D2.7 SYNERGY Framework Architecture including functional, technical and communication specifications v2

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Big Energy Data Value Creation within SYNergetic enERGY-as-a-service Applications through trusted multi party data sharing over an AI big data analytics marketplace

Deliverable nº:	**D2.7**
Deliverable name:	**SYNERGY Framework Architecture including functional, technical and communication specifications v2**
Version:	**1.00**
Release date:	**24/12/2021**
Dissemination level:	**Public**
Status:	**Submitted**
Authors:	**UBITECH**, Suite5, MAG, KBZ, ICCS, COBRA, CUE, ETRA, VTT, CIRCE, VERD, CAV, FVH

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Document history:

Version	Date of issue	Content and changes	Edited by
0.10	05/11/2021	ToC	Dimitris Miltiadou (UBITECH)
0.20	25/11/2021	Example component description	Dimitris Miltiadou (UBITECH)
0.30	30/11/2021	Initial contributions on Sections 2, 3 and 7	UBITECH, Suite5
0.35	2/12/2021	Initial contributions on Sections 4 and 5	UBITECH, Suite5, KBZ, MAG
0.40	3/12/2021	Updated contributions on Sections 3, 4, 5 and 7	UBITECH, Suite5
0.45	06/12/2021	Initial contributions on Section 6	ICCS, COBRA, CUE, ETRA, Suite5, VTT, CIRCE, VERD, CAV, FVH
0.50	07/12/2021	Updated contributions on Sections 4 and 5	UBITECH, Suite5, KBZ, MAG
0.60	10/12/2021	Updated contributions on Sections 3 and 7	UBITECH, Suite5
0.70	13/12/2021	Updated contributions on Section 6	ICCS, COBRA, CUE, ETRA, Suite5, VTT, CIRCE, VERD, CAV, FVH
0.80	17/12/2021	Draft version	UBITECH
0.90	22/12/2021	Reviewed version	CIRCE, MAG
1.00	24/12/2021	Final version	UBITECH

Peer reviewed by:

Partner	Reviewer
CIRCE	Nurseda Yildirim Yurusen
MAG	Nikos Achilleopoulos

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Deliverable beneficiaries:

WP / Task
WP2 / Task 2.4
WP3 / Tasks 3.1, 3.2, 3.3, 3.4, 3.5
WP4 / Tasks 4.1, 4.2, 4.3, 4.4, 4.5
WP5 / Tasks 5.1, 5.2, 5.3, 5.4
WP6 / Tasks 6.1, 6.2, 6.3, 6.4
WP7 / Tasks 7.1, 7.2, 7.3, 7.4

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Table of contents

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Abbreviations and Acronyms

Acronym	Description
ABAC	Attribute-Based Access Control
AI	Artificial Intelligence
API	Application Programming Interfaces
BACS	Building Automation and Control System
BDVA	Big Data Value Association
BPMN	Business Process Model and Notation
CI/CD	Continuous Integration / Continuous Deployment
CIM	Common Information Model
CORS	Cross-Origin Resource Sharing
CRUD	Create Read Update Delete
CSV	Comma-Separated Values
DER	Distributed Energy Resources
DL	Deep Learning
DHW	Domestic Hot Water
DoA	Description of Action (annex I of the Grant Agreement)
DPO	Data Protection Officer
DR	Demand Response
DSO	Distribution System Operator
DSM	Demand Side Management
DSS	Decision Support System
EaaS	Energy-as-a-Service
ESCO	Energy Service Company
GUI	Graphical User Interface
HVAC	Heating, Ventilation and Air Conditioning
JSON	JavaScript Object Notation
KPI	Key Performance Indicators
LV	Low Voltage
M	Month
ML	Machine Learning

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Acronym	Description
MV	Medium Voltage
OPF	Optimal Power Flow
PPP	Power Production Profiles
RES	Renewable Energy Resources
SaaS	Software-as-a-Service
SECAP	Sustainable Energy and Climate Action Plan
SGAM	Smart Grid Architecture Model
SOH	State of Health
TSO	Transmission System Operator
TSV	Tab-Separated Values
VPP	Virtual Power Plant
XML	Extensible Markup Language
WP	Work Package

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Executive summary

The deliverable at hand entitled "D2.7 - SYNERGY Framework Architecture including functional, technical and communication specifications v2" reports the final efforts and the produced results of Task 2.4 "Detailed architecture design, protocols and interfaces specifications for Big Data-enabled Energy Services" of WP2 "Use Cases, Business Requirements and Architecture Design". The main purpose of this final report is to deliver the final version of the overall conceptual architecture of the SYNERGY platform supplemented by the final design specifications of the SYNERGY platform's components. The deliverable is building directly on top of the previous work documented in deliverable D2.6, as well as use cases and requirements analysis, the regulatory framework analysis and the data landscaping activities in WP2 (as documented in D2.2, D2.4 and D2.5, respectively).

Towards this direction, the scope of the current deliverable can be described in the following axes:

- **To deliver the final overall conceptual architecture of the SYNERGY platform**. Towards this end, the deliverable presents the complete and detailed documentation of the three core layers of the architecture, namely the SYNERGY Cloud Infrastructure, the On-Premise Environments and the SYNERGY Energy Apps Portfolio, by documenting their core elements and role in the overall architecture. The deliverable presents also the final list of data-driven services bundles and energy apps which are the core ingredients of the layers of the architecture. In particular, the updated documentation of the eight (8) service bundles and twelve (12) Energy Apps of the SYNERGY platform is presented highlighting their role and overall offerings in the SYNERGY platform. Finally, the final list of components which compose the platform's architecture are presented focusing on their concrete context and positioning within the overall architecture along with the list of roles and users of the SYNERGY platform.
- **To document the updated detailed SYNERGY platform's workflows.** The previously designed platform's workflows have been optimised in order to include the complete list of functionalities which are offered by the SYNERGY platform from the user's perspective. In total, eleven (11) workflows are presented in the form of BPMN (Business Process Model and Notation) diagrams which clearly depict the interactions of the user with the SYNERGY platform, as well as the interactions of various components during the realisation of the designed platform functionalities. The workflows are organised into three main categories,

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

namely the data check-in, the data search and sharing and the data analytics workflows in which all the platform's functionalities are grouped.

- **To provide the updated detailed design specifications of the SYNERGY platform's components.** As the project evolved, the initially designed components received several updates and optimisations. The deliverable presents the complete list of components of the SYNERGY platform which is composed of twenty eight (28) core components in total, plus the ones incorporated on each of the twelve (12) Energy Apps (31 components in total), based on their positioning in the three core layers of the architecture. The updated components description presents their involvement in the respective data-driven services bundle or the Energy Apps and the respective architecture layer. In addition to this, for each component the scope and the role in the platform is documented and the complete list of features of each component is presented. Furthermore, the list of SYNERGY requirements that each component addresses is presented accompanied by the list of technologies that will be leveraged for their implementation. Finally, the updates from the previous version of the SYNERGY platform's architecture are presented.
- **To document the mapping between the SYNERGY platform's architecture and two core reference architectures.** In particular, the deliverable documents the alignment of the SYNERGY platform's architecture with SGAM (Smart Grid Architecture Model) and BDVA (Big Data Value Association) Reference Architectures, presenting how the SYNERGY platform's architecture aspects are mapped to the different aspects of these core reference architectures.

The current deliverable constitutes the final report of Task 2.4 and delivers the final version of the overall conceptual architecture of the SYNERGY platform and the platform's components. The updated documentation supplements the initial documentation provided in deliverable D2.6 by incorporating all the enhancements and optimisations that were introduced as a result of the performed development activities in WP3-WP7 and the analysis of the feedback that has been collected from the SYNERGY stakeholders. The specific deliverable concludes the activities of the specific task.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

1 Introduction

1.1 Purpose of the document

This deliverable presents the work performed in the context of Task 2.4 “Detailed architecture design, protocols and interfaces specifications for Big Data-enabled Energy Services” of WP2 “Use Cases, Business Requirements and Architecture Design” till M24. Its main purpose is to document the final version of the overall conceptual architecture of the SYNERGY platform along with the final specifications of the platform’s components and their respective functionalities.

The deliverable is building on top of the work performed during the first period of project that was documented in deliverable D2.6 where the initial overall conceptual architecture of the SYNERGY platform was documented. Following the extensive analysis of the updated SYNERGY user and business requirements that was performed and documented in deliverable D2.2, the current deliverable translated the results of this analysis into optimisations and enhancements that are introduced in the initial designs of the core components of the SYNERGY platform and their functionalities. These optimisations and enhancements were also propagated to the initial overall conceptual architecture of SYNERGY in order to be enhanced and updated towards the final overall conceptual architecture. The final conceptual architecture is a multi-layered architecture that effectively addresses the requirements of the energy stakeholders in a holistic manner taking into consideration: (a) the updates on the constraints and barriers that have been presented in the SYNERGY deliverable D2.4, (b) the advancements in the energy data landscaping activities that express implicit requirements for the SYNERGY platform as documented in the SYNERGY deliverable D2.5, in addition to the updated use cases and requirements outlined in D2.2. The overall platform’s offerings to each stakeholder of the platform are broken down to the platform’s workflow in which the platform’s distinct components are properly combined and interacting.

Following the same approach as in the first iteration, the design process that was followed for the final version the overall conceptual architecture of the SYNERGY platform was an iterative process, where multiple iterations were performed by the consortium members towards the effective update and optimisation of the initial concrete and solid design specifications of the involved components on each workflow. This results into a series of updates in the designed components and the respective data-driven services bundles that they belong in order to provide the overall set of offerings to the SYNERGY platform.

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

The current deliverable presents the final version of the reference architecture of the SYNERGY platform, as well as the final specifications of the components of the platform, that drive the implementation activities which are currently being performed in the context of WP3, WP4, WP5, WP6 and WP7 of the project. The activities of Task 2.4 concluded on M24 and the produced outcomes will be used by the other work packages as the project evolves.

1.2 Scope of the document

The deliverable D2.7 "SYNERGY Framework Architecture including functional, technical and communication specifications v2" documents the efforts undertaken within the context of Task 2.4 "Detailed architecture design, protocols and interfaces specifications for Big Data-enabled Energy Services" till the completion of the task's activities. It constitutes the final and complete report of the outcomes of this specific task in accordance with the SYNERGY Description of Action on M24 providing the required updates on the previous documentation that was delivered on M12 with regards to the detailed design of the SYNERGY platform's architecture and components' specifications.

In this context, the first objective of the deliverable is to present in detail the final overall conceptual architecture of the SYNERGY platform, highlighting the updates from the previous version as well as the rationale behind these updates. At first, the three main layers of the architecture are presented, documenting their updated context and role in the overall architecture. Following the layers presentation, the updated and final data-driven services bundles, as well as the energy applications, that compile the different layers of the architecture are documents focusing on their context and their overall functionalities. Furthermore, the updated list of components which compose these services bundles and energy applications are presented, describing on their positioning within the overall architecture. Finally, the final overall conceptual architecture presentation is supplemented by the documentation of the different roles and users that are expected to use the SYNERGY platform which remained unchanged.

Following the overall conceptual architecture description, the second objective of the deliverable concerns the updated designs of the user-driven workflows of the SYNERGY platform. The initially designed workflows were updated and enhanced to document the latest platform's functionalities and depict the final interactions of the different roles and users with the platform, as well as the interactions of the platform's components towards the realisation of the final platform's core functionalities around data collection, data search and sharing, and data analytics.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

The third objective of the deliverable is to document the updated and final detailed design specifications of the SYNERGY platform's components. The updated design specifications of each component are organised by the different layers of the architecture, where the components of the involved data-driven services bundles and the energy applications are presented. For each component, the updated description of its role in the architecture is presented along with the final list of features, the addressed SYNERGY requirements as extracted from deliverable D2.2, the technologies that will be exploited for its implementation and the updates from the initial version.

The fourth objective of the deliverable is to document the mapping between the SYNERGY platform's architecture and the SGAM and BDVA Reference Architectures. To this end, the deliverable documents how the SYNERGY platform's reference architecture is aligned with the two reference models, highlighting how their distinct aspects are covered (without any significant updates in respect to D2.6).

The deliverable constitutes the second and final iteration of the design process of the SYNERGY platform's architecture. It documents all the updates introduced from the previous version in the course of development of the SYNERGY platform. It provides the final design specifications of the SYNERGY platform's architecture and of the components of the architecture on M24 and concludes the task's activities in accordance with the SYNERGY Description of Action.

1.3 Structure of the document

This document is structured as follows:

- Section 2 presents the final version of the overall conceptual architecture of the SYNERGY platform by presenting the different layers of the architecture, the data-driven services bundles that are present on each layer and the components of each bundle and each layer. Finally, it presents the platform roles and users of the SYNERGY platform.
- Section 3 presents the updated core workflows of the SYNERGY platform, describing the platform's functionalities and the interactions of the different users with the platform (from the user-oriented perspective of data asset providers and data asset consumers), as well as the interactions between the various components.
- Sections 4, 5 and 6 present the final detailed design specifications of the SYNERGY platform's components for each layer of the architecture, documenting their scope, list of features, the

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

list of addressed requirements and the technologies that will be exploited by each component, along with the updates introduced from the first version.

- Section 7 documents the mapping between to the SYNERGY platform's architecture and the SGAM and BDVA Reference Architectures.
- Section 8 concludes the SYNERGY reference architecture.
- Section 9 provides a list of references.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

2 Overall SYNERGY Platform Architecture Overview

In this section, an overview of the high-level architecture is presented, describing the different layers of the architecture, the components, and the purpose of each component on each layer.

2.1 Layers View

In brief, SYNERGY aims to develop an all-around data platform that builds on state-of-the-art technologies, is driven by the actual needs of the electricity data value chain and turns over a new leaf in the way data sharing and data analytics are leveraged. Taking into consideration the different Use Cases and Requirements of the SYNERGY Project (as reflected in D2.2), the reference architecture of the overall SYNERGY platform (as also described in D2.6) has been conceptually divided in 3 main layers as depicted in Figure 1:

- The **SYNERGY Cloud Infrastructure** that consists of: (a) the **Core Big Data Management Platform**, essentially including the Energy Big Data Platform and the AI Analytics Marketplace which are instrumental for all functionalities that SYNERGY supports at all layers, and (b) the **Secure Experimentation Playgrounds** (SEP) which are realized in the form of dedicated virtual machines that are spawned per organisation to ensure that each electricity data value chain stakeholder is able to execute big data analytics in isolated and secure environments in the SYNERGY cloud infrastructure.
- The **On-Premise Environments** (OPE) which are executed in the energy stakeholders' premises for increased security and trust and can be distinguished in the **server environment** and the **edge environments** that are installed in gateways. The On-Premise Environments are not self-standing, but always communicate with the SYNERGY Cloud Infrastructure to deliver their intended functionality.
- The **SYNERGY Energy Apps Portfolio** that embraces the set of applications addressed to the needs of: (a) DSOs (Distribution System Operators), TSOs (Transmission System Operators) and RES (Renewable Energy Sources) Operators in respect to grid-level analytics for optimized network and asset management services, (b) Electricity Retailers and Aggregators for portfolio-level analytics towards Energy-as-a-Service (EaaS) solutions, (c) Facility Managers and ESCOs

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

(Energy Service Companies) towards building / district-level analytics from the perspective of optimized energy performance management.

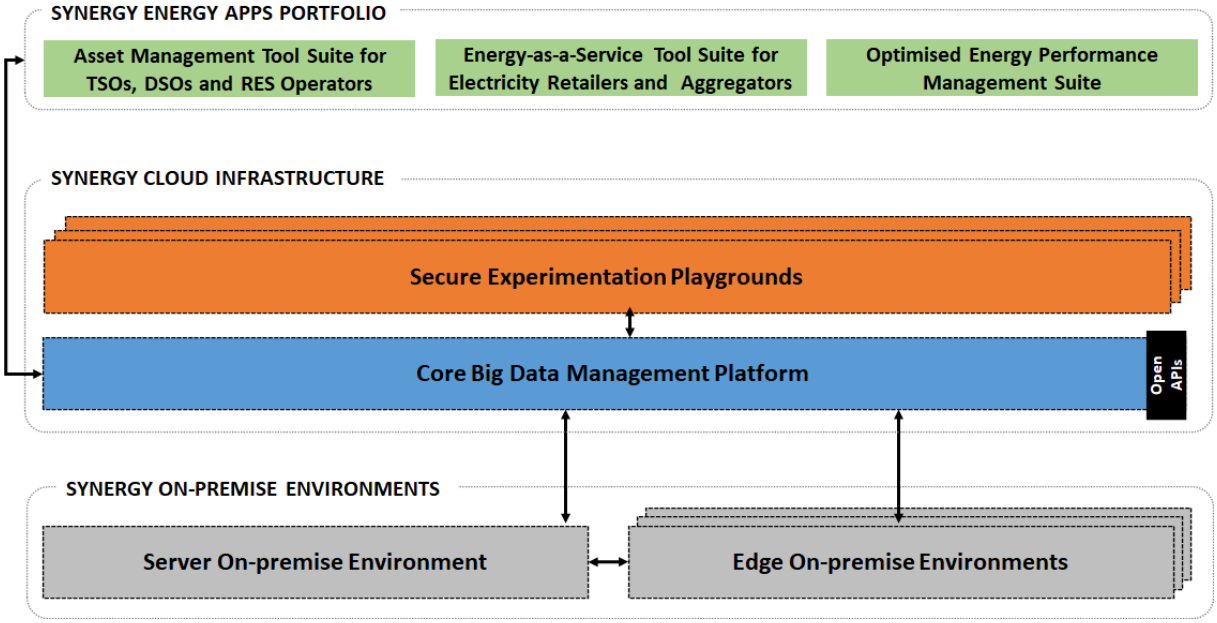


Figure 1: SYNERGY 3-Layered High-level Architecture

2.2 Data Services Bundles View

In order to deliver the intended functionalities towards the different electricity data value chain stakeholders who at any moment may assume the role of data asset providers and / or data asset consumers, the high-level architecture consists of the following data-driven services bundles that have well-defined interfaces to ensure their seamless integration and operation within the SYNERGY integrated platform:

- **Data Collection Services Bundle** which enables the configuration of the data check-in process by the data provider at "design" time in the Core Big Data Management Platform and its proper execution in the SYNERGY Cloud Infrastructure and / or the On-Premise Environments. Different data ingestion, mapping and transformation and cleaning services are invoked to appropriately handle batch, near real-time and streaming data collection.
- **Data Security Services Bundle** that is responsible for safeguarding the data assets in the overall SYNERGY platform (i.e. Core Big Data Management Platform and On-Premise Environments for end-to-end security) through different ways, e.g. by anonymising the sensitive data (from an individual or business perspective), by selectively encrypting the data, and by applying

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

access policies over the data assets that allow a data provider to control who can even view them.

- **Data Sharing Services Bundle**, essentially providing the SYNERGY Core Big Data Management Platform with the functionalities expected from a data and AI analytics marketplace in terms of sharing and trading data assets in a secure and trustful manner, powered by the immutability and non-repudiation aspects that are available in Distributed Ledger Technologies.
- **Data Matchmaking Services Bundle** that delivers exploration and search functionalities (in the SYNERGY Core Big Data Management Platform) over data assets that the data consumers are eligible to view and potentially acquire while providing recommendations for additional data assets of interest or for electricity data value chain stakeholders who could potentially have/create the requested data asset.
- **Data Analytics Services Bundle** which lies at the core of the design of data analytics pipelines including the data manipulation configuration, the basic and baseline (pre-trained) machine learning and deep learning algorithms configuration and the visualization/results configuration, in the SYNERGY Core Big Data Management Platform, while allowing for the execution of the defined pipelines in the Secure Experimentation Playgrounds and the On-Premise Environments.
- **Data Storage Services Bundle** that offers different persistence modalities (ranging from storage of the data assets, their metadata, their indexing, the algorithms and pipelines, the contracts' ledger, etc.) depending on the scope and the type of the data in the SYNERGY Cloud Infrastructure (in the Core Cloud Platform and the Secure Experimentation Playgrounds) and the On-Premise Environments.
- **Data Governance Services Bundle** that provides different features to support the proper coordination and end-to-end management of the data across all layers of the SYNERGY platform (cloud, on-premise).
- **Platform Management Services Bundle** which is responsible for resources management, the security and authentication aspects, the notifications management, the platform analytics, and the Open APIs (Application Programming Interfaces) that the SYNERGY platform provides.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

2.2.1 Services Bundles of the SYNERGY Cloud Infrastructure Layer

As depicted in Figure 2, the SYNERGY Core Cloud Platform (or Core Big Data Management Platform) instantiates all the data services bundles described above while the Secure Experimentation Playgrounds (that are spawn from the SYNERGY Core Cloud Platform for each organisation) only require the Data Collection Services Bundle, the Data Security Services Bundle, the Data Analytics Services Bundle and the Data Storage Services Bundle. The different SYNERGY tasks that are responsible for each data services bundle are also highlighted in the figure.

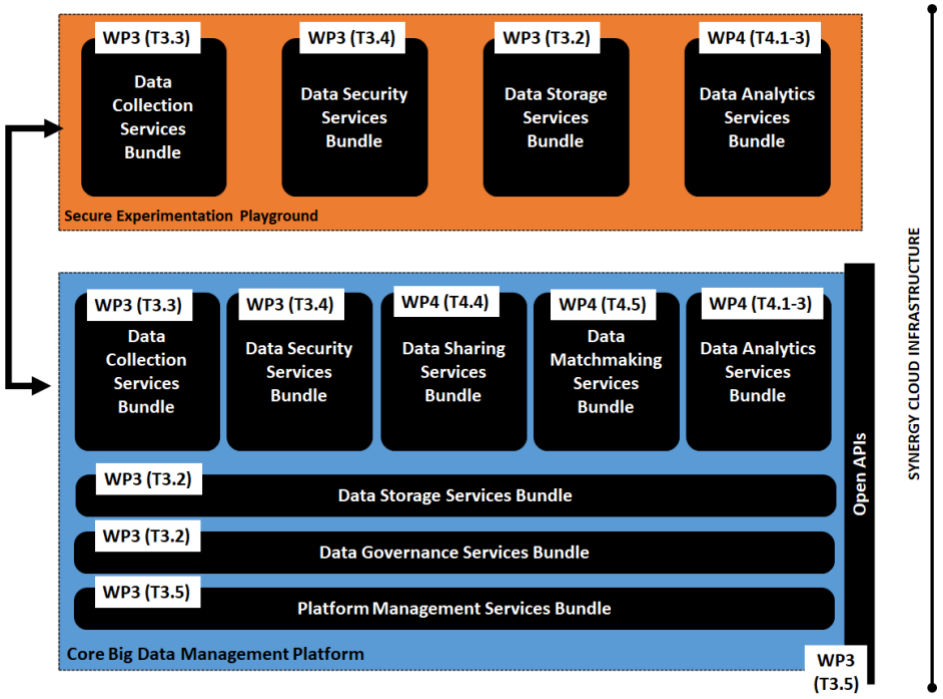


Figure 2: SYNERGY Data Services Bundles in the SYNERGY Cloud Infrastructure

2.2.2 Services Bundles of the SYNERGY On-Premise Environments Layer

As depicted in Figure 3, the SYNERGY On-Premise Environment in its server edition includes the Data Collection Services Bundle, the Data Security Services Bundle, the Data Sharing Services Bundle, the Data Analytics Services Bundle and the Data Storage Services Bundle while the light edge edition anticipates only the Edge Data Collection Services Bundle, and the Edge Data Analytics Services Bundle. The Server On-Premise Environment interacts with the Edge On-Premise Environments and the SYNERGY Cloud Infrastructure that seamlessly orchestrates all services to be executed. The different SYNERGY tasks that are responsible for each data services bundle are again highlighted in the figure.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

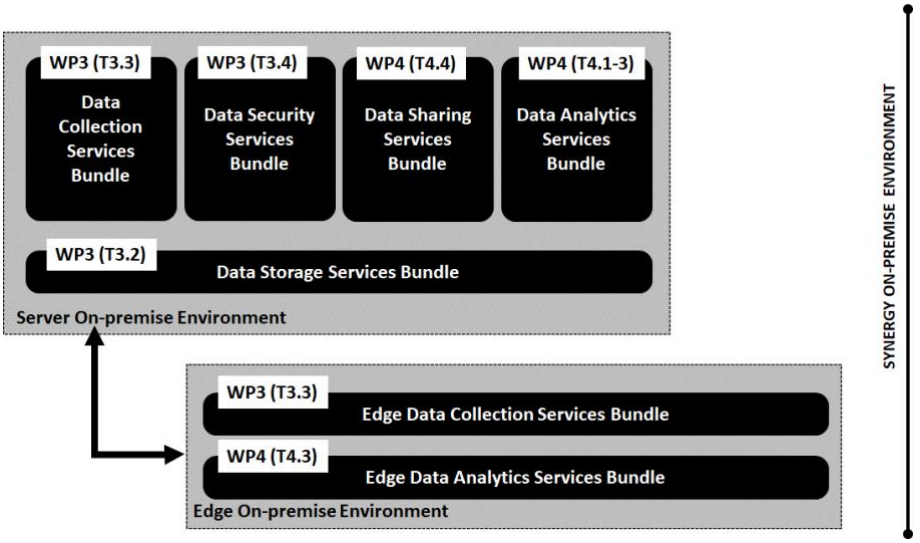


Figure 3: Data Services Bundles in the SYNERGY On-Premise Environments

2.2.3 SYNERGY Energy Apps Portfolio Layer

The SYNERGY Energy Apps interact with the SYNERGY Core Cloud Platform and include in their portfolio: (a) the Infrastructure Sizing and Grid Planning App, the Flexibility-based Network Management App, the RES Performance Monitoring/ Forecasting and Predictive Maintenance App, and the Network Asset Management Optimization App, addressed to DSOs, TSOs and RES Operators; (b) the Retailer Portfolio Analytics and Management App, the Flexibility Analytics and Consumer-Centric DR Optimization App, the Personalized Energy Analytics App, and the DR Smart Contract Management App, addressed to the needs of Electricity Retailers and Aggregators (and through them Prosumers); (c) the Advanced Renovation Support App, the Urban Energy Monitoring and Planning Support App, the Facility Management Analytics, Self-Consumption Optimization and Predictive Maintenance App, and the Building Energy Performance and Smart Readiness Certification App, targeting Facility Managers and ESCOs (and through them Prosumers), as presented in Figure 4.

In respect to D2.6, the Facility Management Analytics, Self-Consumption Optimization and Predictive Maintenance App have been merged into one application that brings together all intended functionality.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

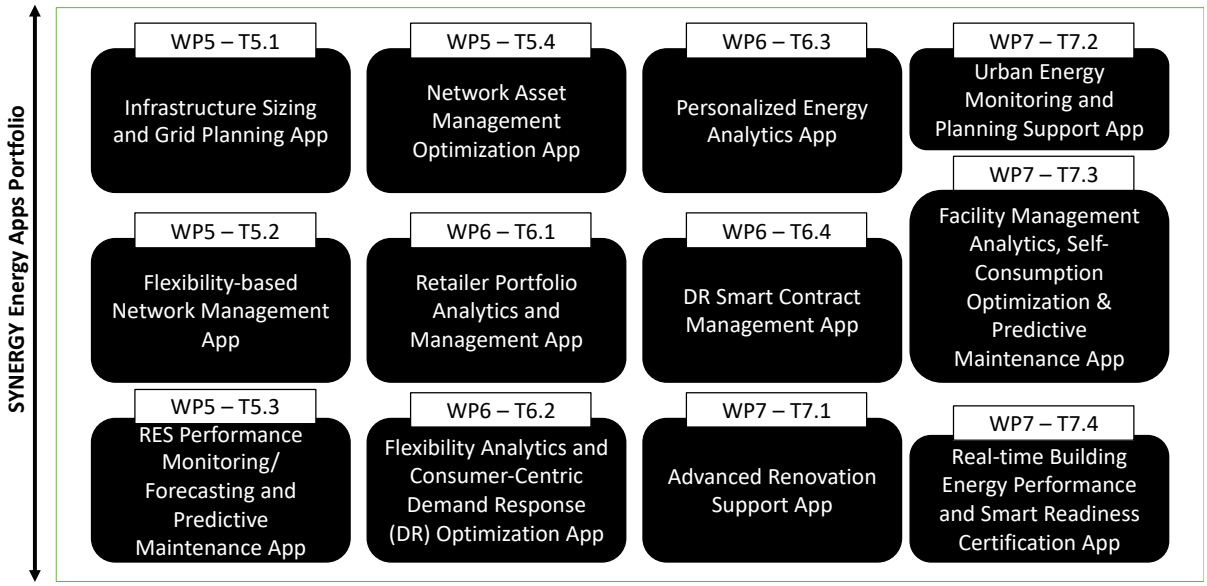


Figure 4: SYNERGY Energy Apps Portfolio

2.3 Components View

The different Data Services Bundles consist of different components and services that deliver their intended functionality and interact in a seamless way across the different layers of the architecture.

2.3.1 SYNERGY Cloud Infrastructure Layer

As depicted in Figure 5, the **SYNERGY Core Big Data Management Platform** (or SYNERGY Core Cloud Platform in abbreviation) is the entry point for any user (as representative of an electricity data value chain stakeholder) in the overall SYNERGY platform. In order to check-in data to the SYNERGY platform, the **Data Handling Manager** (described in section 4.1.1) in the SYNERGY Core Cloud Platform provides the user interfaces to properly configure and manage the data check-in jobs at "design" time, according to the settings and preferences of each data provider for uploading batch data as files; collecting data via 3rd-party applications' APIs, via open data APIs or the SYNERGY Platform's APIs; and ingesting streaming data (through the SYNERGY Platform's mechanisms or through the stakeholders' PubSub mechanisms). Upon configuring the data ingestion step, the data providers need to properly map the sample data they have uploaded to the SYNERGY Common Information Model following the suggestions and guidelines of the **Matching Prediction Engine** (described in section 4.1.2). Optionally, the data providers are able to also configure the cleaning rules, the anonymisation rules and the encryption rules that need to be applied over the data. The **Access Policy Engine** (described in section

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.2.3) provides the opportunity to define access policies based on different attributes in order to fully control which stakeholders can potentially view the specific data asset's details in the SYNERGY platform.

The data check-in job execution is triggered by the **Master Controller** (described in section 4.7.1) according to the schedule set by the data providers and in the execution location they have set (i.e. Cloud Platform or On-Premise Environment). The Master Controller communicates with the Resources Orchestrator (described in section 4.8.1) to ensure the necessary compute and memory resources (esp. in the SYNERGY cloud infrastructure) and orchestrates the appropriate list of services among the **Data Ingestion Service** (described in section 4.1.3), the **Mapping & Transformation Service** (described in section 4.1.4), the **Cleaning Service** (described in section 4.1.5), the **Anonymisation Service** (described in section 4.2.1) and the **Encryption Engine** (described in section 4.2.2), that are invoked in a sequential manner while forwarding them the data check-in job's configuration. The data are stored in Trusted Data Containers in the Data Storage Services Bundle (described in section 4.6) and a set of metadata (in alignment with the SYNERGY metadata schema) are either extracted automatically during the previous steps or defined by the data providers in the Data & AI Marketplace (described in section 4.3.1), and persisted in the Metadata Storage.

The **Data & AI Marketplace** (described in section 4.3.1) is the one-stop shop for energy-related data assets from the electricity data value chain stakeholders as it enables secure and trusted data asset sharing and trading among them. It allows them to efficiently search for data assets of interest through the **Query Builder** (described in section 4.4.1) and provides them with the help of the **Matchmaking Engine** (described in section 4.4.2) with recommendations for data assets or data assets' providers (that may potentially have/create the specific data asset). The Data & AI Marketplace allows data consumers to navigate to the available data assets, preview their offerings and proceed with their acquisition through smart data asset contracts that are created, negotiated, and signed among the involved parties in the **Contract Lifecycle Manager** (described in section 4.3.2) and stored in each step in the **Contracts Ledger**. In order for a signed contract to be considered as active, the respective payment needs to be settled with the help of the **Remuneration Engine** (described in section 4.3.3).

In order for electricity data value chain stakeholders to leverage the potential of data analytics over data that they own or have acquired, the **Analytics Workbench** (described in section 4.5.1) gives them the opportunity to design data analysis pipelines according to their needs and requirements. Such pipelines may consist of different data manipulation functions and pre-trained algorithms that have been created for the needs of the energy domain or simple algorithms that are offered in an out-of-

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

the box manner. The execution settings are defined by the data asset consumers that define when and how the data analysis pipeline should be executed and how the output will be stored. In this context, the **Visualization & Reporting Engine** (described in section 4.5.2) allows the data asset consumers to select, customize and save appropriate visualizations to gain insights into the analytics results, but also to create simple reports to potentially combine results.

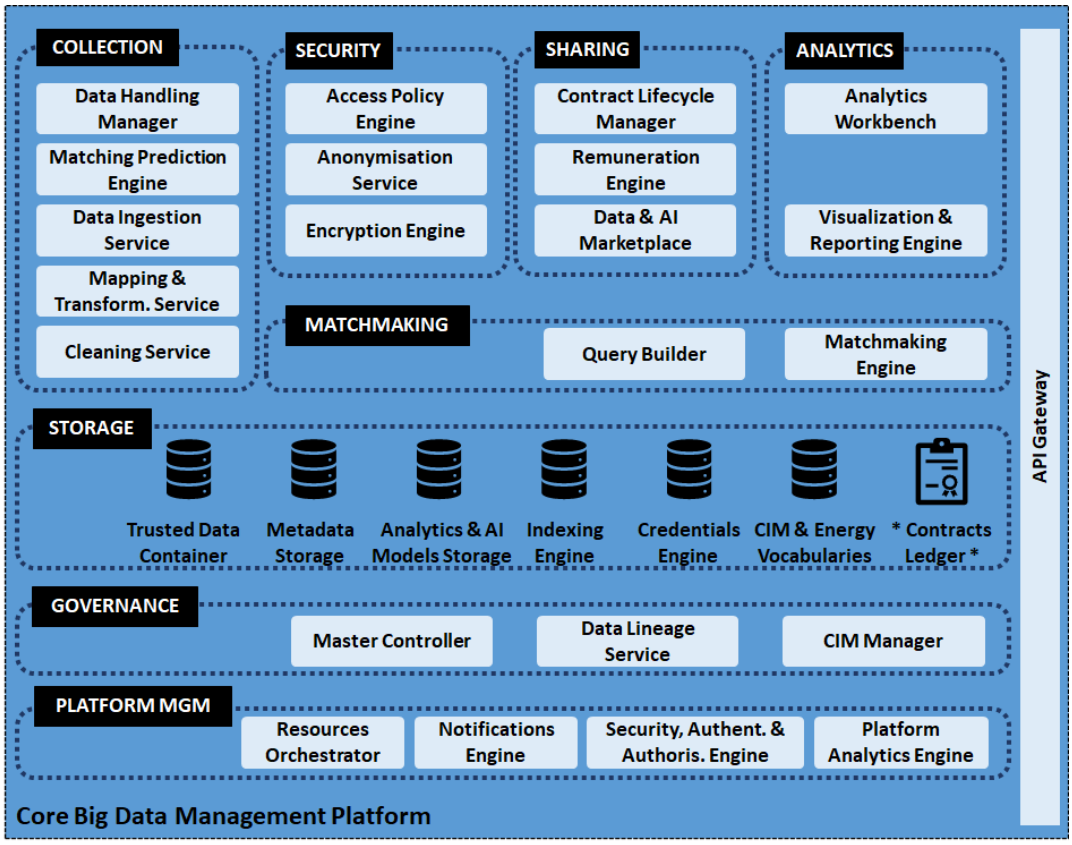


Figure 5: Detailed Component View of the SYNERGY Core Cloud Platform

The API Gateway (described in section 4.8.5) allows the authorised SYNERGY energy applications and any application to retrieve from the SYNERGY platform's Open APIs, the exact raw data or analytics results they need according to filters they are able to set. The overall platform's security, organisation's and user's registration, and authorisation decisions are dependent on the **Security, Authentication & Authorisation Engine** (described in section 4.8.3).

The SYNERGY cloud platform is complemented by the **Data Lineage Service** (described in section 4.7.2) to provide provenance-related views over the data assets; the **Notifications Engine** (described in section 4.8.2) to send notifications about the ongoing processes that are related to a user or organisation; the **Platform Analytics Engine** (described in section 4.8.4) that provides insights into the added value of the data assets in the SYNERGY platform, but also on the overall platform's services

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

progress; and the **CIM Manager** (described in section 4.7.3) that is behind the evolution and propagation of changes of the Common Information Model across the involved services in the whole SYNERGY platform.

The execution of a data analysis job in the SYNERGY Cloud Platform is performed in **Secure Experimentation Playgrounds** which are essentially sandboxed environments that become available per organisation. The data that belong to an organisation or have been acquired by an organisation (based on a legitimate data asset contract) are transferred through the **Data Ingestion Service** based on the instructions provided by the Master Controller, are decrypted upon getting access to the decryption key in the **Encryption Engine** (with the help of the Master Controller and the Security, Authentication & Authorisation Engine), and are stored in Trusted Data Containers. Any data analysis pipeline that needs to be executed is triggered according to the organisation's preferences by the Master Controller that invokes the **Data Manipulation Service** (described in section 4.5.3) and the **Analytics Execution Service** (described in section 4.5.4). The **Secure Results Export Service** (described in section 4.5.5) is responsible to prepare the results for use by the respective organisation in different ways (e.g. as a file, exposing them via an API, sharing them in the Data & AI Marketplace). Finally, the **Data Lineage Service** provides an overview of the relations and provenance of the data assets stored in the Secure Experimentation Playground (as depicted in Figure 6).

