D6.6: Information package for external SMEs to facilitate the design of new applications

This project has received funding from the European Union's Horizon Research and Innovation Actions under Grant Agreement № 101093216.

	Document version:
D6.6: Information package for external SMEs to facilitate the design of new applications	1.0

Project number:	Project Acronym	Project Tittle
101093216	UPCAST Project	UPCAST Project

Contractual Delivery Date:	Actual Delivery Date:	Deliverable Type*-Security*:
M27 (September 2025)	M27 (30 September 2025)	R-PU

*Type: P: Prototype; R: Report; D: Demonstrator; O: Other; ORDP: Open Research Data Pilot; E: Ethics.

**Security Class: PU: Public; PP: Restricted to other program participants (including the Commission); RE: Restricted to a group defined by the consortium (including the Commission); CO: Confidential, only for members of the consortium (including the Commission).

Responsible:	Organization:	Contributing WP:
Nevena Raczko	IDC	WP6

Authors (organization):
Daniela Rao, Giorgio Micheletti (IDC)
Luis Daniel Ibanez, George Konstantinidis (SOTON)
Sofoklis Efremidis, Kostas Kalaboukas (MAG)
Hanene Jemoui (CDR)
George Lioudakis (ABV)
Shanshan Jiang (SINTEF)
Majid Ektesabi (NOK)
Jeremy Decis (DAW)
Kotsifakos Evangelos (LST)

Abstract:
Deliverable D6.6 reports technical information, marketing analysis and business plan guidance to facilitate SMEs and other beneficiaries to leverage UPCAST projects outcomes. It provides: an overview of the UPCAST platform's architecture, the relevant technical information about how plugins can be used and accessed by SMEs and/or other beneficiaries to implement applications dedicated to procurement processes, the results of the European SMEs data-driven strategies survey, useful to better understand and address the European SMEs digital needs, and the business planning guidance for SMEs managers and innovation officers who want to realize the benefits of data spaces, and more broadly, data-sharing platforms at scale.

Keywords:

Platform architecture, APIs, plugin, target users, market opportunities, SMEs data-driven survey, market analysis, SME digital maturity, business plan guidance.

REVISION HISTORY

Revision:	Date:	Description:	Author (Organization)
V0.1	10.06.2025	First version of the DOCUMENT	Daniela Rao, Giorgio Micheletti (IDC)
V0.2	10.07.2025	Second version, adding content in many sections	Daniela Rao (IDC)
V0.3	26.08.2025	Third version, adding content in many sections	Daniela Rao (IDC)
V0.4	05.09.2025	Third version, adding content in many sections	Daniela Rao (IDC)
V0.5	15.09.2025	Adding screenshots and correcting typos	Daniela Rao (IDC)
V0.5	15.09.2025	Adding screenshots and correcting typos	Daniela Rao (IDC)

This project has received funding from the European Union's Horizon Research and Innovation Actions under Grant Agreement Nº 101093216.

More information available at https://upcastproject.eu/

COPYRIGHT STATEMENT

Table of Contents

List of Figures

1 Introduction

UPCAST (Universal Platform Components for Safe Fair Interoperable Data Exchange, Monetisation and Trading) provides a set of universal, trustworthy, transparent and user-friendly data market plugins for the automation of data sharing and processing agreements between businesses, public administrations and citizens. The UPCAST plugins enable actors in the common European data spaces to design and deploy data exchange and trading operations.

The UPCAST plugins are meant to the easily integrated and deployed to data marketplaces to enhance or customize the services they provide to their users. Nevertheless, UPCAST has integrated the plugins to an UPCAST platform that, on one hand can be directly used by data providers and consumers and, on the other hand to be integrated with commercial data sharing platforms.

This document is a blended information package made up of technical information, market analysis and business plan guidance to facilitate SMEs and other beneficiaries to use UPCAST plugins to implement and design innovative procurement applications.

1.1 Purpose of the Document

Since UPCAST aims at engaging SMEs, administrations and citizens by providing a transferability framework, best practices and training to endow users to deploy the new technologies, this document summarizes useful information to facilitate SMEs and other beneficiaries to leverage UPCAST platform and plugins to approach the European data space market.

1.2 Scope of the Document

The scope of this document is to provide:

- an overview of the UPCAST platform's architecture,
- the relevant technical information about how plugins can be used and accessed by SMEs and/or other beneficiaries to implement applications dedicated to procurement processes,
- the results of the European SMEs data-driven strategies survey, we run to better understand and address the European SMEs digital needs,
- the lean business planning guidance for SMEs managers and innovation officers who want to realize the benefits of data spaces, and more broadly, data-sharing platforms at scale.

1.3 Structure of the Document

The Section 2 outlines the key features and components of the UPCAST platform that serve as the basis for evaluating pilot projects, which will use the functions of the entire platform or those of individual plugins to assess the degree to which their requirements are met. The goal is to provide SMEs and other beneficiaries some quick takes about the UPCAST architecture flexible approach, in which plugins (deployed through well-defined Application Programming Interfaces, APIs) may be deployed individually as web services and accessed either through the Provider or Consumer Dashboard, or directly through their interfaces, or be integrated or deployed on the Data Sharing platform and accessed through it.

Section 3 provides key instructions to use and access the UPCAST plugins by SMEs and other Actors of the European \Data Space Economy, with the goal to facilitate the design and implementation of applications dedicated to the procurement agreements between businesses, public administrations, and citizens. For each plugin are summarized key features, how the plugin can be used by SMEs and/or other beneficiaries, users' target benefits and opportunities, and where the plugin can be accessed.

Section 4 provides the outcomes of the European SMEs data-driven strategies, conducted in May 2025 on 1,700 European SMEs with 1-250 employees, useful to understand the European SMEs data-driven strategies and how to address their digital needs with ready-to-use components.

Section 5 provides to SMEs and start-up guidance and suggestions about how to leverage digital tools developed in UPCAST project, to realize the benefits of data spaces, and more broadly, data-sharing platforms at scale**.** For SME or start-up committed to entering the data market space, the creation of a business plan is a crucial step on the road to success. Main components and roadmap of lean business plan canvas for SMEs are summarized to support them to grow leveraging experimental approaches to innovation, building strategic partnerships, and focusing on data-driven decisions, by measuring key metrics to ensure continuous improvement and sustainable value creation.

2 UPCAST platform architecture quick takes

UPCAST provides support for the management, negotiation, and exploitation of resources through a set of plugins that can be installed in Data Marketplaces or other data sharing platforms and can mediate data transactions between data providers and data consumers. In this context, a resource is an abstract entity and can be a dataset, a data operation or an artefact (such as a machine learning model).

UPCAST offers a set of functionalities to assist Dataset (Resource) Provider, and/or the Dataset (Resource) Consumer. Plugins interact with each other and with the marketplace in which they are deployed through well-defined Application Programming Interfaces (APIs). Plugins may be used independently through their APIs, but some of it can be used in combination with each other. To further illustrate, consider the UPCAST architecture, shown in following Figure 3, we logically grouped the plugins used by Dataset Consumers into a "Consumer Dashboard" and those used by Dataset Providers into a "Provider Dashboard". In the middle, a Data Sharing Platform intermediates between Provider and Consumer. A Data Sharing Platform may be a established Data Marketplace (e.g Dawex or Nokia), or by a custom application developed to mediate between Consumers and Providers in a specific context. The Data Sharing Platform implements the authentication of Providers and Consumers and is the repository to which datasets may be advertised by providers and discovered by consumers. Depending on the use case, the Data Sharing platform provides a hosting environment to which UPCAST plugins may be deployed or otherwise integrated.

For instance, the Environmental plugin is shown at the Provider domain, the Consumer domain and also the Data sharing platform. Providers and consumers may want to use this plugin for either annotating the dataset with its environmental impact before advertising it, or, in the case of consumers, to monitor the processing of a dataset and emit corresponding monitoring events. Data sharing platforms may also want to integrate plugins to offer extra functionality to their users.

In this sense, the UPCAST Architecture has a dual interpretation. It can be interpreted as a centralized one in which the Data Sharing Platform serves as the single place of interaction between the Dataset Provider and the Dataset Consumer or a distributed one in which the Data Sharing platform is used for authenticating providers and consumers and interactions between them take place in a peer-to-peer manner through the respective dashboards.

A dataset provider may use the plugins grouped in the Provider Dashboard to produce the specification of a dataset resource to publish in a Data Sharing Platform. In a minimal deployment, the Publishing/Discovery plugin can be used as a Data Sharing Platform. Resource specification includes the annotation of the dataset with plain (type of data, format, creation time, etc.) and semantic metadata, its environmental footprint for its storage by the provider, an estimate of its price, and definition of usage and access policies. Once a dataset resource has been published, it can be discovered by potential consumers.

On their end, the Dataset consumer uses the plugins group to specify a Data Processing Workflow to model the processing they want to do on a dataset. Moreover, the consumer may also define policies that is obliged to abide with, for example internal policies or legal

regulations, and can also estimate the environmental impact of the dataset execution that is modelled by the Data Processing Workflow. After these policies have been specified and prepared with the use of the corresponding plugin, the consumer searches for datasets that meet their criteria. Once a dataset that meets the consumer requirements is discovered, the consumer gets in contact with the dataset provider, and a negotiation takes place between the two. The negotiation is supported by the negotiation plugin and its purpose is for the provider and the consumer to agree on the same terms for the dataset execution, as, on one hand the provider has expressed his usage policies, and, on the other hand, the consumer has expressed his own policies they may be subject to for the execution of the dataset. The negotiation, if successful, results in a contract, which, once agreed by both parties, is secured for later checking compliance with the execution of the Data Processing Workflow (DPW) that uses the dataset that is referred in the contract.

Once a negotiation is completed and a contract has been agreed and signed, execution of the data processing workflow on a dataset can be performed. The first step for the execution to commence is to transfer the dataset from the provider to the consumer space. Once the dataset has been securely transferred, a workflow execution environment is used to carry out the execution. In UPCAST, the execution of a Data Processing Workflow may take place in various execution environments including the consumer's space, the marketplace, a trusted third party, or in the provider itself.

One exception in consumer processing is the case of Federated Machine Learning, in which parts of the execution may take place in the provider's environment, the reason being that analytics processing of classified data may be allowed but the data itself may not be allowed to leave the provider's environment. In this case, the provider needs to provide a hosting and execution environment for containerized FML components to execute, UPCAST contains a Federated Agent component, which abstracts the parts of the execution that need to take place in the provider's space. The Federated Agent component is further highlighted through the surrounding box to indicate that such components should be containerized.

