect will be provided by using the ALIDA framework as part of the solution. Depending on data availability, the Analysis Center will employ various supervised and semi-supervised algorithms for Failure Detection and Identification.

Table 9 summarizes the RE4DY Toolkit components included in the solutions for this pilot, along with the related RA Building Blocks implemented and their corresponding layers.

Table 9: Main RA Components adoption in the pilot

Id	Component name	RA implemented BB (Building Block)	RA Layer
6	CERTH XAI interfaces	Self-service analytics & AI MKTPL	Digital continuity service layer
7	Decentralized data mgmt. analytics	FL services	Digital continuity service layer
9	ALIDA	FL services	Digital continuity service layer
14	KeyCloak	compliance check and assurance	DaaP-toolkits layer

24	Data Container	Data container	DaaP-toolkits layer
		ALL	Integration layer
28	Analysis Center	FL services	Digital continuity service layer

5 Conclusions

This document presented the finalized version of the Digital 4.0 Continuum Reference Framework, consisting of three main elements, each of them playing a critical role within the overall framework.

The Resiliency Framework, grounded in dynamic capabilities theory and IDEF0 functional modelling, has successfully evolved into a comprehensive dashboard easing resilience assessment and development across anticipation, coping, and adaptation stages. This model provides companies with crucial insights into their resilience capabilities and risks, enabling strategic decision-making in dynamic environments.

The Legal Framework addresses critical aspects of data governance and intellectual property within the EU data space. It highlights the evolving definitions of data as a product under EU legislation and emphasizes compliance with the Data Act and Data Governance Act. These regulations underscore the need for transparent data sharing frameworks and fair value distribution, crucial for fostering a trustworthy digital ecosystem. In addition, the work started on the Intellectual Property Rights Ontology (IPRO) and introduced in this document, will continue with focusing on defining the relationship between its classes and object properties. The IPRO is planned to be published alongside the upcoming D2.4 "Digital 4.0 continuum value network industrial agreements" in M36, where it will serve as another enabler of multilateral digital value chains in a smart manufacturing industry context. During its remaining development, the IPRO is planned to be tested in the SSF's Testing and Experimental Facility (TEF). Feedback and insights from this test will be used to iterate upon the IPRO and improve its utility in industrial use cases. Various opportunities to align the IPRO and the Resilience Ontology will be explored, aiming to identify intersection points between these two ontologies to effectively bridge their respective domains.

The Reference Architecture (RA) along with the RE4DY toolkit and its different implementations carried out in the project pilots, serves as a robust data-driven framework tailored for the Digital 4.0 Continuum, integrating specific business needs and technological advancements. Designed across four layers and enhanced by the computing networking continuum, the RA supports decentralized operations and digital continuity, crucial for aligning manufacturing and IT operations effectively.

The consolidation of these three main elements contributes to the definition of the Digital 4.0 Continuum Reference Framework, through which different and challenging objectives can be achieved with a holistic perspective. For example, enabling data spaces and fully leveraging the digital thread in the manufacturing industry means not only providing cutting-edge technological solutions, such as a Reference Architecture and a set of ready-to-use tools, the RE4DY toolkit, easily customizable according to needs, as adopted in the various implementations of the pilots. In fact, in an increasingly dynamic context, where multiple factors—both internal and external, such as geopolitical and health events—can undermine a company's business, it is crucial to have easy-to-use tools to autonomously assess resilience levels, such as the dashboards that can be built with the Resilience Framework. It is also important to address issues such as the data sovereignty, and how to guarantee it, which, if not considered in time, can significantly hinder the full exploitation of emerging business models like data-as-a-product. In this regard, the holistic perspective of this work has continuously aimed to address these challenges, supported by an additional complementary framework dedicated to legal aspects. This Legal Framework has been introduced here and will be further described and consolidated in the upcoming deliverable D2.4.

6 References

Adadi, A. &. (2018). Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access, 6*, 52138-52160. doi: 10.1109/ACCESS.2018.2870052

Ameri, F. S. (2022). Industrial ontologies for interoperability in agile and resilient manufacturing. *International Journal of Production Research, 60*(2), 420-441.

Aplin T., R. A. (2023). The Role of EU Trade Secrets Law in the Data Economy: An Empirical Analysis. *IIC - International Review of Intellectual Property and Competition Law, 54*, 826–858. doi:10.1007/s40319-023-01325-8

Bobev T., e. a. (2023). *White Paper on the Definition of Data Intermediation Services*. doi:10.2139/ssrn.4589987

Boley H, P. A. (2010). RuleML 1.0: the overarching specification of web rules. *International Workshop on Rules and Rule Markup Languages for the Semantic Web.* Springer.