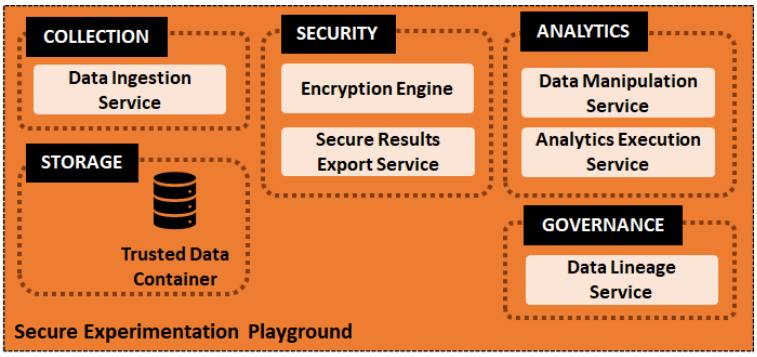


Figure 6: Detailed Component View of the SYNERGY Secure Experimentation Playground

2.3.2 SYNERGY On-Premise Environments Layer

The SYNERGY Server On-Premise Environment is responsible for: (a) preparing the data assets, which an organisation owns, "locally" to ensure end-to-end security (especially when encryption is required in the data check-in job configuration) prior to uploading them in the SYNERGY Core Cloud Platform; (b) preparing and storing the own data assets "locally" in case they are not allowed to even leave a stakeholder's premises; (c) to run analytics "locally" over data that are also stored "locally".

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

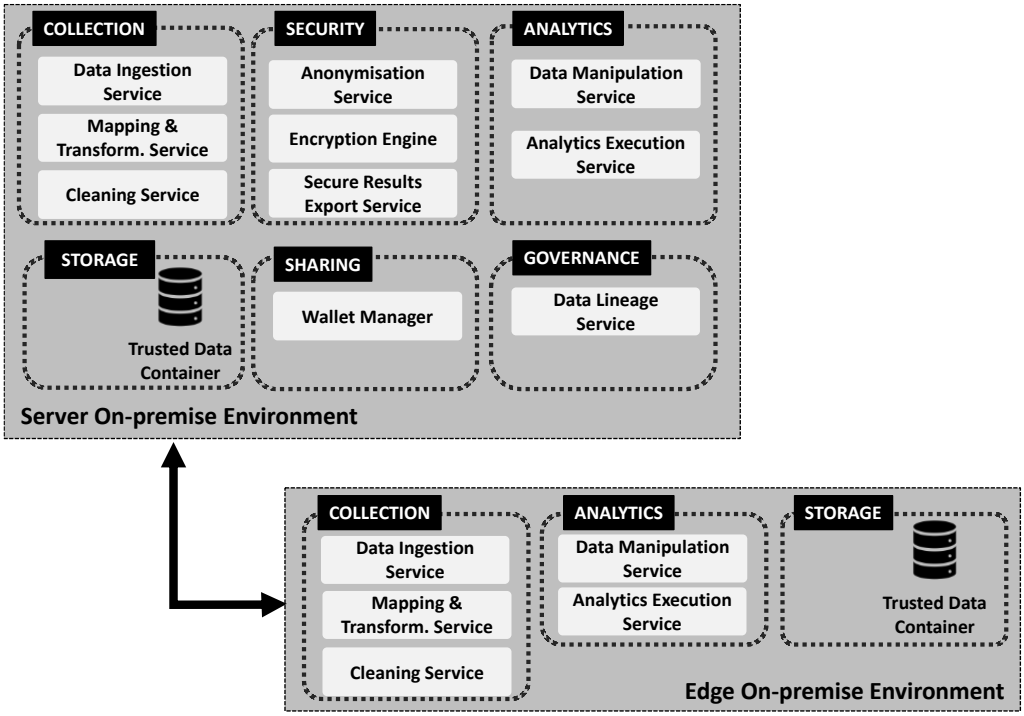


Figure 7: Detailed Component View of the SYNERGY On-Premise Environments

As depicted in Figure 7, according to the instructions received by the Master Controller in the SYNERGY Core Cloud Platform, a data check-in job is executed in the Server On-Premise Environment as follows: the **Data Ingestion Service** (described in section 4.1.3) is responsible for collecting the necessary data, the **Mapping & Transformation Service** (described in section 4.1.4) for processing the data (to ensure their alignment with the CIM), the **Cleaning Service** (described in section 4.1.5) for increasing the data quality, the **Anonymisation Service** (described in section 4.2.1) for handling any personally identifying or commercially sensitive data, and the **Encryption Engine** (described in section 4.2.2) for encrypting the data. Then, the data are either stored locally in the **Trusted Data Container** or transferred to the SYNERGY Core Cloud Platform where they are permanently stored.

In order to execute a data analysis job "locally" in the Server On-Premise Environment, the Master Controller of the SYNERGY Core Cloud Platform appropriately invokes the **Data Manipulation Service** (described in section 4.5.3) and the **Analytics Execution Service** (described in section 4.5.4) to run all necessary steps of the designed pipeline. The results are stored in the Trusted Data Container and can be securely extracted from the **Secure Results Export Service** (described in section 4.5.5) of On-Premise Environment Server Edition.

The **Wallet Manager** (described in section 5.2.1) allows the organisations that have installed the On-Premise Environment Server Edition to securely handle the ledger account and the cryptocurrency

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

funds of their organisation. It is practically used to send payments for smart asset data contracts that allow an organisation to buy data, but also to receive reimbursement for data assets that have been sold by the organisation (especially for derivative data assets in the context of a multi-party smart asset contract). The **Data Lineage Service** again allows a better view of the data asset's provenance.

The **Edge On-Premise Environment** has limited functionalities in respect to the Server On-Premise Environment due to the limited compute, memory and storage capacity it can leverage. It has a light version of: (a) the **Data Ingestion Service,** the **Data Mapping & Transformation Service** and the **Data Cleaning Service** to ensure that a gateway may collect data as part of a data check-in job that has been configured in the SYNERGY Core Cloud Platform, and (b) the **Data Manipulation Service** and the **Analysis Execution Service** that may run limited data analysis pipelines with restrictions.

2.3.3 SYNERGY Energy Apps Portfolio Layer

Stepping on the definition of the Energy Apps of SYNERGY, as presented in section 2.2.3, the following table provides a detailed breakdown of the components that each app will consist in and will be further analysed in chapter 6 of the current deliverable.

Table 1: SYNERGY Energy Apps and Relevant Components

SYNERGY Energy Apps	Relevant Components
Infrastructure Sizing and Grid Planning App	• Network Performance Assessment Manager (6.1.1) • Network Assets Sizing Manager (6.1.2)
Flexibility-based Network Management App	• Network Performance Assessment Manager (6.1.1) • Flexibility-based Network Manager (6.1.3) • DSO-TSO Common Operational Scheduler (6.1.4)
RES Performance Monitoring/ Forecasting and Predictive Maintenance App	• Enhanced Performance Monitor (6.1.5) • Fault Occurrence Inspector and Maintenance Optimizer (6.1.6) • Operational Scheduling Optimizer (6.1.7)
Network Asset Management Optimization App	• Network Asset Health Estimator (6.1.8) • Network Predictive Maintenance Manager (6.1.9)
Retailer Portfolio Analytics and Management App	• Portfolio Pattern Forecasting Engine (6.2.1) • Customer Segmentation Engine (6.2.2) • Wholesale Market Participation Decision Support System (6.2.3)
Flexibility Analytics and Consumer-Centric Demand Response (DR) Optimization App	• Aggregator Portfolio Manager (6.2.8) • Virtual Power Plant (VPP) Configuration Engine (6.2.9)
Personalized Energy Analytics App	• Personalized Energy Analytics Engine (6.2.4) • Smart Home Integration Engine (6.2.5)
DR Smart Contract Management App	• Flexibility Marketplace Search Engine (6.2.10) • Flexibility Contracts Manager (6.2.11)

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

	• Flexibility Settlement and Remuneration Engine (6.2.12)
Advanced Renovation Support App	• AI-boosted Renovation Decision supporting Service (6.3.1) • IDA Indoor Climate and Energy (IDA-ICE) Renovation Analysis Service (6.3.2)
Urban Energy Monitoring and Planning Support App	• Near Real-time City Monitoring and Visualization tool (6.3.3) • Strategic Urban Planning Supporter (6.3.4)
Facility Management Analytics, Self-Consumption Optimization & Predictive Maintenance App	• Facility Management Monitoring Engine (6.3.5) • HVAC Predictive Maintenance Service (6.3.6) • Building-Level Energy Performance Optimisation Manager (6.3.7) • District-level Energy Performance Optimisation Manager (6.3.8)
Real-time Building Energy Performance and Smart Readiness Certification App	• eDECs (enhanced Display Energy Certificates) Calculation Engine (6.3.9) • SRI (Smart Readiness Indicator) Calculation Engine (6.3.10)

2.4 Platform Roles and Users

As mentioned in D2.6, the different roles and users that are expected to use the SYNERGY Big Data Platform & AI Marketplace are classified into different types from: (a) an Organisation Perspective (that embraces the different members of an organisation that are expected to use the SYNERGY Platform), (b) a Platform Perspective (that embraces the core SYNERGY roles that the SYNERGY technical partners will undertake) and (c) the Data Perspective (that presents the different roles which any organisation in the electricity data value chain that owns, wants to provide, and/or wants to consume or use data assets can assume).

The following table lists the identified stakeholders and describes their anticipated role when using the SYNERGY platform.

Table 2: SYNERGY Platform Roles and Users

Type	Role	Description
Organisation Perspective	Manager	The Manager of an Organisation is typically the official legal representative who is authorised to act on behalf of the organisation and is the signatory for all company operational activities, including the asset contracts that will be signed over the SYNERGY platform.
	Technical User	The Technical User of an Organisation is typically an IT manager, developer or database administrator who are knowledgeable how the data are managed, exposed and shared by their organisation's back-end systems.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Type	Role	Description
	Data Scientist / Analyst	The Data Scientist / Analyst of an Organisation has the technical skills to explore, manipulate and analyse data to solve complex problems and derive actionable insights for the organisation's operations.
	Business User	The Business User of an Organisation typically refers to operations, marketing, business development and strategy experts who need to view the results of an analysis (in an interactive way) in order to derive meaningful insights and take proactive decisions in their everyday work.
Platform Perspective	Administrator	The Platform Administrator is responsible for overseeing the uninterrupted availability of the SYNERGY platform, including performing software/hardware upgrades and regular backups while addressing any technical issues that arise.
	CIM Manager	The CIM Manager is responsible for handling the lifecycle of the SYNERGY Common Information Model, e.g. adding new concepts, approving the proposals received (by different organisations) and ensuring the model's consistency.
	DPO	The Data Protection Officer (DPO) is responsible for monitoring the data protection compliance of the SYNERGY platform according to the EU GDPR (General Data Protection Regulation) provisions.
	Help Desk	The SYNERGY Help Desk is intended to provide information and support on the SYNERGY platform functionalities, to the different users and organisations that access the SYNERGY platform.
Data Perspective	Asset Provider	Any organisation in the SYNERGY Platform can act as an Asset Provider, providing data, pre-trained AI modes and/or analytics results to the platform for the purpose of legitimately sharing them with other organisations (as electricity data value chain stakeholders) and / or using them for an analysis.
	Asset Owner	An Asset Owner is a legal entity or natural person creating data, pre-trained AI modes and/or analytics results and/or executing control over them. An organisation acting as Asset Owner automatically assumes the role of the Asset Provider as well in the SYNERGY platform. In cases in which the Asset Owner does not act as the Asset Provider at the same time, the Asset Owner is expected to have authorised (outside the platform) an Asset Provider to make its data, pre-trained AI modes and/or analytics results available to be used by an Asset Consumer.
	Asset Consumer	Any organisation in the SYNERGY Platform can act as an Asset Consumer, acquiring legitimate access to data, pre-trained AI modes and/or analytics results in the SYNERGY Platform (that are typically owned by other organisations, through asset contracts) and using them for further analysis.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

3 SYNERGY Workflows in SYNERGY Cloud Infrastructure & On-Premise Environments

3.1 Introduction

Taking into consideration the advancements in the development of the SYNERGY Integrated Platform in its beta release (documented in D3.4) and in its first, official release (documented in D3.6), the design of the basic workflows of the SYNERGY Reference Architecture that were presented in the SYNERGY Deliverable D2.6, has been revisited to validate how certain core end-to-end functionalities of the SYNERGY platform are delivered. To this end, such workflows focus on how data asset providers and data asset consumers can be supported from a user-oriented perspective along 3 main processes:

- The **data check-in process** allowing data asset providers to make available their data in the SYNERGY Energy Big Data Platform and the AI Analytics Marketplace.
- The **data search and sharing process** allowing data asset consumers to find data assets of interest in the SYNERGY Energy Big Data Platform and the AI Analytics Marketplace and acquire them in a trustful and reliable manner.
- The **data analytics process** allowing data asset providers and consumers to run analytics over their own and the acquired data assets in the SYNERGY Energy Big Data Platform and the AI Analytics Marketplace and gain previously unattainable insights.

For each main workflow, a series of workflows are presented in detail in BPMN diagrams (OMG, 2011) listing the involved components and their interactions in order to perform the desired operations. It needs to be noted that additional components that always run under the hood (e.g. the Security, Authentication and Authorisation Engine, Notification Engine) or are not on the critical path of the above workflows (e.g. Data Lineage Service, CIM Manager, Platform Analytics Engine) have been intentionally omitted from the diagrams that are presented in the following paragraphs.

3.2 Data Check-in Workflow

In order to allow data asset providers to upload their data assets in the SYNERGY platform in a flexible manner through multiple modalities, there is a distinct separation between the design of a data check-in job (described in section 3.2.1) that happens once by a data asset provider and the execution of a

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

data check-in job by the SYNERGY Energy Big Data Platform and the AI Analytics Marketplace (also mentioned as SYNERGY Cloud Platform or SYNERGY platform, for brevity) either in the cloud (described in section 3.2.2) or on-premise (described in section 3.2.3), according to the configuration provided by the data asset provider. As the data collection process is instrumental for the SYNERGY platform, particular attention has been paid on the process how the data asset providers are able to handle any potential failure to upload their data (described in section 3.2.4).

3.2.1 Data Check-in Job Design Workflow

As depicted in Figure 8, at design time, the data asset providers need to perform all preparatory activities to collect the information regarding the data they intend to upload to the SYNERGY platform. Upon accessing the SYNERGY platform, they create a new data check-in job in the Data Handling Manager (described in section 4.1.1), which is stored in the Data Storage Services (described in section 4.6). They need to select which pre-processing steps ranging from Mapping and Transformation (that is a compulsory step) to Cleaning, Anonymisation and Encryption (that are optional steps) should be applied on the data from their ingestion and prior to their permanent storage.

During the configuration of the data check-in job, the data asset providers need to initially configure the data ingestion step, selecting how, from where and when the data will be collected. Then, they need to define the mapping and transformation rules that are applicable on their data, leveraging the predicted mappings to the CIM that the Mapping Prediction Engine has provided. The mapping process is semi-automated since the data asset providers need to manually confirm and update the mappings, as well as propose new concepts for the CIM (in case there is not any appropriate concept for their data in the CIM). The proposed concepts are handled by the model manager in the CIM Manager as part of a separate workflow that was not intended to be detailed here. Upon successfully validating the mappings of the data to the CIM, the data mapping configuration is stored.

If the cleaning step is enabled, the data asset providers need to configure the data cleaning rules. Otherwise, they proceed with the anonymisation step if it is enabled. In that case, they need to configure the data anonymisation rules. Finally, if the encryption step is enabled, they need to configure the encryption rules.

Once the pre-processing rules are finalized, the data asset providers need to set the applicable data access policies in the Access Policy Engine and define the data asset's metadata and licensing details in the Data & AI Marketplace, according to the SYNERGY metadata schema. As soon as all necessary information depending on the desired access level (e.g. if a data asset is intended to be public and

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

available to all, confidential and available only to the data provider's organization, private and available to other organizations if there is an active data asset contract) and the execution location (i.e. whether the job will be executed in the cloud or in a registered server/cloud on-premise environment), the data check-in job execution is ready to start.

In comparison to D2.6, the Data Check-in Job Design Workflow has been updated to take into consideration the preference of the data asset provider for cloud or on-premise execution.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

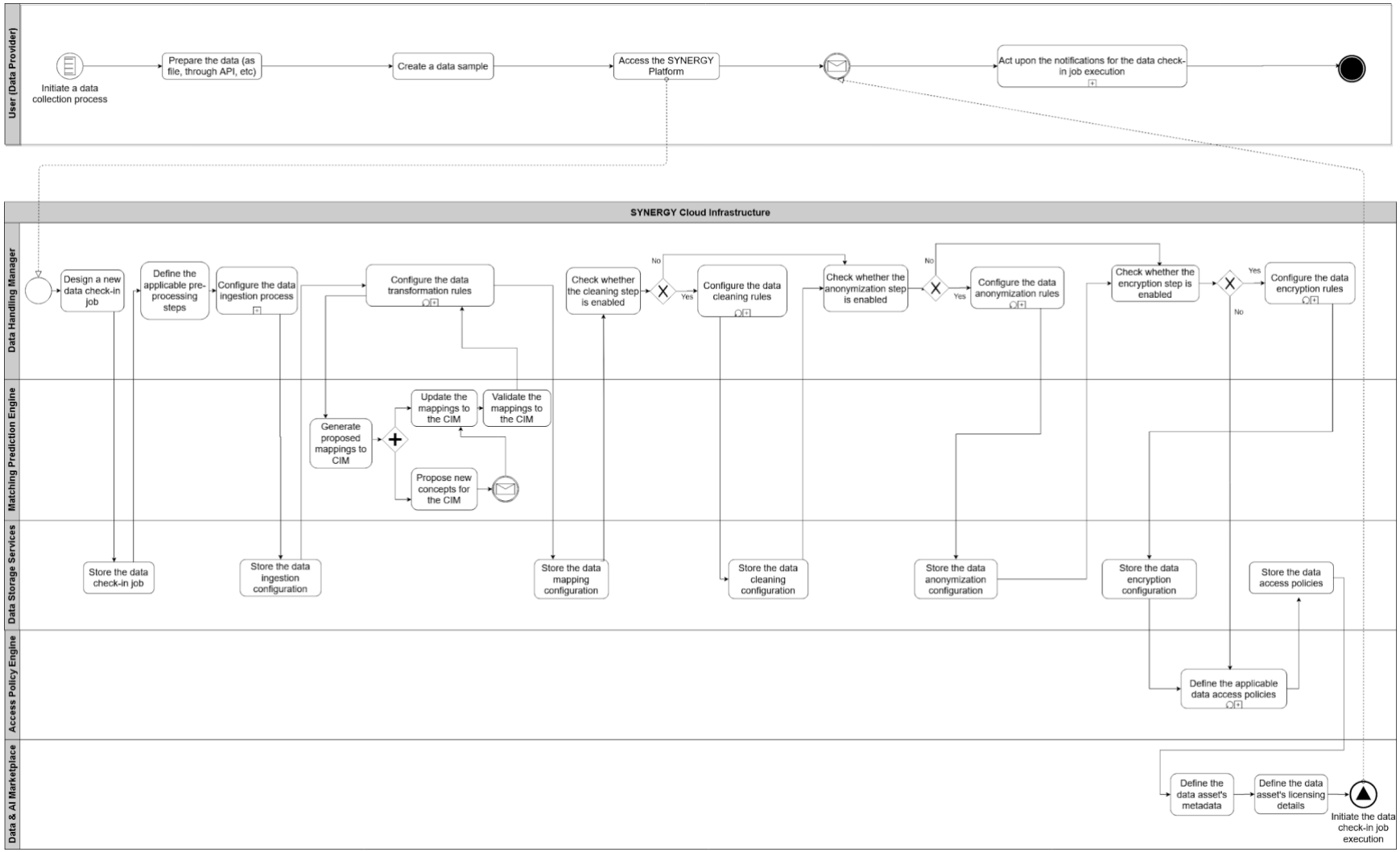


Figure 8: SYNERGY Data Check-in Design Workflow

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

3.2.2 Data Check-in Job Cloud Execution Workflow

Based on the data check-in job configuration that has been provided by the data asset provider and requests the execution of a job in the SYNERGY Cloud Infrastructure, the Master Controller (described in section 4.7.1) appropriately schedules the execution of the specific job while ensuring that the necessary resources are available from the Resources Orchestrator (described in section 4.8.1). When it is the execution time of the data check-in job, the Master Controller triggers the execution of the Data Ingestion Service (described in section 4.1.3) in the SYNERGY Cloud Infrastructure. If the data ingestion process successfully concludes, the data are temporarily stored, otherwise the related error information is collected and the data check-in job fails.

If the data have been properly ingested, the Master Controller triggers the execution of the Mapping and Transformation Service (described in section 4.1.4) in the SYNERGY Cloud Infrastructure, according to the processing schedule the data asset provider set in the data check-in job configuration. If the data mapping process successfully concludes, the transformed data are temporarily stored, otherwise the related error information is collected and the specific step of the data check-in job fails.

If the cleaning step is enabled in the data check-in job configuration, the Master Controller triggers the execution of the Cleaning Service (described in section 4.1.5) in the SYNERGY Cloud Infrastructure. If the data cleaning process successfully concludes, the cleansed data are temporarily stored, otherwise the related error information is collected and the specific step of the data check-in job fails.

If the anonymisation step is enabled in the data check-in job configuration, the Master Controller triggers the execution of the Anonymisation Service (described in section 4.2.1) in the SYNERGY Cloud Infrastructure. If the data anonymisation process successfully concludes, the anonymised data are temporarily stored, otherwise the related error information is collected and the specific step of the data check-in job fails.

If all pre-processing steps have succeeded, the Master Controller initiates the transfer of the data to the Secure Experimentation Playground of the data provider organization.

All feedback messages are collected and the execution of the specific data check-in job concludes until its next iteration for recurring jobs. In case any step of the data check-in process has failed, the feedback messages that have been collected allow the data asset provider to understand what went wrong and address it (as described in section 3.2.4).

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

The full workflow of the data check-in job execution in the SYNERGY Cloud Infrastructure is depicted in Figure 9.

In comparison to D2.6, the Data Check-in Job Cloud Execution Workflow has been updated to remove the execution of the Encryption step that is enabled only in the case of (Server) On-Premise Execution.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

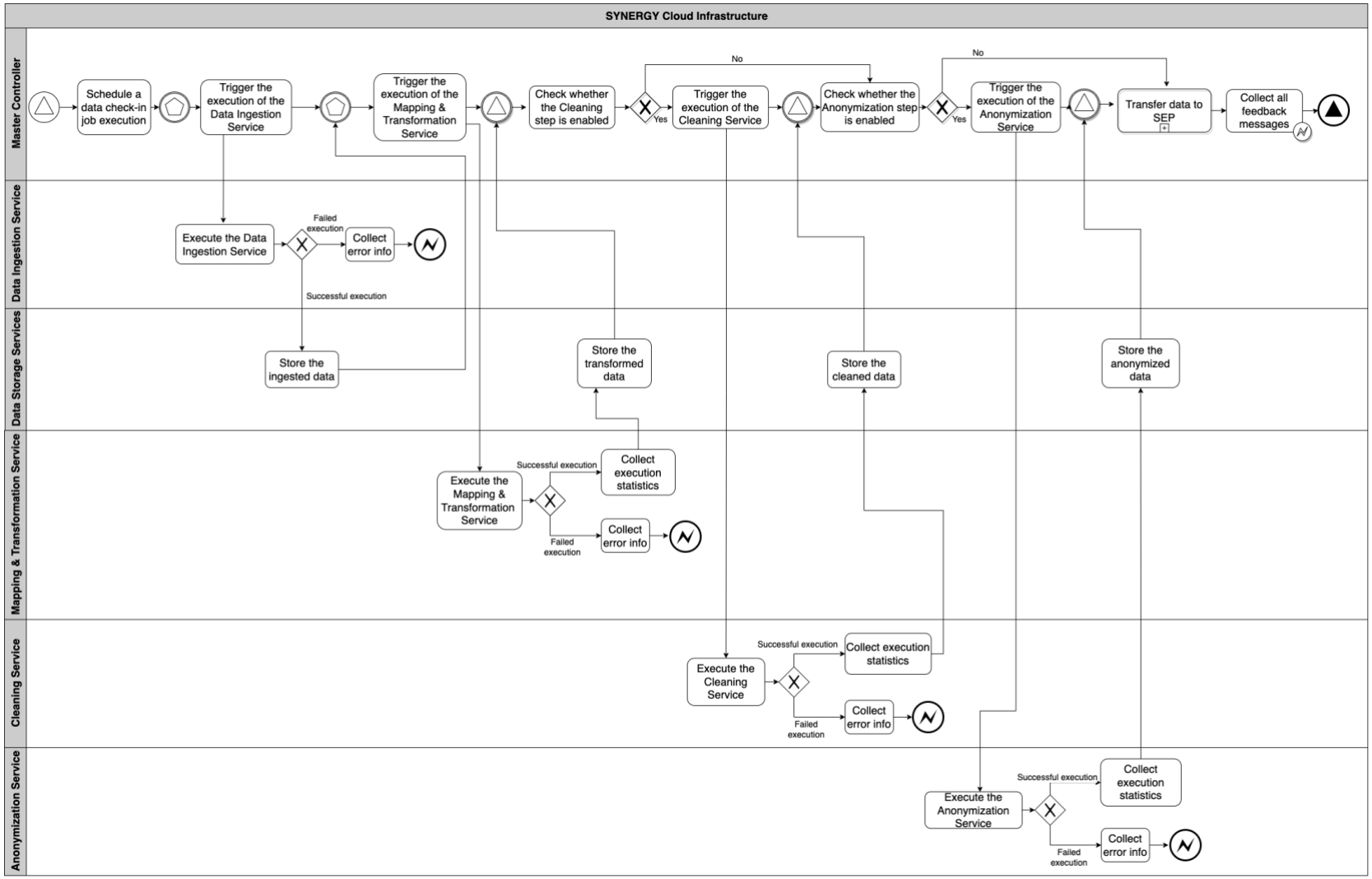


Figure 9: SYNERGY Data Check-in Cloud Execution Workflow

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

3.2.3 Data Check-in Job On-Premise Execution Workflow

Based on the data check-in job configuration that has been provided by the data asset provider and requests the execution of a job in the organization's On-Premise Environment, the Master Controller (described in section 4.7.1) appropriately schedules the execution of the specific job. When it is the execution time of the data check-in job, the Master Controller triggers the execution of the Data Ingestion Service (described in section 4.1.3) in the organization's On-Premise Environment. Following the same philosophy as the cloud execution (described in section 3.2.2), if the data ingestion process successfully concludes, the data are temporarily stored, otherwise the related error information is collected and the data check-in job fails.

If the data have been properly ingested, the Master Controller triggers the execution of the Mapping and Transformation Service (described in section 4.1.4) in the organization's On-Premise Environment, according to the processing schedule the data asset provider set in the data check-in job configuration. If the data mapping process successfully concludes, the transformed data are temporarily stored, otherwise the related error information is collected and the specific step of the data check-in job fails.

If the cleaning step is enabled in the data check-in job configuration, the Master Controller triggers the execution of the Cleaning Service (described in section 4.1.5) in the organization's On-Premise Environment. If the data cleaning process successfully concludes, the cleansed data are temporarily stored, otherwise the related error information is collected and the specific step of the data check-in job fails.

If the anonymisation step is enabled in the data check-in job configuration, the Master Controller triggers the execution of the Anonymisation Service (described in section 4.2.1) in the organization's On-Premise Environment. If the data anonymisation process successfully concludes, the anonymised data are temporarily stored, otherwise the related error information is collected and the specific step of the data check-in job fails.

If the encryption step is enabled in the data check-in job configuration, the Master Controller triggers the execution of the Encryption Engine (described in section 4.2.2) in the organization's On-Premise Environment. The Encryption Engine requests from the Wallet Manager that resides in the data asset provider organization's Server On-Premise Environment to generate an encryption key for the data if a key is not already available for the specific data asset, and requests from the SYNERGY Platform for an asymmetric key to encrypt the symmetric key. Once the keys become securely available and the necessary metadata have been extracted (so as to facilitate the search process), the ciphertext is

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

produced and stored. If the data encryption process fails, the related error information is collected and the specific step of the data check-in job fails.

If all pre-processing steps have succeeded, the Master Controller checks whether cloud storage of the data asset is permitted in the data check-in configuration. If it is indeed allowed, it initiates the transfer of the data from the On-Premise Environment to the Cloud Storage and then to the Secure Experimentation Playground of the data provider organization. In case the data are encrypted, they are decrypted in the organization's Secure Experimentation Playground in order to be readily available for running analytics.

All feedback messages are collected and the execution of the specific data check-in job concludes until its next iteration for recurring jobs. In case any step of the data check-in process has failed, the feedback messages that have been collected allow the data asset provider to understand what went wrong and address it (as described in section 3.2.4).

The full workflow of the data check-in job execution in the SYNERGY On-Premise Environment (always in collaboration with the SYNERGY Cloud Infrastructure) is depicted in Figure 10.

In comparison to D2.6, the Data Check-in Job On-Premise Execution Workflow has essentially remained as originally designed (with the exception of further diving into the mechanism for the encryption keys exchange).

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

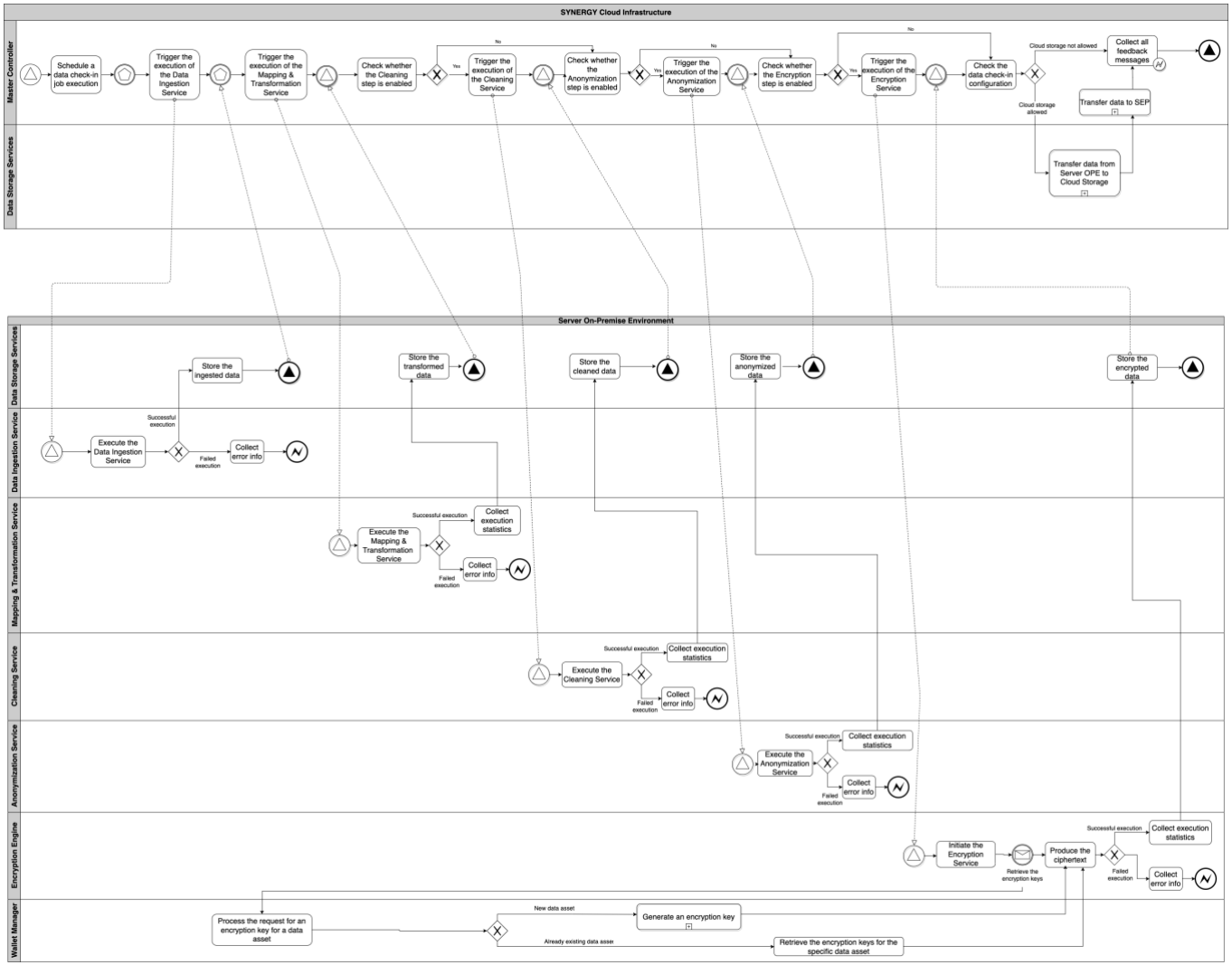


Figure 10: SYNERGY Data Check-in On-Premise Execution Workflow

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

3.2.4 Data Check-in Job Failure Management Workflow

The execution of a data check-in job may fail for a number of reasons, indicatively ranging from erroneous configuration to unavailable services that expose the data from the data asset provider. In order to help the data asset providers understand what went wrong during their data check-in jobs' execution, the Master Controller collects all feedback messages and notifies the respective organization's users (through the Notification Engine that runs under the hood and is not explicitly mentioned in the BPMN diagram in Figure 11).

Once the data asset providers access the SYNERGY platform, they may review the execution details of a data check-in job in the user interface of the Data Handling Manager. If the data ingestion process failed, its configuration is unlocked allowing the data asset providers to update it. In case the performed update in the data ingestion process also affects the data mapping and transformation rules, they may also update them, otherwise they can proceed with the finalization of the data check-in job configuration (keeping the rest of the steps as initially configured).

If the data mapping and transformation process failed, its configuration is unlocked allowing the data asset providers to update it. In case the performed update in the data mapping and transformation rules also affects the data cleaning rules and/or the data anonymisation rules and/or the data encryption rules, they may also update them, otherwise they can proceed with the finalization of the data check-in job configuration (keeping the rest of the steps as initially configured).

Of course, if any of the cleaning, anonymisation, encryption steps failed, their configuration is unlocked allowing the data asset providers to proceed to its revision.

In comparison to D2.6, the Data Check-in Job Failure Management Workflow has essentially remained as originally designed.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

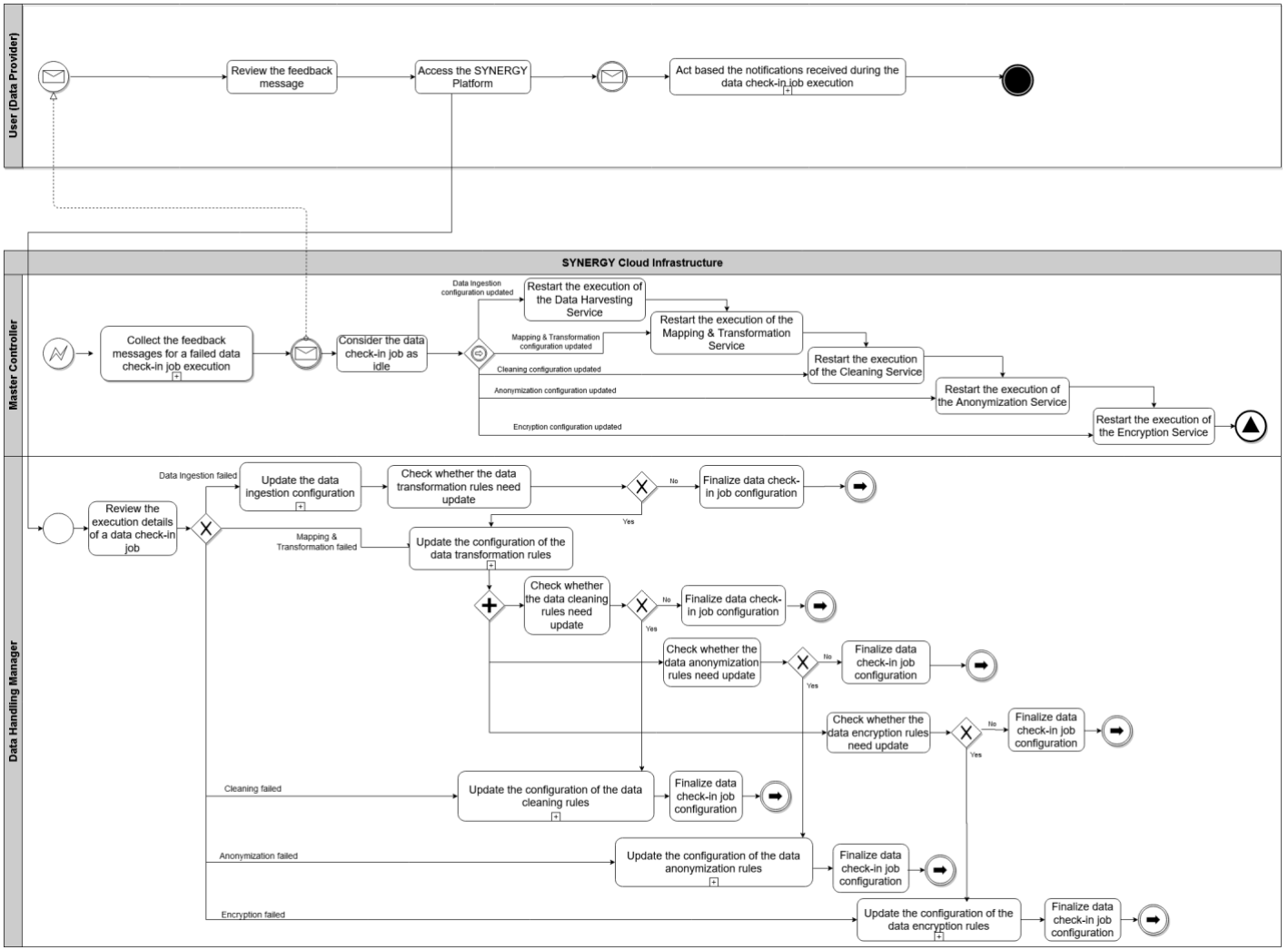


Figure 11: SYNERGY Data Check-in Failure Management Workflow

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

The Master Controller has declared the specific data check-in job as idle, so once the updated data check-in job configuration is finalized, it restarts the execution of the appropriate services: (a) if the data ingestion configuration was updated, the Data Ingestion Service is triggered, followed by the Mapping & Transformation Service, and optionally (depending on the job's configuration) the Cleaning Service, the Anonymisation Service and the Encryption Engine, (b) if the data mapping configuration was updated, the Mapping & Transformation Service is triggered, and optionally followed by the Cleaning Service, the Anonymisation Service and the Encryption Engine depending on the job's configuration, (c) if the data cleaning configuration was updated, the Cleaning Service is triggered, and optionally followed by the Anonymisation Service and the Encryption Engine depending on the job's configuration, (d) if the data anonymisation configuration was updated, the Anonymisation Service is triggered and optionally followed by the Encryption Engine depending on the job's configuration, (d) if the data encryption configuration was updated, the Encryption Engine is triggered.