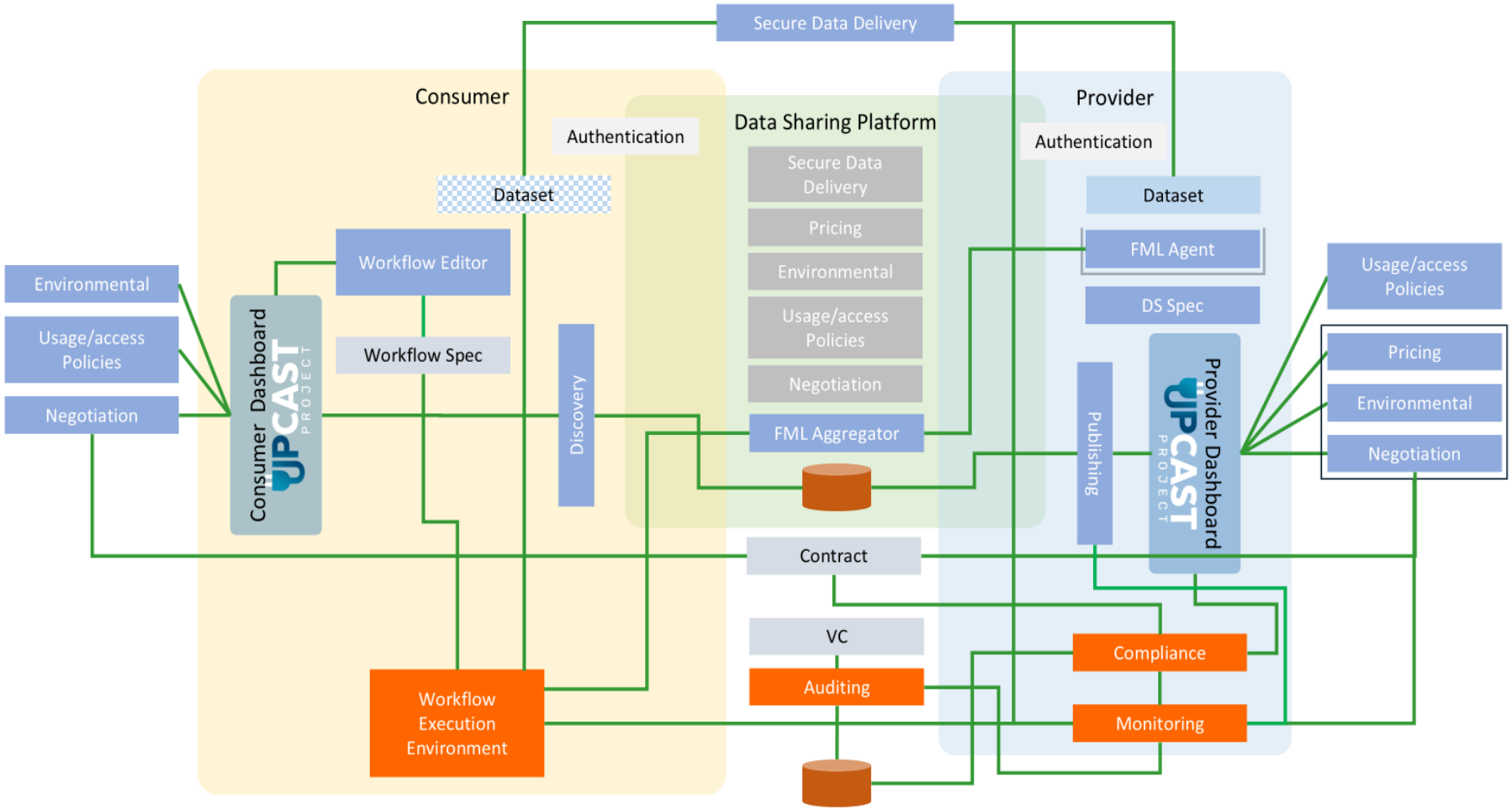


Figure 1: UPCAST Architecture.

3 How plugins can facilitate the design of SME's applications

market plugins for the automation of data sharing and processing agreements between businesses, public administrations, and citizens.

These plugins will enable actors in the common European data spaces to design and deploy data exchange and trading operations guaranteeing:

- automatic negotiation of agreement terms,
- dynamic fair pricing,
- improved data-asset discovery,
- privacy, commercial and administrative confidentiality requirements,
- low environmental footprint,
- compliance with relevant legislation and ethical and responsibility guidelines.

UPCAST provides support for the management, negotiation, and exploitation of resources through a set of plugins that can be installed in Data Marketplaces or other data sharing platforms that can mediate data transactions between providers and consumers. These plugins are:

1. Resource Specification and Profiling
2. Semantic Profiling
3. Publishing and Discovery
4. Privacy and Usage Policy
5. Environmental Impact
6. Pricing
7. Data Processing Workflow
8. Federated Machine Learning
9. Data Integration
10. Negotiation and Contracting
11. Secure Data Exchange
12. Monitoring
13. Compliance
14. Auditing

Following paragraphs provide key instructions to use and access the UPCAST plugins by SMEs and other Actors of the European Data Space Economy, with the goal to facilitate the design and implementation of applications dedicated to the procurement agreements between businesses, public administrations, and citizens.

3.1 Resource Specification and Profiling

This plugin allows the specification and profiling of dataset resources. It involves the tasks of preparing the dataset for publishing it to a marketplace. The preparation tasks involve the annotation of the dataset with metadata, optionally environmental impact and pricing

information, as well as policies for its usage and access.

How the plugin can be used by SMEs and/or other beneficiaries

The resource specification plugin functions are implemented by the UPCAST Provider Dashboard.

The UPCAST Provider Dashboard is an interface which facilitates the whole lifecycle of a dataset in the UPCAST ecosystem, from the perspective of a data provider. It encompasses standalone administrative and monitoring features, like the definition and registration of a dataset, along with functionalities supported by various UPCAST plugins which are integrated on the Dashboard. Thus, it accommodates a seamless process to demonstrate the benefits of UPCAST tools for data providers within a single platform.

The main features of the Provider Dashboard are:

- Resource Specification.
- Dataset Pricing & Environmental Footprint Annotation.
- Definition of Dataset Usage Policies.
- Publishing on Data Marketplaces.
- Contract Negotiations with UPCAST Consumers.
- Monitoring of Workflow Executions under active Contracts.

The plugin can be accessed and used through a web browser that provides a convenient interface to the functions of the plugin.

Users' target benefits and opportunities

The resource specification plugin provides an interface to the functions of several other UPCAST plugins that are integrated into a single platform. Using the plugin, a user (Dataset Provider) can annotate a dataset with metadata that facilitate its subsequent search and transfer to the domain of potential consumers that may be interested in it.
Annotation includes several metadata, like the type of the data, size of data, period in which the data was collected or compiled, its format, and offerings, i.e. instances of the data that can be transferred. Metadata may also include pricing information, information on the environmental footprint of processing the data and semantic annotations based on the semantic content of the data.
To this end, a data provider is powered with all functions to enrich their data and facilitate thus its discovery by potential consumers.

Where the plugin can be accessed

The plugin can be accessed through dev.upcast.maggioli-research.gr/upcast.
The plugin is offered as web service; source code is protected by IPRs and is not available for download.

3.2 Semantic Profiling

The Semantic Profiling (referred also as Semantic Annotation) plugin can generate semantically rich metadata for a dataset for facilitating its discovery. The plugin may be used by dataset providers to annotate their datasets with semantic metadata before publishing

them. Through the Provider Dashboard or even be integrated to data marketplaces for the same purpose.

The semantic profiling plugin is based on LLMDap, an LLM-based data pipeline for data profiling. This pipeline utilizes LLMs for automatic extraction of metadata description for datasets from associated textual documentation (e.g., scientific paper, lab protocol, etc.). See detailed description in D2.3.

How the plugin can be used by SMEs and/or other beneficiaries

As a plugin, LLMDap can be deployed either locally or in the cloud by enterprises and institutions of various sizes, depending on the location of the dataset.

- Local Deployment: SMEs can use LLMDap to facilitate the development of a chatbot-ready internal knowledge base by annotating internal documents with metadata aligned to their specific schemas or ontologies.
- Cloud or Third-Party Deployment: organizations can use LLMDap in the cloud or via platform operators for at least two key use cases: 1) Scaling internal knowledge management for data accessible beyond local servers; 2) Sharing data and related assets—such as documents, models, or protocols—with industry peers or through data marketplaces

Users' target benefits and opportunities

LLMDap is designed to benefit knowledge-based organizations affected by data silos. By reducing the manual labour involved in generating semantic metadata, it facilitates a critical step in creating machine-readable asset documentation. This, in turn, introduces interoperability opportunities on two scales:

- Organizations (e.g., NHRF, e-NIOS): Individuals can more easily improve the discoverability of often unstructured or heterogeneous assets—such as datasets, code, models, lab protocols, and publications—both internally for knowledge management and externally for sharing or commercialization.
- Platform Operators (e.g., Dawex, ArrayExpress, cBioPortal): Data catalog providers can use the plugin to continuously enhance asset specification and discovery by cataloguing previously under-annotated assets and re-categorizing them according to emerging metadata standards.

Where the plugin can be accessed

The code is open source with MIT license: https://github.com/SINTEF-SE/LLMDap or https://github.com/EU-UPCAST/profile

Note: the first repository is open, the second one is so far private. We will sync the two repositories.

3.3 Publishing and Discovery

This plugin provides 2 functionalities:

- the UPCAST Publish Plugin API v2 provides comprehensive dataset management capabilities designed to streamline data publishing workflows. The plugin supports

the complete lifecycle of dataset management within data marketplaces, enabling efficient publishing and discovery processes across multiple platforms.

- the UPCAST Discovery Plugin provides comprehensive search, browse, and discovery capabilities for datasets across multiple marketplaces. The plugin implements intelligent recommendation systems and similarity matching to enhance data discoverability and user experience.

How the plugin can be used by SMEs and/or other beneficiaries

SMEs can use it as a catalogue for datasets they own. It can be used internally, or profiting from the integration with other UPCAST plugins, make it public for data consumers to browse and discover datasets of their interest and potentially start a negotiation for a Data Sharing Agreement. Alternatively, they can use the integration with Data Marketplaces to publish the metadata of their datasets into them to reach a wider audience. Data Marketplace operators can: (A) Onboard datasets from SMEs and organizations using the Publishing plugin as a Data Catalogue and (B) Integrate the advanced Semantic Search capabilities into their own systems.

Users' target benefits and opportunities

SMEs can use it to manage the metadata of their data assets and enable internal search and discovery at a lower cost than full-fledged products aimed at larger organizations. They also have the opportunity to advertise their data assets to data consumers, enabling potential monetization. The UPCAST Data Model also integrates pricing, licensing and access control, facilitating data transactions. Data Marketplaces operators can on the one hand get more customers to advertise in their systems, increasing their offering of datasets. The advanced search capabilities improve the user experience of their existing customers by enabling search beyond keywords. They can also build by enforcing metadata standards on which UPCAST plugins are based.

Where the plugin can be accessed

The plugin can be accessed on https://github.com/EU-UPCAST/discovery-plugin/

3.4 Privacy and Usage Policy

The main functionalities of this plugin (detailed in Deliverable 2.2. and Deliverable 2.4) are:

- Public Repository of Policies and Ontologies: it allows the storage and retrieval of policy objects, and of domain specific ontologies. Policy objects contain details of all make use of domain specific terms, whose definition can be found in the referenced ontologies.
- Policy Editor: it allows users to define policies through an intuitive interface, which considers the selected ontologies to provide relevant suggestions as to which terms to use to define the rules.
- Comparison Engine: it enables policies to be compared, to determine whether they match, or whether they contain conflicting rules. This is implemented as an Application Programming Interface (API) that returns details of any conflicting sets of rules.
- Consent Management: a specialization of this plugin allows for data requests to be

sent to non-technical data owners, who can then login through a dedicated interface to accept, refine or reject the requests that have been sent to them.

- Consumer-side policy engine and Data Processing Workflows. Data consumers can define detailed policies regulating their organization's operations and specify comprehensive Data Processing Workflows to reflect their data acquisition intentions. This plugin enhances interoperability and collaboration through ontological alignment, ODRL integration, and new capabilities for handling datasets, pricing, and energy constraints, enabling seamless policy enforcement across diverse data ecosystems with advanced conflict resolution mechanism.