Buiten, M. d. (2021). *EU Liability Rules for the Age of Artificial Intelligence*. CERRE (Centre on Regulation in Europe).

Calero C., R. F. (2006). *Ontologies for Software Engineering and Software Technology*. Springer Science & Business Media.

Chari A., M. D. (2024). Resilience Compass Navigation through Manufacturing Organization Uncertainty – a Dynamic Capabilities Approach using Mixed Methods. [Manuscript submitted for publication].

Chari, A. e. (2023). Analysing the antecedents to digital platform implementation for resilient and sustainable manufacturing supply chains - An IDEF0 modelling approach. *Journal of Cleaner Production*, 139598.

Drexl, J. (2017). Designing Competitive Markets for Industrial Data – Between Propertisation and Access. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law (JIPITEC), volume 8.*, 257–292.

DuHadway S., C. S. (2017). Understanding risk management for intentional supply chain disruptions: risk detection, risk mitigation, and risk recovery. *Annals of Operations Research, 283*(1-2), 179-198.

Elkington, J. (1998). *Cannibals with Forks: The Triple Bottom Line of 21st Century Business.* New Society Publishers (Gabriola Island, BC, and Stony Creek, CT).

Geiregat, G. (2022). *The Data Act: Start of a New Era for Data Ownership?* doi:10.2139/ssrn.4214704

Gruber, T. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition, 5*(2), 199-220.

Guidotti R., M. A. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys (CSUR), 51*(5), 93. doi:10.1145/3236009

IEEE, S. (1998). Standard for Functional Modeling Language—Syntax and Semantics for IDEF0. Institute of Electrical and Electronics Engineers, Inc.: USA.

Ivanov, D. (2020). Viable supply chain model: integrating agility, resilience and sustainability perspectives-lessons from and thinking beyond the COVID-19 pandemic. *Annals of Operations Research*, 1-21.

Liu Z., M. N. (2018). The Role of Data Fusion in Predictive Maintenance Using Digital Twin. *1949, 1*. doi: 10.1063/1.5031520

Lloyd, J. (2012). *Foundations of logic programming*. Springer Science & Business Media.

Mizoguchi R., I. M. (1998). Towards ontology engineering. *Journal of the Japanese Society for Artificial Intelligence, 13*, 9-10.

Mylly, U.-M. (2024). Trade Secrets and the Data Act. *IIC - International Review of Intellectual Property and Competition Law, 55*, 368–393. doi:10.1007/s40319-024-01432-0

Noy NF, S. M. (2001). Creating semantic web contents with protege-2000. *IEEE Intelligent Systems, 16*(2), 60-71.

Padhi, S. S. (2024). Optimal investment decision for industry 4.0 under uncertainties of capability and competence building for managing supply chain risks. *International Journal of Production Economics, 267*.

Petsiuk, V. e. (2020). Black-box Explanation of Object Detectors via Saliency Maps. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), with the proceedings published in 2020.*, (p. 11438-11447).

Rashid A.H.M., L. S. (2014). Strengthening supply chain risk management for business continuity: a case study approach. *International Journal of Management and Enterprise Development, 13*(3/4).

READI. (2023). *IMF Semantic Resources*. Tratto da https://ns.imfid.org/20230630/

Selvaraju R. R., C. M. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *International Journal of Computer Vision, 128*, 336–359. doi:10.1007/s11263-019-01228-7

Teece, D. a. (1994). The Dynamic Capabilities of Firms- an Introduction. *Industrial and Corporate Change, 3*(3), 537-556.

Thouvenin, F. a.-L. (2021). Data Ownership and Data Access Rights: Meaningful Tools for Promoting the European Digital Single Market. In *Big Data and Global Trade Law* (Cambridge University Press ed., p. 316-339). Burri M. doi:10.1017/9781108919234

UN. (2015). Global Indicator Framework for the Sustainable Development Goals and Targets of the 2030 Agenda for Sustainable Development. Tratto da https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%202020%20review_Eng.pdf

World Bank. (2021). *World Development Report 2021: Data for Better Lives: Spotlight 6.2, The Debate Over Ownership of Personal Data.* doi:10.1596/978-1-4648-1600-0