3.3 Data Search and Sharing Workflow

In order to allow data asset consumers to find and acquire data assets that belong to other electricity data value chain stakeholders, three workflows need to be executed at high level: (a) initially the data consumers need to find the data assets of interest through the Data Search Workflow (described in section 3.3.1) and to acquire the specific data assets. Depending on whether the data asset is a derivative asset and involves a single or multiple data asset providers, the bilateral or the multi-party data asset acquisition workflows come to the foreground (described in sections 3.3.2 and 3.3.3, respectively).

3.3.1 Data Search Workflow

As described in Figure 12, once data asset consumers identify a specific need for data or broadly a data asset, they access the SYNERGY platform and navigate to the Data & AI Marketplace (described in section 4.3.1). They may view only the data assets they are potentially eligible to acquire since the applicable access policies that have been defined by the respective data asset providers have been already resolved with the help of the Access Policy Engine (described in section 4.2.3). They are also able to load a query they had defined in the past or create a new query. In this way, they can search, filter and sort the available data assets according to different parameters offered (through a faceted search functionality) with the help of the Query Builder that translates the user query (described in section 4.4.1). The search query is executed over the metadata and/or the data of the available data

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

assets that are stored in the SYNERGY Data Storage Services, but before the results are prepared by the Query Builder, the applicable access policies need to be enforced by the Access Policy Engine. The search query results that do not violate any access policy are displayed to the data asset consumer in the Data & AI Marketplace and they are complemented by recommendations for data assets or data asset providers by the Matchmaking Engine (described in section 4.4.2).

The data asset consumers may explore the different results that are at their disposal, viewing the profile of the data assets and checking out the available samples. If their data needs are addressed by a single or multiple data assets, they may start the process for acquiring them. Obviously in the case of public data assets, such a process is trivial, but in the case of private data assets, the workflow described in sections 3.3.2 or 3.3.3 needs to be followed.

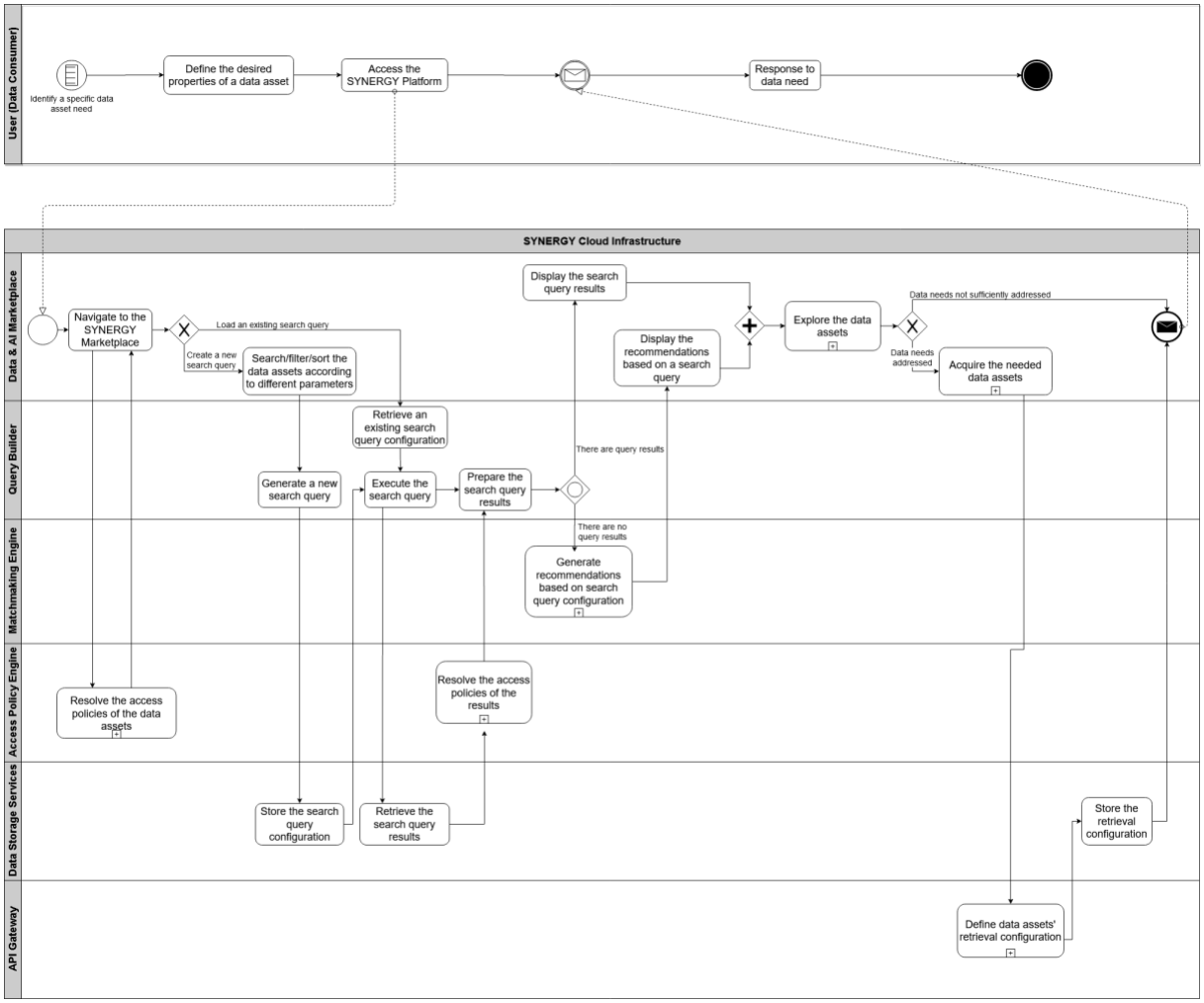


Figure 12: SYNERGY Data Search Workflow

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Once the data asset consumers have an active data asset contract, they are also able to define how they wish to retrieve the data, e.g. as a file, to become available in their Secure Experimentation Playground or through an API. For the latter case, they need to configure the retrieval settings (e.g. which fields exactly they wish to acquire and with what fields acting as query parameters in the case of unencrypted data) in the API Gateway (described in section 4.8.5). The retrieval configuration is stored and can be used at any time by applications owned by the data asset consumers as described in the data and results retrieval workflow in section 3.4.4.

In comparison to D2.6, the Data Search Workflow has essentially remained as originally designed.

3.3.2 Bilateral Data Asset Acquisition Workflow

As depicted in Figure 13, the data asset consumer that has found a data asset of interest in the Data & AI Marketplace is able to initiate a request for quotation for the specific data asset. Such a request is stored and forwarded to the respective data asset provider. Upon receiving a related notification, the data asset provider accesses the Data & AI Marketplace to review the details of the request for quotation, as well as of the organization that made the request, and take a decision on whether it will be accepted. If the request is rejected, the data asset consumer is notified accordingly. If the request is accepted, the data asset provider prepares a draft smart asset contract with the help of the Contract Lifecycle Manager (described in section 4.3.2), that is stored in the blockchain in the Data Storage Services.

The data asset consumer is notified for the draft contract and accesses the Contract Lifecycle Manager through the Data & AI Marketplace interface so as to review the contract's details. If the terms of the contract are acceptable, the data asset consumer shall proceed with the contract signature; if the terms are not acceptable, the data asset consumer can reject the contract; if the terms of the contract are partially acceptable, the data asset consumer shall enter a negotiation process with the data asset provider that iteratively provides counter-offers by each party, expressed through different versions of the smart contract that are stored in the blockchain. Eventually, the negotiation rounds will end when either the data asset provider or the data asset consumer accept or reject the latest version of the smart contract.

It needs to be noted that the signed smart contract is obviously also written in the blockchain and the involved data asset consumer is notified accordingly to proceed with the respective payment. No contract is considered as active until the associated payment fee is addressed in the Remuneration

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Engine (described in section 4.3.3). The data asset consumer may proceed to the payment either through conventional means (e.g. via a bank transfer to the data asset provider) or via the SYNERGY platform's payment mechanism in cryptocurrency (with the help of the Wallet that belongs to each organization). In the case of a conventional payment, the data asset provider needs to confirm the payment in the Remuneration Engine in order for the contract to become active.

Once a data asset contract is marked as paid and active in the blockchain, the respective data asset is by default transferred to the Secure Experimentation Playground of the data asset consumer with the help of the Master Controller that organizes the process.

In order to simplify the interactions in Figure 13, the Wallet Manager of the On-Premise Environment has not been depicted, yet its existence is a precondition for signing any contract in the Contract Lifecycle Manager.

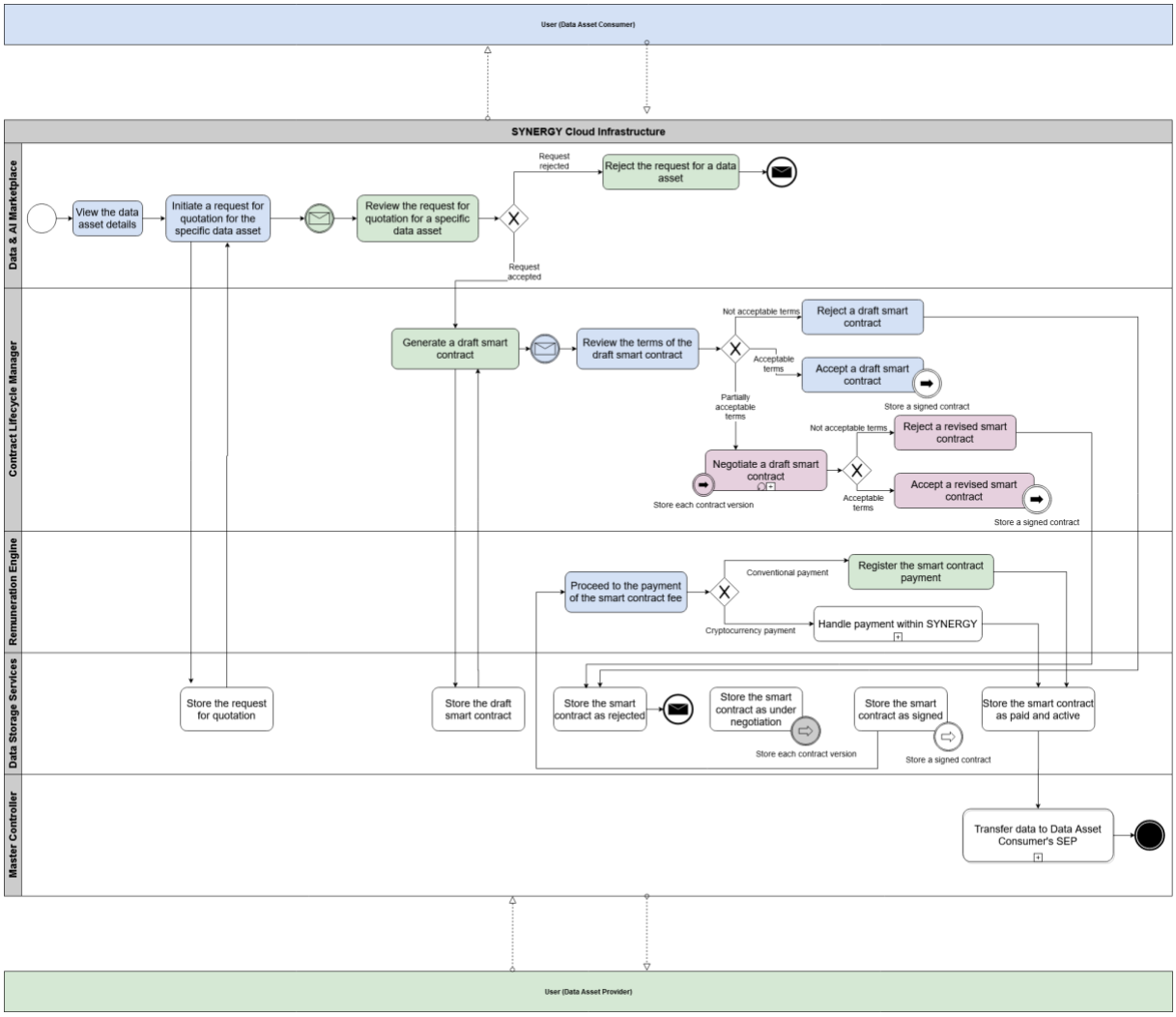


Figure 13: SYNERGY Bilateral Data Acquisition Workflow

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

It needs to be noted that the colours in the swimlanes for the Data Asset Provider and Consumer have been highlighted with colour (although it is not allowed in BPMN) in order to facilitate understanding about who performs each action.

In comparison to D2.6, the Bilateral Data Acquisition Workflow has essentially remained as originally designed.

3.3.3 Multi-party Data Asset Acquisition Workflow

Taking into consideration the fact that the term "data asset" in SYNERGY embraces single or multiple datasets, derivative datasets that have emerged as the outcome of a data analysis pipeline, trained analytics models (along with their training datasets), and analytics results, the data asset consumer often intends to acquire a data asset of interest that belongs to multiple data asset providers in the Data & AI Marketplace. It needs to be noted that in order for such a derivative data asset to be available in the Data & AI Marketplace, derivation contracts need to be in place among all involved data asset providers. Therefore, during the 1st step of the multi-party data asset contract preparation depicted in Figure 14, smart contracts are generated by the data asset providers whose assets are involved in the derivative data asset. Each data asset provider needs to provide the terms of each derivative contract and decide on the appropriate course of action for their organization: (a) accept and sign the draft smart contract, (b) reject the draft smart contract, (c) partially agree with the terms and initiate rounds of negotiation among the involved parties. Each action is always persisted in the blockchain for full traceability of the brokerage process. The Contract Lifecycle Manager is responsible for checking whether there is any data asset provider that has rejected the draft contract. If this is the case, the smart contract that has been signed by some data asset providers is automatically considered as cancelled.

If all data asset providers of a derivative data asset have reached an agreement, then the data asset is available in the Marketplace and visible to all stakeholders for which the accumulative access policies (by all data assets) are satisfied. As depicted in Figure 15, the data asset consumer of the derivative data asset needs to initiate a request for quotation (as in the case of the bilateral data asset acquisition workflow described in section 3.3.2). The request is stored and forwarded to the involved derivative data asset provider. Upon receiving a related notification, the derivative data asset provider accesses the Data & AI Marketplace to review the details of the request for quotation, as well as of the organization that made the request, and take a decision on whether it will be accepted. If the request is rejected by the derivative data asset provider, the data asset consumer is notified accordingly. If the

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

request is accepted by the derivative data asset provider and consumer, the derivative data asset consumer needs to proceed with the payment of the respective fee in the Remuneration Engine. In this case the payment must proceed in the SYNERGY platform and the payment should be performed from the Wallet Manager of the data asset consumer in the cryptocurrency which SYNERGY supports. The data asset providers are then appropriately reimbursed based on the smart contract terms and the Remuneration Engine is responsible for transferring the respective funds to their organization's Wallet Manager.

As soon as the payment and reimbursement process in the Remuneration Engine concludes, the smart consolidated contract is written as paid and active in the blockchain. The respective derivative data asset is by default transferred to the Secure Experimentation Playground of the data asset consumer with the help of the Master Controller that organizes the process. It needs to be noted that it may be, for example, data that need to be securely transferred or a pre-trained analytics model or results acquired through an analytics pipeline.

In order to simplify the interactions in Figures 14 and 15, the Wallet Manager of the On-Premise Environments for each involved organization has not been depicted, yet its existence is a precondition for signing any contract in the Contract Lifecycle Manager. It needs to be noted that the colours in the swimlanes for the Data Asset Provider and Consumer have been highlighted with colour (although it is not allowed in BPMN) in order to facilitate understanding about who performs each action.

In comparison to D2.6, the Multi-Party Data Acquisition Workflow has been significantly revised in order to: (a) remove complexity from the contract preparation time of a derivative data asset (that previously included negotiation at different levels); (b) ensure that the data asset providers from which a derivative data asset emerged remain in control of whether the derivative data asset appears in the Data & AI Marketplace (through another type of smart contracts, entitled derivation contracts). The previously single Multi-Party Data Acquisition Workflow has been split to the derivation part among the derivative data asset providers, and the acquisition part that directly involves the data asset consumer and the derivative data asset provider, and indirectly all data asset providers who get reimbursed according to the derivation contracts terms.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

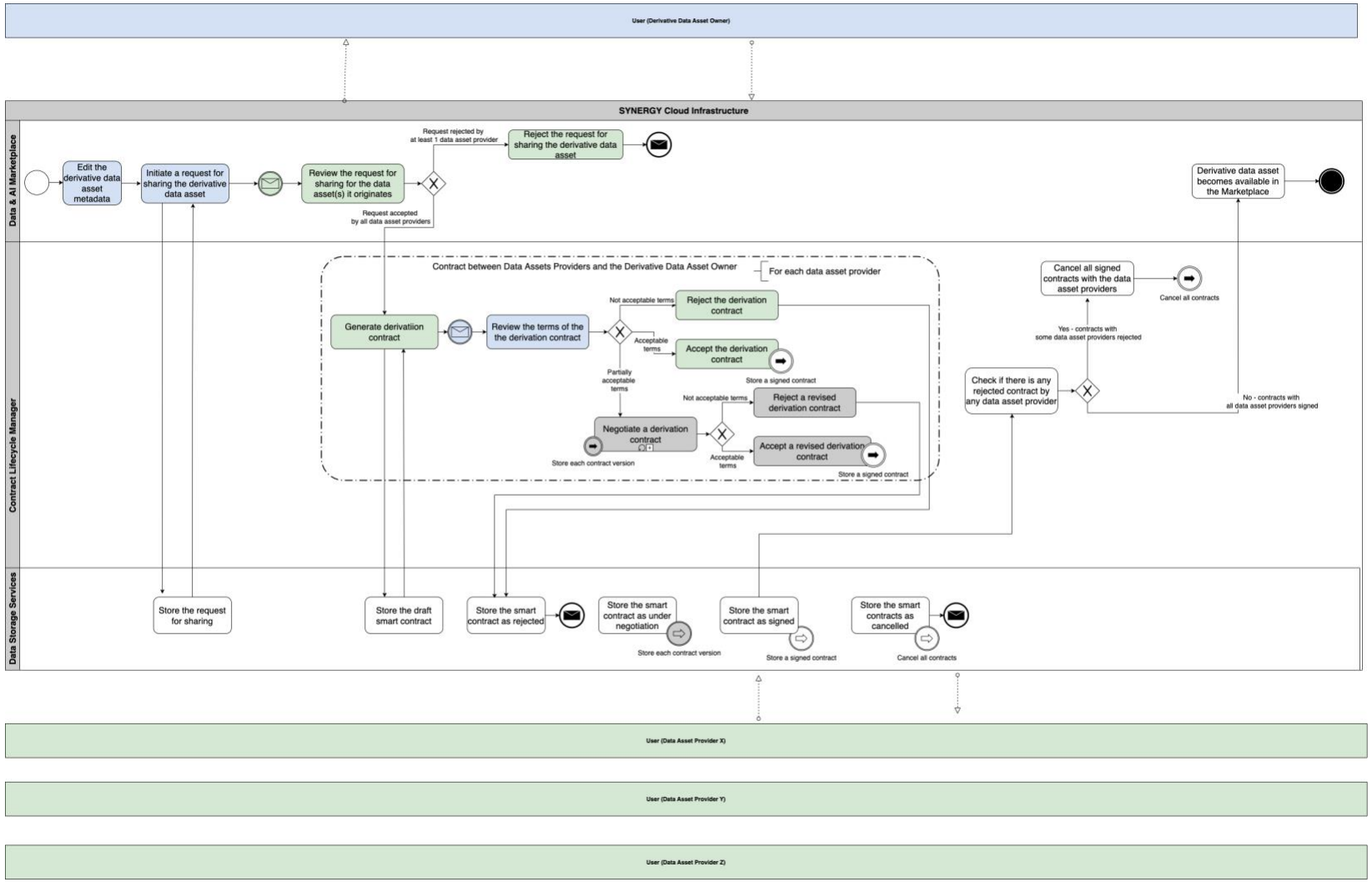


Figure 14: SYNERGY Multi-party Data Acquisition Workflow – Derivation Contracts for availability in Marketplace

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

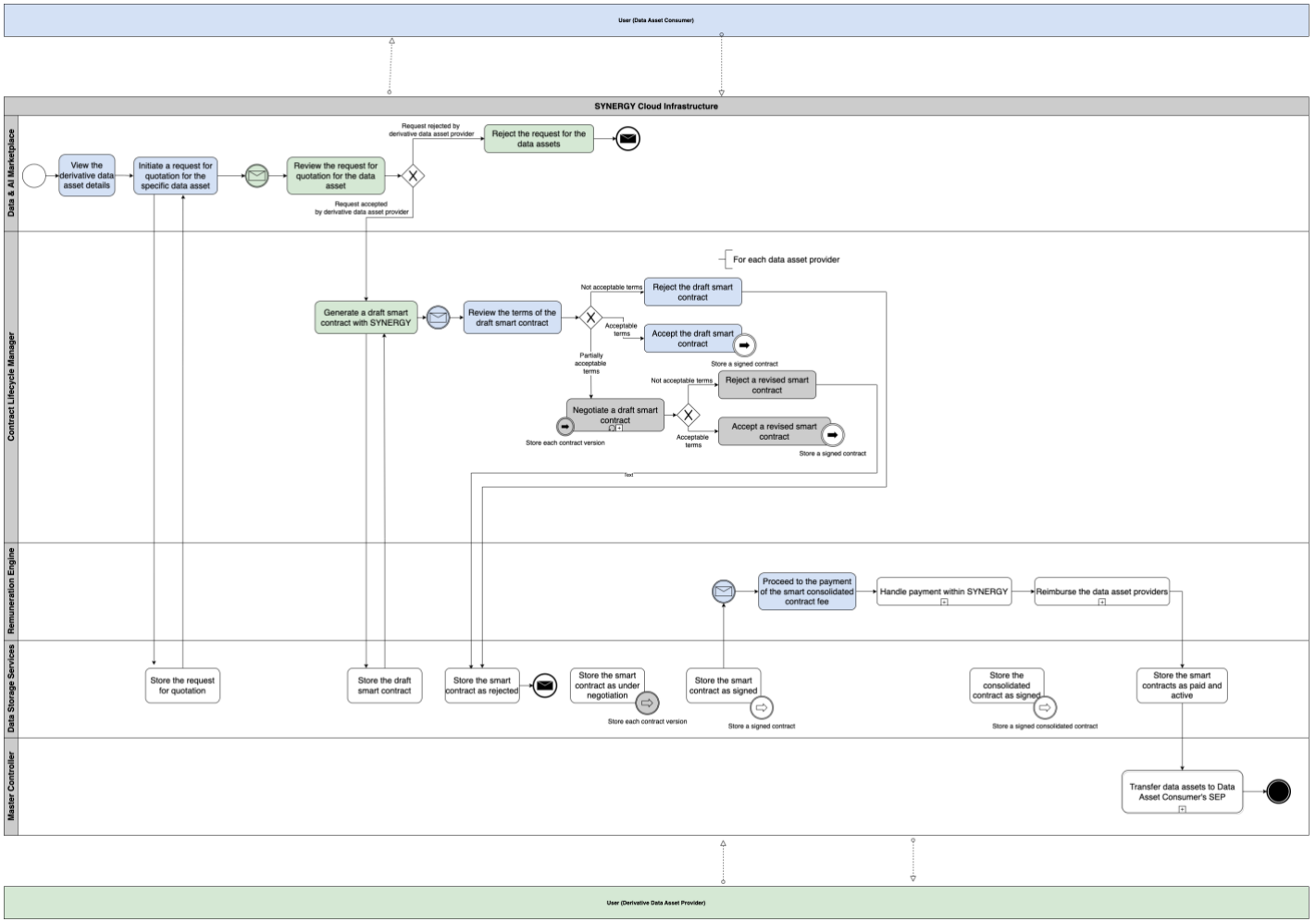


Figure 15: SYNERGY Multi-party Data Acquisition Workflow – Acquisition Contract of a Derivative Data Aset

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

3.4 Data Analytics Workflow

In order to allow data asset consumers to run any analytics over their own data assets (that have been uploaded in the SYNERGY platform as described in section 3.2) or data assets that they have acquired from other electricity data value chain stakeholders (as described in section 3.3), there is a distinct separation between the design of a data analysis job (described in section 3.4.1) that happens once by a data asset consumer and the execution of a data analysis job either in the organization's Secure Experimentation Playground in the cloud (described in section 3.4.2) or on-premise (described in section 3.4.3) that occurs according to the schedule provided by the data asset consumer in the job's configuration. The raw data but also the analytics results become available to the SYNERGY energy apps (described in detail in section 6) and to any application that belongs to the respective electricity data value chain stakeholders through the retrieval workflow (described in section 3.4.4).

3.4.1 Data Analysis Job Design Workflow

As soon as data asset consumers have concretely identified a business problem that needs to be addressed, they may access the Analytics Workbench (described in section 4.5.1) in the SYNERGY Cloud Platform to design a new data analysis pipeline. Such a pipeline consists of different types of blocks that need to be properly placed in the pipeline and configured:

i. The data asset consumer needs to select the data input blocks and can include them in the pipeline once their licensing details are resolved with the help of Data & AI Marketplace (e.g. whether derivation is allowed). Of course, whenever needed, the Data & AI Marketplace addresses the Contract Lifecycle Manager to retrieve any applicable smart contract's details.

ii. The data asset consumer needs to select what data manipulation functions need to be applied on the data, how and with what order in a complex data manipulation block or in multiple data manipulation blocks.

iii. The data asset consumer needs to configure what data analytics models need to be applied on the data, ranging from the basic algorithms that are offered in an out-of-the box manner to the baseline models that have been pre-trained for specific business problems, in the data analytics models blocks.

iv. The data asset consumer needs to define the output block that refers to how the results should become available. In the Visualization & Reporting Engine (described in section 4.5.2), the data

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

asset consumer shall configure the visualization and reporting settings if the results are to be consumed in the SYNERGY Cloud Platform. With the help of the API Gateway (described in section 4.8.5), the data asset consumer may configure the results retrieval by external applications through the SYNERGY Open APIs.

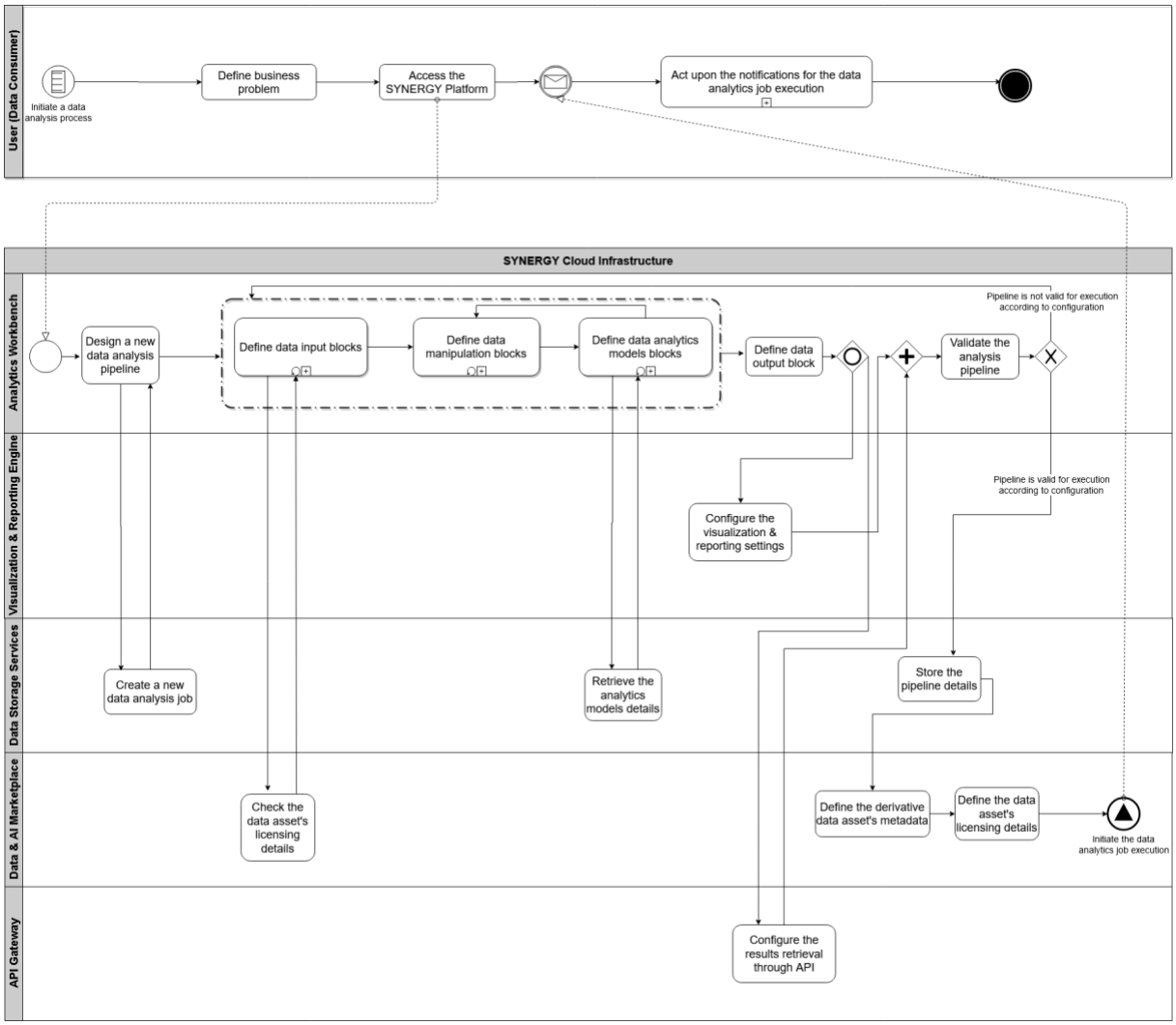


Figure 16: SYNERGY Data Analytics Design Workflow

As depicted in Figure 16, the data asset consumers can place as many data manipulation blocks and data analysis models blocks they wish and in the order they want them to be executed, considering that the output of one block acts as the input to its following block. Before finalising the pipeline configuration, they need to provide the validate the analysis pipeline consistency, especially with regard to the execution location preferences they have expressed. This validation check is quite important to prevent execution failures to the extent it is possible since there are computation limitations in the On-Premise Environments, especially in their edge edition, and it may be impossible

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

to run complex or computationally demanding pipelines. If the analysis pipeline is not valid for execution according to the configuration, the data asset consumers need to revise the pipeline, otherwise they need to define the metadata and licensing details of derivative data asset (that shall emerge as the outcome of the analysis) in the Data & AI Marketplace. Once all necessary details have been provided, the data analysis job is ready for execution.

In comparison to D2.6, the Data Analytics Design Workflow has essentially remained the same.

3.4.2 Data Analysis Job Cloud Execution Workflow

As depicted in Figure 17, if a data analysis job is configured to be executed in the SYNERGY Cloud Infrastructures and in particular in the Secure Experimentation Playgrounds that are reserved for each organization, the Master Controller appropriately schedules the pipeline execution. Once it is time for the analysis job execution, it orchestrates the execution of each block in the pipeline and triggers the appropriate services for execution in the organization's Secure Experimentation Playground.

For data input blocks, the Master Controller checks whether the necessary data are already available in the Secure Experimentation Playground. If they are not available, it initiates the complex process of transferring them to the Playground, decrypting and storing them (with the involvement of the appropriate SYNERGY components, e.g. the Encryption Engine and the Data Storage Services in the Playground).

For data manipulation blocks, the Master Controller triggers the execution of the Data Manipulation Service (described in section 4.5.3) and communicates with the Resources Orchestrator (described in section 4.8.1) to ensure the necessary computation and memory resources are available in the Secure Experimentation Playground (otherwise it dynamically scales them). The Data Manipulation Service is executed and if it successfully concludes, the data manipulation results are temporarily stored, otherwise the related error information is collected and the specific data manipulation block fails.

For data analytics blocks, the Master Controller triggers the execution of the Analytics Execution Service (described in section 4.5.4) and communicates with the Resources Orchestrator (described in section 4.8.1) to ensure the necessary computation and memory resources are available in the Secure Experimentation Playground (otherwise it dynamically scales them). The Analytics Execution Service is executed and if it successfully concludes, the data analytics results are temporarily stored, otherwise the related error information is collected and the specific data analytics block fails.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

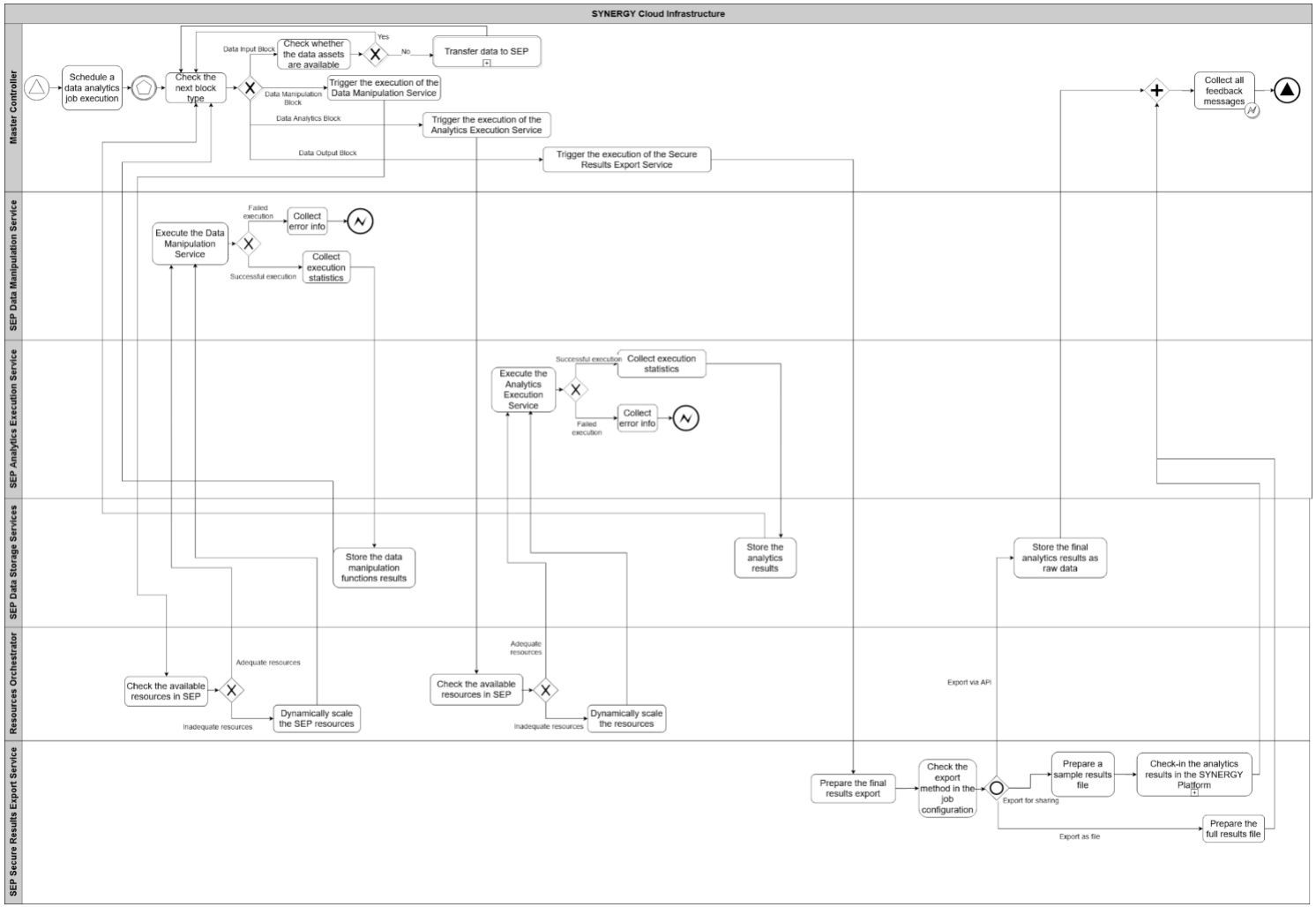


Figure 17: SYNERGY Data Analytics Cloud Execution Workflow

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

For data output blocks, the Master Controller triggers the execution of the Secure Results Export Service (described in section 4.5.5) that is responsible for preparing the final results and for taking appropriate action depending on the export method selected by the data asset consumer in the data analysis job configuration. If the results are to be visualized in the Visualization & Reporting Engine or to be exported via API in the API Gateway for consumption by external applications, the raw results are stored in the Secure Experimentation Playground's Data Storage Services. If the results need to be exported as a file, then the file is prepared and a secure link to download it from the Secure Experimentation Playground becomes available. If the results are to be shared with other electricity data value chain stakeholders in the SYNERGY Cloud Platform, a sample results file is prepared and the data check-in process that has been described in section 3.2 is broadly followed.

In case of any error in the execution of any of the blocks, appropriate feedback messages are collected to appropriately inform the data asset consumer.

In comparison to D2.6, the Data Analytics Job Cloud Execution Workflow has essentially remained the same.

3.4.3 Data Analysis Job On-Premise Execution Workflow

If a data asset consumer has opted for the On-Premise Execution of a data analysis job, a similar process to the cloud execution described in section 3.4.2 is followed.

As depicted in Figure 18, the Master Controller (residing in the SYNERGY Cloud Infrastructure) appropriately schedules the pipeline execution. Once it is time for the data analysis job execution, it orchestrates the execution of each block in the pipeline and triggers the appropriate services for execution in the organization's Server On-Premise Environment.