How the plugin can be used by SMEs and/or other beneficiaries

The Policy Editor module can be used to help both technical and non-technical personnel of an organization to effectively write usage constraints as ODRL policies based on general and domain specific vocabularies.
The Consent Manager module can be used by an organization that acts as Data Controller to solicit consent from Data Subjects that don't have the willingness to use tools designed for data sharing between organizations. Consent is then formatted as ODRL Policies, ensuring interoperability with other UPCAST plugins and Data Spaces in general.

Users' target benefits and opportunities

Policy editor's main benefit is facilitating the generation of machine-readable usage policies that can then be used by other semi-automated tools such as UPCAST's own Negotiation and Compliance plugins. Lowering the barrier for writing these policies means more data assets can be annotated with usage constraints, and more precisely thanks to the business knowledge these users possess.
The Consent Manager module gives the opportunity of collecting consent in a machine-readable format that can be verified by other tools, facilitating the assurance of compliance along the Data Value Chain. An example using UPCAST tools: Compliance with consent can be checked at Processing Design stage with the DPW plugin, while at execution stage can be checked with the Compliance plugin.

Where the Plugin can be accessed

The plugin can be accessed on https://github.com/EU-UPCAST/policy-editor-ui and https://github.com/EU-UPCAST/upconsent

3.5 Environmental Impact

The environmental plugin delivers two core features: the environmental impact estimator and the environmental impact monitoring tool. Together, these tools offer complementary approaches to promoting energy-efficient data processing workflows (DPWs) and storage, helping to reduce carbon emissions while supporting compliance with energy regulations (Deliverable 4.1).

How the plugin can be used by SMEs and/or other beneficiaries

SMEs and other beneficiaries can use the environmental plugin in two main ways:

- Can estimate energy consumption for their data processing workflows and datasets stored at idle using the environmental impact estimator REST API, which provides

predictions of energy usage.

- Can monitor real-time power usage during the execution of data processing workflows with the environmental impact monitoring tool, which continuously collects and displays live power consumption data.

Users' target benefits and opportunities

By using the environmental plugin, organizations gain the ability to track and optimize energy use, reduce costs, and support sustainability efforts. This supports lower carbon emissions, compliance with energy regulations, and maximized efficiency. Additionally, detailed monitoring and custom saving scenarios empower users to make data-driven, environmentally responsible decisions.

The estimator enables forecasting and planning for future energy consumption, while the real-time monitoring tool provides ongoing oversight. This modular solution can be integrated easily, allowing businesses to select features that fit their needs.

Where the Plugin can be accessed

Any SME that wants to use our tool needs to contact us at:

https://forms.gle/4jeEXAHBsXFtiAFAA

3.6 Pricing

The pricing plugin delivers three core features: static pricing, dynamic pricing, and optimal bidding. These features offer distinct methodologies for determining dataset value, each addressing different market and operational needs (see Deliverable 4.1).

How the plugin can be used by SMEs and/or other beneficiaries

SMEs and other beneficiaries can use the pricing plugin in three main ways:

- Can estimate dataset prices based on key properties such as volume, data type, category, and other intrinsic attributes using the static pricing REST API.
- Can predict and adjust dataset prices over time by factoring in current market conditions and other influential variables via the dynamic pricing REST API.
- Can calculate and receive recommendations for bidding strategies that effectively balance the likelihood of winning with maximizing expected returns, using the optimal bidding REST API.

Users' target benefits and opportunities

The pricing plugin offers organizations a range of methodologies to optimize dataset value for diverse business needs. Static pricing uses competitor benchmarks, while dynamic pricing adapts to demand fluctuations to maximize revenue. Optimal bidding, using game theory to strategically secure deals. Together, these approaches enable organizations to support various data monetization models, from fixed catalogue sales to adaptive marketplace exchanges and strategic auctions.
This flexibility ensures successful and profitable data monetization across different scenarios by allowing companies to adjust data prices to shifting market conditions and strategic goals.

Where the Plugin can be accessed

Any SME that wants to use our tool needs to contact us at:https://forms.gle/r8KkW77ELp5Y77WQ8

3.7 Data Processing Workflow

This plugin enables the specification of UPCAST DPWs, thereby also serving as the means through which data consumers state their intentions for the data they seek to acquire. These intentions are derived from jointly considering a variety of aspects, including: the processing operations intended to be performed; the entities in direct or indirect control of their execution; the attributes of the acquired data and the conditions under which any processing and exchange is meant to take place; the stated purposes that the DPW in question is intended to serve.

Furthermore, the DPW plugin allows for the automatic verification and re-engineering of DPW models, considering on the one hand the internal policies of an organization and, on the other hand, the constraints defined by the data providers of acquired datasets. In this context, it receives input from the consumer-side PUC that undertakes the reasoning over the policies and constraints and caters for conflict resolution thereof.

Finally, the DPW plugin provides the functionality for the automatic translation of DPW models to executable workflows. To this end, it incorporates a module that undertakes the translation of workflow models to NextFlow scripts, thereby allowing for their execution in operational environments. The translation takes into consideration additional UPCAST needs, such as the injection of appropriate functionality in the scripts in order to send monitoring events to the Monitoring plugin (Deliverable 4.1).

How the plugin can be used by SMEs and/or other beneficiaries

SMEs and other beneficiaries can use the DPW plugin to visually design their data processing workflows, specifying their elements in a comprehensive, fine-grained manner, and generate an executable instance with no or low coding. They can also validate the compliance of workflow models against both organisational rules (formally specified using the consumer-side Privacy and Usage Control plugin) and constraints defined by the data providers, and even automatically re-engineer the models, in order to become compliant with the underlying rules. Further, through UPCAST integrations, they may leverage various aspects of UPCAST functionality, in order to discover, negotiate and acquire datasets from respective providers.

Users' target benefits and opportunities

With the use of the DPW plugin, data consumers benefit from the functionality of specifying and validating data processing workflows, discovering and negotiating datasets, and incorporating automatically data providers' constraints in their pipelines. Automated verification and re-engineering reduce compliance risk and provide auditable evidence for contracts, thereby improving trust among data providers and consumers. Additionally, the DPW plugin provides opportunities for organisations with limited resources, especially SMEs, to implement comprehensive compliance programmes through alignment of their processes with the underlying requirements, whereas data brokers and dataspace operators can offer their users added-value services towards faster time-to-data.

Where the Plugin can be accessed

The tool is characterised by proprietary licence, so the source code is not publicly available. It is offered as a service at: https://upcast.ict-abovo.gr.

3.8 Federated Machine Learning

A usage scenario of FML in the context of UPCAST is as follows. A Data Consumer needs to train a Machine Learning model. They discover in a Data Marketplace multiple Data Providers that their own data could be useful to train the model. However, the usage constraints of Data Providers prohibit the transfer of data outside of their premises. Providers are still interested in doing business with the Data Consumer and are willing to use their own computational resources to collaborate with the Consumer without having to transfer data (Deliverable 4.1).

How the plugin can be used by SMEs and/or other beneficiaries

Federated Machine Learning (FML) has two parts. One which handles the orchestration and should be deployed on the server (or marketplace) and one which is client agent and should be deployed on the data provider side. The Data provider infrastructure should be able to handle ML model training on the data locally and he has an open internet face connection with a particular port. These two parts should be to communicate.

Users' target benefits and opportunities

There are multiple benefits of FML:

- Data Privacy: Raw data never leaves the local devices, protecting user privacy.
- Regulatory Compliance: Helps in complying with data protection regulations like GDPR.
- Reduced Data Transfer: Only model updates are shared, reducing bandwidth usage.
- Access to Diverse Data: Enables learning from a wide range of data sources without centralization.
- Continuous Learning: Models can be updated frequently with fresh, real-world data.
- Scalability: Can handle large numbers of participants and huge amounts of decentralized data.
- Reduced Central Storage Need: No need for a large central data repository on data marketplace or data consumer side.
- Collaboration: Allows multiple organizations to collaborate without sharing sensitive data.

Where the Plugin can be accessed

There are 2 dockers that compose file that could be installed. It is offered at: https://github.com/EU-UPCAST/OpenAPISpecification/tree/main/fml_plugin

3.9 Data integration

The Data Integration and Exchange Operation is an end-to-end approach to forward or backward-chaining. The framework packages several tools featuring writers, translators, parsers, and integrators to combine and seamlessly transfer between dataset implementations. By providing a unified framework, we may leverage the system's strengths while avoiding their weaknesses by mixing and matching the optimal systems for each task. Furthermore, we make it easier on the user to answer a query by combining algorithms from both approaches (Deliverable 4.1).

How the plugin can be used by SMEs and/or other beneficiaries

Data Consumers can use it to align schemas of datasets they have acquired as part of Data Processing Workflows. The plugin allows the integration to be "virtual", useful for when the datasets are accessible through a query interface; or "materialized", useful for dataset files that need to be processed in a pipeline fashion. Data Providers may also use it independently of Data Processing Workflows to prepare combinations of datasets as a single Data Product. Data Marketplace operators can offer integration as a feature to allow users to explore how different datasets can be integrated or mashed before buying or acquiring them.

Users' target benefits and opportunities

The benefit for Data Consumers is that they can streamline the acquisition of multiple data sources because the plugin facilitates their integration as part of a Data Processing Workflow. Data Providers have the opportunity of creating novel Data Products faster and at a lower cost. Data Marketplace Operators can improve the experience of their Data Consumer customers, and at the same time create new monetization opportunities for their Data Provider customers: a Consumer that can see how two different datasets integrate is more likely to buy, compared to the scenario where they need to buy them before trying to integrate them.

Where the plugin can be accessed

The plugin is available on https://github.com/EU-UPCAST/data-integration-plugin

3.10 Negotiation and contracting

The UPCAST Negotiation and Contracting plugin facilitates seamless and policy-compliant data transactions between providers and consumers by managing the complete negotiation lifecycle from initiation to contract finalization. The plugin addresses inherent differences in processing intentions, usage constraints, and pricing expectations through a structured, machine-readable contract framework built on standards such as ODRL and DPV (Deliverable 4.1).

How the plugin can be used by SMEs and/or other beneficiaries

SMEs that sell data can integrate the plugin in their systems to enable negotiation of usage policies, price and environmental constraints. Conversely, SMEs that buy or have the need to access third party data can invite their providers to negotiate the terms of access. Data Marketplace Operators can offer it as an additional service for their users.

In all cases described above, the plugin may also be used to generate a Machine-Readable contract that can be semi-automatically processed by other tools, e.g., the Compliance and Auditing UPCAST plugins.

Users' target benefits and opportunities

Reduces time spent on manual back-and-forth negotiations and the number of errors and inconsistencies in contract terms. Facilitates the handling of many data agreements simultaneously, and the tracking of negotiation history, versioning, and approvals. The semi-automated capabilities are useful to accelerate negotiation cycles, speeding up drafting and reviewing agreements and onboarding of new providers or consumers. It also enables the attachment of human-readable versions to ensure collaboration with non-technical users.

Finally, the plugin implements and extends the IDSA Negotiation protocol, hence, a further benefit is alignment to established standards, providing the opportunity of interoperation with other tools.

Where the plugin can be accessed

The plugin is available on https://github.com/EU-UPCAST/negotiation-plugin

3.11 Secure Data Exchange

The Secure Data Exchange Plugin is an open source component relying on a open architecture with Trust as its anchor, able to interact both with other UPCAST plugins, and with other data connectors already existing in the markets. It:

- Allow data provider to perform data exchange with a trusted Safe data product Transfer Plugin,
- Interconnect distributed data connectors to perform data exchange under the supervision of a data space orchestrator.
- Enable organizations to design and manage data products from multiple data source in their own environment,
- Facilitate deployment of the plugin with cloud agnostic and low consumption solution,
- Provide advanced tracing and telemetry metrics to analyze performance, execution and compliance of Data Transfer.