For data input blocks, the Master Controller checks whether the necessary data are available in the Server On-Premise Environment. If they are not available (e.g. in cases of data acquired from another electricity data value chain stakeholder), it initiates the complex process of transferring them if there is an active data contract and its terms permit it (that is again checked with the help of the Contract Lifecycle Manager).

For data manipulation blocks, the Master Controller triggers the execution of the Data Manipulation Service (described in section 4.5.3). If the Data Manipulation Service is successfully executed, the data manipulation results are temporarily stored, otherwise the related error information is collected and the specific data manipulation block fails.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

For data analytics blocks, the Master Controller triggers the execution of the Analytics Execution Service (described in section 4.5.4). If the Analytics Execution Service is successfully executed, the data analytics results are temporarily stored, otherwise the related error information is collected and the specific data analytics block fails.

As it cannot be guaranteed that the necessary computation and memory resources are available in the On-Premise Environment, it is expected that execution failures for data manipulation blocks and data analytics blocks will occur due to insufficient resources.

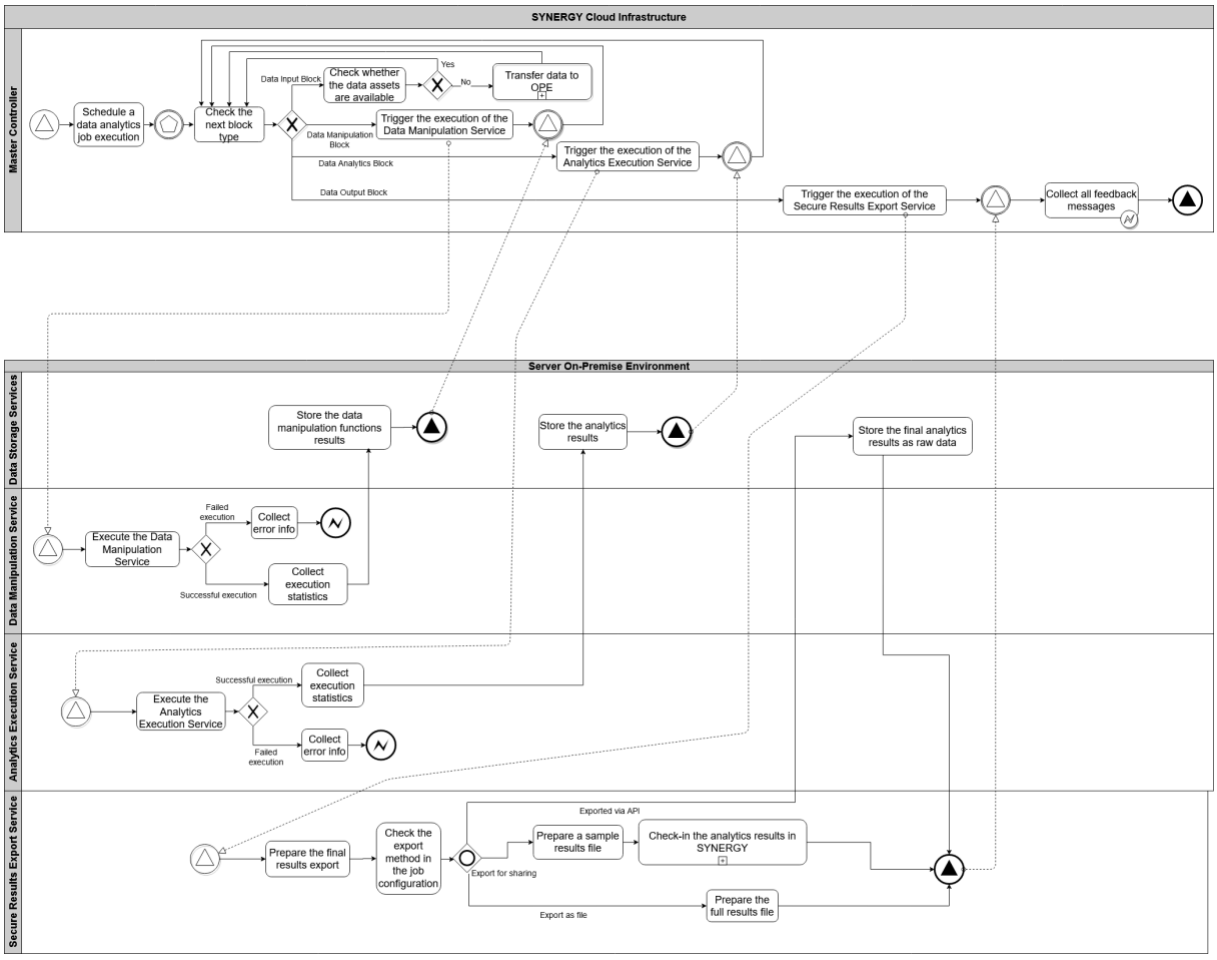


Figure 18: SYNERGY Data Analytics On-Premise Execution Workflow

For data output blocks, the Master Controller triggers the execution of the Secure Results Export Service (described in section 4.5.5) that prepares the final results and takes appropriate action depending on the export method selected by the data asset consumer in the data analysis job configuration. If the results are to be visualized in the Visualization & Reporting Engine or to be exported via API in the API Gateway for consumption by external applications, the raw results are stored in the Data Storage Services. If the results need to be exported as a file, then the file is prepared

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

and a secure link to download it becomes available. If the results are to be shared with other electricity data value chain stakeholders in the SYNERGY Cloud Platform, a sample results file is prepared and the data check-in process that has been described in section 3.2 is broadly followed.

In case of any error in the execution of any of the blocks, appropriate feedback messages are collected to appropriately inform the data asset consumer.

It needs to be noted that the execution in the Edge On-Premise Environment is similar to the process described in Figure 18 for the Server On-Premise Environment, with the exception that data are not transferred to the Edge On-Premise Environment from the Cloud Platform, but they are directly available from the source through the Data Ingestion Service.

In comparison to D2.6, the Data Analytics Job On-Premise Execution Workflow has essentially remained the same.

3.4.4 Applications' Data & Results Retrieval Workflow

As depicted in Figure 19, when an application (e.g. a SYNERGY Energy App or an external application belonging to an electricity data value chain stakeholder) wants to retrieve a specific data asset – in the form of datasets or derivative analytics results - from the overall SYNERGY platform (i.e. Cloud Platform, Secure Experimentation Playgrounds, On-Premise Environments), it needs to call the pre-defined retrieval query that has been configured as described in sections 3.3.1 and 3.4.1. It utilises the query parameters defined in the query configuration in order to properly filter the results and only retrieve the data "slice" that is relevant in each instance.

The API Gateway is responsible for initially checking the validity of the request. If the request is unauthorized or invalid, the respective error message is returned to the application, otherwise it decomposes the retrieval query to the datasets it refers to based on the configuration provided by the respective application owner. Both the applicable access policies for the involved datasets are checked with the help of the Access Policy Engine and the existence of active data asset contracts is confirmed by the Contract Lifecycle Manager. If access to the query results should be granted, the Query Builder executes the specific query and retrieves the specific data slice that is requested based on the initial query parameters for unencrypted data that are stored in the Data Storage Services; in case the data are encrypted, filtering the query results is obviously not possible. It needs to be noted that the requested data may originate from one or multiple data assets. For analytics results, the Secure Results Export Service is triggered to provide the respective derivative data.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

The API Gateway collects all responses from the different services and if the results are available for the specific application and the specific request, they are returned. Otherwise the failure is properly handled by returning an error code or partial results, if available, depending on the preferences expressed in the retrieval query configuration.

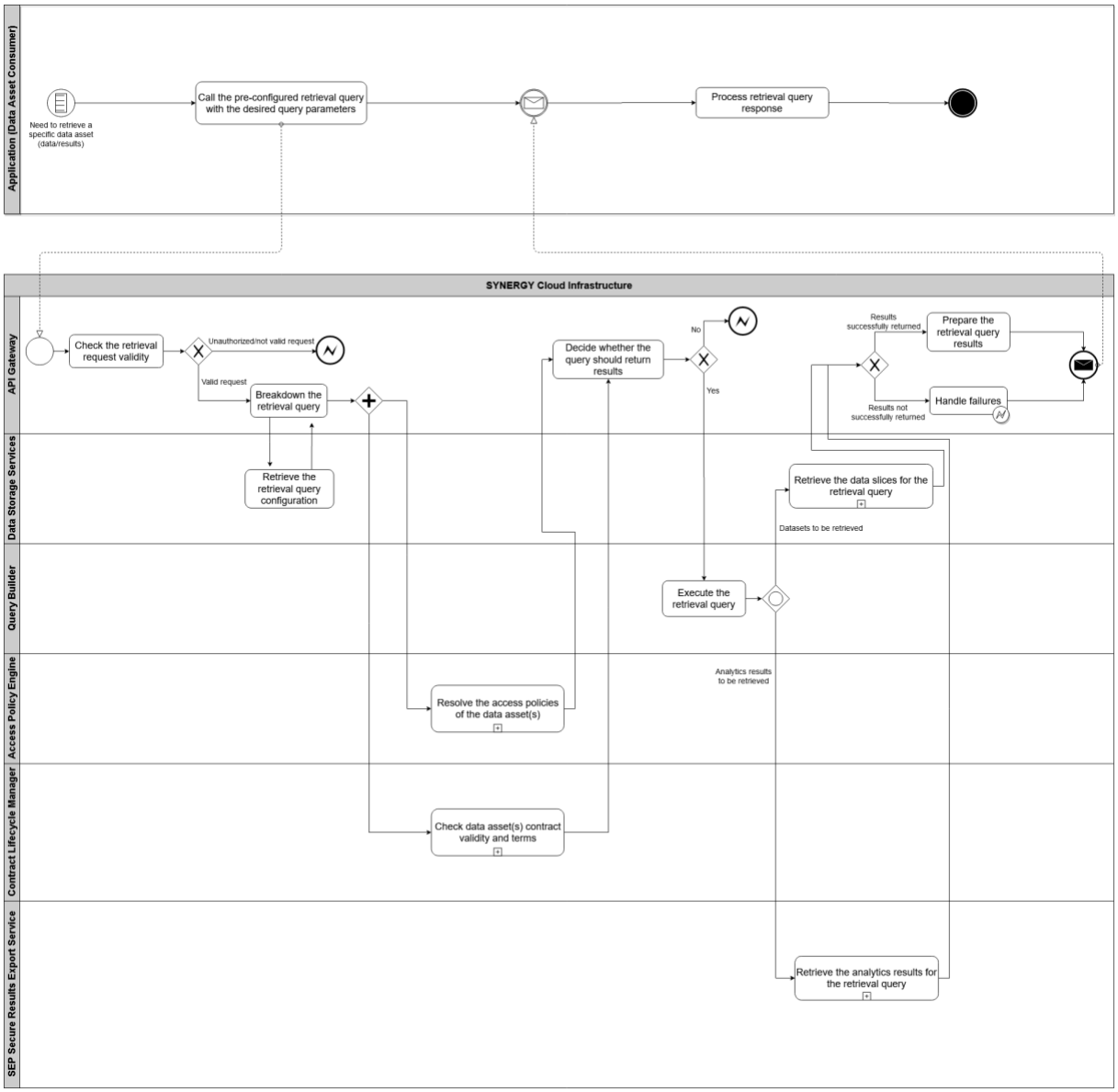


Figure 19: SYNERGY Applications' Data & Results Retrieval Workflow

In comparison to D2.6, the Applications' Data & Results Retrieval Workflow has essentially remained the same.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4 SYNERGY Cloud Infrastructure Components

4.1 Data Collection Services Bundle

The Data Collection Services Bundle that is involved in the data check-in process in the SYNERGY Cloud Infrastructure (as defined in section 3.2) consists of: (a) the Data Handling Manager, (b) the Matching Prediction Engine, (c) the Data Ingestion Service, (d) the Mapping & Transformation Service, (d) the Cleaning Service, as described in the following paragraphs.

4.1.1 Data Handling Manager

4.1.1.1 Component Overview

The Data Handling Manager is responsible for the fully-fledged definition of the data check-in processes by the different stakeholders in order to ingest energy-related data from various sources in the SYNERGY Big Data Platform. The data check-in process practically refers to the process whereby data providers from the electricity data value chain announce how, when and what data they own or are legitimate to provide, shall arrive at the SYNERGY Big Data Platform. Such a data check-in process is fully configurable by the data provider through guided and intuitive user interfaces, and dictates the terms and settings under which the timely and successful retrieval of data shall be performed in an end-to-end manner, namely: (a) the data ingestion method and its associated parameters, (b) the data mapping and processing requirements, (c) the data cleaning rules, (d) the data anonymisation rules, (e) the data encryption rules, and (f) the data storage modality.

The Data Handling Manager is considered as an instrumental component within the Data Collection Services Bundle in SYNERGY as it gathers, stores, and handles the necessary configurations (in the user interface) in order for the data ingestion, data mapping and transformation, data cleaning, data anonymisation and data encryption services (described in Sections 4.1.3, 4.1.4, 4.1.5, 4.2.2, 4.2.3) to be properly executed, according to the preferences of each data provider.

4.1.1.2 List of Features

The list of features offered by the Data Handling Manager (DHM) include:

- **DHM_1 - Step-by-step definition of the data ingestion configuration in an intuitive manner**: The Data Handling Manager allows electricity data value chain stakeholders, acting as data asset providers, to define the parameters of the data ingestion process, according to their preferences

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

and needs. Indicatively, they are able to define the ingestion method (i.e. Direct File(s) Upload, Energy Stakeholders' APIs, Open Data APIs, SYNERGY Platform APIs, Streaming Data Ingestion via the SYNERGY PubSub mechanism, Streaming Data Ingestion via the Energy Stakeholders' own PubSub mechanisms), and set the retrieval schedule, as well as the part of the sample data that should be stored.

- **DHM_2 - Step-by-step data mapping configuration in an intuitive manner**: The Data Handling Manager shall provide an intuitive user interface to the electricity data value chain stakeholders, in order to navigate through the mapping predictions (provided by the Matching Prediction Engine), providing the necessary mapping and transformation details or updating them to the correct concepts of the SYNERGY Common Information Model (CIM).
- **DHM_3 - Step-by-step data cleaning configuration in an intuitive manner**: The Data Handling Manager gives the opportunity to the electricity data value chain stakeholders to improve the overall quality and added value of their data assets, by declaring how incomplete, incorrect, inaccurate or irrelevant parts of the data should be handled.
- **DHM_4 - Step-by-step data anonymisation configuration in an intuitive manner**: The Data Handling Manager allows the electricity data value chain stakeholders to: (a) semi-automatically check for any "privacy-risky" columns within their datasets taking into account the SYNERGY Common Information Model (CIM), and (b) define appropriate anonymisation rules for any potentially "identifying", "quasi-identifying" and "sensitive" columns/fields in their data.
- **DHM_5 - Step-by-step data encryption configuration in an intuitive manner**: With the help of the Data Handling Manager, the different asset providers that intend to upload their data to the SYNERGY Platform are able to set their encryption parameters, selecting which part of the dataset (in terms of selected columns/fields) will be indexed. As encryption is supported in transit for end-to-end data security, the whole dataset is typically encrypted for the transfer from the On-Premise Environment of a stakeholder to the Secure Experimentation Playground in the SYNERGY Platform.
- **DHM_6 - Data Storage configuration in a flexible and informed manner**: With the help of the Data Handling Manager, the different asset providers that intend to upload their data to the SYNERGY Platform are able to configure: (a) the intermediate, non-permanent data storage for processing their data (either on the SYNERGY Cloud infrastructure or in their own On-Premise Environment); and (b) the permanent data storage where their data will be persisted (either on the SYNERGY Cloud infrastructure or in their own On-Premise Environment).

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **DHM_7 - Secure handling of advanced authentication aspects, while protecting sensitive data**: When an API-based data check-in job or a streaming data check-in job is initiated in the Data Handling Manager, the authentication aspects need to be defined for establishing a reliable connection with the SYNERGY platform. As authentication is typically handled in different ways, the data asset provider shall define the applicable type of authentication and provide the necessary parameters (e.g. tokens, username and password for custom login) during the ingestion configuration.
- **DHM_8 - Lifecycle management of data check-in configurations**: The user-defined data collection configuration is effectively created and stored in a configuration file for the specific data check-in job. For an already finalized configuration, updates by the data asset provider are available under certain conditions: (a) in order to update already checked-in data assets, (b) in order to update selected aspects of the data check-in configuration, e.g. the periodic data retrieval schedule can be revised if the data check-in job is not running at the specific moment or to restart a completed API-based / streaming data based job), (c) in order to delete a data check-in job (considering that the already checked-in data assets will not be able to be updated in the future).

4.1.1.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 3: Mapping of the Data Handling Manager to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
DHM_1	WP3_005, WP3_006, WP3_010, WP3_019, WP3_020, WP3_021, WP3_022, WP3_023, WP3_024, WP3_025, WP3_026, WP3_027, WP3_092
DHM_2	WP3_032, WP3_033, WP3_034, WP3_035
DHM_3	WP3_004, WP3_038, WP3_039, WP3_040
DHM_4	WP3_042, WP3_043
DHM_5	WP3_044, WP3_045, WP3_074
DHM_6	WP3_012, WP3_014, WP3_015, WP3_068, WP3_069
DHM_7	N.A.
DHM_8	WP3_029, WP3_030, WP3_031

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.1.1.4 Exploited Technology Stack

In order to provide its intended functionalities, the Data Handling Manager will build on state-of-the-art technologies, namely: (a) in the back-end layer, the Nest (NodeJS[1]) web framework for delivering efficient, reliable and scalable server-side applications as required by the SYNERGY Platform, (b) in the front-end layer, VueJS[2] and TailwindCSS[3] for custom SYNERGY front-end design; (c) in the data storage layer, PostgreSQL[4] (as the relational database for the data check-in job configuration storage), MinIO[5] (as the data lake for the intermediate data management) and Vault[6] (as the secure database for sensitive and secret parameters).

4.1.1.5 Updates from Draft Architecture

The scope and functionalities of the Data Handling Manager remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, changes have been introduced in the following features in order to be consistent with the development progress in the respective component:

- *DHM_5 - Step-by-step data encryption configuration in an intuitive manner*: The encryption configuration concerns the data asset provider's preferences for indexing a specific field in the final architecture, rather than focusing on the selection of which fields shall be encrypted (as described in D2.6). Such a decision was taken as significant performance improvement was noticed in the required encryption/decryption time of a fully encrypted file in comparison to a partially encrypted file (with only few encrypted columns).
- *DHM_8 - Lifecycle management of data check-in configurations*: The options available when a data asset provider updates a data check-in job have been revisited and streamlined, taking into consideration the dependencies between a data check-in job and its resulting data asset and to ensure that there are no inconsistencies in the data stored.

[1] https://nestjs.com/
[2] https://vuejs.org/
[3] https://tailwindcss.com/
[4] https://www.postgresql.org/
[5] https://min.io/
[6] https://www.vaultproject.io/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.1.2 Matching Execution Engine

4.1.2.1 Component Overview

All datasets ingested in the SYNERGY platform need to conform to a Common Information Model (CIM), in particular the SYNERGY CIM (as defined in D3.1), in order to gain a common domain-specific understanding over the data, facilitate the application of all data-related services of the platform and enable the provision of advanced services, indicatively including semantically enhanced search functionalities across data from various sources and domain-dependent contextualisation of data processing. The Matching Prediction Engine is responsible for defining the way each dataset is mapped to the concepts of the CIM and exporting this definition in a configuration file that will be used by the Mapping & Transformation Service to perform the required transformations on the ingested data to be conformant to the model. Considering that the CIM foresees specific types, formats and measurement units (where applicable) for its concepts and fields, the defined mapping configuration may require the definition of the applicable measurement units on numerical data, as well as the specification of the exact date formats and time zones used in the data to allow the appropriate conversions to be performed. It should be noted that by mapping a dataset to the CIM, it is implicitly semantically annotated and enriched with additional properties and also relationships among the dataset concepts and among the current datasets and other SYNERGY assets can be identified.

The availability of the CIM is a precondition for the Mapping Prediction Engine in order to ensure that the ingested data will be properly ingested and mapped. To this end, the Mapping Prediction Engine collaborates with the CIM Manager and retrieves the latest CIM version that is persisted in the SYNERGY Data Storage.

4.1.2.2 List of Features

The component's features include:

- **MPE_1 - Generation of proposed mapping from the dataset fields to the CIM concepts**: The Matching Prediction Engine employs various techniques in order to compute the best possible mapping from each of the dataset fields to a concept of the CIM. Linguistic-based matching on the concept names, also leveraging concept descriptions and provided relevant terms, and sample-based matching using trained concept prediction models, when available, are among the utilised approaches. The mapping calculation service thus returns the best candidate(s) for each dataset concept (called mapping prediction), together with a probability score denoting the prediction confidence.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **MPE_2 – Updates and refinements to the generated mapping:** The Matching Prediction Engine allows the user to manually edit the automatically generated mapping in order to correct mis-matched concepts, provide additional required information (e.g. specify the measurement unit of the input data so that they can be correctly transformed to the respective measurement unit used in the SYNERGY CIM) and select the appropriate concept when no mapping prediction is available for a field.
- **MPE_3 – Easy exploration of the CIM concepts:** In order to allow the user to verify the correctness of the generated mapping prediction, but also to complement and refine it, the Matching Prediction Engine allows the user to browse through the CIM hierarchy and offers information for each concept.
- **MPE_4 – Mapping validation**: As explained, the final mapping configuration will be used to perform the necessary transformations on the data being ingested. In order to eliminate the possibility of data corruption, e.g. when a field of the dataset is of different type from the model field and a casting step would be required, the Matching Prediction Engine performs an integrity check on the mapping configuration and generates warnings for cases that either require user validation or cannot be handled.
- **MPE_5 – Mapping template export, update and re-use:** Upon user request, a completed and validated mapping configuration can be exported as a set of mapping instructions, i.e. actions that need to be performed on a dataset in order to render it compliant to the CIM. This configuration is also called mapping template, as it can be re-used across datasets without requiring the mapping process to be repeated. As an example, when ingesting data from APIs, the mapping process will only be needed once and the same mapping template can be then used every time new data are retrieved. Additionally, it may be the case that a data provider offers different datasets of the same structure, therefore requiring a single mapping template for all, or very similar templates that can be created with small modifications to the same template.
- **MPE_6 – Mapping template revision once a new CIM version is available:** In case the mapping configuration is ongoing by a data provider and a new major version of the Common Information Model becomes available from the CIM Manager, the Matching Prediction Engine cross-checks the changes introduced between the versions and whether they affect the incomplete mapping. In the case of non-backwards compatible changes in the CIM and depending on whether they concern concepts or fields, there are 2 options: (a) the mapping is fully removed by the Matching

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Prediction Engine and the data provider has to re-start the mapping configuration; (b) the mapping is partially removed by the Matching Prediction Engine and the data provider has to correct the mapping configuration for some fields only.

4.1.2.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 4: Mapping of the Matching Prediction Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
MPE_1	WP3_001, WP3_004, WP3_032, WP3_070, WP3_093
MPE_2	WP3_001, WP3_004, WP3_032, WP3_033, WP3_034, WP3_035, WP3_071, WP3_072, WP3_093
MPE_3	WP3_004, WP3_033
MPE_4	WP3_001, WP3_034, WP3_071, WP3_072, WP3_093
MPE_5	WP3_029, WP3_030
MPE_6	WP3_001

4.1.2.4 Exploited Technology Stack

In order to provide its intended functionalities, the Matching Prediction Engine will build on state-of-the-art technologies, namely: (a) in the back-end layer, the NestJS (NodeJS) web framework, (b) in the front-end layer, VueJS and TailwindCSS for custom SYNERGY front-end design; (c) in the data storage layer, PostgreSQL (for storing the mapping configuration), and Elasticsearch[7] (for retrieving the latest CIM version).

4.1.2.5 Updates from Draft Architecture

The scope and functionalities of the Matching Prediction Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6.

[7] https://www.elastic.co/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.1.3 Data Ingestion Service

4.1.3.1 Component Overview

The Data Ingestion Service is responsible for the collection of data into the SYNERGY platform according to the data check-in job configuration performed through the Data Handling Manager. The service handles various ingestion methods to support stakeholders needs, e.g. providing datasets as files, through APIs and using streaming (PubSub) mechanisms. Depending on the selected method, different configuration options are available, e.g. scheduling when data collection should be performed through APIs, authentication aspects, the applicable connection error handling strategy and pagination aspects. Since the API and the PubSub data ingestion involve sensitive parameters (e.g. API keys, tokens, PubSub connection details), they are stored with care, in an encrypted form in the Data Storage layer.

The Data Ingestion Service is triggered for execution on demand by the Master Controller. The ingested data are stored in a temporary object storage for increased reliability and roll-back options in case of failure in the next pre-processing steps (Mapping & Transformation Service, Cleaning Service, Anonymisation Service, Encryption Service).

4.1.3.2 List of Features

The list of features offered by the component are as follows:

- **DIS_1 - Flexible configuration of data ingestion process:** Although the configuration of the data collection is performed through the Data Handling Manager as part of a data check-in job configuration, the available options and the backend services to support them correspond to functionalities of the Data Ingestion Service.
- **DIS_2 - Data collection from files:** The Data Ingestion Service allows for data retrieval from files in formats that can be processed (e.g. csv, tsv, json, xml) or that should be stored as-is (e.g. other types).
- **DIS_3 - Data collection from APIs:** The Data Ingestion Service supports data retrieval from 3rd-party APIs (that are exposed by the electricity data value chain stakeholders' systems or the SYNERGY energy apps), from Open Data APIs and from the SYNERGY Platform's own APIs (to allow stakeholders who do not have any APIs to push data to the platform).
- **DIS_4 - Data collection from PubSub mechanisms:** The Data Ingestion Service supports streaming data ingestion through PubSub mechanisms hosted in the SYNERGY platform

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

(providing the connection details and the topic the stakeholders should use to push data to the platform) and through any PubSub mechanisms already available in the stakeholders' premises (providing the connection details and the topic the Data Ingestion Service should use to collect data).

- **DIS_5 - Selection, pre-processing and storage of data payload subset:** Following the data collection, the Data Ingestion Service will process the data to keep the part that corresponds to the fields selected by the user during the data check-in job (or everything in case the user did not specify any particular fields) and will temporarily store them in an object store in an appropriate way to be consumed by the subsequent services in the pipeline, as configured during data check-in in the Data Handling Manager.
- **DIS_6 – Secure and reliable data transfer:** The Data Ingestion Service is invoked to efficiently transfer the datasets an organisation owns or has legitimately acquired (according to an active smart contract in the Contract Lifecycle Manager) to the Secure Experimentation Playground. The service invocation occurs as always from the Master Controller that checks whether the necessary data input is available at the time a data analysis job is scheduled to run.

4.1.3.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 5: Mapping of the Data Ingestion Service to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
DIS_1	WP3_005, WP3_006, WP3_010, WP3_019, WP3_020, WP3_021, WP3_022, WP3_023, WP3_030, WP3_065, WP3_091
DIS_2	WP3_005, WP3_020, WP3_022, WP3_023
DIS_3	WP3_006, WP3_019, WP3_024, WP3_025, WP3_026, WP3_027, WP3_030
DIS_4	WP3_010, WP3_021
DIS_5	WP3_005, WP3_030, WP3_091, WP3_093, WP3_094, WP3_094
DIS_6	N.A.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.1.3.4 Exploited Technology Stack

In order to provide its intended functionalities, the Data Ingestion Service will build on state-of-the-art technologies, namely: (a) in the back-end layer, the Flask micro web framework[8], the Kafka[9] distributed stream-processing software platform as the PubSub mechanism for streaming data collection, and the RabbitMQ[10] message broker system for sending the appropriate feedback messages to the Master Controller; (b) in the data storage layer, PostgreSQL (as the relational database for retrieving the configuration settings of a data check-in job), MinIO (as the data lake for temporarily storing the ingested data) and Vault (as the secure database for sensitive and secret parameters).

4.1.3.5 Updates from Draft Architecture

The scope and functionalities of the Data Ingestion Service remain essentially the same as presented in the SYNERGY Deliverable D2.6.

4.1.4 Mapping & Transformation Service

4.1.4.1 Component Overview

The Mapping & Transformation Service is responsible for performing on datasets the transformations defined in the matching configuration which is generated by the Mapping Prediction Engine, i.e. for bringing the dataset being ingested into the appropriate form as foreseen by the SYNERGY CIM and defined in a previously created mapping template. Considering that the CIM foresees specific types, formats and measurement units (where applicable) for its concepts, the defined configuration may apply, apart from field renaming, actions such as casting data types, adapting data to foreseen measurement units, specifying date formats to allow the appropriate conversions to be performed.

Data mapping constitutes an important part of the data check-in process and is always the second service invoked by the Master Controller (after the Data Ingestion Service), acting as an enabler for many of the functionalities offered by other SYNERGY components and services, as explained in the corresponding sections. The transformed data are stored in a temporary object storage for increased reliability and roll-back options in case of failure in the next pre-processing steps (Cleaning Service, Anonymisation Service, Encryption Service).

[8] https://flask.palletsprojects.com/en/1.1.x/
[9] https://kafka.apache.org/
[10] https://www.rabbitmq.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.1.4.2 List of Features

The component's features include:

- **MTS_1 – Mapping of ingested data to the SYNERGY CIM:** The Mapping & Transformation Service offers a backend mapping service that executes the defined mapping configuration (mapping template). This execution includes renaming the dataset's fields (e.g. column names of tabular data, field names of json data etc.) according to the mapped CIM concept names.
- **MTS_2 – Transformation of ingested data to comply with the SYNERGY CIM:** Depending on the data type and the accompanying CIM provisions, the Mapping & Transformation Service performs calculations to bring the dataset values to the CIM measurement units, reformats data (e.g. transforming datetime fields to the CIM foreseen format and to the UTC time zone if time zone info is not included in the data), casts data types and applies any other transformation required to execute the defined mapping configuration.
- **MTS_3 – Insights into the results of the transformation rules performed over the ingested data:** Every time the Mapping & Transformation Service is executed, it collects certain metrics about the changes and transformations that were actually performed on the ingested data. The data provider may view the mapping and transformation report for the last run (e.g. in case data are collected from APIs or PubSub mechanisms), but also for all successful runs.
- **MTS_4 – Clear failure indications of transformation rules:** In case a set of transformation rules fails to execute successfully, the Mapping & Transformation Service highlights the transformation rules that created problems and in case any critical errors are made in the mapping, it fails as a whole. In this way, the data provider will be able to decide whether any of the mapping and transformation rules should be revised before re-executing the Mapping & Transformation Service.

4.1.4.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 6: Mapping of the Mapping & Transformation Service to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
MTS_1	WP3_001, WP3_032
MTS_2	WP3_001, WP3_004, WP3_016, WP3_032, WP3_034, WP3_071, WP3_072

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Feature	Related Requirements (D2.2)
MTS_3	N.A.
MTS_4	N.A.

4.1.4.4 Exploited Technology Stack

In order to provide its intended functionalities, the Mapping & Transformation Service will build on state-of-the-art technologies, namely: (a) in the back-end layer, the Flask micro web framework, Pandas[11] for handling the data (upon bringing them to a tabular format) and the RabbitMQ message broker system for sending the appropriate feedback messages to the Master Controller; (b) in the data storage layer, PostgreSQL (as the relational database for retrieving the mapping configuration settings of a data check-in job), and MinIO (as the data lake for temporarily storing the transformed data).

4.1.4.5 Updates from Draft Architecture

The scope and functionalities of the Mapping & Transformation Service remain essentially the same as presented in the SYNERGY Deliverable D2.6.

4.1.5 Cleaning Service

4.1.5.1 Component Overview

The Cleaning Service is responsible for ensuring that data ingested into SYNERGY are accurate and complete, according to the rules defined by the data provider during the data check-in configuration. Removal or correction of incomplete, inconsistent, improperly formatted or otherwise incorrect data increases data quality and leads to effortless reusability and more trustworthy extraction of insights. As such, the specific service is responsible for providing the functionalities needed to clean the data, ranging from simple value substitutions, reformatting and duplicate removal, to more advanced tasks such as outlier detection and substitution. Furthermore, through monitoring the cleaning rules execution, the service can provide valuable insights to the data provider and also offer quality assurances to prospective data consumers.

The Data Cleaning Service is triggered for execution on demand by the Master Controller. The cleansed data are stored in a temporary object storage for increased reliability and roll-back options in case of failure in the next pre-processing steps (Anonymisation Service, Encryption Service).

[11] https://pandas.pydata.org/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.1.5.2 List of Features

The list of features offered by the Cleaning Service are as follows:

- **CS_1 - Flexible configuration of cleaning rules:** Although the configuration of the data cleaning is performed through the Data Handling Manager as part of a data check-in job configuration, the available options and the backend services to support them correspond to functionalities of the Cleaning Service and therefore the provided parameterization constitutes an important feature of the service.
- **CS_2 - Data cleaning rules execution**: The core functionality of the service is the provision of data cleaning methods. During the cleaning configuration process, the user defines constraints that the data must adhere to and corrective actions to handle the violation of the constraints. This configuration translates to a set of cleaning rules which are executed by invoking the corresponding data cleaning methods. The Cleaning Service provides numerous data validation options depending on the data type of each column/field in a dataset, such as allowed value ranges, uniqueness constraints, mandatory constraints (for handling missing data values), regular expression patterns, and outliers identification. It also offers two main types of corrective measures: dropping entries and replacing values, the latter offering more options as to how the new value is generated, either provided by the user (i.e. a fixed value) or dynamically calculated based on the data (i.e. mean, min, max, previous values). The combination of a validation option and a corrective action for a certain field forms a cleaning rule, and all cleaning rules are executed on the transformed data (from the Mapping & Transformation Service) resulting in a clean dataset.
- **CS_3 - Feedback from the executed cleaning rules:** When executed, the defined cleaning rules will affect some of the ingested dataset's entries, resulting either in dropping entries or in altering the values in certain fields. During the execution of the rules, statistics measuring how many times a cleaning rule was fired (i.e. its corrective action was applied) are collected. The data provider may thus consult these numbers as a cross-checking mechanism both for the properties of the data that he/she is making available in SYNERGY and the defined cleaning rules. As an example, a rule being fired too often could indicate either an unforeseen issue with the ingested data or a misconfigured cleaning rule – both cases would require the provider's attention and thus the cleaning service facilitates spotting the issue. Furthermore, such statistics may act as data quality indicators. In case a cleaning rule has resulted into a failure for the whole

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

cleaning service, the data provider will have a clear view of the rule(s) that created the problem in order to correct them.

- **CS_4 – Easy testing of defined cleaning rules on sample data:** Due to the importance of the cleaning step and its intrusive nature, i.e. the fact that data may be removed or irreversibly altered, the Cleaning Service provides testing functionalities operating on user defined data samples to support the users in properly configuring the cleaning rules.

4.1.5.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 7: Mapping of the Cleaning Service to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
CS_1	WP3_035, WP3_038, WP3_040, WP3_065
CS_2	WP3_016, WP3_029, WP3_030, WP3_037, WP3_038
CS_3	WP3_038, WP3_039
CS_4	WP3_093

4.1.5.4 Exploited Technology Stack

In order to provide its intended functionalities, the Cleaning Service will build on state-of-the-art technologies, namely: (a) in the back-end layer, the Flask micro web framework, Pandas for applying the cleaning methods (upon bringing them to a tabular format) and the RabbitMQ message broker system for sending the appropriate feedback messages to the Master Controller; (b) in the data storage layer, PostgreSQL (as the relational database for retrieving the cleaning configuration settings of a data check-in job), and MinIO (as the data lake for temporarily storing the cleaned data).

4.1.5.5 Updates from Draft Architecture

The scope and functionalities of the Cleaning Service remain essentially the same as presented in the SYNERGY Deliverable D2.6.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.2 Data Security Services Bundle

The Data Security Services Bundle in the SYNERGY Cloud Infrastructure consists of: (a) the Anonymisation Service, (b) the Encryption Engine, (c) the Access Policy Engine, as described in the following paragraphs.

4.2.1 Anonymisation Service

4.2.1.1 Component Overview

Anonymisation is one of the ways in which SYNERGY safeguards data assets against unintended disclosure of personal or corporate information. As briefly described in section 4.1.1, one of the steps of a data check-in job configuration is the definition of the anonymisation actions that should be performed, in particular which parts of the data (i.e. which fields) should be anonymized prior to making the data available within SYNERGY. The Anonymisation Service, as part of the SYNERGY Data Security Services Bundle, is responsible for informing the data providers about potentially sensitive information within the datasets and for providing the functionalities to anonymise this information, as well as any field of their data that they consider as containing any identifying information (individually or in conjunction with other fields).

The Anonymisation Service is triggered for execution on demand by the Master Controller. The anonymised data are stored in a temporary object storage for increased reliability and roll-back options in case of failure in the next pre-processing steps (Encryption Service). Since the Anonymisation Service needs the whole dataset at its disposal otherwise any grouping within the data and overall the anonymised data will be totally inconsistent, certain restrictions shall be applied in the configuration of the Anonymisation Service (e.g. for streaming data or small volumes of data that are gradually acquired through APIs).

4.2.1.2 List of Features

The list of features offered by the component are as follows:

- **AS_1 - Identification of fields that need anonymisation:** The anonymisation service (through the UI provided in the Data Handling Manager) helps the data providers to identify any dataset fields that hold information that is either sensitive or identifying personal/corporate information (ranging between identifying fields and quasi-identifier fields that jointly help identify personal/corporate information).

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **AS_2 - Flexible configuration of anonymisation rules:** Although the configuration of the data anonymisation process is performed through the Data Handling Manager as part of a data check-in job configuration, the available options and the backend services to support them correspond to functionalities of the Anonymisation Service and therefore the provided parameterization flexibility constitutes an important feature of the service.
- **AS_3 - Data anonymisation rules execution**: Depending on the type of the field (sensitive, identifying, quasi-identifying) and the data type, the Anonymisation Service allows the data provider to define the anonymisation method that should be applied in an easy-to-understand manner with examples. Indicative methods that are supported under the k-anonymity algorithm include: generalization methods to create arithmetic intervals or categories for numeric fields, and masking methods for string fields. It should be noted that it is not within the scope of the SYNERGY project to offer a fully-fledged data anonymisation tool, but to provide targeted data anonymisation functionalities that can be configured and used by the prospective data providers.
- **AS_4 - Insights extracted from the executed anonymisation rules:** Considering that data anonymisation is, as it was the case with the cleaning process described in section 4.1.5, an intrusive process which significantly and irreversibly affects the underlying data, it is important to ensure that the data provider understands how the defined rules will affect the data. Therefore, the service offers a utility function which is used to compute data loss caused by the anonymisation process, aiming to help the provider find the right balance to safeguard sensitive information without rendering the data unusable. In case the data loss by the Anonymisation Service (in order to achieve the desired k-anonymity level) is higher than the threshold defined by the data provider, the Anonymisation Service will fail on purpose in order to give the chance to the data provider to check again the anonymisation rules that have been defined.

4.2.1.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 8: Mapping of the Anonymisation Service to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
AS_1	WP3_042, WP3_043
AS_2	WP3_041, WP3_043

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Feature	Related Requirements (D2.2)
AS_3	WP3_029, WP3_030, WP3_041, WP3_065, WP3_093
AS_4	N.A.

4.2.1.4 Exploited Technology Stack

In order to provide its intended functionalities, the Anonymisation Service will build on state-of-the-art technologies, namely: (a) in the back-end layer, the Flask micro web framework, Pandas for applying the anonymisation techniques (upon bringing them to a tabular format) and the RabbitMQ message broker system for sending the appropriate feedback messages to the Master Controller; (b) in the data storage layer, PostgreSQL (as the relational database for retrieving the anonymisation configuration settings of a data check-in job), and MinIO (as the data lake for temporarily storing the anonymised data).

4.2.1.5 Updates from Draft Architecture

The scope and functionalities of the Anonymisation Service remain essentially the same as presented in the SYNERGY Deliverable D2.6.

4.2.2 Encryption Engine

4.2.2.1 Component Overview

SYNERGY offers data encryption functionalities to data providers as additional mechanisms to safeguard their data and eliminate the possibility of unauthorised data access and/or data leakage of any type. The component that undertakes this task is the Encryption Engine, which is part of the Data Security Services Bundle. Proprietary datasets can be therefore encrypted prior to leaving the owner's premises (in the Server On-Premise Environment), according to the owner's needs and requirements. The Encryption Engine provides flexibility in the definition of the encryption configuration to be applied, enabling modular control over the data contents that should be encrypted. Specifically, the engine allows the data provider to encrypt the full dataset or select the parts of a dataset that should be indexed, i.e. the contents that correspond to specific dataset's fields/columns for which search will be enabled. In collaboration with the Security, Authentication & Authorisation Engine, the Encryption Engine also offers the required key sharing and data decryption services to allow secure sharing of datasets between data providers and consumers and rightful data processing in the Secure Experimentation Playgrounds and/or the Server On-Premise Environments. It needs to be noted that for performance reasons, certain restrictions are to be applied in the configuration of the encryption

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

process (e.g. for streaming data or small volumes of data that are gradually acquired through APIs, but near real-time access is required).

4.2.2.2 List of Features

The list of features offered by the component are as follows:

- **EE_1 - Flexible configuration of encryption rules:** Although the configuration of the data encryption rules (that essentially refer to the indexing preferences) is performed through the Data Handling Manager as part of a data check-in job configuration, the available options and the backend services to support them correspond to functionalities of the Encryption Engine and therefore the provided parameterization flexibility constitutes an important feature of the component.
- **EE_2 - Data encryption rules execution**: The Encryption Engine provides symmetric encryption mechanisms for the data and is responsible for both the generation of the encryption key and the actual encryption of the underlying data so that the data provider will be in full control of who can access the information.
- **EE_3 – Computations on data prior to encryption for search purposes:** Part of the encryption process configuration is to define which dataset fields should be used for data discoverability purposes, even if the actual field values will be encrypted. In order to provide such functionality, the Encryption Engine will calculate some predefined statistics on the fields that are marked as to-be encrypted, but searchable. These calculations depend on the type of the field, e.g. for numeric data, extracting minimum and maximum values or total number of null values could be such a computation. The extracted information will be made available to other SYNERGY services responsible for searching within assets and for providing insights into asset contents to prospective consumers.
- **EE_4 - Data decryption:** The engine also provides the services needed to decrypt the data in the Secure Experimentation Playgrounds (and the On-Premise Environment if on-premise download is foreseen) so that they can be used by the users that are authorised to access them. In this case, the Encryption Engine ensures together with the Master Controller that the decryption key becomes available in order to decrypt and access the underlying information. The key for the symmetric data encryption is asymmetrically encrypted and shared between the involved parties.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **EE_5 – Key revocation:** A key revocation process needs to be followed when a user (organisation) should no longer have access to the data (e.g. based on an expired data sharing contract). In that case, the data that data consumers should not have at their disposal anymore are removed from their organisation's Secure Experimentation Playground. A relevant notification is also sent to inform the data consumers that they should remove the specific data asset from their On-Premise Environment (in case the contract's terms had allowed them to download the data "locally"), otherwise legal action may be taken by the respective data asset providers. The key revocation process is currently elaborated considering the overall SYNERGY project needs.

4.2.2.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 9: Mapping of the Encryption Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
ES_1	WP3_012, WP3_014, WP3_044, WP3_045, WP3_073, WP3_074
ES_2	WP3_012, WP3_014, WP3_029, WP3_030, WP3_044, WP3_045, WP3_073
ES_3	WP3_074
ES_4	WP3_046, WP3_075, WP3_076
ES_5	WP3_075, WP3_076

4.2.2.4 Exploited Technology Stack

In order to provide its intended functionalities, the Encryption Engine will build on state-of-the-art technologies, namely: (a) in the back-end layer, the NestJS (NodeJS) web framework, and the RabbitMQ message broker system for sending the appropriate feedback messages or signals for the key exchange to the Master Controller; (b) in the data storage layer, PostgreSQL (as the relational database for retrieving the encryption configuration settings of a data check-in job), and MinIO (as the data lake for temporarily storing the encrypted data prior to their permanent storage in the Data Storage Services Bundle).

4.2.2.5 Updates from Draft Architecture

Although the scope and functionalities of the Encryption Engine generally remain the same as presented in the SYNERGY Deliverable D2.6, changes have been introduced in: (a) the key sharing

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

mechanism and (b) the options available in the encryption rules (as per field/column encryption is not allowed based on the development advancements to increase performance).

4.2.3 Access Policy Engine

4.2.3.1 Component Overview

As part of the Data Security Services offered to increase trust of the electricity data value chain stakeholders in the overall SYNERGY platform, the Access Policy Engine provides mechanisms that allow data providers to define in a flexible and fine-grained way the access rules that should be applied in the form of authorisation policies for permitting or denying access requests on their data assets within SYNERGY. The Access Policy Engine allows the user to define rules in a user-friendly manner both for allow policies and deny policies and appropriately combine and order any foreseen exceptions so as to achieve the desired behaviour. Access to data assets in the Access Policy Engine is regulated through Attribute-Based Access Control (ABAC) policies, that allow the data providers to protect and share their data assets, even when they do not have any prior knowledge of the potential individual data consumers in the overall SYNERGY platform. For example, a data provider may define an access policy for a specific data asset that "no organisation that is a DSO and is from Greece can access the data asset". Such access policies can be updated and/or deleted only by the data provider as an organisation and any potential changes need to be immediately enforced.

A proper separation of concerns between policy definition and policy enforcement is effectively ensured in the Access Policy Engine. For example, whenever data consumers search for data assets in the Data & AI Marketplace, the Access Policy Engine evaluates the applicable access policies and decides whether all potential results that the Query Builder has returned should be visible. In this way, it is ensured that the only data assets that data consumes can view in the Data & AI Marketplace are data assets which their organisation is eligible to acquire (through the Contract Lifecycle Manager).

4.2.3.2 List of Features

The list of features offered by the Access Policy Engine are as follows:

- **APE_1 - Flexible definition, configuration, and update of data asset access policies:** The Access Policy Engine offers a graphical user interface through which the users can easily create complex access policy rules for their assets. These rules may depend on properties of the asset, properties of the requestor (e.g. the type of organisation they belong to) and also contextual properties of the request (e.g. the time the request was performed or the country it originated from). Rules

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

may define when access is granted or when it is denied, and they can be combined using Boolean logic to form complex rules. The rules are stored and can be easily changed through the provided interface by the data asset provider.

- **APE_2 - Enforcement of asset access policies:** The Access Policy Engine is responsible for applying the defined access policies when a request is performed to access a data asset that has been checked-in in SYNERGY. The allow/deny decision is taken by the Access Policy Engine in a performant manner as it affects not only the search results that appear in the Data & AI Marketplace, but also the data that are to be retrieved through the API Gateway (esp. for public data assets). It should be noted that the scope of the access policies is different from the terms defined in a data sharing contract and enforced by the Contract Lifecycle Manager.

4.2.3.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 10: Mapping of the Access Policy Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
APE_1	WP3_053, WP3_04, WP3_055, WP3_056, WP3_057, WP3_061, WP3_081, WP3_082, WP3_083, WP3_084
APE_2	WP3_055, WP3_056, WP3_057, WP3_061, WP3_084

4.2.3.4 Exploited Technology Stack

In order to provide its intended functionalities, the Access Policy Engine will build on state-of-the-art technologies, namely: a) in the back-end layer, the Nest (NodeJS) web framework and the Casbin[12] authorisation library, (b) in the front-end layer, VueJS and TailwindCSS for custom SYNERGY front-end design; (c) in the data storage layer, PostgreSQL (as the relational database for persisting the access policies per data asset).

4.2.3.5 Updates from Draft Architecture

The scope and functionalities of the Access Policy Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6.

[12] https://casbin.org/en/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.3 Data Sharing Services Bundle

The Data Sharing Services Bundle in the SYNERGY Cloud Infrastructure consists of: (a) the Data & AI Marketplace, (b) the Contracts Lifecycle Manager, (c) the Remuneration Engine, as described in the following paragraphs.

4.3.1 Data & AI Marketplace

4.3.1.1 Component Overview

The Data & AI Marketplace aims at effectively addressing data sharing in the SYNERGY platform, as it provides the interface for all functionalities related to discovering assets, browsing through them and engaging into data asset sharing agreements, either as a data asset provider or a data asset consumer, or in certain cases assuming both roles in the context of multi-party contracts. The term data asset in this context stresses the fact that the SYNERGY marketplace is not limited to data acquisition, but also handles other assets around data, mainly AI pre-trained models targeting domain problems and results of data analytics.

The Data & AI Marketplace plays an instrumental role in the SYNERGY vision for interconnected datasets and collective intelligence, as it allows the electricity data value chain stakeholders to easily locate data assets of interest, under clearly defined licenses, and to effortlessly express their interest, negotiate with providers and proceed to acquire them. The SYNERGY CIM and its metadata schema significantly facilitate apart from the assets' discoverability, the acquisition and consumption process (i.e. the asset's usage) as all assets conform to the same standards and vocabularies, which accelerates identifying the most suitable assets for their needs.

From a UI-perspective, the Data & AI Marketplace encapsulates functionalities regarding asset search (in collaboration with the Query Builder), provision of recommendations (in collaboration with the Matchmaking Engine), assets browsing, as well as actions required to draft and sign smart asset sharing contracts (as supported by the Contract Lifecycle Manager), in order to provide a seamless experience to the users, even though the underlying services are implemented by other components which will be described in subsequent sections.

4.3.1.2 List of Features

The list of features offered by the Data & AI Marketplace are as follows:

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **DAIM_1 - Clear definition of data asset metadata:** The Data & AI Marketplace is responsible for the definition, storage, retrieval, and update of all metadata, according to the SYNERGY metadata schema. Depending on the type and the intended purpose of the data asset (e.g. if it intended to be shared or not), SYNERGY allows for customized metadata that need to be filled in to facilitate search for the specific data asset.
- **DAIM_2 - Handling of licenses, IPR handling and pricing schemes for a data asset:** The Data & AI Marketplace is responsible for the definition, storage, retrieval and update of licensing, IPR and pricing information around available data assets. Specifically, the data providers define the appropriate license, IPR details and pricing scheme for their data using the available SYNERGY options which range from static uniform pricing to more dynamic schemes, e.g. depending on the number of rows to be retrieved in case of data assets. Drafting custom licenses that correspond to the specific needs of a data asset provider is supported while predefined license templates, which can be reviewed and directly assigned to data assets, can be also used to accelerate and simplify the process for the data asset providers.
- **DAIM_3 – Data asset search and recommendations:** The details of these features will be examined in the sections of the corresponding underlying services (i.e. the Query Builder and the Matchmaking Engine), but these functionalities are also tightly linked to the marketplace experience and therefore are perceived as part of the marketplace.
- **DAIM_4 – Data asset browsing:** The Data & AI Marketplace offers different views allowing the users (acting as data asset consumers) both to: (a) browse through numerous data assets (e.g. datasets, pre-trained algorithms, analytics results/reports) shown in the form of a catalogue, e.g. when results of a search query are presented, and (b) examine each data asset independently and view more details. In the second case, the marketplace offers diverse views depending on whether the data asset is owned by the organisation, acquired by the organisation through SYNERGY or is neither owned nor acquired.
- **DAIM_5 – Management of data asset sharing contracts:** To offer a complete marketplace experience, the Data & AI Marketplace also provides (again from a user perspective) functionalities around initiating a data asset acquisition process, accepting/ denying such requests, drafting, negotiating over, and signing data sharing contracts (as dictated by the Contract Lifecycle Manager). The Data & AI Marketplace also requires a different type of contracts, derivative contracts (that are also handled by the Contract Lifecycle Manager), to be already in place in order to include the derivative data assets (e.g. pre-trained AI models,

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

analytics results) that inevitably involve many data asset providers. Performing the payment foreseen by a contract can also be seen as part of this process, although it is again handled by the Remuneration Engine.

- **DAIM_6 – Acquisition of multiple data assets at once through a cart:** The Data & AI Marketplace allows data asset consumers to gather all data assets of interest in their cart and, when they are ready, to acquire them all at once or proceed to separate contracts.

4.3.1.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 11: Mapping of the Data & AI Marketplace to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
DAIM_1	WP3_003, WP3_048, WP3_084, WP3_085, WP4_104
DAIM_2	WP3_003, WP3_048, WP3_049, WP3_051, WP3_052, WP3_080, WP3_084, WP3_085, WP4_104
DAIM_3	WP4_106, WP4_107, WP4_108, WP4_109, WP4_110, WP4_111, WP4_112
DAIM_4	WP3_085, WP3_103, WP4_104, WP4_105, WP4_106, WP4_107, WP4_108, WP4_113, WP4_115, WP4_118, WP4_119, WP4_013, WP4_014, WP4_019 - WP4_096, WP4_165, WP4_166
DAIM_5	WP3_103, WP4_001, WP4_003, WP4_007, WP4_008, WP4_009, WP4_011, WP4_098, WP4_099, WP4_102, WP4_103, WP4_122, WP4_123, WP4_124, WP4_125, WP4_162

4.3.1.4 Exploited Technology Stack

In order to provide its intended functionalities, the Data & AI Marketplace will build on state-of-the-art technologies, namely: (a) in the back-end layer, the Nest (NodeJS[13]) web framework, (b) in the front-end layer, VueJS[14] and TailwindCSS[15] for custom SYNERGY front-end design; (c) in the data storage layer, Elasticsearch (for performant search over the actual data and/or metadata).

4.3.1.5 Updates from Draft Architecture

The scope and functionalities of the Data & AI Marketplace remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, changes have been introduced

[13] https://nestjs.com/
[14] https://vuejs.org/
[15] https://tailwindcss.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

with regard to the multi-party contracts in order to be consistent with the development progress in the respective component:

- **DAIM_5 – Management of data asset sharing contracts:** In order to increase trust of the electricity data value chain stakeholders, derivation contracts have been introduced and need to be signed among the involved stakeholders as a prerequisite in order for a derivative data asset to appear in the Data & AI Marketplace.
- **DAIM_6 – Acquisition of multiple data assets at once through a cart:** It is a newly introduced functionality to improve the user experience when a data asset consumer intends to acquire many different data assets quicker.

4.3.2 Contract Lifecycle Manager

4.3.2.1 Component Overview

The Contract Lifecycle Manager is responsible for all operations around smart data contracts for data asset sharing and trading. The SYNERGY smart contracts are utilised to handle data and data-enabled assets' sharing agreements. The machine-processable format of smart contracts, the support for all steps of contract drafting, the secure storage of the contracts in a distributed ledger and the effective enforcement of contract terms for assets used within the SYNERGY tools and services, significantly facilitate, and accelerate the data asset sharing process.

4.3.2.2 List of Features

The list of features offered by the Contract Lifecycle Manager are as follows:

- **CLM_1 - Smart contract drafting and update:** When prospective data asset consumers identify a data asset of interest within the SYNERGY Data & AI Marketplace that belongs to another organisation, they may address a request to acquire the asset to its provider, as explained, as long as their organisation are eligible to acquire it according to the applicable access policies (that are resolved by the Access Policy Engine). If the data asset provider does not reject the request, a draft smart contract is created, with terms stemming from the data asset's metadata (e.g. license and pricing attributes). The involved parties may then enter a phase of negotiation over the contract terms, proposing, performing and accepting/rejecting changes on the terms. The process may terminate in either party rejecting the contract or both parties accepting it and signing it, in which case it enters a pre-valid state, pending payment to be performed (in the Remuneration Engine) in order to become valid. It should be stressed that the Contract Lifecycle

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Manager also supports multi-party smart contract drafting, negotiation and signing functionalities, i.e. not limited to two involved stakeholders each time, in order to address cases of data asset consumers intending to acquire derivative data assets or multiple data assets at once. All involved services for these actions over smart contracts are provided by the Contract Lifecycle Manager, which also handles the communication with the SYNERGY distributed ledger in charge of keeping the smart contracts.

- **CLM_2 - Smart contracts validity (status) check:** The Contract Lifecycle Manager is responsible for the communication with the SYNERGY distributed ledger to check whether a specific contract is valid and active. This service is invoked for example when a user attempts to access a purchased dataset, in which case the current component will be first asked to ensure there is an active asset contract that permits granting access to the requestor.
- **CLM_3 – Enforcement of smart contract terms:** Apart from checking and reporting on the contract's status, the Contract Lifecycle Manager is also responsible for assessing attempted asset manipulation/handling/processing actions against the contract's terms to ensure that no terms are violated and thus the actions defined in the Analytics Workbench should be allowed. It should be noted that not all contract terms are enforceable within the overall SYNERGY platform (i.e. some terms may refer to obligations not relevant to the way the consumer treats the asset within SYNERGY, but in on-premise tools, outside the SYNERGY On-Premise Environments, that cannot be controlled by SYNERGY).
- **CLM_4 – Downloading smart contracts:** All stakeholders involved in a specific smart asset sharing contract can download the contract terms as pdf files locally in order to cross check them with their legal team, but also for archiving purposes.
- **CLM_5 – Easy renewal and/or extension of smart contracts:** The data asset consumers are able to extend the duration of a smart contract (prior or after their expiry), as well as to add more fields of the original data asset or extend the filtered period for the already acquired fields in the case of datasets.

4.3.2.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 12: Mapping of the Contract Lifecycle Manager to the SYNERGY Requirements

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Feature	Related Requirements (D2.2)
CLM_1	WP4_001, WP4_002, WP4_003, WP4_004, WP4_005, WP4_006, WP4_010, WP4_097, WP4_102, WP4_103
CLM_2	WP4_115, WP4_116, WP4_117, WP4_118, WP4_119, WP4_127, WP4_128
CLM_3	WP3_065, WP3_075, WP3_076, WP4_100, WP4_115, WP4_116, WP4_117, WP4_118, WP4_119, WP4_127, WP4_128
CLM_4	WP4_098
CLM_5	WP4_101

4.3.2.4 Exploited Technology Stack

In order to provide its intended functionalities, the Contract Lifecycle Manager will build on state-of-the art technologies, namely: (a) in the back-end layer, the Nest (NodeJS) web framework, and (b) in the blockchain layer, the Ethereum[16] distributed platform.

4.3.2.5 Updates from Draft Architecture

The overall scope and functionalities of the Contract Lifecycle Manager remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, the feature CLM_5 (Easy renewal and/or extension of smart contracts) has been added in order to address an underlying need of the stakeholders (that emerged from the experimentation with the beta release of the SYNERGY Integrated Platform).

4.3.3 Remuneration Engine

4.3.3.1 Component Overview

The Remuneration Engine acts as the underlying payment system of the data asset contracts in the SYNERGY platform as it is responsible for performing and confirming the payments foreseen by a smart asset sharing contract, using the selected SYNERGY cryptocurrency, which at the early design time is ETH, the Ethereum's native cryptocurrency. Such an engine is practically an intermediary layer between the data asset consumer and the different data asset providers for processing the cryptocurrency payments which also involves recording the transaction on the SYNERGY distributed ledger for multi-party data asset contracts. It needs to be noted that in cases of bilateral data asset contracts that involve a single data asset provider and consumer, the Remuneration Engine may act

[16] https://ethereum.org/en/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

on the background allowing for a conventional payment between the 2 parties (through a bank account, a credit card, etc.), yet such a payment needs to be confirmed in the SYNERGY platform by the data asset provider and written in the SYNERGY distributed ledger in order for the respective data asset contract to be considered as active.

The Remuneration Engine is also responsible for calculating: (a) the remuneration amount that should be paid to the different data asset providers (as they define a pricing scheme for their assets, according to their preferences) for the purchase of a specific asset through the SYNERGY marketplace, based on a signed smart contract; (b) the smart contract fee that should be reserved for the SYNERGY platform and the SYNERGY distributed ledger for facilitating the data asset transaction. In order to eliminate any volatility risk, the Remuneration Engine reimburses the different data asset providers through cryptocurrency payments in their own wallets (instead of exchanging them for flat currencies).

4.3.3.2 List of Features

The list of features offered by the Remuneration Engine (RE) include:

- **RE_1 – Management of frictionless payments (in a cryptocurrency) for multi-party data asset contracts based on their pricing terms:** As the data consumer needs to fulfil the payment for a multi-party smart contract to the SYNERGY platform, the Remuneration Engine handles not only the specific payment, but also the payments that need to be made to the involved data asset providers. All payments are made in the selected cryptocurrency by SYNERGY (which, at the design time, is ETH, the Ethereum's native cryptocurrency) in the wallets of the respective platform users (as described in section 5.2.1) immediately.
- **RE_2 – Calculation of remuneration payments in multi-party data asset contracts:** As the data assets that are shared in the Data & AI Marketplace may be derivative works from many asset providers, it is important for the Remuneration Engine to properly estimate the amount that should be paid to such providers based on the smart contract's terms. Such a remuneration also involves the SYNERGY platform that reserves a small fee (as explicitly mentioned in the smart contract) for facilitating the transaction and for writing all steps of the smart contract in the SYNERGY distributed ledger.
- **RE_3 – Confirmation of offline payments in bilateral smart contracts:** In order to increase trust to the SYNERGY platform for bilateral transactions that only involve a data consumer and a data provider, the Remuneration Engine will allow for conventional payments (e.g. through bank transfer to an account of the data provider, or via the credit card of the data consumer).

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

However, such payments are not performed through the Remuneration Engine, but offline in respect to the SYNERGY platform. In order for the contract to be activated and the data consumer to obtain access to the respective data asset, the data provider needs to confirm the payment and this transaction is recorded in the SYNERGY distributed ledger.

- **RE_4 – Credit payment to the data asset consumers in case the multi-party smart contract terms are violated:** In case any of the involved parties in a multi-party contract for a data asset fail to uphold the contract terms and it was not within the control of the SYNERGY platform to prevent it, the Remuneration Engine should take all necessary actions to ensure that part of the funds is returned from the wallet of the defaulting data asset provider(s) to the wallet of the data asset consumer.

4.3.3.3 Mapping with SYNERGY Requirements

WP4_099 is the only requirement that refers to the broader scope of the Remuneration Engine functionalities.

4.3.3.4 Exploited Technology Stack

In order to provide its intended functionalities, the Remuneration Engine will build on state-of-the-art technologies, namely: (a) in the back-end layer, the Nest (NodeJS[17]) web framework, and (b) in the blockchain layer, the Ethereum[18] distributed platform.

4.3.3.5 Updates from Draft Architecture

The scope and functionalities of the Remuneration Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6.

4.4 Data Matchmaking Services Bundle

The Data Matchmaking Services Bundle in the SYNERGY Cloud Infrastructure consists of: (a) the Query Builder, (b) the Matchmaking Engine, as described in the following paragraphs.

[17] https://nestjs.com/
[18] https://ethereum.org/en/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.4.1 Query Builder

4.4.1.1 Component Overview

The Query Builder is the cornerstone of the Data & AI Marketplace, as it provides asset discoverability and exploration functionalities that in turn constitute the enablers for asset sharing transactions. The Query Builder allows SYNERGY users (acting as data asset consumers) to search for data assets of interest, browse the results and explore the selected data assets in more detail in order to identify useful candidates for acquisition. The services offered by the Query Builder extend beyond typical search functionalities found in data marketplaces, as the component's search scope is not limited to data but searches across all SYNERGY assets, i.e. also AI models for the energy domain, analysis results and reports and offers advanced search powered by semantics and relations provided by the SYNERGY Common Information Model. Furthermore, mechanisms are introduced to handle the complexity caused by the fact that datasets may be uploaded as (partially) encrypted in the SYNERGY Data & AI Marketplace.

4.4.1.2 List of Features

The list of features offered by the component are as follows:

- **QB_1 – Intuitive asset search supporting both keyword-based queries and faceted search:** In order to offer flexibility in the way users search for and discover data assets, the Query Builder enables query generation using both a flexible free-text search and filtering on the assets' information. Specifically, the Query Builder enables users to search based on the assets' metadata (which conform to the SYNERGY metadata schema), their structure (i.e. for data this refers to the concepts of the CIM to which the original asset fields have been mapped), their contents (when this is allowed by the data provider), and meta-information regarding the contents (e.g. some statistical information for the distribution of numeric fields which have been encrypted and cannot be directly accessed for search). Faceted search is also supported across selected facets most commonly used to identify interesting assets.
- **QB_2 – Asset search extending beyond data:** Although this feature is implicitly included in the previous one, it should be stressed that the query definition mechanisms that enable the more straightforward process of data search, also support other types of assets, including AI models and analytics reports. This poses additional challenges in offering a useful and seamless querying and exploration functionality.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **QB_3 – Execution of search queries and provision of results:** The query configuration provided by the data asset consumers is translated by the Query Builder to a query to the SYNERGY Storage Services Bundle (actually its indexing service), and results that match the query are returned, processed and provided to the user in an intuitive way. Appropriate accompanying information for each data asset is provided to facilitate the user in quickly locating the most promising results. When no results are returned by the search, the Query Builder will invoke the Matchmaking Engine to retrieve some suggestions instead (either for assets or for other SYNERGY organisations that could act as potential data asset providers), if available.
- **QB_4 – Search history and search query update and re-execution:** There are some queries for assets that the users may want to re-execute in the future either unchanged or slightly altered, especially if the advanced filtering functionalities have been used to avoid repeating steps. Furthermore, apart from queries resulting in data asset results, some of the advanced querying functionalities may retrieve actual data contents as results, in which case saving and re-executing the query (e.g. to get updated or slightly modified data subsets) is extremely useful. Therefore, users may store the search queries that they want and update them or use them as needed.

4.4.1.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 13: Mapping of the Query Builder to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
QB_1	WP4_106, WP4_107, WP4_108
QB_2	WP4_107
QB_3	WP4_106, WP4_107, WP4_108, WP4_109
QB_4	N.A.

4.4.1.4 Exploited Technology Stack

In order to provide its intended functionalities, the Query Builder will build on state-of-the-art technologies, namely: (a) in the back-end layer, the NestJS (NodeJS) web framework, (b) in the data storage layer, PostgreSQL (for storing the query configuration), and Elasticsearch (for executing the search queries effectively and efficiently).

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.4.1.5 Updates from Draft Architecture

The scope and functionalities of the Query Builder remain essentially the same as presented in the SYNERGY Deliverable D2.6.

4.4.2 Matchmaking Engine

4.4.2.1 Component Overview

The Matchmaking Engine is responsible for recommending data assets to the users (i.e. data asset consumers) that might be interesting for them. The recommendations include existing energy-related datasets and potential combinations of the datasets that can be used by the user, but may be also expanded to data asset providers.

The Matchmaking Engine is considered as a fundamental component within the Matchmaking Services Bundle in SYNERGY that analyses the previous interactions between users and data assets such as searching a data asset (through the Query Builder), navigating to the data asset's details (in the Data & AI Marketplace), or purchasing a data asset (through the Contract Lifecycle Manager), which can be used as indicators of future decisions and seeks to predict the preferences of a user in choosing a data asset in the future. The results of the recommendation process will be used either to provide implicit or explicit suggestions to the data asset consumers based on the prediction of user interests or to re-rank the results of data asset search process (as presented in the Data & AI Marketplace):

- Prediction approach: The Matchmaking Engine aims to predict the score for a user-item combination meaning the interest of a user to a data asset or the relation with a specific data asset provider. This approach requires training data such as the history of dataset search and selection of users, that indicates user preferences for data assets.
- Ranking approach: The Matchmaking Engine aims to recommend similar data assets or data asset providers to the users according to their profile similarity or similarity of provided datasets and analytics services. In this case the recommendation will use similarity measures or neighbourhood approaches to find the top items for a user or determine the top users for a particular item.

4.4.2.2 List of Features

The list of features offered by the Matchmaking Engine are as follows:

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **MME_1 – Settings of the user preferences for recommendations:** The Matchmaking Engine allows the users to configure the parameters that can direct the recommendation process. For example, the user can choose if they want to receive personalised or generic recommendations. For the personalised recommendations, the users can determine preferred data concepts from the Common Information Model and/or select their favourite data assets.
- **MME_2 – Providing data asset suggestions to the users:** The Matchmaking Engine will recommend relevant datasets and the organisations that have such data assets relevant to the users according to their search history or the datasets and analytics services that they provide.
- **MME_3 – Provide data provider suggestions to the users:** The Matchmaking Engine recommends to the users a list of data asset providers that may potentially create a relevant data asset in case no results are found for existing datasets.
- **MME_4 – View the list of recommendations:** The users will get the results of the recommendation process as a list of suggested data assets (e.g. datasets, or data asset providing organisations). The SYNERGY Data & AI Marketplace allows the users to browse a sample of the recommended data sets, request for obtaining the datasets or add them to their favourite lists that will be reused as feedback for the future personalised recommendations.

4.4.2.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 14: Mapping of the Matchmaking Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
MME_1	WP4_104, WP4_109, WP4_113
MME_2	WP4_111
MME_3	WP4_109, WP4_110
MME_4	WP4_105, WP4_109, WP4_110, WP4_111, WP4_112, WP4_113

4.4.2.4 Exploited Technology Stack

The Matchmaking engine will be developed by using the state-of-the-art algorithms for recommendation. The core recommendation algorithms will be implemented by using PySpark and

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

PySpark.ML[19] that is a set of high-level APIs that help users create and tune practical machine learning pipelines. This library provides the facilities of developing algorithms for Big Data environment for both batch and streaming data using the open-source distributed general-purpose cluster-computing framework of Apache Spark[20]. The algorithms of the library specially for the ones which are in the alpha release will be complimented by using ScikitLearn machine learning library. A set of complimentary libraries such as Pandas for data manipulation and analysis and Flask for development of APIs will be utilised. The component will be dependent on Elasticsearch for exploiting the indexes and search results provided by Indexing Engine and Query Builder.

4.4.2.5 Updates from Draft Architecture

The scope and functionalities of the Matchmaking Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6. Since no changes are required from the draft architecture, no specific updates are made to the Matchmaking Engine features.

4.5 Data Analytics Services Bundle

The Data Analytics Services Bundle in the SYNERGY Cloud Infrastructure consists of: (a) the Analytics Workbench, (b) the Visualization & Reporting Engine, (c) the Data Manipulation Service, (d) the Analytics Execution Service, (e) the Secure Results Export Service, as described in the following paragraphs.

4.5.1 Analytics Workbench

4.5.1.1 Component Overview

The Analytics Workbench is practically the interface through which the electricity data value chain stakeholders (acting as data asset consumers) can leverage the added value that data analytics can bring through the offerings of the SYNERGY Data Analytics Services Bundle. It addresses the needs of different types of users (i.e. data scientists, technical users, business users) in terms of data processing, from executing simple data manipulation functions (e.g. filters and aggregations) and applying

[19] spark.apache.org/docs/latest/api/python/pyspark.html
[20] spark.apache.org

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

machine learning (ML) and deep learning (DL) models for energy data analytics, to creating visualisations and reports to highlight insights extracted from datasets and from analytics processes.

The Analytics Workbench provides an intuitive user interface that guides the user through the creation of complete data analysis pipelines at "design" time, from selecting and configuring the input data, to the step-by-step processing and exporting the results in the desired form, i.e. as data, as visualisation, as report. Data manipulation and analysis functions are provided in the form of easily configurable blocks that can be combined to form a data analysis pipeline. The different data analysis algorithms that the user can select range from basic algorithms offered by popular frameworks to customised, pre-trained algorithms for specific energy problems (as described in the SYNERGY Deliverable D4.2).

Through the Analytics Workbench, the execution of the defined data analysis pipelines is also configured in order to define where and when it will be planned to run in the Data Manipulation Service and the Analytics Execution Service (to be triggered in the Secure Experimentation Playground or on the On-Premise Environment of an organisation) and the generated outputs are exported or saved according to the configuration (through the Secure Results Export Service), or directly consumed through a connected visualization/report. In case a data asset consumer intends to share the results of an analysis, the related metadata need to be provided in the Data & AI Marketplace.

4.5.1.2 List of Features

The list of features offered by the Analytics Workbench are as follows:

- **AW_1 – Definition of re-usable and customisable data manipulation blocks:** The Analytics Workbench wraps the functions provided by the Data Manipulation Service in configurable and chainable pipeline blocks, allowing the user to select the block that offers the desired functionality, parameterise it and combine it with other data manipulation blocks to form a series of actions to be performed on a dataset (or on numerous datasets).
- **AW_2 – Application of re-usable and customisable energy analytics models based on pre-trained algorithms:** The Analytics Workbench allows the users to select among available ML/DL models that target specific problems of the energy domain, configure and apply them to their own data. These models correspond to the SYNERGY baseline algorithms and are tailored to the industry needs.
- **AW_3 – Definition of re-usable and customisable energy analytics models blocks based on basic algorithms:** For users that want to have more flexibility in implementing their own models, the workbench offers some generic algorithms (i.e. not targeting the energy domain, but generic

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

problems such as classification, clustering, regression, etc.) that can be trained and configured in a more flexible way, but with much more effort to be spent on experimentation.

- **AW_4 – Re-usable and customisable data input blocks:** To allow the users to use the aforementioned data manipulation and analysis blocks, appropriate input blocks are also provided, which can be easily configured through the UI and combined with the other blocks through a drag and drop functionality that allows the user to easily combine blocks in a pipeline. The input blocks indicate own data assets that an organisation has at its disposal as the rightful owner, but also acquired data assets for which there is an active data asset contract (according to the Contract Lifecycle Manager).
- **AW_5 – Re-usable and customisable data output blocks:** The Analytics Workbench allows the user to define output blocks according to their own needs in order to visualize the results through custom diagrams that are created and saved with the help of the Visualization & Reporting Engine, to download the results as a file or to retrieve the results via the SYNERGY Platform's Open APIs (in the API Gateway).
- **AW_6 – Configuration and validation of data analysis pipelines:** After creating and connecting a series of blocks in a pipeline (through a drag and drop approach), the user needs to provide the execution details of the defined pipeline. The execution of a pipeline can be practically scheduled to run once, at specific intervals or upon a triggering event in different locations (e.g. in the Secure Experimentation Playground of an organisation, on a Server or Edge On-Premise Environment). Depending on the configuration provided, different validation checks will be performed to ensure that the pipeline is properly configured or alert the user for any issues. In the background, the configured pipeline is sent to the Master Controller which will ensure that the underlying functions are properly sent to the Data Manipulation Service and the Analytics Execution Service, either of a Secure Experimentation Playground or an On-Premise Environment to be executed according to schedule, on the resources provisioned by the Resources Orchestrator.
- **AW_7 – Update of data analysis pipelines:** Once a data analysis pipeline is finalised and has transitioned to the so-called "execution" time (upon having run at least once), the user is able to revise a limited part of the configuration in order to leave some flexibility for updates, but without compromising the consistency of the results that have been already provided by the Data Manipulation Service and / or the Analytics Execution Service.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **AW_8 – Insights into the execution of data analysis pipelines:** Every time the Data Manipulation Service and the Analytics Execution Service are executed, they collect certain metrics about the outcome and different performance aspects of the data analysis pipeline that are displayed in the Analytics Workbench. The data asset consumer may view the report for the latest run, but also for all successful runs. In case a data analysis pipeline fails to execute successfully, the Analytics Workbench highlights the blocks that created problems and in case any critical errors are detected, the pipeline fails as a whole. In this way, the data asset consumer will be informed about the blocks and execution parameters that should be revised and re-execute the Data Manipulation Service and / or the Analytics Execution Service.
- **AW_9 – Registration of externally trained compatible ML/DL models:** The Analytics Workbench allows the users that have trained ML/DL models (targeting specific problems of the energy domain) outside the SYNERGY Platform yet in compatible libraries, to register them in the Platform. Such models can be applied in analytics pipelines in the same way as models trained in the SYNERGY Platform, and can be shared through the Data & AI Marketplace (as long as their metadata is provided and they are based on open data or data coming from the organization that registers the model).