How the plugin can be used by SMEs and/or other beneficiaries

The plugin covers the two main scenarios of data exchange:

- **One to one data exchange**: In this scenario, the *Safe data product Transfer Plugin* is working as a request/response proxy to exchange data between a data provider and a data acquirer. Once terms are negotiated and contract established, the data acquirer will be allowed to request data product transfer securely to a trusted data destination or consume data product within application directly.

- **One to many data exchange**: in this scenario, the *Safe data product Transfer Plugin* is working as a broadcast proxy to exchange data from a data provider to multiple data acquirers. Based on a data product delivered by the data provider, the plugin will automatically transfer this data product to all active subscribers based on their access

& usage rights in the contract negotiated. Data acquirer will negotiate a data contract before subscribing to the data product delivery process.

Users' target benefits and opportunities

Data providers

Using the Publishing plugin, data providers should be able to publish data products in the marketplace with additional metadata. Once data products are synchronized with the marketplace, data providers are able to create data products . Data products include data contracts composed with data policies and datasets. These policies will be replicated into the *Safe data product Transfer Plugin.*

Data Acquirers

Data Acquirers will be able to search data products with the discovery plugin. Once the data product is identified, the provider and acquirer agree with the terms of the contract such as price, policies, compliance. When the acquisition is complete through a trusted data transaction, the agreed data contract is synchronized with *Safe data product Transfer Plugin* to make it ready for incoming Data Transfer requests from Acquirer. The *Safe data product Transfer Plugin* is in charge of securely performing the data transfer enforcing agreed policies and data access restrictions with strong authentication.

Where the plugin can be accessed

The plugin is available in this Github repository.

3.12 Monitoring

The monitoring plugin of the UPCAST platform includes updates to the data model of the monitoring events and the automatic injection of monitoring code to the data processing workflow.

The Monitoring plugin is used for two purposes: (a) for collecting events that are generated form the use of other UPCAST plugins so as to later be able to perform auditing of the use of these plugins for compliance to legislation in force, and (b) for collecting events that are generated from the processing of a dataset to check compliance of the execution with the terms of the contract that has been agreed between a dataset provider and a dataset consumer (Deliverable 4.1).

How the plugin can be used by SMEs and/or other beneficiaries

Emission of events from other UPCAST plugins is performed as part of the function of those plugins. A client (SME, other beneficiary) can use the plugin for emitting events during the processing of a dataset and send them to topic UPCAST.

A typical Python client is like the following one (credentials are shown only for illustration purposes):

```
prod = KafkaProducer(bootstrap_servers=BROKER_URL,
                     security_protocol='SASL_SSL',
                     sasl_mechanism='PLAIN',
```

```
sasl_plain_username='UPC4ST-2025',
sasl_plain_password='3.14159',
key_serializer=lambda k: k.encode('utf-8'),
value_serializer=lambda v: json.dumps(v). encode('utf-8')
```

For a client to use the plugin and be able to send events, they have to declare the IP addresses from which the events will be generated to Maggioli for the whitelisted. After the username and password have to be agreed with the Maggioli technical staff events can be sent to topic UPCAST for the Monitoring plugin to collect them.

Users' target benefits and opportunities

The Monitoring plugin is used to collect events from the processing of a dataset. A user of the plugin can benefit from its functions to prove compliance with the terms of the contract that a dataset consumer has agreed with a dataset provider. The collected monitoring events are sent to the Compliance plugin also stored for auditing.

Where the Plugin can be accessed

The plugin can be accessed through https://dev.upcast.maggioli-research.gr/.

The plugin is offered as web service; source code is protected by IPRs and is not available for download.

3.13 Compliance

The compliance plugin is responsible for checking the compliance of a dataflow execution against the rules of the contract that has been agreed between the dataset provider and the dataset consumer. It receives as input the rules of the contract and the stream of events that come from a dataset workflow execution and flags any violations of the rules. In fact, the compliance plugin generates information message for compliant executions and alert messages for non-compliant executions.

How the plugin can be used by SMEs and/or other beneficiaries

The compliance plugin is designed to integrate and cooperate with the execution monitoring plugin (Deliverable 4.1).

Emission of events from other UPCAST plugins is performed as part of the function of those plugins. A client (SME, other beneficiary) can use the plugin for emitting events during the processing of a dataset.

Users' target benefits and opportunities

Users of the compliance plugin are dataset providers who benefit from the Compliance plugin as they receive real time information for the compliance or non-compliance of a dataset execution by a consumer based on the rules of the contract that has been agreed between the

two (dataset provider and consumer). The Compliance plugin generates in rea time alerts for executions that violate any of the rules of the contract as the dataset execution progresses. It is then up to the dataset provider to decide how to handle the violation.

Where the plugin can be accessed

The plugin cannot be accessed directly as it runs as a daemon service at the Maggioli Environment (https://dev.upcast.maggioli-research.gr/).

It is used indirectly through the Monitoring plugin and it sends is output to a set of configurable endpoints. Its source code is protected by IPRs and is not available for download.

3.14 Auditing TO DO DAWEX

UPCAST Auditing and validation plugin is responsible for attesting that the execution of a dataset, referenced by a unique contract-id and execution-id, is complete and for generating a verifiable proof of the result of the execution of a Data Processing Workflow on a dataset. This proof is called a Verifiable Credential (VC), which encapsulates all recorded events of an execution, certifying both its validity and compliance. This credential can be used as proof of execution adherence, ensuring transparency and auditability.

How the plugin can be used by SMEs and/or other beneficiaries

A company (Verifier) wants to ensure that a candidate (Holder) actually holds a degree issued by a University (Issuer). The company uses the Auditing and validation plugin to receive proof in the form of a Verifiable Presentation. Based on the signature and the trust in the issuer, the company or auditor can make a decision with formal proofs that the process has happened in accordance with the conditions agreed in the contract and negotiation plugin.

Users' target benefits and opportunities

The Validation Plugin enhances user confidence in the Data Space by ensuring that every contract execution is compliant, verifiable, and tamper-proof. It seamlessly integrates with the Negotiation Plugin to receive finalized contracts and with the Compliance Plugin to collect certified monitoring data. Through an automated process, it transforms execution logs into Verifiable Credentials, eliminating manual work and reducing errors. These credentials act as immutable proof of compliance, available for providers, consumers, and auditors alike.

For users, this means full auditability, transparency, and reliability in data transactions. By embedding compliance checks directly into the workflow, the Validation Plugin creates a trusted environment where data exchange becomes safer, faster, and easier to verify.

Where the plugin can be accessed

The plugin is available on **Github** : Verification Plugin Project

4 The European SMEs data-driven strategy

Digital technologies play a pivotal role in maximising the benefits of a data-driven society and a data-based economy but need to be solidly anchored to trustworthy, compliant, privacy-preserving and environmentally sustainable data sharing methods, architectures, and processes.

European Small and Medium Enterprises (SMEs) from 1 to 250 employees are at a critical juncture in digital transformation of European society, as these organizations represent more than 95% of all EU businesses and employ around two-thirds of the private-sector workforce.

The digitization stage of European SMEs varies significantly across countries, sectors, and business sizes, but in general, most European SMEs are engaged in digital transformation, and to be able to use data to inform and improve their business processes could significantly enhance competitiveness.

Understanding the current adoption levels, challenges, and perceived value of data-driven strategies and the SMEs required assets is essential for developing tailored solutions that align with their specific needs and constraints.

4.1 European SMEs data-driven strategy survey: main results

To give UPCAST project community's and other beneficiaries a set of data and information useful to understand the European SMEs data-driven strategies and how to address their digital needs with ready-to-use components, in May 2025 we run a **survey on 1,700 European SMEs with 1-250 employees** (see sample description in Appendix).

Based on this, the survey was geared to understand:

- level of innovation culture and digital maturity,
- top business priorities and challenges for the next 12 months,
- top IT investments and challenges for the next 12 months,
- level of maturity on online presence,
- primary technological approach used for procurement and customer relations and sales processes,
- main challenges encountered when implementing data-driven strategy projects,
- approach to acquiring or enhancing digital and data analysis skills,
- areas on which the SME most need support to increase its digital maturity,
- most critical data sources for organization's key business outcomes,
- data and analytics top priorities for the next 12 months.

In following pages, we provide the most relevant survey results that UPCAST project's community and other beneficiaries can leverage to:

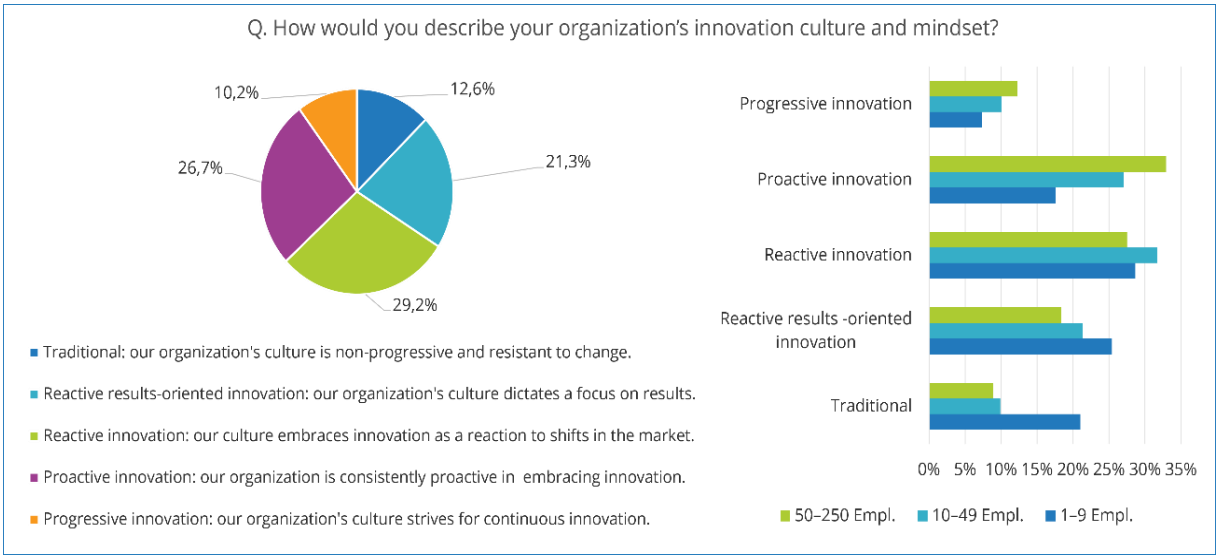
- refine the plugin capabilities tailored for SMEs, fine tune demos, training and approaches to better answer to specific needs of this type of customers.
- feed communications and dissemination with key information to improve the visibility to the results of the project.
- better identify exploitation strategy and go-to-market opportunities across the European data space.

Innovation culture and digital maturity

More than one-third of interviewed SMEs describe their companies' innovation culture and mindset as "proactive" or "progressive", as the innovation approach drive their organizations strategy and actions widely.

Although, 50% SME considers their innovation culture “reactive” or “reactive results-oriented”, as they innovate as a reaction to a market shift or focusing only some specific activities to obtain results, while the remaining 13% of SME maintains a “traditional” culture not oriented to introduce innovation in their business strategy and processes.