4.5.1.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 15: Mapping of the Analytics Workbench to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
AW_1	WP4_117, WP4_126, WP4_127, WP4_129, WP4_133, WP4_134, WP4_135
AW_2	WP4_013, WP4_014, WP4_019 - WP4_096, WP4_165, WP4_166
AW_3	WP4_015, WP4_016, WP4_017, WP4_018, WP4_130, WP4_131, WP4_132
AW_4	WP4_134
AW_5	WP4_120, WP4_060- WP4_064
AW_6	WP4_117, WP4_124, WP4_125, WP4_136, WP4_137, WP4_138, WP4_139, WP4_141 - WP4_157, WP4_159
AW_7	N.A.
AW_8	WP4_140
AW_9	N.A.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.5.1.4 Exploited Technology Stack

In order to provide its intended functionalities, the Analytics Workbench will build on state-of-the art technologies, namely: (a) in the back-end layer, the Nest (NodeJS) web framework and the Suite5 Enterprise Analytics Suite; (b) in the front-end layer, VueJS and TailwindCSS for custom SYNERGY front-end design; (c) in the data storage layer, PostgreSQL (as the relational database for retrieving the configuration settings of the data analysis job), MinIO (as the repository for the pre-trained models).

4.5.1.5 Updates from Draft Architecture

The scope and functionalities of the Analytics Workbench remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, a new functionality, namely AW_9 – Registration of externally trained compatible ML/DL models, has been added in order to address the feedback received from data scientists (concerning the inclusion of models that have been created and trained externally to the Platform) without compromising the SYNERGY Platform's security.

4.5.2 Visualisation & Reporting Engine

4.5.2.1 Component Overview

The Visualisation & Reporting Engine is an integral part of the Data Analytics Services Bundle, as it gives to the data asset providers and consumers the ability to visually gain insights from specific data assets and/or analytics results. Such an engine allows users to create custom visualizations of interesting aspects of a particular dataset they want to dive into.

The Visualisation & Reporting Engine works closely with the Analytics Marketplace as the configuration of a visualization is essentially part of an output block of a data analysis pipeline, as well as with the Analytics Execution Service and the Secure Results Export Service to ensure that the results are properly displayed in the preferred, customized diagrams the user has selected. The Visualisation & Reporting Engine can also display the provenance information that is available from the Data Lineage Service in customised visualizations.

Finally, it needs to be noted that visualisations on their behalf also constitute data assets in the SYNERGY context and can also be shared through the Data & AI Marketplace.

4.5.2.2 List of Features

The list of features offered by the Visualisation & Reporting Engine are as follows:

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **VRE_1 - Support for built-in charts:** The Visualisation & Reporting Engine offers numerous predefined charts which can be easily configured and fed with data to provide meaningful and aesthetically pleasing visualisations with minimum user effort. It allows displaying not only the results of a data analysis pipeline, but also the raw data of a dataset and the lineage details (from the Data Lineage Service).
- **VRE_2 – Visualization through multiple charts over the data analytics pipelines:** The Visualisation & Reporting Engine supports the configuration of multiple visualisation charts per analytics job in order to provide different views over the analytics results.
- **VRE_3 – Saving and exporting visualisations:** All assets created using the Visualisation & Reporting Engine can be saved both as static objects (i.e. as image or pdf) and as dynamic objects (that can be updated and added upon an analytics pipeline is finalized and executed).

4.5.2.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 16: Mapping of the Visualization & Reporting Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
VRE_1	WP4_160, WP4_162, WP4_163, WP4_164
VRE_2	WP4_120, WP4_136, WP4_137, WP4_149
VRE_3	WP4_121, WP4_123, WP4_160

4.5.2.4 Exploited Technology Stack

In order to provide its intended functionalities, the Visualization & Reporting Engine will build on state-of-the art technologies, namely: (a) in the back-end layer, the Nest (NodeJS) web framework; (b) in the front-end layer, VueJS and TailwindCSS for custom SYNERGY front-end design, as well as Kibana and / or Cube.js as an open-source data visualization dashboards used to build business intelligence tools; (b) in the data storage layer, PostgreSQL (as the relational database for retrieving the output configuration settings of the data analysis job), HDFS (as the distributed file system).

4.5.2.5 Updates from Draft Architecture

The scope and functionalities of the Visualization & Reporting Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6 with regard to the visualization charts. In respect to the draft architecture, the feature VRE_2 (Visualization through custom reports over the data and their

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

derivatives) related to the reporting functionality has been revised as the creation of reports implied the user-driven generation of dashboards that is not considered as feasible (within the SYNERGY project implementation). Therefore, changing the focus of VRE_2 to the configuration of multiple visualisation charts per analytics pipeline was considered as prudent.

4.5.3 Data Manipulation Service

4.5.3.1 Component Overview

The Data Manipulation Service offers various data manipulation functionalities, including filters, conditional column creations, aggregations, merging with other datasets, performing mathematical computations, removing null values, etc. The purpose of the Data Manipulation Service is to help users bring their datasets in a format more suitable for analysis and/or visualisations and reporting. Depending on the technical background of the user and the scope of a particular data manipulation process, the utilised Data Manipulation Service functions may range from simple filters to complex transformations and serve the needs of users that want to create a simple statistics chart as well as those who are preparing the data to be used as input to ML/DL algorithms in the Analytics Workbench. The Data Manipulation Service leverages the CIM mapping to adapt the options provided to the user for each functionality to the dataset fields. It can be executed upon being invoked by the Master Controller in the Secure Experimentation Playground (and the Server On-Premise Environment).

4.5.3.2 List of Features

The list of features offered by the Data Manipulation Service are as follows:

- **DMS_1 – Definition of modular data manipulation functions:** The main offering of the Data Manipulation Service is a set of easily re-usable and chainable functions that implement numerous, commonly used data manipulation/preparation tasks, while hiding any underlying complexities. The input data are expected in a tabular format (as dataframes constitute the most commonly used data structure in data analysis) and the supported actions include filters, aggregations, merges (join operations), column creation, column dropping, conditional row dropping, various datetime and string transformations, datetime manipulation, rolling averages, data contextualisation/enrichment through domain vocabularies, etc.
- **DMS_2 – Creation of configurable data manipulation function templates:** Each of the available data manipulation functions exposes parameters that can be configured by the users to increase re-usability. A configuration template with common properties can be used to configure

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

functions in a consistent way and also allow them to be chained and executed together. This provides an additional way to execute data manipulation functions, i.e. instead of using them directly, to configure them through json templates. The Data Manipulation Service can then use the templates to configure and invoke the functions to produce and return the resulting dataset. This execution method is used when a user creates complex pipelines through the Analytics Workbench.

- **DMS_3 – Extraction of summary statistics for each data manipulation function applied in a dataset:** The Data Manipulation Service can provide insights into the resulting step-by-step dataset emerging as an output from each data manipulation function outputs, as well as an indicative preview (for a small sample) in order for the data asset consumer to directly understand the changes introduced in the data.

4.5.3.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 17: Mapping of the Data Manipulation Service to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
DMS_1	WP3_004, WP4_126, WP4_127, WP4_128, WP4_129, WP4_133, WP4_134, WP4_135, WP4_141, WP4_144, WP4_154
DMS_2	WP3_004, WP4_126, WP4_127, WP4_128, WP4_129, WP4_133, WP4_134, WP4_135, WP4_141, WP4_144, WP4_154
DMS_3	N.A.

4.5.3.4 Exploited Technology Stack

In order to provide its intended functionalities, the Data Manipulation Service will build on state-of-the-art technologies, namely: (a) in the back-end layer, the Flask micro web framework, the Spark and Pandas frameworks, and the RabbitMQ message broker system for sending the appropriate feedback messages to the Master Controller; (b) in the data storage layer, PostgreSQL (as the relational database for retrieving the configuration settings of the data manipulation functions within a data analysis job), MinIO (as the data lake for temporarily storing the transformed data).

4.5.3.5 Updates from Draft Architecture

The scope and functionalities of the Data Manipulation Service remain essentially the same as presented in the SYNERGY Deliverable D2.6.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.5.4 Analytics Execution Service

4.5.4.1 Component Overview

The Analytics Execution Service is responsible for executing data analytics jobs based on configuration files generated by the Analytics Workbench in the Secure Experimentation Playgrounds (or the On-Premise Environment on its server or edge edition), depending on the execution settings that were provided by the user. The Analytics Execution Service practically allows the users to execute diverse data processing pipelines, combining blocks that act as wrappers for the analytics models of the Analytics Workbench. The Analytics Execution Service will use the provided configuration to invoke the defined analytics functions, i.e. to apply ML/DL models to the specified input data according to the configuration parameters and export the results, as foreseen by the configuration.

The Analytics Execution Service is triggered by the Master Controller according to the execution schedule set in the data analysis job configuration and is allocated with the necessary compute and memory resources from the Resources Orchestrator. The results of each job are saved and can be retrieved with the help of the Secure Results Export Service.

4.5.4.2 List of Features

The list of features offered by the Analytics Execution Service are as follows:

- **AES_1 – Implementation, configuration and application of predefined, pre-trained ML/DL models for energy problems:** Baseline algorithms that tackle specific problems of the energy domain (e.g. demand forecasting), are packaged in the form of configurable functions that can be applied on user defined datasets (that need to conform to specific formats according to each model's configuration). Each of the available models exposes certain parameters that allow them to be configured through the Analytics Workbench, as explained, so that the execution service can retrieve the source code that corresponds to a particular model, apply the configuration as needed and perform the analysis in the defined execution engine.
- **AES_2 – Application of generic algorithms from popular ML/DL libraries:** Apart from the aforementioned domain-trained models that are readily applicable, the Analytics Execution Service offers generic lower-level analysis functionalities that enable users to design and implement their own solution, e.g. using classification, clustering, regression etc. algorithms. The Analytics Execution Service is responsible for invoking the right functions, which in the generic algorithms' case will also involve training and extended testing and evaluation functions.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **AES_3 – Unified multi-purpose analytics API:** The Analytics Execution Service offers a simple API to the Analytics Workbench so that all analysis blocks are handled in the same way, using a consistent configuration template, while hiding the underlying complexity of transforming the configuration file to actual coding blocks to be executed using different ML/DL libraries and potentially in different execution engines.
- **AES_4 - Feedback from the executed analytics job:** During the execution of the defined data analytics pipeline, summary statistics about the analytics execution are generated and displayed to the user. In case a specific analytics block has resulted into a failure for the whole pipeline execution, the data asset consumer will have a clear view of the problem in order to correct it and re-execute the specific pipeline.

4.5.4.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 18: Mapping of the Analytics Execution Service to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
AES_1	WP4_013, WP4_014, WP4_019 - WP4_096, WP4_165, WP4_166
AES_2	WP4_015, WP4_016, WP4_017, WP4_018, WP4_130, WP4_131, WP4_132
AES_3	WP4_138, WP4_144
AES_4	N.A.

4.5.4.4 Exploited Technology Stack

In order to provide its intended functionalities, the Analytics Execution Service will build on state-of-the-art technologies, namely: (a) in the back-end layer, the NestJS (NodeJS) web framework, the Spark, scikit-learn and Dask frameworks (amongst which the prioritization for the early release of the Analytics Execution Service will depend on the SYNERGY energy apps' needs that are currently under elaboration), and the RabbitMQ[21] message broker system for sending the appropriate feedback messages to the Master Controller; (b) in the data storage layer, PostgreSQL[22] (as the relational database for retrieving the configuration settings of the data analysis job), MinIO (as the repository for storing the models).

[21] https://www.rabbitmq.com/
[22] https://www.postgresql.org/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.5.4.5 Updates from Draft Architecture

The scope and functionalities of the Analytics Execution Service remain essentially the same as presented in the SYNERGY Deliverable D2.6.

4.5.5 Secure Results Export Service

4.5.5.1 Component Overview

The Secure Results Export Service is triggered in the Secure Experimentation Playgrounds (and the On-Premise Environment) in order to prepare the results of a data analysis pipeline for export and different uses. It allows the electricity data value chain stakeholders to bring the requested data results' "slice" in the appropriate form (and format, if applicable) and make the exportable payload available to be forwarded to the SYNERGY Core Cloud Platform or to other applications (i.e. SYNERGY energy applications or 3rd-party applications through the API Gateway).

4.5.5.2 List of Features

The list of features offered by the Secure Results Export Service (SRES) include:

- **SRES_1 – Securely export data analytics results as a file:** The Secure Results Export Service allows a data consumer to extract the results of an analysis pipeline that is executed once or every time it is executed as a file (e.g. in csv format) that becomes available through a temporary link in an encrypted or unencrypted form.
- **SRES_2 – Securely export data analytics results through an API:** The Secure Results Export Service allows a data consumer to extract the stored results of an analysis pipeline that is executed once or every time it is executed through an API endpoint with the help of the API Gateway of the SYNERGY platform. In this case, the data are expected to be typically exposed in an unencrypted form.
- **SRES_3 – Securely export data analytics results intended to be shared in the SYNERGY platform:** In case a data consumer intends to share the analytics results in the Data & AI Marketplace, the Secure Results Export Service initially extracts a sample of the data analytics results in order to be used to check-in such derivative data in the SYNERGY platform (through the Data Handling Manager). In parallel, it prepares the full analytics' results payload as a file that will be directly retrieved from the Data Ingestion Service when the respective data check-in job is executed.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.5.5.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 19: Mapping of the Secure Results Export Service to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
SRES_1	WP4_116, WP4_121, WP4_151
SRES_2	WP4_152
SRES_3	WP4_150

4.5.5.4 Exploited Technology Stack

In order to provide its intended functionalities, the Secure Results Export Service will build on state-of-the-art technologies, namely: (a) in the back-end layer, the Flask micro web framework and the RabbitMQ message broker system for sending the appropriate feedback messages to the Master Controller; (b) in the data storage layer, MongoDB (as the NoSQL database for storing the results), and MinIO (as the data lake for temporarily storing the results data as a file).

4.5.5.5 Updates from Draft Architecture

The scope and functionalities of the Secure Results Export Service remain essentially the same as presented in the SYNERGY Deliverable D2.6.

4.6 SYNERGY Data Storage Services Bundle

4.6.1 Component Overview

The SYNERGY Data Storage Services Bundle is responsible for persisting a plethora of data assets, along with their associated metadata, and the jobs-related data in a secure and reliable manner in the SYNERGY Cloud Infrastructure. The SYNERGY Data Storage Services Bundle is not an actual component per se, but the term refers to all storage and indexing tools used in the context of all the SYNERGY provided platform services, mainly pertaining to:

- Storing the configuration of the data check-in jobs (involving the Data Handling Manager, the Data Ingestion Service, the Mapping & Transformation Service, the Cleaning Service, the Anonymisation Service, the Encryption Engine, the Master Controller, and the Resources Orchestrator) and the configuration of the data analysis jobs (involving the Analytics

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Workbench, the Visualization & Reporting Engine, the Data Manipulation Service, the Analytics Execution Service, the Secure Results Export Service, the Master Controller and the Resources Orchestrator) to ensure that all steps performed over the data are traceable.

- Storing the different data assets (i.e. datasets, analytics models, analytics results and reports) along with their metadata, in order to be available for all the components and services in the SYNERGY platform.
- Storing information regarding the data assets contracts in a hybrid smart contract form in an immutable manner in a distributed ledger.
- Secure, encrypted storage of sensitive data (e.g. sensitive parameters for the API calls, such as tokens, API keys, usernames and passwords).
- Storing log-related information for the SYNERGY platform operation and usage, users' and organisations' data and all administrative information required for the smooth operations of the Core Cloud Platform, the Secure Experimentation Playgrounds and the On-Premise Environments.
- Persisting the SYNERGY Common Information Model in its different versions along with its associated generic and energy-specific vocabularies.

Depending on the type of information that is to be stored in the SYNERGY Platform and the way it will be retrieved, different storage and indexing tools are foreseen to accommodate the varying needs.

4.6.2 List of Features

The SYNERGY Data Storage Services Bundle offers various functionalities through the diverse storage solutions it provides and encapsulates, whose purpose is examined in the corresponding sections of the components that store and retrieve data from the SYNERGY storage layer. For completeness purposes, the various functionalities, some of which are offered through the same storage, are presented below. It should be noted that appropriate indexing mechanisms are also available in each case depending on the intended data retrieval functionalities.

- **DSS_1 – Storage of Datasets in Trusted Data Containers:** Once the data assets are appropriately checked-in and handled, the processed data along with the processed data sample are permanently stored in an inherently scalable database, allowing for the optimized management of big data. The derivate datasets that emerge as the results of a data analytics job are also persisted in trusted data containers.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **DSS_2 – Metadata Storage for SYNERGY Data Assets:** Along with the processed data assets, the Data Storage Services Bundle receives the accompanying metadata for storage, creating the appropriate links between the stored data assets and their metadata information.
- **DSS_3 – Storage of AI models:** Since the Analytics Workbench (along with the Analytics Execution Service) will offer a set of basic and pre-trained algorithms, a dedicated storage space is reserved for analytics models (e.g. trained algorithms) that essentially contains containers/code, as well.
- **DSS_4 – Storage of all data check-in and data analysis jobs configurations:** As described, many of the SYNERGY services generate and export configuration templates which are executed by different backend services in different locations. Such configurations are stored either for logging purposes or in order to be cloned, updated and re-executed as needed.

- **DSS_5 - Temporary object storage of intermediate files**: During the execution of the different services involved in the data check-in jobs and the data analytics jobs, intermediate data files are created for temporary storage, enabling the pause of a job (in case of a failure) and its continuation at a later stage.
- **DSS_6 – Storage of sensitive data and credentials:** The Data Storage Services Bundle provides a dedicated database to secure, persist and tightly control access to different sensitive parameters (e.g. tokens, API keys, usernames and passwords) as well as to effectively manage such secret codes.
- **DSS_7 – Storage of the SYNERGY CIM and energy vocabularies:** As the Common Information Model is crucial for different services and components of the overall SYNERGY platform, particular attention is paid to the effective storage for the model concepts, fields and hierarchies.
- **DSS_8 – Storage in a distributed Contracts Ledger:** The Data Storage Services Bundle expands over a distributed ledger which is responsible for storing the smart data asset sharing contracts in a way that ensures non-repudiation and full traceability while maintaining the privacy of the involved stakeholders and the confidential details (e.g. about the contract price).
- **DSS_9 – Centralized storage for all platform's operational data:** The Data Storage Services Bundle anticipates storage for additional types of data not explicitly mentioned in the previous items of the list, mainly including platform data, information about users and organisations, metadata from the platform's operations etc.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.6.3 Mapping with SYNERGY Requirements

Although there are certain requirements that explicitly refer to storage functionalities, (in particular WP3_002, WP3_012, WP3_013, WP3_014, WP3_015, WP3_031, WP3_047, WP3_065, WP3_068, WP3_086, WP3_097, WP3_099, WP4_006, WP4_137, WP4_150, WP4_151, WP4_161), as is evident from the described functionalities of the current component (or more accurately set of storage solutions used across the various SYNERGY components), the ability to securely store data (here used in a broader scope and not only referring to data assets) is a prerequisite for all SYNERGY operations and as such all WP3 and WP4 requirements can be indirectly linked to the SYNERGY Storage.

4.6.4 Exploited Technology Stack

The SYNERGY Data Storage Services Bundle builds on state-of-the-art data storage and indexing technologies including: PostgreSQL as the platform's relational database, MongoDB as the NoSQL database for datasets, Elasticsearch as the search optimization and indexing engine, MinIO as the data lake, Vault for the secrets and sensitive data storage, Ethereum as the distributed ledger technology, HDFS as the distributed file system.

4.6.5 Updates from Draft Architecture

The scope and functionalities of the Data Storage Services Bundle remain essentially the same as presented in the SYNERGY Deliverable D2.6. The only change in respect to the draft architecture is related to the replacement of Gitlab[23] from MinIO due to the latest functionalities of the Analytics Workbench.

4.7 Data Governance Services Bundle

The Data Governance Services Bundle in the SYNERGY Cloud Infrastructure consists of: (a) the Master Controller, (b) the Data Lineage Service, (c) the CIM Manager, as described in the following paragraphs.

[23] https://about.gitlab.com

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.7.1 Master Controller

4.7.1.1 Component Overview

The Master Controller is instrumental for the orchestration of the different services that are executed in the SYNERGY Core Cloud Platform, the Secure Experimentation Playgrounds and the On-Premise Environments. Depending on the configuration that has been defined by the data provider (for data check-in processes) and the data consumer (for data analytics processes), the Master Controller is responsible for triggering the relevant services and providing them with the exact information that is needed to perform their tasks. To this end, the Master Controller effectively communicates with the Resources Orchestrator to: (a) run the Data Ingestion Service, the Data Mapping & Transformation Service, the Cleaning Service, the Anonymisation Service and the Encryption Engine during the execution of a data check-in process in the SYNERGY Core Cloud Platform and/or the On-Premise Environments; (b) run the Data Manipulation Service and the Analytics Execution Service during the execution of data manipulation and analytics processes in the Secure Experimentation Playgrounds and the On-Premise Environments.

In parallel, the Master Controller acts as a scheduler that automates the execution of the different services related with the data check-in and analytics processes, according to the recurring schedule that has been set by the data asset provider or the data asset consumer, respectively. The Master Controller is also responsible for coordinating the necessary exchange of keys involved in the encryption/decryption services of the Encryption Engine that are executed in the Secure Experimentation Playgrounds and/or the On-Premise Environments.

4.7.1.2 List of Features

The list of features offered by the Master Controller (MC) include:

- **MC_1 – Cloud orchestration of services adapted to the needs of each data check-in job:** The Master Controller discovers and places the necessary services that have been configured for each data check-in job in the form of an end-to-end data collection pipeline (e.g. data ingestion, mapping, cleaning, anonymisation and encryption), according to the preferences of the respective data provider, in an execution cluster in the cloud (as configured by the Resources Orchestrator). The different data check-in jobs are to be executed once or for a specific time period according to a pre-defined schedule based on the user-defined configuration files.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **MC_2 – On-premise orchestration of services adapted to the needs of each data check-in job:** The Master Controller discovers and places the necessary services that have been configured for each data check-in job in the form of an end-to-end data collection pipeline (e.g. data ingestion, mapping, cleaning, anonymisation and encryption), according to the preferences of the respective data provider, in an execution cluster in the On-Premise Environment. The different data check-in jobs are to be executed once or for a specific time period according to a pre-defined schedule based on the user-defined configuration files, and store the respective data payload locally (on-premise) or in the cloud (in the SYNERGY platform). With regard to the necessary resources, the Master Controller instructs the provisioning of a threshold of memory and computation resources for the execution of the different steps in the data check-in job.
- **MC_3 – Cloud orchestration of services adapted to the needs of each data analysis job:** The Master Controller discovers and places the necessary services that have been configured for each data analysis job in the form of an end-to-end data analysis pipeline (e.g. data transfer, decryption, manipulation, and analytics), according to the preferences of the respective data consumer, in an execution cluster in Secure Experimentation Playgrounds (as configured by the Resources Orchestrator). The different data analysis jobs are to be executed once or upon a triggering event or for a specific time period, according to a pre-defined schedule based on the user-defined configuration files.
- **MC_4 – On-premise orchestration of services adapted to the needs of each data analysis job:** The Master Controller discovers and places the necessary services that have been configured for each data analysis job in the form of an end-to-end data analysis pipeline (e.g. data transfer, decryption, manipulation, and analytics), according to the preferences of the respective data consumer, in an execution cluster in the On-Premise Environment. The different data analysis jobs are to be executed once or upon a triggering event or for a specific time period according to a pre-defined schedule based on the user-defined configuration files, and store the respective results locally (on-premise) or in the cloud (in the SYNERGY platform). With regard to the necessary resources, the Master Controller instructs the provisioning of a threshold of memory and computation resources for the execution of the different steps in the data analysis job.
- **MC_5 – Managing the secure transfer of data from/to the Secure Experimentation Playgrounds**: The Master Controller is responsible for coordinating the data transfer from the Cloud Storage or the On-Premise Environment to the Secure Experimentation Playground of an organisation for data assets they own or have legitimately acquired.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **MC_6 – Monitoring and managing the services execution status**: The Master Controller is responsible for efficiently handling the status of the different services among the different SYNERGY components and for managing any error that occurs during their execution, mediating the collection of a number of feedback messages. In case a service has failed, the Master Controller allows for unlocking its configuration in order to be revised in the Data Handling Manager by the respective data asset provider/consumer and schedule its re-execution.

4.7.1.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 20: Mapping of the Master Controller to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
MC_1	WP3_022, WP3_023, WP3_024, WP3_025, WP3_026, WP3_027, WP3_029, WP3_030, WP3_031, WP3_094, WP3_095
MC_2	WP3_077, WP3_094, WP3_095
MC_3	WP4_098, WP4_143, WP4_145, WP4_146, WP4_147, WP4_148, WP4_153, WP4_155
MC_4	WP4_098, WP4_143, WP4_145, WP4_146, WP4_147, WP4_148, WP4_153, WP4_156, WP4_157, WP4_159
MC_5	WP3_078, WP3_079, WP4_098, WP4_158
MC_6	WP3_095

4.7.1.4 Exploited Technology Stack

In order to provide its intended functionalities, the Master Controller will build on different state-of-the-art, open source technologies, namely: (a) in the broader back-end layer, the Nest (NodeJS) web framework for delivering efficient, reliable and scalable server-side applications as required by the SYNERGY Platform, the Kubernetes[24] portable platform for managing containerized workloads and services in different Kubernetes clusters, and the RabbitMQ message broker system that sends the appropriate messages/triggers to the respective services; (b) in the data storage layer, PostgreSQL (as the relational database for the jobs' configuration storage), MinIO (as the data lake for the intermediate data management) and Vault (as the secure database for sensitive and secret parameters).

[24] https://kubernetes.io/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.7.1.5 Updates from Draft Architecture

The scope and functionalities of the Master Controller remain essentially the same as presented in the SYNERGY Deliverable D2.6.

4.7.2 Data Lineage Service

4.7.2.1 Component Overview

The Data Lineage Service provides insights into the way a dataset undergoes changes within the overall SYNERGY platform, keeping track of all actions performed on it that result in alterations in its form, structure, content or metadata. The service offers provenance information in an intuitive way, effectively capturing the data assets' evolution, as well as their interactions, dependencies, and combinations when it comes to generating derivative work, either in the form of new data, data analysis or visualisations and reports.

The component retrieves information generated and stored by various SYNERGY components and services regarding high-level actions performed on datasets and compiles diverse views, each one informative from a different perspective, e.g. monitoring data provenance for transparency and showing derivative assets' evolution to help identify issues with outdated data. The Data Lineage Service is actually a monitoring component that depending on the assets for which it is invoked and the requesting user's access rights, can be leveraged to provide different types of insights.

4.7.2.2 List of Features

The core features of the Data Lineage Service are as follows:

- **DLS_1 – Retrieval and presentation of dataset changes:** One of the most useful views that can be created by the Data Lineage Service shows the evolution of a particular data asset within SYNERGY, in terms of alterations to its contents (as these are perceived through high-level logged actions, e.g. daily updates), structure and metadata. It should be noted that the view does not show exactly how the asset changed, e.g. it cannot know that a certain value was changed, but that a data update action was performed by its provider, resulting in a new version of the original asset.
- **DLS_2 – Retrieval and presentation of connections among data assets:** The Data Lineage Service focuses on connections among data assets discovering any potentially "hidden" links and dependencies and may be used to explore common assets directly or indirectly used to generate two different analytics reports.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.7.2.3 Mapping with SYNERGY Requirements

The only requirement that refers to this service is WP3_091, but its usage is related to broader monitoring and provenance functionalities needed in the overall SYNERGY platform.

4.7.2.4 Exploited Technology Stack

In order to provide its intended functionalities, the Data Lineage Service will build on state-of-the art technologies, namely: (a) in the back-end layer, the NestJS (NodeJS) web framework; (b) in the data storage layer, PostgreSQL (for storing provenance related information).

4.7.2.5 Updates from Draft Architecture

The scope and functionalities of the Data Lineage Service remain essentially the same as presented in the SYNERGY Deliverable D2.6, especially taking into consideration that it was not prioritized for the beta release or release 1.00.

4.7.3 CIM Manager

4.7.3.1 Component Overview

As described in section 4.1.2 and 4.1.4, the data assets ingested in the SYNERGY platform need to conform to a Common Information Model (CIM). The CIM Manager is the component responsible for the functionalities related to the model's lifecycle management, i.e. creation, specification, update and deprecation of the model and its concepts, as well as the accompanying services that facilitate the above actions and ensure the model's integrity and sustainability. It should be stressed that the component's purpose is not limited to the initial definition of the model but allows knowledge representation to evolve following the needs of the electricity data value chain. The model is not a static reference structure and an exhaustive list of concepts and relationships describing thoroughly all potentially relevant data cannot be drafted, thus an evolution mechanism is foreseen to ensure backwards-compatibility to the extent it is possible and migration of all changes in the CIM in all services and components that use it.

The primary user of the CIM Manager is the model administrator, i.e. an administrator of the SYNERGY platform with model editing privileges. Defining and editing the model properties affects the knowledge representation to which all data assets are mapped and are therefore operations not expected to be performed often and may only be performed by the platform's administrator (with domain knowledge), as explained. The significance of the task and the way it affects almost all SYNERGY

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

services also explains why a dedicated component is required in order to ensure that appropriate control mechanisms will be put in place to safeguard the model robustness, consistency and validity as it evolves over time.

4.7.3.2 List of Features

The list of features offered by the component are as follows:

- **CIMM_1 – Definition, update and deprecation of model concepts, fields and hierarchies:** The component offers all CRUD (create-read-update-delete) operations for the concepts of the SYNERGY Common Information Model along with their properties, which are governed by strict rules to ensure the consistency of the overall model. The component's user interface guides the user through the definition of new nodes (concepts, properties) in a way that ensures connection to the existing nodes (i.e. concepts and properties cannot exist disconnected from the rest of the model) and completeness of the provided information (e.g. there cannot be concepts without name or properties without parent concepts). Validation checks are fired when update and/or deletion attempts are made to ensure that the underlying knowledge representation is not broken and that the existing assets that reference the changing concepts are not negatively affected in any way.
- **CIMM_2 – User-driven suggestions of new model concepts or updates in the existing CIM:** As explained, the intended user of the CIM Manager is a privileged user, i.e. a domain expert with model manager rights. However, the electricity data value chain stakeholders that leverage the SYNERGY offerings, may also identify a need to extend or otherwise alter the underlying CIM and considering their understanding of the domain needs, this is a reality that should not be ignored. However, it is not possible to allow changes to be made freely in the CIM and therefore a controlled suggestion mechanism is foreseen by the CIM Manager.
- **CIMM_3 – Informed decision over CIM suggestions:** To allow the CIM Manager user to handle the aforementioned suggestions for extensions or alterations to the SYNERGY CIM, the CIM Manager presents the suggestions to the model manager in an intuitive way, along with results from validation checks to facilitate the decision whether a suggestion should be accepted or rejected. The services to perform the rejection or approval are also included in the CIM Manager's functionalities.
- **CIMM_4 – Navigation to the CIM concepts, properties, hierarchies and relationships:** Intuitive exploration of the CIM is a prerequisite for its maintenance and evolution, especially as it

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

exponentially grows in terms of total number of concepts and the complexity of relationships. The CIM Manager provides a user interface that allows the model manager to quickly search and browse through the model, review its structure and relationships, dive into specific concepts and get detailed information for each of them and in general drive the model's evolution as it is considered appropriate.

4.7.3.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 21: Mapping of the CIM Manager to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
CIMM_1	WP3_001, WP3_004, WP3_036, WP3_067, WP3_096, WP3_104
CIMM_2	WP3_028
CIMM_3	WP3_036, WP3_067, WP3_096, WP3_104
CIMM_4	WP3_011, WP3_096, WP3_104

4.7.3.4 Exploited Technology Stack

In order to provide its intended functionalities, the CIM Manager will build on state-of-the art technologies, namely: (a) in the back-end layer, the NestJS (NodeJS) web framework; (b) in the data storage layer, PostgreSQL and Elasticsearch (for storing the CIM-related information).

4.7.3.5 Updates from Draft Architecture

The scope and functionalities of the CIM Manager remain essentially the same as presented in the SYNERGY Deliverable D2.6.

4.8 Platform Management Services Bundle

The Platform Data Management Services Bundle in the SYNERGY Cloud Infrastructure consists of: (a) the Resources Orchestrator, (b) the Notifications Engine, (c) the Security, Authentication & Authorisation Engine, (d) the Platform Analytics Engine, (e) the API Gateway, as described in the following paragraphs.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.8.1 Resources Orchestrator

4.8.1.1 Component Overview

The Resources Orchestrator aims at provisioning the appropriate resources for the uninterrupted SYNERGY platform operation and the successful execution of all services related to different data check-in and data analysis jobs in a secure and isolated manner, whenever needed. It is responsible for spawning the necessary Secure Experimentation Playgrounds that are considered as sandbox environments dedicated to each organisation for securely decrypting their data, while allocating the appropriate resources in terms of memory, compute and storage capacity, taking into account the needs of the different analysis jobs that are to be executed in an isolated manner.

The Resources Orchestrator works very closely with the Master Controller to ensure dynamic deployment of the services and scaling of the resources allocated per service and job over the available virtualized infrastructure. Whenever applicable, the overall available resources in SYNERGY shall be leveraged and flexibly shared depending on the dynamic configuration for a job's execution to provide high-quality service (under the hood) to the electricity data value chain stakeholders that access the SYNERGY platform either as data asset providers or consumers. In this way, the Resources Orchestrator ensures that data check-in and data analysis jobs will not fail to run due to any resource limitation and at the same time, resources will not be wasted in jobs with low memory or compute requirements.

It needs to be noted that the Resources Orchestrator has full control over the SYNERGY Cloud infrastructure, but can only exercise limited influence over the data check-in and analysis jobs that are to be executed on the On-Premise Environments (in their server or gateway editions) through the Master Controller.

4.8.1.2 List of Features

The list of features offered by the Resources Orchestrator (RO) include:

- **RO_1 – Management of resources in the SYNERGY Cloud infrastructure:** The Resources Orchestrator is responsible for the automated deployment and monitoring of the containerised services and the overall components of the SYNERGY Cloud Platform based on templates to describe the configurations of the necessary cloud computing resources and their dependencies. In this context, the Resources Orchestrator automatically manages load balancing, scaling and failure management (e.g. restarting a container that failed without encountering any data loss) while providing pre-built rollout and rollback mechanisms with almost zero-downtime.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **RO_2 – Dynamic calculation and allocation of resources to the data check-in jobs in the SYNERGY Core Cloud Platform:** The Resources Orchestrator is tasked to dynamically set and allocate resources for the execution of each data check-in job, taking into consideration the associated data volume and the applicable services' needs (i.e. the Data Ingestion Service, the Data Mapping & Transformation Service, the Cleaning Service, the Anonymisation Service and the Encryption Engine).
- **RO_3 – Spawning the SYNERGY Secure Experimentation Playgrounds:** Once an organisation is eligible and approved for access in the SYNERGY platform, the Resources Orchestrator is responsible for the creation of the respective Secure Experimentation Playground and for the isolated management of the respective storage resources. The Resources Orchestrator practically ensures persistence of the unencrypted data of organisations in their own Secure Experimentation Playground that is only accessible to execute analytics jobs for their organisation's members. The allocated storage resources shall scale depending on the data volume that needs to be stored per organisation.
- **RO_4 – Dynamic scaling of resources for running the data analysis jobs in the SYNERGY Secure Experimentation Playgrounds:** The Resources Orchestrator is tasked to dynamically set and allocate memory and compute resources for the execution of each data analysis job in the SYNERGY Secure Experimentation Playground of each organisation, taking into consideration the applicable data manipulation and analytics services' needs (under the Analytics Execution Service).

4.8.1.3 Mapping with SYNERGY Requirements

The only requirement explicitly referring to the orchestrator functionalities is WP3_100, yet the component has a significant role in ensuring smooth operation of the SYNERGY cloud platform and SEPs.

4.8.1.4 Exploited Technology Stack

In order to provide its intended functionalities, the Resources Orchestrator will build on the Kubernetes portable platform for managing containerized workloads and services in different

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Kubernetes clusters, as well as on Docker[25] for containerization of the different services and components of the SYNERGY platform.

4.8.1.5 Updates from Draft Architecture

The scope and functionalities of the Resources Orchestrator remain essentially the same as presented in the SYNERGY Deliverable D2.6.

4.8.2 Notifications Engine

4.8.2.1 Component Overview

The Notifications Engine intends to notify the different SYNERGY platform users (as defined in section 2.4) about certain events that occur in the platform and concern them or their organisation, according to the preferences they have set. The Notification Engine provides real-time information for the progress of ongoing data check-in jobs and/or data analysis jobs at specific milestones without overwhelming the users with irrelevant and redundant communications. In addition, notifications are triggered towards the managers of an organisation regarding the progress of data asset brokerage activities and the overall organisation aspects. Notifications are sent both within the SYNERGY platform and through email in order to ensure that important information (e.g. regarding data assets contracts) is not delayed to be viewed by the respective electricity data value chain stakeholders.

4.8.2.2 List of Features

The list of features offered by the Notifications Engine (NE) include:

- **NE_1 – Evaluate different incoming events in the SYNERGY Platform to find the interested users/organisation:** The Notifications Engine will consume events that come from different SYNERGY components (e.g. Master Controller, Contract Lifecycle Manager) in order to identify the appropriate organisations and/or users that are the intended recipients.
- **NE_2 – Issue notifications for different events in the SYNERGY Platform:** The Notifications Engine will generate the relevant notifications for the progress of the execution of a data check-in job (in terms of successful completion each time it is executed, or details for any failure), the progress of the execution of a data analysis job (in terms of successful completion once it is

[25] https://www.docker.com

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

executed, or details for any failure), the progress of the data asset brokerage processes (e.g. when a draft contract is available, whenever it has been negotiated back'n'forth among the involved stakeholders, when it has been signed or rejected, when it has been settled and becomes active).

- **NE_3 – Deliver messages to users with appropriate content per event in the SYNERGY Platform:** The Notifications Engine will deliver the relevant notifications for the progress of the execution of a data check-in job or a data analysis job, as well as the progress of the data asset brokerage processes through the SYNERGY platform and/or via email.
- **NE_4 – Management of notifications by the users in the SYNERGY Platform:** The users may view the different notifications, act upon them or delete them. In addition, they may set their preferences about what notifications they wish to receive and control what mode of communication to use at every event (e.g. platform notification, email). The manager of an organisation receives additional notifications related to the organisation and configures the data asset brokerage notifications.

4.8.2.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 22: Mapping of the Notifications Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
NE_1	WP3_088, WP3_090, WP3_095
NE_2	WP3_088, WP3_090, WP3_095
NE_3	WP3_088, WP3_090, WP3_095
NE_4	WP3_087, WP3_089

4.8.2.4 Exploited Technology Stack

In order to provide its intended functionalities, the Notifications Engine will leverage the Server-Sent Events (SSE) technology enabling a client to receive automatic updates from a server via HTTP connection and NestJS as the selected NodeJS framework.

4.8.2.5 Updates from Draft Architecture

The scope and functionalities of the Notification Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.8.3 Security, Authentication & Authorisation Engine

4.8.3.1 Component Overview

The Security, Authentication & Authorisation Engine is responsible for different aspects and layers of security across the Core Cloud Platform, the Secure Experimentation Playgrounds and the On-Premise Environments. It acts as the overall SYNERGY platform's identity provider allowing for reliable registration of organisations and users, verifying an organisation and authenticating the platform's users (as an organisation's members). In collaboration with the API Gateway, the Security, Authentication & Authorisation Engine verifies the API keys for external applications (either the SYNERGY energy applications or 3rd-party applications), while generating tokens for the secure data exchange between the internal platform components.

The Security, Authentication & Authorisation Engine also undertakes the verification of the On-Premise Environments' integrity, authenticating their configuration to the SYNERGY Core Cloud Platform. As long as their integrity is ensured, the Security, Authentication & Authorisation Engine oversees the secure exchange of decryption keys for data assets between the On-Premise Environments of different stakeholders (who have a respective active data asset contract), and between the On-Premise Environment and the Secure Experimentation Playground of the same organisation (in case the specific data asset is only available in an encrypted form).

4.8.3.2 List of Features

The list of features offered by the Security, Authentication & Authorisation Engine (SAAE) include:

- **SAAE_1 – Manage identity information for the SYNERGY platform's users and organisations:** As any identity provider, the Security, Authentication & Authorisation Engine will create and manages identity information for the users belonging to organisations that are eligible to gain access to the SYNERGY platform. It is also responsible for the registration of organisations and for the verification of the information they have submitted regarding their legal entity and representative (i.e. manager who will sign the data assets' contracts in the SYNERGY platform).
- **SAAE_2 – Provide fine-grained authentication and authorisation services to the SYNERGY platform's components and services:** The Security, Authentication & Authorisation Engine provides authentication and authorisation services to relying applications within the SYNERGY platform for defining appropriate policies per case and generating and managing the respective tokens. The Security, Authentication & Authorisation Engine offers single sign-on functionality

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

not only to the different components and services of the SYNERGY Platform, but also to the SYNERGY energy apps.

- **SAAE_3 – Establish integrity and trustworthiness across the different layers of the SYNERGY platform:** The Security, Authentication & Authorisation Engine provides different mechanisms to verify the integrity and origin of the data exchanged between the On-Premise Environments and the Core Cloud Platform, as well as between the Core Cloud Platform and the Secure Experimentation Playgrounds.
- **SAAE_4 – Facilitate the secure exchange of keys across the different layers of the SYNERGY platform:** In the case when encrypted data are involved from the Encryption Engine, the Security, Authentication & Authorisation Engine coordinates the secure exchange of decryption keys between the On-Premise Environment and the Secure Experimentation Playground of the data asset provider in order to execute analytics in a way that ensures end-to-end security. It is also responsible to handle the keys revocation once a data asset contract expires.

4.8.3.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 23: Mapping of the Security, Authentication & Authorisation Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
SAAE_1	WP3_013, WP3_061, WP3_062
SAAE_2	WP3_013, WP3_086, WP4_011
SAAE_3	WP3_013, WP3_086, WP3_101, WP3_102
SAAE_4	WP4_011

4.8.3.4 Exploited Technology Stack

In order to provide its intended functionalities, the Security, Authentication & Authorisation Engine will leverage the Keycloak, an open-source identity provider; and NestJS as the selected NodeJS web framework.

4.8.3.5 Updates from Draft Architecture

The scope and functionalities of the Security, Authentication & Authorisation Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, changes have been introduced in SAAE_4 (Facilitate the secure exchange of keys across the different layers of

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

the SYNERGY platform) to streamline its implicit functionalities to the need for end-to-end security and the support for data encryption in transit between the SYNERGY Server On-Premise Environment and the Secure Experimentation Playground of the data provider.

4.8.4 Platform Analytics Engine

4.8.4.1 Component Overview

The Platform Analytics Engine is responsible for generating customised insights regarding the added value of different data assets and the usage of the different SYNERGY services by collecting, aggregating, and visualizing the respective statistics to build the complete picture. Due to the large number and broad scope of the SYNERGY services, these insights may range from statistics on asset searches and purchases through the Data & AI Marketplace, overview of resources utilisation for analytics jobs per organisation or across all Secure Experimentation Playgrounds, overview of failed data check-in jobs, etc.

The Platform Analytics Engine is addressed to: (a) the data asset providers who want to understand the impact and popularity of their data assets as well as get an overview of their data check-in jobs and data analysis jobs, (b) the SYNERGY platform's administrators who need to get a deeper understanding of the use of the SYNERGY services and ensure their smooth operation.

It needs to be noted that the Platform Analytics Engine works in a non-intrusive manner for the users as only anonymised data on the platform usage are collected, in accordance with the GDPR provisions.

4.8.4.2 List of Features

The list of features offered by the Platform Analytics Engine (PAE) include:

- **PAE_1 – Monitoring the SYNERGY Services:** The Platform Analytics Engine allows the platform's administrators to constantly monitor the SYNERGY services and ensure their smooth operation, timely detect issues and proceed to corrective actions or plan maintenance activities, identify popular services (which could indicate a need to scale-up or extend the relevant offerings) or common failure points (indicating a need for improvements).
- **PAE_2 – Understand the data asset's use per organisation:** The Platform Analytics Engine allows the organisations to gain a better view of: (a) the resources and typical issues associated with their data check-in jobs and their data analysis jobs, (b) the data assets that they have uploaded and how they have been perceived by other organisations (in terms of purchases, searches,

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

downloads), (c) the data assets they have acquired and how they have been used in the SYNERGY platform.

4.8.4.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 24: Mapping of the Platform Analytics Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
PAE_1	WP3_063, WP3_094
PAE_2	WP3_094

4.8.4.4 Exploited Technology Stack

In order to provide its intended functionalities, the Platform Analytics Engine will build on the Nest (NodeJS) web framework and PostgreSQL and Elasticsearch as the data storage mechanisms of the overall SYNERGY platform.

4.8.4.5 Updates from Draft Architecture

The scope and functionalities of the Platform Analytics Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6, especially taking into consideration that it was not prioritized for the beta release or release 1.00.

4.8.5 API Gateway

4.8.5.1 Component Overview

The API Gateway is the single-entry point for any application (including the SYNERGY Energy Apps described in Section 6) to retrieve energy data they are legitimate to access from the SYNERGY Platform, as well as analytics results from the respective organisation's Secure Experimentation Playground. The API Gateway practically handles all requests to the SYNERGY Open Application Programming Interfaces (APIs) and either proxies/routes them to the appropriate service or fans them out to multiple services of the SYNERGY platform. To this end, the API gateway acts as a reverse proxy to accept all API calls, fetch and aggregate data from the various services required to respond to the call, and return the appropriate result.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

The API Gateway essentially allows the 3rd-party applications that belong to energy data value chain stakeholders, as well as the SYNERGY Energy Apps, to reduce the number of necessary requests/roundtrips to retrieve the exact data slice they need from: (a) a single dataset for which they have an active contract according to the Contract Lifecycle Manager and the Access Policy Engine, (b) multiple data assets for which they have active contracts and that have been joined and synchronised in the data manipulation part of a data analysis job (configured in the Analytics Workbench and executed with the help of the Data Manipulation Service and the Analytics Execution Service), and (c) analytics results that the specific application is eligible and authorised to retrieve (in collaboration with the Access Policy Engine and the Secure Results Export Service). It needs to be noted that the API Gateway displays a slightly different behaviour with many limitations, depending on whether the data are stored in an encrypted or unencrypted form in the SYNERGY core cloud platform, in the Secure Experimentation Playgrounds or the On-Premise Environments, in order not to compromise the end-to-end security of the private data at any point.

4.8.5.2 List of Features

The list of features offered by the API Gateway (APIG) include:

- **APIG_1 – User-driven configuration of the raw and derivative energy data retrieval:** The API Gateway allows the developers of authorised applications to configure the retrieval of data originating from a single dataset, multiple datasets or analytics results, through the SYNERGY Open APIs by defining which exact fields of the data they need, which fields also represent API request parameters that they intend to use to filter the data and what API method they intend to use. If the intended retrieval is legitimate (based on active contracts' terms), the API Gateway creates a unique identifier, provides instructions for the endpoints to be used and provides a test API functionality to quickly test them to check what results they retrieve.
- **APIG_2 – Direct access to raw and derivative energy data for the authorised applications:** The API Gateway allows the authorised applications to retrieve data through the SYNERGY Open APIs by using the same API method they had selected in the retrieval configuration, providing the unique identifier and the selected retrieval parameters. At any moment an authorised SYNERGY application requests for data, the API Gateway breaks it into multiple requests, routes them to the right places (e.g. the active contracts' terms are enforced with the help of the Contract Lifecycle Manager, the applicable access policies are also resolved in the Access Policy Engine, the data are retrieved from the SYNERGY storage and the analytics results from the Secure Results Export Service of the Secure Experimentation Playgrounds) and produces a response.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **APIG_3 – Management of failures:** Depending on the service that may be unresponsive or unavailable, the API Gateway determines whether a partial response or an error is preferable to be returned to the authorised SYNERGY energy application and any authorised 3rd party application.

- **APIG_4 – Management of API keys**: The API Gateway creates and manages API keys for the different applications (either created in SYNERGY or used in the different electricity data value chain stakeholders) that have retrieval scope and a specific expiry date. To this end, in collaboration with the Security, Authentication & Authorisation Engine, the API Gateway applies authorisation and access control.

- **APIG_5 – Handling of concurrent API calls:** The API Gateway handles the different tasks that are involved in accepting and processing concurrent API calls, including traffic management, CORS (Cross-Origin Resource Sharing) support, pagination, throttling, monitoring, and version management.

4.8.5.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 25: Mapping of the API Gateway to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
APIG_1	WP4_152
APIG_2	WP4_145
APIG_3	N.A.
APIG_4	N.A.
APIG_5	N.A.

4.8.5.4 Exploited Technology Stack

In order to provide its intended functionalities, the API Gateway will build on the Nest (NodeJS) web framework for delivering efficient, reliable and scalable server-side applications as required by the SYNERGY Platform.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

4.8.5.5 Updates from Draft Architecture

The scope and functionalities of the API Gateway remain essentially the same as presented in the SYNERGY Deliverable D2.6.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

5 SYNERGY On-Premise Environments

5.1 Introduction

The SYNERGY On-Premise Environments are stand-alone environments that are intended to be installed either in a server or in gateways that reside on-premise for the electricity data value chain stakeholders. They are designed to follow the Master/Slave paradigm that has been introduced by the GitLab Runner[26] with runners being practically the agents that run the CI/CD (Continuous Integration/Continuous Deployment) jobs that come from GitLab. To this direction, the SYNERGY On-Premise Environments will have a direct communication with the SYNERGY Core Cloud Platform (through the Master Controller) and will be able to run data check-in jobs and data analysis jobs locally using Docker containers (note: Docker is used to package all services and components of the SYNERGY platform). The SYNERGY On-Premise Environments will allow running multiple jobs concurrently and will download the latest docker containers of the services that are to be executed from the SYNERGY Cloud Platform.

The different services that will be available in the SYNERGY Server On-Premise Environment include: (a) the Data Ingestion Service (described in Section 4.1.3), the Mapping & Transformation Service (described in Section 4.1.4), the Cleaning Service (described in Section 4.1.5), the Anonymisation Service (described in Section 4.2.1), the Encryption Engine (described in Section 4.2.2), the Data Manipulation Service (described in Section 4.5.3), the Analytics Execution Service (described in Section 4.5.4), the Secure Results Export Service (described in Section 4.5.5) and the Data Lineage Service (described in Section 4.7.2). The different services that will be available in the SYNERGY Edge On-Premise Environment include: the Data Ingestion Service (described in Section 4.1.3), the Mapping & Transformation Service (described in Section 4.1.4), the Cleaning Service (described in Section 4.1.5), the Data Manipulation Service (described in Section 4.5.3), and the Analytics Execution Service (described in Section 4.5.4). As the functionalities of the above services have been already extensively described and do not display any particular deviations from their “cloud” version, it was not considered necessary to reiterate them in this section. Different validation checks will be performed during the configuration of a data check-in job (in the Data Handling Manager) and of a data analysis job (in the Analytics Workbench) to ensure that there are not any limitations in executing the applicable services

[26] https://gitlab.com/gitlab-org/gitlab-runner

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

in the On-Premise Environments in case the data asset provider has opted for an on-premise execution.

As depicted in Figure 7, the On-Premise Environments are intended to only execute different services on demand by the Master Controller and will only provide a minimal user interface. The Edge On-Premise Environment is essentially a headless version of the Server On-Premise Environment that is optimized for execution in different gateways.

The indicative requirements that apply to the On-Premise Environments are: WP3_014, WP3_015, WP3_044, WP3_047, WP3_059, WP3_069, WP3_075, WP3_077, WP3_098, WP3_102, WP4_098, WP4_121, WP4_125, WP4_157.

In terms of technologies, the SYNERGY On-Premise Environments will be written in Go (based on the open source Gitlab Runner[27]) and will be distributed as single binary without any other requirements to run in Linux, macOS and Windows operating systems for the Server edition and in ARM in addition to x64 or i386 for the Edge edition.

5.2 Additional Components in respect to the SYNERGY Cloud Platform

5.2.1 Wallet Manager

5.2.1.1 Component Overview

The Wallet Manager is a supplementary component offered in the SYNERGY On-Premise Environments in order to allow the platform users (acting as data asset consumers and / or providers) to create a blockchain account (including their blockchain address) and interact with their blockchain account. It practically allows them to sign contracts that are written by the Contract Lifecycle Manager in the SYNERGY distributed ledger at any stage for full traceability of the data brokerage process. The Wallet Manager also lets the platform users read their balance and send funds for settling the data asset contracts they have signed (with the help of the Remuneration Engine).

The Wallet Manager will not be tied to any existing wallet provider. Since it will store a copy of the user's wallet in an encrypted form in the SYNERGY On-Premise Environment, it will leave the selection of a wallet type open to the users. For example, they may opt for physical hardware wallets that are

[27] https://gitlab.com/gitlab-org/gitlab-runner

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

the most secure option as they keep their crypto offline, or for mobile applications, web wallets and desktop applications to manage the funds from a mobile device, via a web browser or from a server, respectively.

5.2.1.2 List of Features

The list of features offered by the Wallet Manager (WM) include:

- **WM_1 – Secure creation and storage of a wallet:** Through the Wallet Manager, the managers of an organisation in the SYNERGY overall platform are able to create and store their organisation's wallet that is a prerequisite for creating and signing data asset contracts. Such a wallet is always stored in an encrypted form and may be stored in the On-Premise Environment for increased security.
- **WM_2 – Import of an existing wallet:** In case an organisation has already at its disposal an Ethereum wallet, the Wallet Manager allows the managers of the specific organisation to import and use it.
- **WM_3 – Management of funds in the wallet:** The Wallet Manager displays the organisation's balances, transaction history and provides the organisation's managers with an easy and reliable way to send/receive funds for data asset contracts.
- **WM_4 – Encryption/Decryption keys storage:** The Wallet Manager will also store the keytray responsible for storing and revoking encryption/decryption keys for the different datasets that were checked in in SYNERGY as encrypted data.
- **WM_5 – Export of the wallet:** In case an organisation needs to utilize its Ethereum wallet in another application (e.g. in the SYNERGY energy apps), the Wallet Manager allows the managers of the specific organisation to export it.

5.2.1.3 Mapping with SYNERGY Requirements

There are no requirements with a direct relation to the Wallet Manager per se even though different requirements related to the Contract Lifecycle Manager and the Remuneration Engine could be associated with it as well.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

5.2.1.4 Exploited Technology Stack

In order to provide its intended functionalities, the Wallet Manager will not create a new decentralised application, but will build on existing wallet providers (such as MetaMask) for checking and transferring funds.

5.2.1.5 Updates from Draft Architecture

The scope and functionalities of the Wallet Manager remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, the only change is the addition of the WM_5 (Export of the wallet) feature to address the need for exporting the wallet of an organization in order to be used by the Flexibility Contracts Manager (described in section 6.2.11) and its Wallet Manager (as mentioned in section 6.2.13).

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6 SYNERGY Energy Apps Portfolio

6.1 Advanced Grid-level Analytics for Optimized Network and Asset Management Services and Applications

6.1.1 Network Performance Assessment Manager

6.1.1.1 Component Overview

The *Network Performance Assessment Engine (NPAE)* is responsible for the simulation of the operation of the network and assesses the state of the network in terms of performance, and power quality metrics. The assessment involves the evaluation of the electricity network steady state aspects using power flow calculations. Congestion issues and operational constraints, such as voltage violations, line/transformer over-loadings are considered in the assessment. The component involves the assessment of various operating conditions, incorporating the analysis of stochastic demand and generation profiles.

6.1.1.2 List of Features

The list of features offered by the Network Performance Assessment Engine (NPAE) include:

- **NPAE_1 - Generation of demand and production scenarios:** The Network Performance Assessment Engine allows the integration of uncertainty aspects into the network analysis. Considering the integration of varying renewable resources and stochastic loads, uncertainty is represented by the generation of a finite set of sample scenarios for the demand and production sources integrated into the network or by the inclusion of their probability distributions.
- **NPAE_2 - Power flow calculation:** The Network Performance Assessment Engine allows the calculation of the state of the system in normal operating conditions. Specific demand and generation snapshots can be studied, as well as seasonal time series profiles. The pf computation includes corrective and preventive control policies related to congestion detection (voltage and thermal violations). The optimization shall include the minimization of generation cost of intact system state. Static power flow and network calculations are performed via the utilization of pandapower package, which combines the data analysis library pandas and the power flow solver PYPOWER. Power flow analysis is performed in timeseries,

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

for all the generated scenarios, as a chronological probabilistic power flow simulation (Monte Carlo).

- **NPAE_3 - Identification of critical operational events:** The estimation of probability of assets overloading (loading above the nameplate rating) is calculated for standard test scenarios (e.g. min-max generation) and for a Monte Carlo simulation of the generated scenarios. In those circumstances, critical nodes and lines of the infrastructure are determined, critical loading situations are identified and critical operational events that can cause DER curtailment, load shedding, reserve deployment can be identified Also, the estimation of probability of assets overloading (loading above the nameplate rating) can be calculated for estimated peak loads when the occurrence of peak load is higher than the assets capacity.
- **NPAE_4 - Detection of congestion problems:** Results coming from the power flow analysis, such as loading of the components (lines) are compared with the maximum current loading of the component to identify and report line overloadings. Similarly, the resulting voltage magnitude from power flow calculation is compared with maximum and minimum accepted voltage levels of the network to identify and report overvoltages/undervoltages.
- **NPAE_5 - KPIs calculation - basic status indicators:** The status indicators give an overview on the service efficiency of the network operation. Indicative KPIs include frequency of congestions, size of congestions, capacity reinforcement needs. Also, maximum hosting capacity, to help identify the range of new DER integration that can be absorbed in existing infrastructure. In addition, this specific feature will indicate flexibility needs, reinforcement needs and the related cost benefit metrics. The Network Performance Assessment Engine shall provide a KPIs visualization feature.
- **NPAE_6 - Data ingestion:** Allow and build interaction with the SYNERGY platform to access any data, measurements, grid models stored in the platform. Workflows have to be constructed, contracts have to be signed with DSOs and retailers, and finally the appropriate data retrieval queries have to be configured to allow data ingestion through the SYNERGY platform.

6.1.1.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 26: Mapping of the Network Performance Assessment Engine to the SYNERGY Requirements

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Feature	Related Requirements (D2.2)
NPAM_1	WP5_055
NPAM_2	N.A.
NPAM_3	N.A.
NPAM_4	N.A.
NPAM_5	WP5_006
NPAM_6	WP5_027, WP5_028, WP5_029, WP5_030, WP5_031, WP5_032, WP5_033, WP5_034, WP5_035, WP5_036, WP5_037, WP5_038, WP5_039, WP5_040, WP5_045, WP5_093, WP5_094, WP5_095, WP5_096, WP5_097, WP5_098, WP5_099, WP5_123, WP5_124, WP5_125, WP5_126, WP5_001, WP5_002

6.1.1.4 Exploited Technology Stack

In order to meet the requirements of the platform, the Network Performance Assessment Engine will be developed on state-of-the-art technologies, namely: (a) in the back-end and the front-end layer, the React framework for serving the components as an efficient, secure and scalable REST API, (b) in the data storage layer, mySQL Server as the relational database for local power grid calculations, (c) in the application layer, Python programming language and its related libraries (pandapower etc.) for fast and accurate power grid calculations.

6.1.1.5 Updates from Draft Architecture

The scope and functionalities of the Network Performance Assessment Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, changes have been introduced in the following features in order to be consistent with the development progress in the respective component:

- *NPAE_2 - Power flow calculation:* Fault scenarios, such as specific transformer or line failures, N-1 criterion have not been eventually included in the analysis, since it has been decided that these features are out of the scope of this engine and similar studies are carried out within the asset management application.

6.1.2 Network Assets Sizing Manager

6.1.2.1 Component Overview

The *Network Assets Sizing Engine (NASE)* is responsible for providing solutions regarding optimal planning and sizing of the network assets. Solutions shall involve optimal capacity for new substations and lines in case further reinforcements are needed, as well as optimal sizing of new demand

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

connections or DERs. The NASE is a complementary component to the NPAE, exploiting results such as reinforcement and flexibility needs to propose operational planning solutions in the mid-term (up to a month ahead) and in the long term (up to a year) for the provision of critical services such as congestion management in medium voltage distribution grids.

6.1.2.2 List of Features

The list of features offered by the Network Assets Sizing Engine (NASE) include:

- **NASE_1 - List operational planning alternatives and horizon:** Depending on the outcomes of the Monte Carlo analysis and the severity of the events, custom decisions can be made for a network, understanding on whether the grid issue is long standing, short term, irregular (happening in a rare occurrence) or periodic. The suggested planning alternatives are line reinforcement, substation reinforcement, capacity coming from flexible resources, new DER hosting capacity. It is suggested that flexible capacity is a short and medium-term solution, whereas reinforcement and new DER installations solutions are long term.
- **NASE_2 - Mid – term flexibility planning:** In case flexibility resources is an available option within network operation, the identified capacity needs (resulting from congestions that have been reported in Monte Carlo power flow calculations) are resolved through the utilization of flexible resources. The flexible capacity needed for the mid-term is identified per network bus (MW of flexible capacity needed per bus).
- **NASE_3 - Long term expansion planning:** The Network Assets Sizing Engine shall facilitate proper capacity sizing for substations, lines as well as new demand connections or DERs to ensure an optimum expansion. Load distribution especially during peak demand periods, forecasted load growth, anticipated penetration of EVs, storage and RES technologies and additional flexibility capacity means enabled by demand side response can be considered. Additionally, capital costs and depreciation rate of grid assets shall also be considered to allow the selection of appropriate conductor's cross section areas to guarantee normal operation and back up capacity needed. The identified capacity needs (resulting from line congestions that have been reported in Monte Carlo power flow calculations) indicate the line reinforcements that are needed.
- **NASE_4 - KPIs calculator – performance indicators:** The identified capacity and reinforcement needs will be associated to appropriate costs (cost of utilizing flexibility, cost of reinforcement) to facilitate the user take a planning decision.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **NASE_5 - Data ingestion:** Allow and build interaction with the SYNERGY platform to access any data, measurements, grid models stored in the platform. Workflows have to be constructed, contracts have to be signed with DSOs and retailers, and finally the appropriate data retrieval queries have to be configured to allow data ingestion through the SYNERGY platform.

6.1.2.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 27: Mapping of the Network Assets Sizing Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
NASM_1	WP5_019
NASM_2	N.A.
NASM_3	WP5_044, WP5_046, WP5_047, WP5_076, WP5_077, WP5_078, WP5_079, WP5_080, WP5_081, WP5_082
NASM_4	WP5_142, WP5_128,
NASM_5	WP5_093, WP5_094, WP5_095, WP5_096, WP5_097, WP5_098, WP5_099, WP5_119, WP5_120, WP5_126

6.1.2.4 Exploited Technology Stack

In order to meet the requirements of the platform, the Network Assets Sizing Engine will be developed on state-of-the-art technologies, namely: (a) in the back-end and the front-end layer, the React framework for serving the components as an efficient, secure and scalable REST API, (b) in the data storage layer, mySQL Server as the relational database for local power grid calculations, (c) in the application layer, Python programming language and its related libraries (pandapower etc.) for fast and accurate power grid calculations.

6.1.2.5 Updates from Draft Architecture

The scope and functionalities of the Network Performance Assessment Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, changes have been introduced in the following features in order to be consistent with the development progress in the respective component:

- NASE_2 - Mid – term flexibility planning and NASE_3 - Long term expansion planning: Fault scenarios, such as specific transformer or line failures, N-1 criterion have not been eventually

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

included in the analysis, since it has been decided that these features are out of the scope of this engine and similar studies are carried out within the asset management application.

6.1.3 Flexibility-based Network Manager

6.1.3.1 Component Overview

The *Flexibility-Based Network Manager* is responsible for assisting network operators to perform their short-term planning activities by utilizing available flexibility. The manager involves the process of data ingestion regarding flexibility availability from flexibility providers, identification of aggregated flexibility in substation and feeders' level, calculation of flexibility deficit periods and calculation of flexibility requirements for network operators to facilitate their short-term planning and needs for congestion management and balancing.

6.1.3.2 List of Features

The list of features offered by the Flexibility-Based Network Manager (FBNM) include:

- **FBNM_1 - Data exchange between flexibility service providers and network operators:** FBNM_1 will ingest flexibility related forecast data (flexibility availability, DER production, DR flexibility) to identify available flexibility on substation and feeder level. The respected data will become available to the application via the SYNERGY platform as an output result from WP6 applications (Flexibility Analytics and Consumer-Centric DR Optimization Application and Portfolio Pattern Forecasting Engine). The aggregated DER available flexibility shall be available to network operators at feeder and substation level to help them plan their control actions in case of system needs.

- **FBNM_2 - Identification of requirements for flexibility:** Static power flow and network calculations are performed via the utilization of pandapower package, which combines the data analysis library pandas and the power flow solver PYPOWER. Power flow analysis is performed in timeseries, utilizing sample forecasts for DER generation and energy demand located in the sample network under study.

  Power flow results are compared against network and operational constraints (line capacity limits and voltage violation limits) and the requirements for flexible capacity are calculated so that resulting network congestions would have been avoided if that flexible upward/downward capacity existed. The requirements for flexibility are formulated as

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

products, in the sense that flexibility is characterized as upward or downward requirement with a respective duration of the service needed. Finally, flexibility requirements are compared against the available flexibility provided by the portfolio managers to identify and calculate flexibility deficit.

6.1.3.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 28: Mapping of the Flexibility-based Network Manager to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
FBNM_1	WP5_048, WP5_093, WP5_094, WP5_095, WP5_096, WP5_097, WP5_098, WP5_099, WP5_148, WP5_145, WP5_144, WP5_143
FBNM_2	N/A
FBNM_3	WP5_093, WP5_094, WP5_095, WP5_096, WP5_097, WP5_098, WP5_099, WP5_147, WP5_146

6.1.3.4 Exploited Technology Stack

In order to meet the requirements of the platform, the Flexibility-based Network Manager will be developed on state-of-the-art technologies, namely: (a) in the back-end and the front-end layer, the React framework for serving the components as an efficient, secure and scalable REST API, (b) in the data storage layer, mySQL Server as the relational database for local power grid calculations, (c) in the application layer, Python programming language and its related libraries (pandapower etc.) for fast and accurate power grid calculations.

6.1.3.5 Updates from Draft Architecture

The scope and functionalities of the Network Performance Assessment Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, changes have been introduced in the following features in order to be consistent with the development progress in the respective component:

- *FBNM_3 Definition of flexibility products*: This feature has been combined with FBNM_2 feature, essentially merging the features into one. Definition of flexibility services and the corresponding standardized flexibility products.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.1.4 DSO-TSO Common Operational Scheduler

6.1.4.1 Component Overview

The *DSO-TSO Common Operational Scheduler* is responsible for providing a common interface for the DSO and the TSO to facilitate common operational scheduling, considering the flexibility requirements for both actors for congestion management needs, balancing needs and other relevant ancillary services. The scheduler also proposes an appropriate coordination scheme clarifying the sequence of actions between the two operators and identifies the definite short term operational scheduling for both actors.

6.1.4.2 List of Features

The list of features offered by the Flexibility-Based Network Manager (FBNM) include:

- **DSO-TSO_1 - DSO-TSO coordination scheme:** The DSO-TSO Common Operational Scheduler also describes the pre-agreed coordination scheme between the DSO and the TSO regarding daily system operational activities, pre-agreed processes with respect to market design and system security, regarding procurement of ancillary services for balancing and congestion management from resources located in the distribution grid. The coordination scheme between the DSO and the TSO has been decided and it is based on the 'local ancillary services market' scheme proposed in the literature where priorities are shifted towards the DSO. That is, the DSO has the priority over flexibility located at the distribution network. The TSO can only contract DER indirectly, after the DSO has aggregated these resources and has transferred them to the TSO ancillary services market.
- **DSO-TSO_2 - Flexibility Conflict Identifier:** Requirements for flexibility are calculated in the functionality FBNM_2 for the next 1 to 72 hours, respectively for the DSO and the TSO. These are the necessary inputs to be utilized within the flexibility conflict identifier to identify not only the hours that the DSO and the TSO have conflicting requests but also the amount of capacity which is located in the DSO network and required by both operators.
- **DSO-TSO_3 -Short-term operational scheduling:** The feature calculates the final short term operational flexibility scheduling for the DSO and the TSO, incorporating the flexibility conflict identifier outputs and implementing the priority strategy algorithm for the next 72 hours.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.1.4.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 29: Mapping of the DSO-TSO Common Operational Scheduler to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
DSO_TSO_1	N/A
DSO_TSO_2	N/A
DSO_TSO_3	WP5_011

6.1.4.4 Exploited Technology Stack

In order to meet the requirements of the platform, the DSO-TSO Common Operational Scheduler will be developed on state-of-the-art technologies, namely: (a) in the back-end and the front-end layer, the React framework for serving the components as an efficient, secure and scalable REST API, (b) in the data storage layer, mySQL Server as the relational database for local power grid calculations, (c) in the application layer, Python programming language and its related libraries (pandapower etc.) for fast and accurate power grid calculations.

6.1.4.5 Updates from Draft Architecture

The scope and functionalities of the Network Performance Assessment Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6.

6.1.5 Enhanced Performance Monitor

6.1.5.1 Component Overview

The Enhanced Performance Monitor (EPM) will be the control centre of the performance of a PV plant. It will consist of several dashboards showing the main production parameters and specific Key Performance Indicators (KPIs) defined and used to easily visualize the performance of the plant. The structure of the EPM will have different layers going from the whole plant to the component level. In this way, the user will have the possibility of choosing different detail levels with a unified view. Four levels are foreseen, namely: Full plant, Generation Unit (Inverter level), String box and panel. The monitoring, reporting and analytical capabilities of the EPM will be similar for all these detail levels.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Regarding the EPM capabilities, they will be deployed using three different dashboards at each detail level, *Operational* for real time and/or short-term monitoring, *Strategic* for long term reporting and *Analytical* for identifying trends and check the historical performance.

6.1.5.2 List of Features

The list of features offered by the Network Component Health Estimator includes:

- **EPM_1 – Collecting information function**: this function will aim to continuous collect and aggregate all information and variables coming from different data sources such as the Plant SCADA database, meters, weather stations and locally deployed sensors.
- **EPM_2 – Operational dashboard:** this dashboard will provide the visualization of the ongoing parameters and KPIs in real time or with a short-time horizon. This will allow to verify and control the performance of the asset in the day-to-day business. It will provide the possibility of displaying information included in the EPM at different detail levels.
- **EPM_3 – Strategic dashboard**: this will be the reporting tool of the EPM, mainly for long-term visualizations. It will allow to summarize the performance over set time frames: past month, quarter, or year, for example. Specific user defined time frames will also be possible. The reporting will be performed at the different detail levels included in the EPM.
- **EPM_4 – Analytical dashboard**: this dashboard will help to identify trends, compare them with multiple variables and create predictions, and targets, which can be implemented in the operation and maintenance of the plant. The analysis will be possible at all the detail levels included in the EPM.

6.1.5.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 30: Mapping of the Enhanced Performance Monitor to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
EPM_1	WP5_008, WP_071, WP5_073, WP5_088, WP5_089, WP_090, WP5_091, WP5_092
EPM_2	WP5_008, WP_071, WP5_073, WP5_088, WP5_089, WP_090, WP5_091, WP5_092
EPM_3	WP5_008, WP_071, WP5_073, WP5_088, WP5_089, WP_090, WP5_091, WP5_092
EPM_4	WP5_008, WP_071, WP5_073, WP5_088, WP5_089, WP_090, WP5_091, WP5_092

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.1.5.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[28], packaging each of the components as an independent docker image.

The resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[29] for backend services, Meteor[30] for front-end/back-end communication and reactive data delivery to clients, and React[31] for reusable and highly modular front-end development.

In order to host component-specific data, MongoDB[32] database is proposed.

6.1.5.5 Updates from Draft Architecture

The scope and functionalities of the Data Handling Manager remain essentially the same as presented in the SYNERGY Deliverable D2.6, and no relevant changes have been updates from the draft architecture.

6.1.6 Fault Occurrence Inspector and Maintenance Optimizer

6.1.6.1 Component Overview

This Fault Occurrence Inspector and Maintenance Optimizer (FOIMO), using the information collected by the EPM, aims to introduce new approaches for predictive maintenance of PV plants. The core of the system will be a set of statistical and Machine Learning (ML) algorithms ingesting data from the plant and looking for trend deviations, misclassifications, uncorrelations, etc. revealing component failures at the plant detail levels defined in the EPM. Control charts and reliability indexes including, when possible, the probability of occurrence will allow the operator to anticipate possible failures in the components. Two dashboards, for facilitating visualization and reporting, will be developed under FOIMO.

[28] https://www.docker.com/
[29] https://nodejs.org/en/
[30] https://www.meteor.com/
[31] https://reactjs.org/
[32] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.1.6.2 List of Features

The list of features offered by the FOIMO include:

- **FOIMO_1 – Statistical and Machine Learning core**: Here, all the data ingested by the EPM will be analysed using a combination of ad-hoc developed algorithms. A combination of methodologies and algorithms will be used in order to optimise the proactive detection of equipment failures in function of the available information and the component. They will go from simple statistical trends to complex machine learning and probabilistic algorithms. The algorithms will allow to define control charts and reliability indexes showing the health status and the probability of fault occurrence (and possibly cascading effects) of the components, thus pointing out to predictive maintenance actions that need to be performed over specific plant assets and components.
- **FOIMO _2 – Operational dashboard:** this dashboard will provide the visualization of the control charts and reliability indexes at the different detail levels defined in the EPM. This will allow to component failures in the day-to-day business.
- **FOIMO _3 – Strategic dashboard**: this will be the reporting tool of the FOIMO, mainly for long-term visualizations. It will allow to summarize the detected failures and to identify long-term trends indicating the evolution of the health of the components. The reports will be obtained at the different detail levels included in the EPM.

6.1.6.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the FOIMO.

Table 31: Mapping of the Fault Occurrence Inspector and Maintenance Optimizer to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
FOIMO_1	WP5_008, WP5_010, WP5_071, WP5_072, WP5_073, WP5_088, WP5_089, WP_090, WP5_091, WP5_092, WP5_123, WP5_124
FOIMO_2	WP5_008, WP5_010, WP5_071, WP5_072, WP5_073, WP5_088, WP5_089, WP_090, WP5_091, WP5_092, WP5_123, WP5_124
FOIMO_3	WP5_008, WP5_010, WP5_071, WP5_072, WP5_073, WP5_088, WP5_089, WP_090, WP5_091, WP5_092, WP5_123, WP5_124

6.1.6.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[33], packaging each of the components as an independent docker image.

The resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[34] for backend services, Meteor[35] for front-end/back-end communication and reactive data delivery to clients, and React[36] for reusable and highly modular front-end development.

In order to host component-specific data, MongoDB[37] database is proposed.

6.1.6.5 Updates from Draft Architecture

The scope and functionalities of the Fault Occurrence Inspector and Maintenance Optimizer remain essentially the same as presented in the SYNERGY Deliverable D2.6, and no relevant changes have been updates from the draft architecture.

6.1.7 Operational Scheduling Optimizer

6.1.7.1 Component Overview

The Operational Scheduling Optimizer (OSO) is responsible for calculating the power production forecasts (for different time horizons) for a determined large PV plant. This power forecasts will consider the actual state of health of the installation, estimated by the App through the Enhanced Performance Monitor and the Fault Occurrence Inspector and Maintenance Optimizer module.

The Operational Scheduling Optimizer will also communicate with the functions Wholesale Market Participation Decision Support System and Portfolio Pattern Forecasting Engine to send this advanced power forecasts to demand aggregators or Retailers, so that they can evaluate the amount of energy to be allocated in the wholesale markets (day-ahead, intraday, ancillary services) or through long term PPA agreements with third parties.