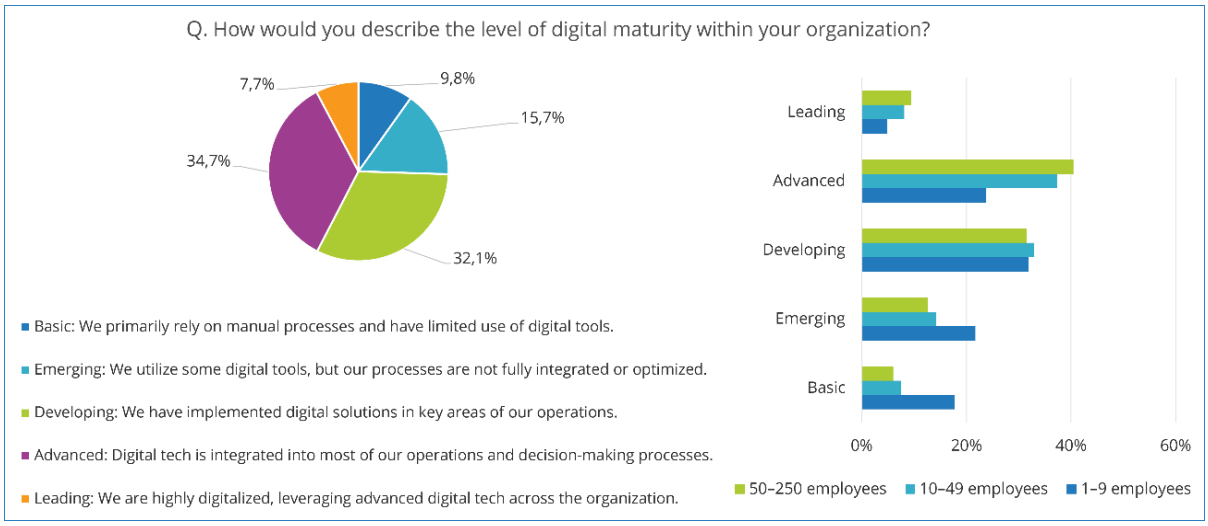
The smaller SMEs (1-9 empl.) have the most conservative approach, while most of largest SMEs (10-250 empl.) are equally confirming their mid-stage innovation culture approach, distributing themselves between “proactive” and “reactive” categories.



Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 2 – SMEs perceived innovation culture & mindset stage

Two-thirds of SMEs describe their companies’ level of digital maturity as "advanced" or "developing," having implemented technologies in several areas but taking time to extend them to digitize all processes. Despite 8% of SMEs saying their organizations are at the "leading" level, 26% of SMEs reported that they are at the "emerging" or "basic" digital maturity levels. Differences between company size are significant — 40% of large SMEs (50–250 employees) are in the "advanced" phase, while small companies (1–9 employees) are at the more at “emerging” and “basic” levels.



Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 3 – SMEs perceived digital maturity level

Business priority and challenges

Improving efficiency, productivity, and workflow management are the top business priorities of 35% of SMEs over the next year. Improving customer experience and cash flow priorities follow, but they are the most important for small SMEs (1–9 empl.). In about 30% of the largest SMEs (50–250 employees), enhancing innovation, improving competitive position, and expanding business area and launching new products are also included in the top 3 priorities.

Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 4 – SMEs top business priorities in next 12 months

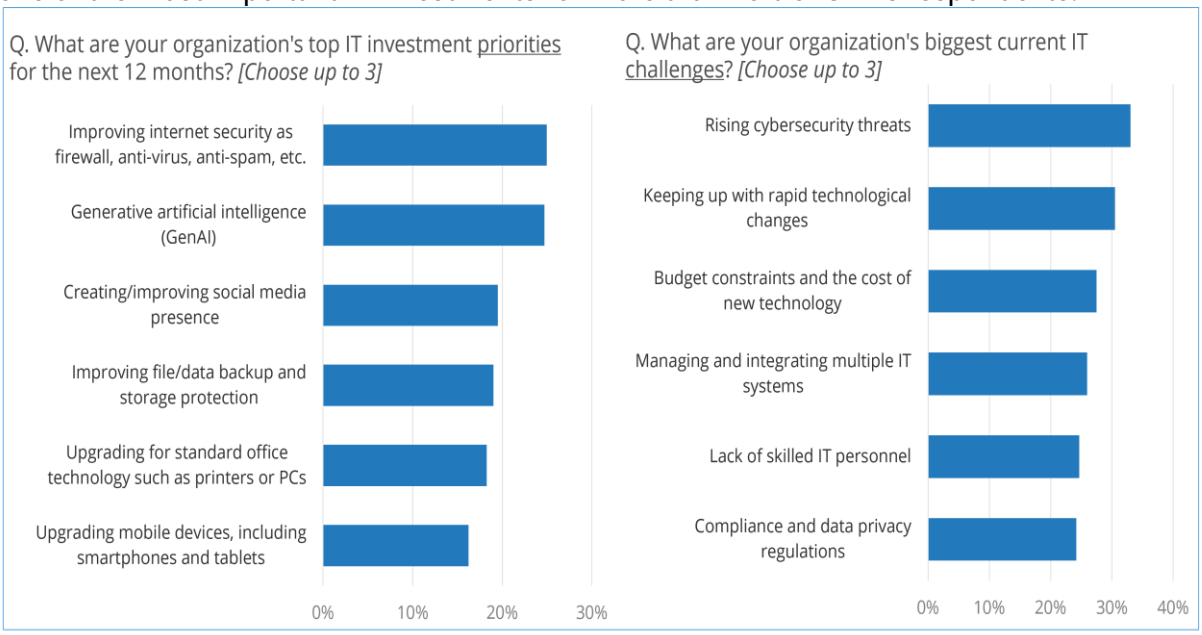
Keeping up with regulatory compliance, maintaining consistent cash flow, and addressing the lack of data into operations and customer activities are the most important business challenges. The main challenge for SMEs with 10–250 empl.is staying aligned with regulatory compliance and changes, while for small SMBs (1–9 empl.), it is maintaining consistent cash flow. These small companies also consider lack to capital to grow as a major challenge.

Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 5 – SMEs top business challenges in next 12 months

IT investments priority and challenges

Internet security, GenAI, and social media presence improvement are the top IT investment priorities, while the main IT challenges are related to cybersecurity risk, the speed of technological changes, and budget constraints. Security aspects are considered the most important investments and challenges for the largest SMEs (50–250 employees). GenAI is one of the most important IT investments for more than 20% of SMEs respondents.

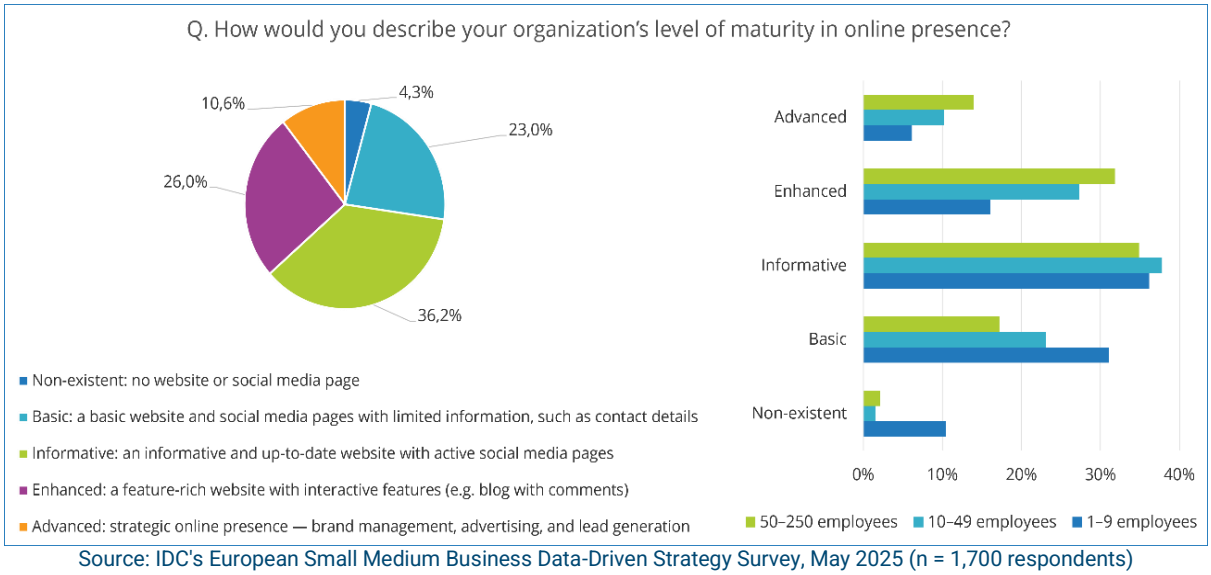


Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 6 – SMEs top IT investments and challenges

Online presence maturity level

37% of SMEs have high level of maturity in online presence ("enhanced" or "advanced") with interactive websites, social media and advanced usage of digital marketing tools. Another 36% of SMEs is at "informative" level with informative website and social media pages about their company profile and main activities, while 23% of SMEs have a "basic" level with limited information. Only 4% of SMEs don't have website or social media presence and these companies mainly belong to smaller sizes.

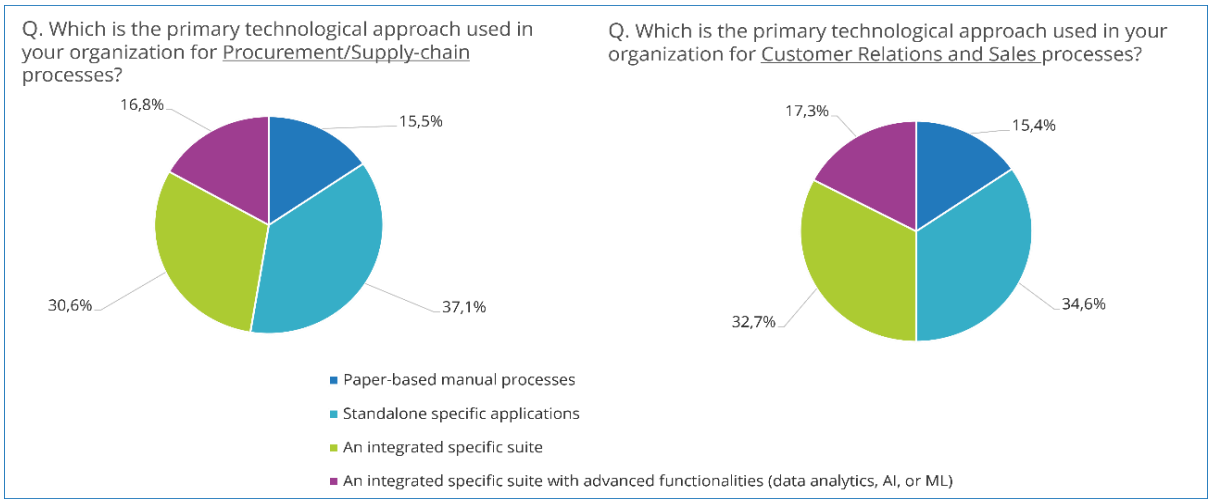


Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 7 – SMEs online presence maturity level

Technological approach to procurement/supply-chain and customer relations processes

On both processes almost half of SMEs have an advanced technological approach with integrated specific software suites and in about 17% of cases also advanced functionalities such as data analytics, AI, or ML. Although more than 35% of other SMEs are managing these processes with standalone specific suites, and the remaining 15% of companies are continuing to work on paper-based processes.



Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

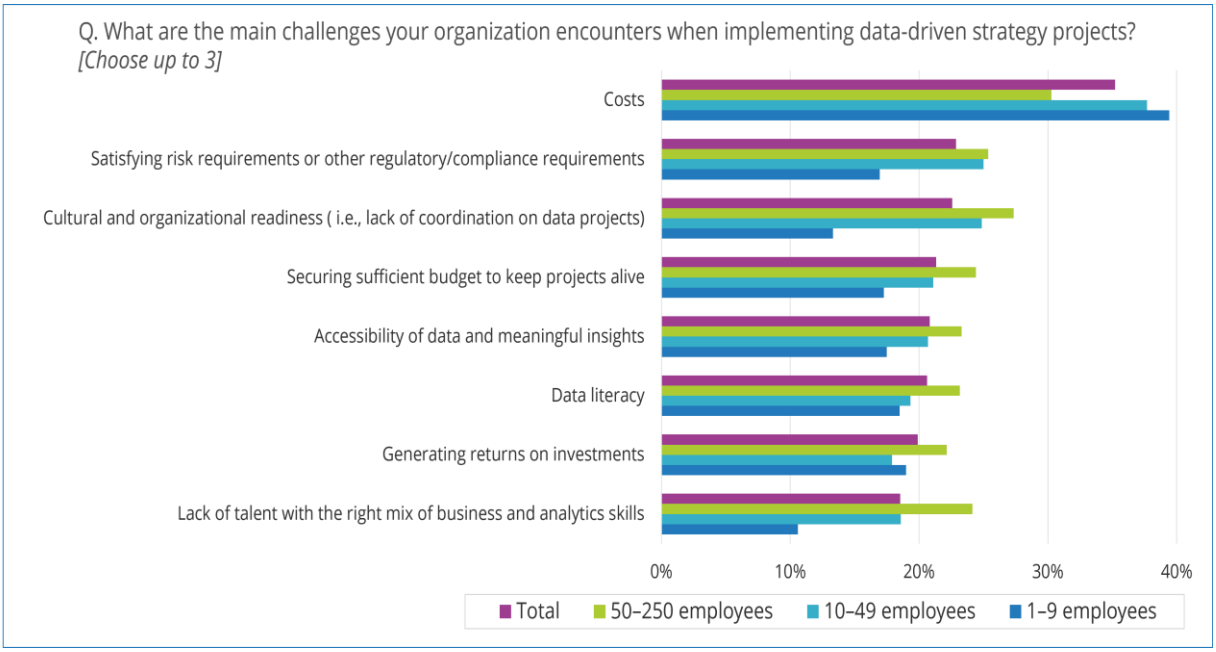
Figure 8 – SMEs digitization processes approaches

Main challenges encountered when implementing data-driven projects

For more than 30% of SMEs, and mainly for the smaller companies, costs are the main challenges to implement data-driven projects.

Regulatory requirements, cultural and organizational coordination, budget concerns on the long term, and accessibility to data and insights are mentioned as challenges by more than 20% of respondents.

SMEs with 50-250 employees are more sensitive to these challenges and to the lack of skills.

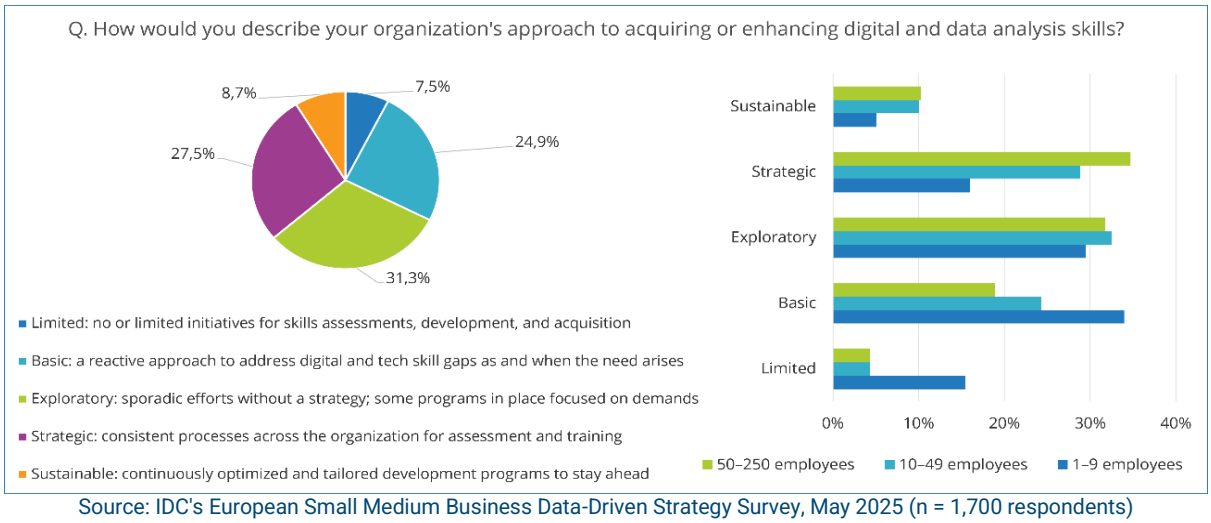


Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 9 – SMEs challenges to implement data-driven projects

Approach to acquiring or enhancing digital and data analysis skills

36% of SMEs have an advanced approach to acquire or enhancing digital and data analysis skills (“strategic” or “sustainable”), only the 9% of them continuously develop and optimize programs to stay ahead. 31% of SME have an “exploratory” approach based on sporadic efforts with some programs in place, while a remaining 33% of companies (mainly belonging to the smaller sizes) have a “basic” or “limited” approaches with limited initiatives launched when the needs arise.



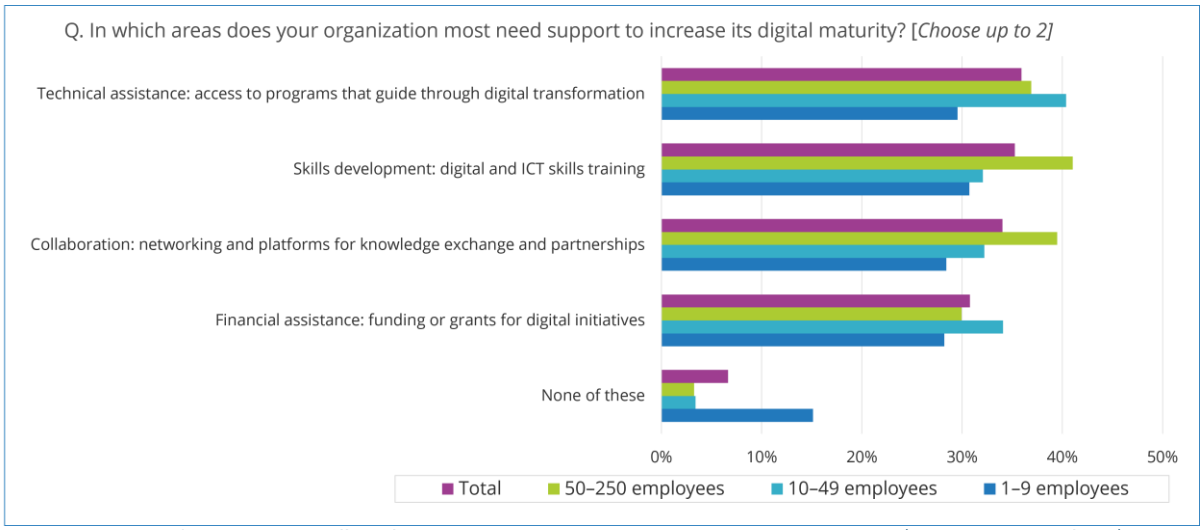
Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 10 – SMEs approach to acquiring digital and data skills

Needs of support to increase digital maturity

More than 90% of SMEs need technical support and guide to digital transformation (36%), digital and ICT training development (35%), collaboration tools to facilitate the knowledge exchange and partnerships (34%), and financial assistance for digital initiatives (31%).

Although while these needs are clearly perceived by almost all SME with more than 10 employees, 15% of smaller companies don't see them. Largest SMEs feel the skills development and collaboration needs, while mid-size SMEs perceive more the need of technical assistance through their DX roadmap.



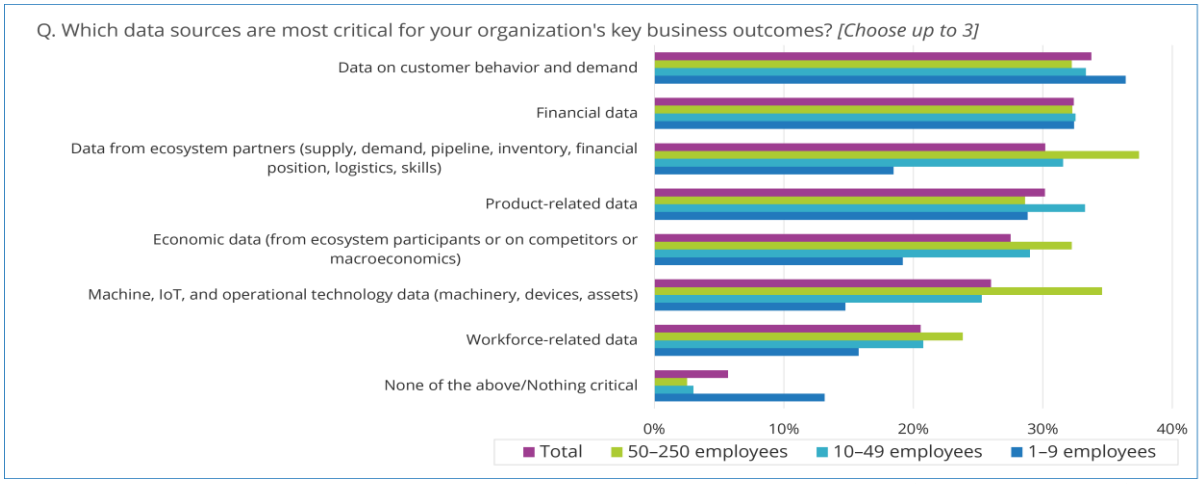
Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 11 – Areas where SMEs need support to increase digital maturity

Most critical data sources for achieving key business outcomes

More than one-third of SMEs consider data about customer behaviour and demand the most critical data sources to obtain relevant business outcomes. Financial data (as information about fundings, income statement, cash flow statement) are mentioned as most critical data for more than another third of SMEs. Data from the partner's ecosystem and product-related data are mentioned with a frequency of 30%.

Only 13% of smaller SMEs didn't perceive critical data sources to act successfully. Almost 40% of largest SMEs (50-250 empl.) consider as most critical data sources information from partners ecosystem and data from machines, IoT, operational, devices, assets.

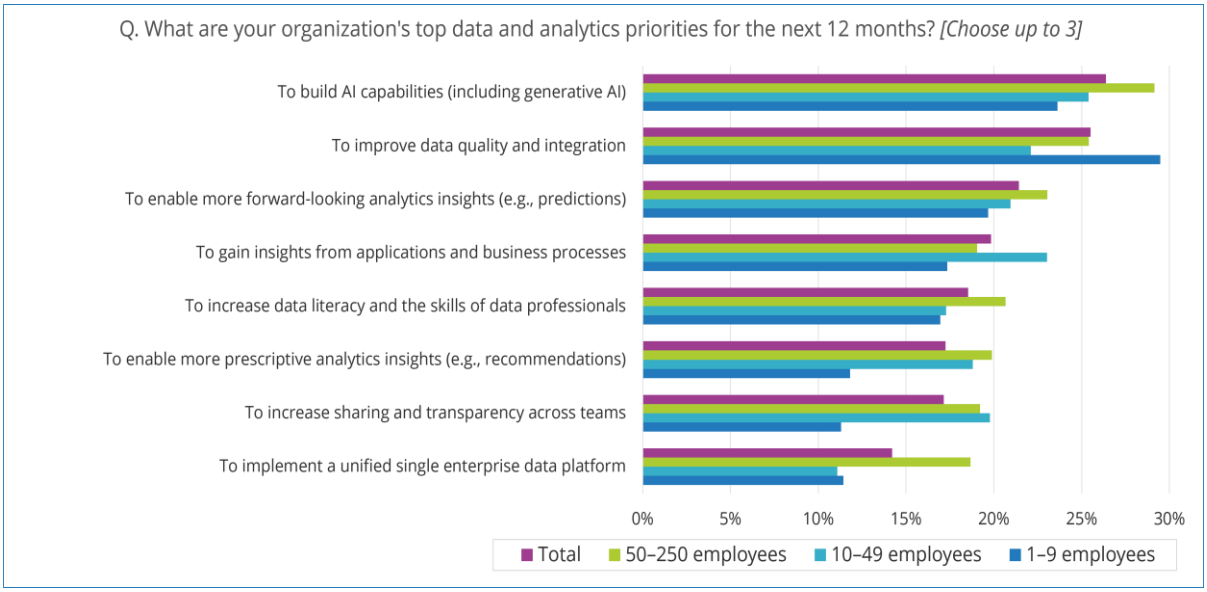


Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 12 – Most critical data sources for achieving key business outcomes

Data and analytics top priorities

Building AI capabilities and improving data quality and integration are the main data and analytics priorities (more than 25% of answers).
Smaller SMEs prioritize quality and integration, largest SMEs (50-250 empl.) focus the AI capabilities and mid-size companies give to insight from apps and business processes the priority.



Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 13 – SMEs Data and analytics priorities

4.2 European SMEs data-driven strategies: level of adoption and priorities

The survey highlights that SMEs are in the mid-stage of their data-driven strategies adoption, as:

- In more than one-third of SMEs innovation culture drive their organizations strategy and actions widely and two-thirds of SMEs have implemented technologies in several areas but taking time to extend them to digitize all processes.
- Improving efficiency, productivity, and workflow management are the top business priorities, while keeping up with regulatory compliance, maintaining consistent cash flow, and addressing the lack of data into operations and customer activities are the most important business challenges.
- Costs, regulatory requirements, cultural and organizational coordination are the main challenges perceived to implement data-driven projects.
- More than one-third of SMEs have an advanced approach to acquire or enhancing digital and data analysis skills, another third has an approach based on sporadic efforts with some programs in place, and only the remaining third of companies adopts approaches with limited initiatives.
- Most of SMEs say they need technical support and guide to digital transformation, digital and ICT training development, collaboration tools to facilitate the knowledge exchange and partnerships, and financial assistance to launch digital initiatives.
- More than one-third of SMEs consider data about customer behaviour and demand the most critical data sources to obtain relevant business outcomes, while financial data are mentioned as most critical data for more than another third of SMEs.
- Building AI capabilities and improving data quality and integration are the main data and analytics priorities in the next year.

5 SMES business plan guidance

Data sharing, monetization and trading platforms that enable actors in common European data spaces to collaboratively negotiate, improve and enforce data sharing contracts automatically, providing dynamic fair pricing mechanisms while implementing energy-efficient data exchange, ensuring privacy, confidentiality and legislation compliance and adhering to ethical and responsibility guidelines.

Consistently address this objective, we consider important to give SMEs and start-up some guidance and suggestions about how to leverage digital tools developed in UPCAST project to realize the benefits of data spaces, and more broadly, data-sharing platforms at scale.

Unlike large enterprises, which must have very detailed traditional business plans, SMEs need a lean and flexible business plan, tailored to facilitate SMEs to present a clear and realistic strategy for launching or growing the business in the European data market space.

For a SME or a start-up committed to entering the data market space, the creation of a business plan is a crucial step on the road to success. Knowing how to capture the interest of investors, customers, partners and talents makes it possible to concretize their expansion strategies.

5.1 Guidance for SMEs to grow in the European Data Space

SMEs managers and start-up who want to realize the benefits of data spaces, and more broadly, data-sharing platforms at scale, should:

- Building a resilient business innovation culture. Business leaders should foster an environment of trust and open communication in which employees feel valued and heard. They encourage adaptability by promoting continuous learning and innovation, ensuring the team can pivot in response to challenges. By setting clear goals and providing the necessary resources, these leaders empower their teams to take ownership and make informed decisions. They also prioritize mental and emotional well-being, recognizing that a healthy workforce is more capable of handling stress and uncertainty.
- Be ready to make complete and sudden changes. This is the practice of seeing an emerging idea or technology in a completely different direction. In practice, this occurs when a great idea is developed by a team that makes good brainstorming and has the potential to produce excellent outcomes, but before going forward, is better to verify and check the market readiness, internal and internal costs, risks and all other business plan indicators, and avoid that the exuberance of a great idea blinds people from seeing evidences that can negatively influence the economic results.
- Practice continuous learning. The ubiquitous phrase, "What got you here won't keep you here," is instinctually understood or learned by innovative organizations. What makes the innovations comes from not just an understanding of something new learned, but from the ability to make logical leaps from one discipline of thinking, industry, product, tool, tech, or code to another.
- Take advantage of funding programs, training programs, sandbox environments, and other government innovation vehicles that can help their organizations accelerate collaboration with SMEs. This approach allows familiarity with local regulations and certification requirements, and knowledge of local economic context.

- Observe lessons learned from EU data spaces projects and other emerging initiatives, such as data embassies, that governments globally are looking to set up in physical locations outside of their countries to create copies of key registries that could be used in case of a major availability incident in the country or vice-versa to offer other governments and private businesses safe locations.
- Define strategic goals including data sharing as a critical pillar to multiply the value of data in a trusted and efficient manner, in alignment with business priorities and regulatory compliance requirements.

Working on these directions will allow SMEs to grow leveraging experimental approaches to business model innovation, building strategic partnerships, and focusing on data-driven decisions by measuring key metrics to ensure continuous improvement and sustainable value creation.

5.2 Lean business plan step-by-step guide

A lean business plan and step-by-step guide to elaborate it can offer SMEs the following advantages:

- Simplicity: focus on main relevant business concepts, simple for readers to understand the values.
- Efficiency: concise description of key metrics and performance results, to give the readers a complete overview at a glance.
- Actionable insights: planned actions, impact evaluation with revenue streams, financial projections, funding ask, call to action, target market, will be crucial to attract investors and secure funding.

In addition, a lean business plan can be easily updated as new information and developments occur and encouraging an iterative and experimental approach to the market.

A lean business plan consists of 9 building blocks covering key aspects of the business model (see Figure 14 below):

1. Problem
2. Customer segment
3. Unique value proposition
4. Solution
5. Channels
6. Revenue streams
7. Cost structure
8. Key metrics
9. Competitive advantage

The step-by-step guide describe content and data the SME should include in each building block, focusing on questions to be prepared to answer when presenting its business plan:

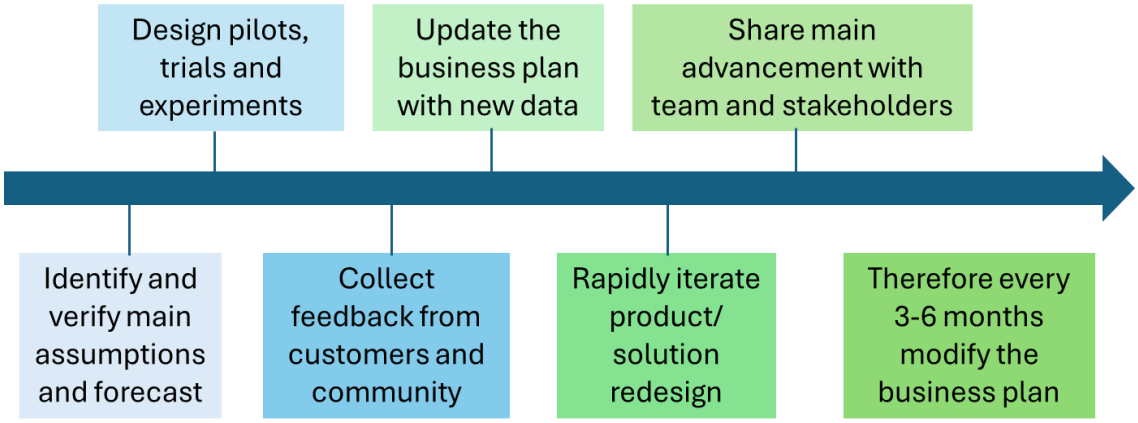
1. **Problem Statement:** the top 3 problems or needs that the product/service intends to solve for target customers should be listed here. It is also helpful to indicate existing alternatives, i.e., how customers currently address these problems. Aspects to consider:
   - What are the customers' main needs?
   - Which activities are most difficult or wasteful for them?
   - What needs are not currently met by the offerings on the market?
2. **Customer segments:** in this section the main target customer groups should be identified, with a focus on early adopters. It is essential to be as specific as possible, avoiding overly general definitions. Define your target market (size, segments, trends), customer personas and behaviours. Key questions to answer:
   - Who are our ideal customers?
   - What characteristics do they have in common?
   - Who might adopt our solution first (early adopters)?
3. **Unique value proposition:** this represents the heart of the business model, describing effectively why the proposed solution is unique and deserves customers' attention. Clearly define the market problem or gap you are solving, support with research, data, or quotes from customers or industry experts and show why existing solutions are insufficient. Elements to include:
   - Key customer benefits
   - Differentiation from competitors
   - Reasons why your product/service is better
4. **Innovative solution:** outline the 3 main features of the product/service that solve the identified problems. Describe your product/service and what makes it innovative (e.g. tech, IP, business model, process); show development status (prototype, MVP, ready-to-scale) and include visuals or diagrams if helpful. Key questions to answer:
   - What features are essential?
   - How does the solution solve customer problems?
   - What elements are most valuable to users?
5. **Marketing & Sales**: indicate how you intend to reach customers, with a go-to-market plan, customer acquisition strategy, marketing tactics, sales channels and strategic partnerships. Aspects to be evaluated:
   - Online and offline channels
   - Inbound and outbound strategies
   - Strategic partnerships
6. **Key Metrics**: identify 3-5 key indicators to measure the success and progress of the business in the early stages. Examples of relevant metrics:
   - Customer acquisition rate
   - Retention and churn rate
   - Customer acquisition cost
   - Lifetime value
   - Market Analysis

7. **Competitive advantage:** identify the factors that enable you to maintain a competitive and lasting advantage over competitors. Possible competitive advantages:
   - Intellectual property
   - Being part of an ecosystem/community
   - Economies of scale
   - Brand and reputation
8. **Cost Structure:** list the main cost items required to start up and run the business, both fixed and variable. Typical costs to consider:
   - Product development
   - Marketing and customer acquisition
   - Infrastructure and technology
   - Personnel
9. **Revenue streams:** describes main revenue sources of the business, specifying revenue model (e.g., SaaS, licensing, B2B sales), pricing strategy and scalability of the business model. Elements to include:
   - Revenue models (e.g. subscriptions, direct sales, freemium)
   - Indicative prices
   - 3–5 years of projections (income statement, cash flow, balance sheet).
   - Break-even analysis
   - Funding required and use of funds
   - Exit strategy (for investors).

The lean business plan should not be considered as a static document, but as a dynamic tool to be constantly updated. It could be a valuable tool for innovators who want to maximize the success for their projects, but its effectiveness depends on the ability to use it dynamically and iteratively, constantly testing their assumptions through experiments and customer feedback (see Figure 15), to arrive at a business model that is truly validated by the market.

Source: IDC elaboration on several sources, 2025

Figure 14 – Lean business plan model steps



Source: IDC elaboration on several sources, 2025

Figure 15 – Lean business plan creation process

5.3 Business Plan Guidance key take-away

SMEs managers and innovation officers who want to realize the benefits of data spaces, and more broadly, data-sharing platforms at scale, should:

- Building a resilient business innovation culture.
- Be ready to make complete changes,
- Practice continuous learning,
- Be able to take advantage of funding programs, and other support initiatives,
- Observe lessons learned from EU data spaces projects and other emerging initiatives,
- Define strategic goals and data sharing approach.

SMEs should design a lean and flexible business plan, tailored to facilitate them to present a clear and realistic strategy for launching or growing the business.

A lean business plan consists of 9 building blocks covering key aspects of the business model (see Figure 14).

1. Problem
2. Customer segment
3. Unique value proposition
4. Solution
5. Channels
6. Revenue streams
7. Cost structure
8. Key metrics
9. Competitive advantage

SMEs can adopt a step-by-step to define actions a goal in each building block, focusing on questions to be prepared to answer when presenting its business plan to investors and business communities.

6 Conclusions

Digital technologies and AI are transforming businesses and public administrations as well as changing consumer experiences and society. European organizations will need to increasingly work with partners across their ecosystems to bring together the volume and granularity of data that is necessary for digitize their processes and exploit benefits.

One of the pillars of the European Union strategy is the bold vision "to create a single European data space — a genuine single market for data, open to data from across the world — where personal as well as non-personal data, including sensitive business data, is secure and businesses also have easy access to an almost infinite amount of high-quality industrial data, boosting growth and creating value, while minimizing the human carbon and environmental footprint."

The emerging data spaces architectural and governance archetypes such as open data repositories, data marketplaces, and trusted data networks are arising offering compelling content and tools that software developers, SMEs and start-up and innovative business leaders can leverage.

Nonetheless, evolving regulatory framework and technological developments, that create new modalities for users to access and share data, are amplifying governance, semantic, technical interoperability, and legal interoperability challenges.

As data spaces is everyday more complex, organizations, providers, part the data space ecosystem, have to increasing and focus their efforts on:

- **Data governance and control investments**. Data governance refers to the exercise of authority and control over the management of data. The purpose of data governance is to increase the value of data and minimize data-related cost and risk. Governance is critical in the context of data spaces because of the multilateral nature of the ecosystem.
    - Providers have to agree and comply with sets of common rules for data sharing to prevent infringement fundamental rights such as privacy or dominant companies from exerting unilateral control, thereby safeguarding the ability to generate mutual benefits, while also adhering to digital sovereignty policies and guidelines.
    - Governments and private organizations collaborate to develop, disseminate, and maintain common, reusable governance policies and guidelines for federated management, sharing, and exploitation of data, while ensuring security and privacy.
    - Policymakers (as EU Commission) work with the private sector to make available simple guides that transcribe important aspects of regulation into practical examples of business, governance, and legal capabilities in the context of cross-organization and cross-industry data-sharing use cases. At the very least, these guidelines should embed data protection principles and fair market/competition and IP protection principles that align with legislation. At the data-sharing platform level, these policies, principles, and guidelines should be translated into contractual terms and conditions for user identification and access management, data classification, data provisioning SLAs, billing, anonymization, and so forth.
- **Common reusable technological architectural, coding, and semantic blueprints development**. International initiatives as, the Sitra Rulebook, the IDSA Rulebook, the EU DSSC blueprints, and Gaia-X, and also EU funded projects as Open DEI, FIWARE and UPCAST offer a starting point for these technical building blocks.

- **Guidelines and principles embedded into automated protocols within the data-sharing platform infrastructure** to streamline the multiple required verification processes and ensure accountability along the compliance chain as a whole; for instance, by embedding capabilities to automate data classification rules that trigger deployment in sovereign or non-sovereign environments or automating assignment of tasks in case of shared responsibilities for security updates.

To effectively act in these above directions, UPCAST developed a set of activities focused to advance the state-of-the-art of data sharing and trading, provide the tools to capture data sharing agreements/contracts in a formal, machine-processable representation and technology for their algorithmic enforcement, capture pricing and complex contractual requirements, and encode clauses of relevant legislation (such as GDPR).

With the integrated platform and several plugins, an extended knowledge of the maturity level of SMEs and a set of simple instruction to capture the interest of investors, customers, partners and talents UPCAST projects can accelerate the SMEs data processing agreement processes dramatically and help them to monetize data they create and manage, to grow in the European data space.

7 References

- D1.3 UPCAST: Updated project concept and architecture," 2024
- D4.1 UPCAST "Integrated Solution", 2025

ACRONYMS

Acronym	Explanation
AI	Artificial Intelligence
API	Application Programming Interface
CPU	Central Processing Unit
DPV	Data Privacy Vocabulary
DPW	Data Processing Workflow
DSL	Domain-Specific Language
EIO	Environmental Impact Optimiser
GDPR	General Data Protection Regulation
GUI	Graphical User Interface
HE	Homomorphic Encryption
HPC	High Performance Computing
IDSA	International Data Spaces Association
LLM	Large Language Model
MVP	Minimum Viable Product
NLP	Natural Language Processing
ODRL	Open Digital Rights Language
PDP	Policy Decision Point
PMP	Policy Management Point
PUC	Privacy and Usage Control
RC	Resource Consumer
RP	Resource Provider
UI	User Interface
WMO	Workflow Model Ontology

1 ANNEX 1: IDC IDC'S EUROPEAN SMALL MEDIUM BUSINESS DATA-DRIVEN STRATEGY SURVEY - METHODOLOGY AND SAMPLE STRUCTURE

IDC's European Consumer & Small Medium Business Research team conducted a survey to identify the level of digital maturity, data usage and stage of their data-driven strategies of European SMEs across 13 European countries, segmented by company size and sectors.

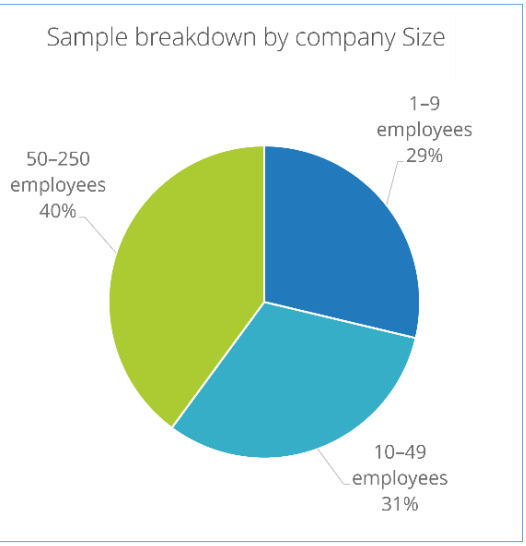
The field survey was run in May 2025, with CAWI (Computer Aided Web Interview) data collection methodology, to SMEs screened for size, sector, and usage of fixed and/or mobile telecommunications line, registered as business contract.

The survey targeted 1,700 organizations distributed in 13 Countries with 1–250 employees distributed in 3 company sizes, and belonging to all vertical sectors (see figure below)

Target respondents were: IT and LOB respondents, managers, DMs, and influencers for their organizations' decisions related to IT (hardware, software, IT services) and telecommunications (fixed and/or mobile line).

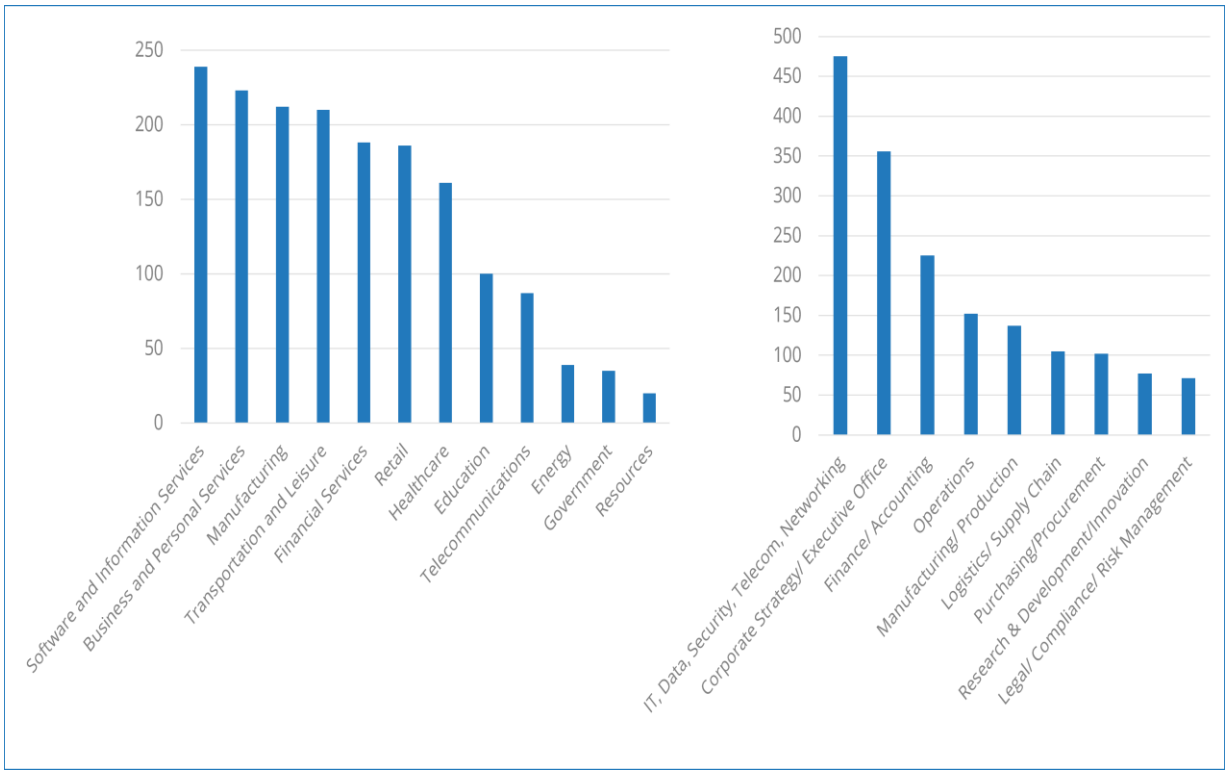
The sampling methodology was based on quotas by country and size (number of employees) and the results were weighted by country, company size and sector.

Country	Sample by Country (N. of companies)	%
Belgium	50	2.9
Czech Republic	100	5.9
Denmark	30	1.8
Finland	30	1.8
France	200	11.8
Germany	250	14.7
Greece	100	5.9
Italy	200	11.8
Netherlands	50	2.9
Poland	200	11.8
Spain	200	11.8
Sweden	40	2.4
U.K.	250	14.7
Total	**1,700**	**100**



Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 16 – Survey sample structure by Country and company size.



Source: IDC's European Small Medium Business Data-Driven Strategy Survey, May 2025 (n = 1,700 respondents)

Figure 17 – Survey sample structure by Industry and Respondent's role