[33] https://www.docker.com/
[34] https://nodejs.org/en/
[35] https://www.meteor.com/
[36] https://reactjs.org/
[37] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Finally, the proposed power production profiles (PPP) will be transferred to the RES Operators, so that they can validate if the PPP are technically feasible. Once the bid is accepted, the RES operator will generate appropriate control signals and communicate them, through the App, to DR aggregators or Retailers.

6.1.7.2 List of Features

- **OSO_1 – Large PV Power forecast engine:** based on the global SOH (State of Health) and availability of the PV Park, this function will provide more accurate active/ reactive power forecasts, for different timelines (hourly, daily, yearly).
- **OSO_2 - Validation of power schedules function:** this function will help RES operators to validate if the Power Production Profiles proposed by DR aggregators or Retailers are technically feasible. In case the plant cannot assume such production profiles, this function will provide a feasible alternative.
- **OSO_3 – Data exchanger with Wholesale Market Participation Decision Support System App:** for the optimized energy bidding and output sharing in wholesale markets, OSO_3 will share the required inputs (power forecast) to identify portions of the energy forecasted that could be included in the Day ahead market and in the other markets using short term forecasts.
- **OSO_4 – Data exchanger with Portfolio Pattern Forecasting Engine App:** for the Optimized energy bidding and output sharing in PPAs, OSO_4 will share the required inputs (power forecast) to identify portions of the energy forecasted that could be included in long-term PPAs with energy retailers.

6.1.7.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 32: Mapping of the Operational Scheduling Optimizer to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
OSO_1	WP5_008, WP5_010, WP5_071, WP5_072, WP5_073, WP5_088, WP5_089, WP_090, WP5_091, WP5_092, WP5_123, WP5_124
OSO_2	WP5_008, WP5_010, WP5_071, WP5_072, WP5_073, WP5_088, WP5_089, WP_090, WP5_091, WP5_092, WP5_123, WP5_124
OSO_3	WP5_008, WP5_010, WP_011, WP5_026, WP5_053, WP5_071, WP5_072, WP5_073, WP5_088, WP5_089, WP_090, WP5_091, WP5_092, WP5_123, WP5_124,

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Feature	Related Requirements (D2.2)
OSO_4	WP5_008, WP5_010, WP5_052, WP5_067, WP5_071, WP5_072, WP5_073, WP5_083, WP5_088, WP5_089, WP_090, WP5_091, WP5_092, WP5_121, WP5_123, WP5_124, WP5_131

6.1.7.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[38], packaging each of the components as an independent docker image.

The resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[39] for backend services, Meteor[40] for front-end/back-end communication and reactive data delivery to clients, and React[41] for reusable and highly modular front-end development.

In order to host component-specific data, MongoDB[42] database is proposed.

6.1.7.5 Updates from Draft Architecture

The scope and functionalities of the Operational Scheduling Optimizer remain essentially the same as presented in the SYNERGY Deliverable D2.6, and no relevant changes have been updates from the draft architecture.

6.1.8 Network Asset Health Estimator

6.1.8.1 Component Overview

The Network Asset Health Estimator component will make use of the SYNERGY Platform to analyse a variety of network asset data sources towards the estimation of metrics that provide indications about the status of the different assets subject to maintenance, thus complementing Maintenance

[38] https://www.docker.com/
[39] https://nodejs.org/en/
[40] https://www.meteor.com/
[41] https://reactjs.org/
[42] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Management Systems with valuable information that will allow network operators to schedule the maintenance actions optimally.

6.1.8.2 List of Features

The list of features offered by the Network Component Health Estimator includes:

- **NCHE_1 – Picture analysis**: available network asset pictures managed by the network operator, including infrared pictures, will be analysed by Computer Vision algorithms in order to extract relevant metrics, such as temperature, person presence or smoke presence. These metrics will be used to enrich the available information of the assets under maintenance.
- **NCHE_2 – Metrics visualization**: for all available metrics, this component will provide appropriate graph visualizations to support the maintenance operator work.
- **NCHE_3 – Asset health score calculation**: by taking into account the specification of each type of asset under maintenance and the available data, a high-level health score will be provided, allowing the maintenance operator to rank the assets accordingly and use this information in the maintenance scheduling.
- **NCHE_4 – Presentation of results by means of VR/AR systems**: Virtual and Augmented Reality visualizations of the information managed by this component will be explored in order to provide maintenance operators with novel means of accessing information, which will enhance their user experience.

6.1.8.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 33: Mapping of the Network Asset Health Estimator to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
NCHE_1	WP5_137, WP5_138, WP5_139
NCHE_2	WP5_059, WP5_141
NCHE_3	WP5_060, WP5_066
NCHE_4	WP5_149

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.1.8.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[43], packaging each of the components as an independent docker image.

The resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications), complemented by a VR/AR application for maintenance operators.

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[44] for backend services, Meteor[45] for front-end/back-end communication and reactive data delivery to clients, and React[46] for reusable and highly modular front-end development.

In order to host component-specific data, MongoDB[47] database is proposed.

6.1.8.5 Updates from Draft Architecture

The scope and functionalities of this component remain the same as presented in the SYNERGY Deliverable D2.6.

6.1.9 Network Predictive Maintenance Manager

6.1.9.1 Component Overview

The Network Predictive Maintenance Manager will analyse a variety of network component data sources, including those metrics provided by the Network Component Health Estimator, with the objective of complementing Maintenance Management Systems with estimations of the probability to failure of single network assets in different time frames. This information will support network maintenance operators in the optimal scheduling of the maintenance actions.

6.1.9.2 List of Features

The list of features offered by the Network Predictive Maintenance Manager includes:

[43] https://www.docker.com/
[44] https://nodejs.org/en/
[45] https://www.meteor.com/
[46] https://reactjs.org/
[47] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **NPM_1 – Registration of assets subject to predictive maintenance**: the component will offer an UI that will allow maintenance operators to register the assets under maintenance, together with the required technical details and related datasets.
- **NPM_2 – Probability scores of occurrence of failures**: bringing together all data linked to a particular network asset, the component will estimate the probability of failure within different time ranges.
- **NPM_3 – Prioritization of maintenance tasks**: the different metrics available for each network asset will be brought together with the technical maintenance specifications to assign relative maintenance priorities to each of the network assets under maintenance.
- **NPM_4 – Track of maintenance actions**: in order to support the maintenance operations follow-up, the component will allow maintenance operator to register all necessary information to allow maintenance actions follow-up and the calculation of maintenance-related KPIs.

6.1.9.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 34: Mapping of the Network Predictive Maintenance Manager to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
NPM_1	WP5_134
NPM_2	WP5_063
NPM_3	WP5_062
NPM_4	WP5_065

6.1.9.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[48], packaging each of the components as an independent docker image. The resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

[48] https://www.docker.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[49] for backend services, Meteor[50] for front-end/back-end communication and reactive data delivery to clients, and React[51] for reusable and highly modular front-end development. In order to host component-specific data, MongoDB[52] database is proposed.

6.1.9.5 Updates from Draft Architecture

The scope and functionalities of this component remain the same as presented in the SYNERGY Deliverable D2.6.

6.2 Portfolio-level Analytics for Energy-as-a-Service (EaaS) Applications for Electricity Retailers and Aggregators

6.2.1 Portfolio Pattern Forecasting Engine

6.2.1.1 Component Overview

The Portfolio Pattern Forecasting Engine is the component in charge of interfacing the SYNERGY platform in order to obtain, analyse and present to the retailer with a clear view of both the detected behaviour patterns of its portfolio of customers (in energy demand and surplus terms), as well as the expected evolution of these patterns in different future time scopes.

6.2.1.2 List of Features

The list of features offered by the Portfolio Pattern Forecasting Engine includes:

- **PPFE_1 – Next 6 hours portfolio demand forecast**: detailed forecast with 1 hour resolution, useful for decision support on participation of intra-day energy market (handling deviations).
- **PPFE_2 – Day-ahead portfolio demand forecast**: detailed forecast with 1 hour resolution, useful for decision support on participation of day-ahead energy market.

[49] https://nodejs.org/en/
[50] https://www.meteor.com/
[51] https://reactjs.org/
[52] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **PPFE_3 – Long-term portfolio demand forecast**: detailed forecast with 1 month resolution forecasting 1 year ahead, useful for decision support on business planning and establishment of long-term energy delivery contracts with production partners.
- **PPFE_4 – Next 6 hours portfolio energy exported forecast**: detailed forecast with 1 hour resolution, useful for decision support on participation of intra-day energy market (handling deviations).
- **PPFE_5 – Day-ahead portfolio energy exported forecast**: detailed forecast with 1 hour resolution, useful for decision support on participation of day-ahead energy market.
- **PPFE_6 – Long-term portfolio energy exported forecast**: detailed forecast with 1 month resolution forecasting 1 ahead, useful for decision support on business planning and establishment of long-term energy delivery contracts with production partners.

6.2.1.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 35: Mapping of the Portfolio Pattern Forecasting Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
PPFE_1	WP6_010
PPFE_2	WP6_010
PPFE_3	WP6_129
PPFE_4	WP6_076
PPFE_5	WP6_075
PPFE_6	WP6_078

6.2.1.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[53], packaging each of the components as an independent docker image.

[53] https://www.docker.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[54] for backend services, Meteor[55] for front-end/back-end communication and reactive data delivery to clients, and React[56] for reusable and highly modular front-end development.

In order to host component-specific data, MongoDB[57] database is proposed.

6.2.1.5 Updates from Draft Architecture

The scope and functionalities of this component remain the same as presented in the SYNERGY Deliverable D2.6.

6.2.2 Customer Segmentation Engine

6.2.2.1 Component Overview

The Customer Segmentation Engine is the component in charge of retrieving, analysing and representing all analytics that help retailers and aggregators to clearly understand the behaviour of its customer portfolio with regards to different combinations of socioeconomic characteristics, with the ultimate objective of making it possible to offer more customized and attractive services to those.

In addition, this component will provide the retailer with all necessary insights to enable the appropriate implementation of dynamic tariff schemes that will provide novel mechanisms to handle intraday market deviations.

6.2.2.2 List of Features

The list of features offered by the Customer Segmentation Engine includes:

- **CSE_1 – Elasticity profiles**: both individual and cluster-based elasticity profiles will be calculated as a basic information block necessary to implement appropriate dynamic tariff schemes

[54] https://nodejs.org/en/

[55] https://www.meteor.com/

[56] https://reactjs.org/

[57] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **CSE_2 – Socioeconomic profiling**: helps retailers to appropriately understand the composition of its customer portfolio, identify similarities among them and classifies them according to different combinations of socioeconomic characteristics in order to enable more beneficial and effectively targeted actions. In addition, this profiling will also enable the provision of detailed information to the customers.
- **CSE_3 – Demand profiling and classification**: in order to facilitate the understanding, prevision and management of the energy patterns of the customers, those will be automatically classified under groups of similar members (according to their demand profiles) which will be more easily handled and addressed by the retailer
- **CSE_4 – Energy exported profiling and classification**: similar to CSE_3, based on the exported energy (production surplus) profiles
- **CSE_5 – Elasticity profiling**: all customers exposed to dynamic tariff schemes, will be profiled and grouped to similar peers according to their elasticity properties (reaction to price-based signals)

6.2.2.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 36: Mapping of the Customer Segmentation Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
CSE_1	WP6_111
CSE_2	WP6_024
CSE_3	WP6_092, WP6_094
CSE_4	WP6_093, WP6_095
CSE_5	WP6_114

6.2.2.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[58], packaging each of the components as an independent docker image.

Resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[59] for backend services, Meteor[60] for front-end/back-end communication and reactive data delivery to clients, and React[61] for reusable and highly modular front-end development.

In order to host component-specific data, MongoDB[62] database is proposed.

6.2.2.5 Updates from Draft Architecture

The scope and functionalities of this section remain the same as presented in the SYNERGY Deliverable D2.6.

6.2.3 Wholesale Market Participation Decision Support System

6.2.3.1 Component Overview

The Wholesale Market Participation Decision Support System is the component entailing all necessary analytics to help retailers to clearly understand and predict the energy requirements of its customer portfolio and the possible deviations, as the basis to address the different time frames of the wholesale market, with the final objective of addressing longer-term markets with high accuracy, thus reducing to the extent possible the participation in shorter-term markets, which usually imply higher costs.

Moreover, this module will explore the potential of novel dynamic tariff schemes as a mechanism to address wholesale market deviations by means of implicit DSM strategies.

6.2.3.2 List of Features

The list of features offered by the Wholesale Market Participation Decision Support System includes:

[58] https://www.docker.com/

[59] https://nodejs.org/en/

[60] https://www.meteor.com/

[61] https://reactjs.org/

[62] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **WMPDSS_1 – Wholesale market deviations forecast**: by using all available real-time data, both owned by the retailer and by the DSO/TSO, this module will provide the estimation of the actual aggregated demand and production of the customer portfolio for the short-term future (next 6 hours).
- **WMPDSS_2 – Adjustments for intra-day market participation**: the information provided by WMPDSS_1 will be compared with the data from the wholesale market participation in order to provide near-real-time accurate deviation metrics which need to be minimised by the retailer in order to avoid economic penalties.
- **WMPDSS_3 – Dynamic tariff calculation**: will enable a novel mechanism to address wholesale market deviations, which consists on triggering implicit DSM actions by dynamically changing the customer tariff prices for the next hours in an hourly basis. Elasticity profiles of customer segments will be analysed and used to calculate the optimum prices that will contribute to minimise the detected deviations.
- **WMPDSS_4 – Communication of dynamic tariffs to customers**: due to the live nature of the dynamic tariffs, relevant mechanisms need to be developed to make those available to end-customers in real-time, including Machine-To-Machine and Machine-To-Human channels.
- **WMPDSS_5 – Simulation of dynamic tariffs**: retailer operators will be able to manually test the effect of different dynamic tariff pricing schemes on particular time frames via simulation capabilities

6.2.3.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 37: Mapping of the Wholesale Market Participation Decision Support System to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
WMPDSS_1	WP6_012, WP6_081
WMPDSS_2	WP6_013
WMPDSS_3	WP6_019
WMPDSS_4	WP6_020, WP6_021, WP6_022, WP6_083
WMPDSS_5	WP6_114

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.2.3.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[63], packaging each of the components as an independent docker image.

Resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[64] for backend services, Meteor[65] for front-end/back-end communication and reactive data delivery to clients, and React[66] for reusable and highly modular front-end development.

In order to host component-specific data, MongoDB[67] database is proposed.

6.2.3.5 Updates from Draft Architecture

The scope and functionalities of this section remain the same as presented in the SYNERGY Deliverable D2.6.

6.2.4 Personalized Energy Analytics Engine

6.2.4.1 Component Overview

The Personalized Energy Analytics Engine is the component in charge of interfacing the SYNERGY platform in order to obtain, analyse and present to the retailer's end customers with clear and easily understandable insights of their energy-related behaviour, including useful features such as peer-comparison and energy and cost-saving tips.

6.2.4.2 List of Features

The list of features offered by the Personalized Energy Analytics Engine includes:

[63] https://www.docker.com/

[64] https://nodejs.org/en/

[65] https://www.meteor.com/

[66] https://reactjs.org/

[67] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **PEAE_1 – Active power histograms per supply point**: in order to help customers to understand their active power peaks, when and with which frequency they occur. This information is valuable to identify home assets that consume excessive power, and to adjust the contracted capacity appropriately
- **PEAE_2 – Energy demand time distribution**: graphical representation of the demand over a selectable time-frame axis
- **PEAE_3 – Self-consumption**: for those customers with metered production, calculation of self-consumption ratios (percentage of production that is locally consumed) over a selected time period
- **PEAE_4 – Self-sufficiency**: for those customers with metered production, calculation of self-sufficiency ratios (percentage of demand covered by local production) over a selected time period
- **PEAE_5 – Comparison of demand between relevant time periods**: overlaid graphical representation of the demand of two selected time periods, accompanied of relevant indicators, which will allow customers to observe differences over relevant time periods.
- **PEAE_6 – Customer peer comparison**: provides customer KPI comparison against baseline (clusters as provided by the Customer Segmentation Engine component). KPIs will include demand, production, CO2 equivalent emissions, energy cost, self-consumption and self-sufficiency
- **PEAE_7 – Demand and production forecasts:** personalized forecasts according to historical data and contextual parameters (type of day and weather)
- **PEAE_8 – Deviation of demand and production against forecasts/baseline:** alert mechanism to trigger customer attention whenever the energy usage differs significantly from the expected behaviour
- **PEAE_9 – Provide alternative options for energy behaviour changes:** based on the different clusters of customers provided by the Customer Segmentation Engine component, provides hints that will allow the customer to move towards a customer group with better energy usage KPIs
- **PEAE_10 – Target-monitoring:** alert mechanism of deviations against consumption target set by user

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.2.4.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 38: Mapping of the Personalized Energy Analytics Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
PEAE_1	WP6_065
PEAE_2	WP6_066
PEAE_3	WP6_067
PEAE_4	WP6_068
PEAE_5	N.A.
PEAE_6	WP6_069
PEAE_7	N.A.
PEAE_8	N.A.
PEAE_9	WP6_026, WP6_070
PEAE_10	N.A.

6.2.4.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[68], packaging each of the components as an independent docker image.

Resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[69] for backend services, Meteor[70] for front-end/back-end communication and reactive data delivery to clients, and React[71] for reusable and highly modular front-end development.

[68] https://www.docker.com/

[69] https://nodejs.org/en/

[70] https://www.meteor.com/

[71] https://reactjs.org/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

In order to host component-specific data, MongoDB[72] database is proposed.

6.2.4.5 Updates from Draft Architecture

The scope and functionalities of this section remain the same as presented in the SYNERGY Deliverable D2.6.

6.2.5 Smart Home Integration Engine

6.2.5.1 Component Overview

The Smart Home Integration Engine is the component implements all necessary functions to enable Intelligent control and automation features of Smart Home assets, towards facilitating the establishment of sustainable and long-lasting energy behaviors, while reducing intrusiveness for energy consumers.

6.2.5.2 List of Features

The list of features offered by the Smart Home Integration Engine includes:

- **SHIE_1 – Comfort profiles**: all available context data per customer (smart assets setpoints, temperature, humidity, illuminance…) will be used to build individual comfort profile models, which will be further on considered on the automatic action calculations
- **SHIE_2 – Automatic control schedules for HVAC devices**: based on the comfort-centric flexibility profiles, the component will provide the optimal schedule for the considered Heating, Ventilation and Air Conditioning assets towards maximizing energy- and cost-savings. End-users will be kept in the loop, by offering a selection of different operation modes focused on human-centric energy cost minimization, comfort maximization or manual scheduling.
- **SHIE_3 – Automatic control schedules for lighting devices**: based on the comfort-centric flexibility profiles, the component will provide the optimal schedule for the considered lighting assets towards maximizing energy- and cost-savings. Similarly to SHIE_2, end-users will be kept in the loop, by offering a selection of different operation modes.
- **SHIE_4 – Automatic control schedules for DHW devices**: based on the comfort-centric flexibility profiles, the component will provide the optimal schedule for the considered Domestic Hot

[72] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Water assets towards maximizing energy- and cost-savings. Similarly to SHIE_2, end-users will be kept in the loop, by offering a selection of different operation modes.

- **SHIE_5 – Control signals communication to customers**: in order to enable this features in those cases when assets are not telecontrollable, signals will be communicated to asset owners by means of appropriate Machine-To-Human channels
- **SHIE_6 – Control signals communication to flexibility assets**: whenever the assets allow telecontrol, signals will be communicated to those assets by means of appropriate Machine-To-Machine channels
- **SHIE_7 – Detect presence and absence**: based on consumption patterns and available IoT data, presence and absence of persons in the monitored space will be inferred and used in the relevant comfort profile models

6.2.5.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 39: Mapping of the Smart Home Integration Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
SHIE_1	N.A.
SHIE_2	WP6_062, WP6_071, WP6_073, WP6_101
SHIE_3	WP6_062, WP6_072, WP6_073, WP6_102
SHIE_4	WP6_062, WP6_103
SHIE_5	WP6_063
SHIE_6	WP6_064
SHIE_7	N.A.

6.2.5.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[73], packaging each of the components as an independent docker image.

Resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[74] for backend services, Meteor[75] for front-end/back-end communication and reactive data delivery to clients, and React[76] for reusable and highly modular front-end development.

In order to host component-specific data, MongoDB[77] database is proposed.

6.2.5.5 Updates from Draft Architecture

The scope and functionalities of this section remain the same as presented in the SYNERGY Deliverable D2.6.

6.2.6 Market Positioning Optimization Tool

6.2.6.1 Component Overview

The Market Positioning Optimization Tool is a component with the objective of providing the RES plant operator/aggregator a comprehensive analysis of its situation in the wholesale markets and key suggestions to improve and optimize different KPIs such as revenue obtained from operation. The component is in charge of displaying different indicators based on data coming from European electricity platforms, exchanges and RES plants generation forecast provided by other SYNERGY applications. The aim is to let the facility operator know which hypothetical strategies will result in an optimization of the final revenue obtained by allocating its production in the different wholesale markets (day-ahead, intra-day, continuous) and in flexibility markets.

[73] https://www.docker.com/

[74] https://nodejs.org/en/

[75] https://www.meteor.com/

[76] https://reactjs.org/

[77] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.2.6.2 List of Features

The list of features offered by the Market Positioning Optimization Tool includes:

- **MPOT_1 - Day-ahead price forecast**: time series prediction of the hourly prices of the day-ahead market based on regression methods displayed prior to the publication of results on D-1
- **MPOT_2 - Intra-day price forecast**: time series prediction of the hourly prices of the multiple sessions of intra-day market based on regression or classification methods displayed prior to the publication of results on D-1 and before contract session close.
- **MPOT_3 - Continuous intra-day price forecast**: time series prediction of the hourly prices of the intra-day continuous market based on regression or classification methods displayed prior publication of results on D-1 and before contract session close.
- **MPOT_4 - Indicators summary:** detailed set of indicators representing reliability of the above predictions and suggested guidelines for optimal operation.

6.2.6.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 40: Mapping of the Market Positioning Optimization Tool to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
MPOT_1	WP6_109, WP6_083, WP6_109
MPOT_2	WP6_109, WP6_083, WP6_109
MPOT_3	WP6_109, WP6_083, WP6_109
MPOT_4	WP6_109, WP6_083, WP6_109

6.2.6.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[78], packaging each of the components as an independent docker image.

[78] https://www.docker.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[79] for backend services, Meteor[80] for front-end/back-end communication and reactive data delivery to clients, and React[81] for reusable and highly modular front-end development.

In order to host component-specific data, MongoDB[82] database is proposed.

6.2.6.5 Updates from Draft Architecture

The specific component was introduced in this final version of the architecture as an outcome of the work performed in WP6. The complete details of the components are described in the previous paragraphs.

6.2.7 PPA Optimization Tool

6.2.7.1 Component Overview

The PPA Optimization Tool will help demand aggregators to estimate the optimal sizing of the generation power plant to reduce the price that an energy community would pay once the IRR of the project has been set. The inputs of the tool are historical data of consumptions and meteorology, so the forecasts of generation and consumptions would be done. The temporal horizon of those forecasts will be established in a range of 10 to 15 years. From that data, a series of synthetic generation and consumption series will be created and further used to, depending on energetic balances and economic parameters of the PV power plant, establish PPA prices and associated savings according to the target IRR. The optimal sizing of the PV power plant will be the one that provides the minimum PPA price for established percentile (e.g. P50).

6.2.7.2 List of Features

The list of features offered by the Smart Home Integration Engine includes:

[79] https://nodejs.org/en/

[80] https://www.meteor.com/

[81] https://reactjs.org/

[82] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **POT_1 – Long-term day-ahead price forecast**: time series prediction of the hourly prices of the day-ahead market for the next 15 years.
- **POT_2 – PPA Optimization**: based on the defined scenario characteristics, the optimum characteristics for a PPA are computed, including size of the PV Power Plant, PPA price and relevant expected effects for the portfolio (self-consumption rates, cost savings…)

6.2.7.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 41: Mapping of the PPA Optimization Tool to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
POT_1	WP6_109, WP6_083, WP6_109
POT_2	N.A

6.2.7.4 Exploited Technology Stack

The design of this component will follow a micro-services approach, where different atomic features will be developed in will defined software modules offering an API for enabling interactions with external components or submodules. The proposed technology to facilitate the micro-services-based deployment is Docker[83], packaging each of the components as an independent docker image.

Resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

With regards to the software libraries and platforms foreseen for the development, those include NodeJS[84] for backend services, Meteor[85] for front-end/back-end communication and reactive data delivery to clients, and React[86] for reusable and highly modular front-end development.

In order to host component-specific data, MongoDB[87] database is proposed.

[83] https://www.docker.com/

[84] https://nodejs.org/en/

[85] https://www.meteor.com/

[86] https://reactjs.org/

[87] https://www.mongodb.com/

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.2.7.5 Updates from Draft Architecture

The specific component was introduced in this final version of the architecture as an outcome of the work performed in WP6. The complete details of the components are described in the previous paragraphs.

6.2.8 Aggregator Portfolio Manager

6.2.8.1 Component Overview

The Aggregator Portfolio Manager aims to facilitate aggregators to continuously have a clear picture and management of their portfolio's performance, in order to be able to deliver the required flexibility to network operators considering the flexibility characteristics of their customers (e.g. duration of flexibility activation, flexibility amount that can be provided, contractual restrictions, location of events/ flexibility, timing of events/ flexibility availability, etc).

In this context, the Aggregator Portfolio Manager shall incorporate the results of fine-grained analytics (defined in the Analytics Workbench as described in section 4.5.1) and enable spatio-temporal segmentation and characteristics-based classification of flexibility assets as well as definition of appropriate clusters of flexibility assets to address evolving requirements of network operators (TSOs and DSOs) for the provision of ancillary services.

6.2.8.2 List of Features

The list of features offered by the Aggregator Portfolio Manager (APM) include:

- **APM_1 – Multi dimensional flexible source classification based on operational and business parameters:** The Aggregator Portfolio Manager should facilitate aggregators to perform analysis of their portfolio of flexibility sources taking into account their operational, i.e., nominal capacity, flexibility amount that can be provided as well as business (contractual related) characteristics.
- **APM_2 – Multi dimensional flexible source classification based on spatial parameters:** The Aggregator Portfolio Manager should facilitate aggregators to perform location-based analysis of their flexibility sources on the way to provide location-based energy services to the network operators.
- **APM_3 – Multi dimensional flexible source classification based on temporal parameters:** The Aggregator Portfolio Manager should facilitate aggregators to perform time-based analysis of

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

their flexibility sources on the way to optimally balance their flexibility bids in the different flexibility markets and schemas.

- **APM_4 – User-defined classification of flexibility sources:** The end user (aggregator) of the Aggregator Portfolio Manager should be able to set his/her manual settings based on which dynamic clustering of flexibility sources is applied.

The end user (aggregator) of the Aggregator Portfolio Manager should be able to (a) store the results of the classification analysis as presented above, (b) get insights about the performance of each cluster and (c) manually modify/edit the list of flexible sources/ assets that are part of each cluster.

6.2.8.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 42: Mapping of the Aggregator Portfolio Manager to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
APM _1	WP6_131, WP6_133 (WP6_050), WP6_134 (WP6_052), WP6_135, WP6_136, WP6_137, WP6_138, WP6_139, WP6_140 (WP6_051), WP6_141 (WP6_126), WP6_142 (WP6_125), WP6_143, WP6_144 (WP6_055), WP6_145, WP6_146, WP6_155, WP6_156, WP6_157, WP6_158, WP6_159
APM _2	WP6_147 (WP6_127), WP6_148
APM _3	WP6_149, WP6_150, WP6_151, WP6_152, WP6_154
APM _4	WP6_160, WP6_161

6.2.8.4 Exploited Technology Stack

In order to provide its intended functionalities, the Aggregator Portfolio Manager component builds on state-of-the-art technologies, namely: (a) in the back-end layer, the Nest (NodeJS) web framework as a mature framework for delivering efficient, reliable and scalable server-side applications, (b) in the front-end layer, VueJS for custom SYNERGY front-end design and analytics results visualizations; (c) in the data storage layer, depending on the applicable data assets' licenses (that determine whether local storage will be eventually allowed), PostgreSQL (as the relational database for the operational data of the APM, e.g. management of the energy and flexibility related properties), Elasticsearch (as the indexing engine to facilitate efficient clustering and segmentation of data).

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.2.8.5 Updates from Draft Architecture

The scope and functionalities of the Aggregator Portfolio Manager remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, changes have been introduced in the following features in order to be consistent with the development progress in the respective component:

- APM_1 – Multi dimensional flexible source classification based on operational parameters: The aggregators are able to perform analysis of their portfolio of flexibility sources taking into account not only flexible asset operational characteristics but also business-related characteristics as defined at the contract established with the flexible asset owners through the flexibility marketplace application.
- Management of the groups of flexibility sources: The end user (aggregator) of the Aggregator Portfolio Manager should be able to (a) store the results of the classification analysis as presented above, (b) get direct access to the performance of each cluster and (c) manually modify/edit the list of flexible sources/ assets that are part of each cluster.

6.2.9 Virtual Power Plant (VPP) Configuration Engine

6.2.9.1 Component Overview

The role of the VPP Configuration Engine is to provide advanced Decision Support System (DSS) functionalities towards creating ad-hoc dynamic Virtual Power Plants (VPP) considering the type of service requested by the network operator and the flexibility characteristics of its underlying portfolio. In addition, the VPP Configuration Engine continuously monitors the performance of the VPP and re-configures on the fly the initial VPP once a flexibility source has unexpectedly withdrawn from the VPP during the evolution of a flexibility provision event.

6.2.9.2 List of Features

The list of features offered by the VPP Configuration Engine include:

- **VPP_1 – Provision of ancillary services to TSOs/DSOs:** The VPP Configuration Engine to facilitate aggregators to optimally bid their flexibility potential for different ancillary services, addressing the needs of network operators.
- **VPP_2 – Flexible source clustering to VPPs:** The VPP Configuration Engine to incorporate fine grained optimization techniques to match the available flexible sources of the aggregators with

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

the demand for ancillary services as triggered by the network operator. The performance of this optimization process is also visible to the business stakeholder of the tool (Aggregator).

- **VPP_3 – Automatic reconfiguration of VPPs:** On the way to perform control of flexibility sources in a robust and reliable manner that ensures no penalization on the aggregator side, the VPP Configuration Engine to incorporate (near) real time optimization techniques to support on-the - fly reconfiguration of their portfolio during the evolution of different flexibility provision events.
- **VPP_4 – User-defined reconfiguration of VPPs:** Along with the automatic reconfiguration of VPP, the VPP Configuration Engine shall enable aggregators to take the final decision about their portfolio operational strategy. The end user of the VPP Configuration Engine, through an intuitive GUI, has always the alternative to define and set the portfolio flexibility strategies of interest, bypassing any automatic configuration performed by the engine.

6.2.9.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 43: Mapping of the VPP Configuration Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
VPP _1	WP6_162 (WP6_132), WP6_163, WP6_164, WP6_165, WP6_166, WP6_167, WP6_168, WP6_169, WP6_170, WP6_171
VPP _2	WP6_172, WP6_173
VPP _3	WP6_174, WP6_175, WP6_176, WP6_178, WP6_179
VPP _4	WP6_177

6.2.9.4 Exploited Technology Stack

In order to provide its intended functionalities, the VPP Configuration Engine component will build on state-of-the-art technologies, namely: (a) in the back-end layer, the Nest (NodeJS) web framework as a mature framework for delivering efficient, reliable and scalable server-side applications, (b) in the front-end layer, VueJS for custom SYNERGY front-end design to present the results of the optimization process; (c) in the data storage layer, PostgreSQL (as the relational database) to store the results of the VPP configuration and the flexibility related strategies. The data analytics and optimizations libraries to be used in the VPP Configuration Engine will be aligned with the frameworks and libraries that will be supported in the Data Analytics Services Bundle (Section 4.5) in the SYNERGY Cloud Platform.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.2.9.5 Updates from Draft Architecture

The scope and functionalities of the VPP Configuration Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, changes have been introduced in the following features in order to be consistent with the development progress in the respective component:

- VPP_2 – Flexible source clustering to VPPs: Apart from the results of the optimization process (to match the available flexible sources of the aggregators with the demand for ancillary services as triggered by the network operator), the end user of the tool has access on the ex-post results to get insights about the actual performance of the selected groups of flexible assets.

6.2.10 Flexibility Marketplace Search Engine

6.2.10.1 Component Overview

The role of the Flexibility Marketplace Search Engine is to enable aggregators to search in real time for a pool of flexibility sources according to specific details about their operational characteristics. On the other side, the flexibility source owners shall have quick, easy and free to choose access to a market framework that will facilitate their active participation and collaboration in evolving flexibility-based market schemas while also being able to properly configure the characteristics and parameters of their flexible assets in the marketplace.

6.2.10.2 List of Features

The list of features offered by the Flexibility Marketplace Search Engine (FMS) include:

- **FMS_1 – Prosumers to report assets availability in the marketplace**: By entering into the Flexibility Marketplace Search Engine, the flexibility source owners should be able to provide information about the flexibility sources. Details about the operational characteristics of the devices may be provided along with the marketplace configuration parameters where (s)he is prompted to set time schedules for the marketplace availability.
- **FMS_2 – Aggregators to search for the available flexibility sources in the marketplace:** Once the flexible source owners report their assets in the SYNERGY platform, these are made available to the aggregators (based on the data asset agreements made in the Data & AI Marketplace with the Blockchain-enabled DR Smart Contract Monitoring, Handling, Settlement and Remuneration

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Platform). Aggregators shall be provided with a tool to facilitate their search over the pool of available flexibility sources, by providing specific criteria of interest, i.e., type of flexibility source, time availability, service type, activation period etc.

6.2.10.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 44: Mapping of the Flexibility Marketplace Search Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
FMS_1	WP6_180, WP6_181, WP6_182, WP6_183, WP6_184, WP6_185, WP6_186, WP6_209
FMS_2	WP6_187, WP6_188, WP6_189, WP6_190, WP6_191, WP6_192, WP6_193, WP6_207

6.2.10.4 Exploited Technology Stack

In order to provide its intended functionalities, the Flexibility Marketplace Search Engine will build on state-of-the art technologies, namely: (a) in the back-end layer, the Nest (NodeJS) web framework as a mature framework for delivering efficient, reliable and scalable server-side applications (b) in the front-end layer, VueJS for custom front-end design for the flexibility marketplace; (c) in the data storage layer, Elasticsearch (to facilitate search over the flexibility sources available in the market platform) and PostgreSQL to store the marketplace participation configuration parameters.

6.2.10.5 Updates from Draft Architecture

The scope and functionalities of the Flexibility Marketplace Search Engine remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, changes have been introduced in the following features in order to be consistent with the development progress in the respective component:

- FMS_1 – Prosumers to report assets availability in the marketplace: A more detailed marketplace configuration process has been considered where the flexible asset owner is prompted to set time schedules as parameters for the asset availability at the flexibility marketplace.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.2.11 Flexibility Contracts Manager

6.2.11.1 Component Overview

The role of the Flexibility Contracts Manager is to manage the contractual process among aggregators and flexibility source owners. Aggregators are able to make a contractual offer to a flexible source owner by specifying the terms and conditions for the enrolment of their flexible assets to the aggregator's portfolio. In parallel, the flexibility source owners should be able to select the offer (by opting-in to a contractual process with an aggregator) and further accept, reject or negotiate (with the aggregator) the terms and conditions of the offer towards reaching an agreement. Eventually, a binding contract of flexibility will be established as reflected in a smart contract written in the blockchain.

6.2.11.2 List of Features

The list of features offered by the Flexibility Contracts Manager (FCM) include:

- **FCM_1 – Aggregators to offer draft contracts to the prosumers for their available flexibility:** The Flexibility Contracts Manager shall facilitate aggregators to pool flexibility sources under a potential contract and further set the technical and financial details about this draft contract, establishing that way an offer for the exploitation of the selected flexibility sources.
- **FCM_2 – Prosumers to be able to accept/reject/negotiate contractual terms:** The Flexibility Contracts Manager shall provide the mechanisms to the flexibility source managers to get an overview of the different offers triggered by the aggregators, evaluate the contractual terms – both operational and financial – and further accept, reject or negotiate the contractual terms of the offer.
- **FCM_3 – Aggregators to provide counteroffers during the negotiation phase:** In the case there is a rejection or negotiation of an offer by a flexibility source owner, the aggregator shall be able to provide a counter offer via the Flexibility Contracts Manager as part of the bilateral negotiation of the contract terms and conditions among the different business entities.
- **FCM_4 – Aggregators and prosumers to receive notifications about the status of the contractual process:** The platform should incorporate a notification feature in order to facilitate the different business entities to receive updates about the status of the negotiation in real time. Once a new contractual offer is triggered by the aggregator or an update during the negotiation phase takes place, the responsible entity is notified about this update in order to act accordingly.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.2.11.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 45: Mapping of the Flexibility Contracts Manager to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
FCM _1	WP6_194, WP6_195, WP6_196, WP6_197, WP6_198, WP6_199, WP6_210
FCM _2	WP6_200, WP6_201, WP6_202, WP6_203, WP6_225, WP6_226
FCM _3	WP6_204, WP6_205, WP6_208, WP6_225, WP6_226
FCM _4	WP6_206

6.2.11.4 Exploited Technology Stack

In order to provide its intended functionalities, the Flexibility Contracts Manager will build on state-of-the art technologies, namely: (a) in the back-end layer, the Ethereum blockchain and the Nest (NodeJS) web framework, (b) in the front-end layer, VueJS for custom front-end design for the contracts management part; (c) in the data storage layer, the contracts will be stored in the blockchain, but also in a relational database (PostgreSQL).

6.2.11.5 Updates from Draft Architecture

The scope and functionalities of the Flexibility Contracts Manager remain the same as presented in the SYNERGY Deliverable D2.6.

6.2.12 Flexibility Settlement & Remuneration Engine

6.2.12.1 Component Overview

Following the contractual process, a fair and transparent settlement and remuneration of each contract should be supported by the Blockchain-enabled DR Smart Contract Management Application. The entities of the market platform should be sure that any flexibility transaction is settled on the basis of an accurate, realistic and objective measurement and verification process based on the actual status quo of demand at each specific time period. Towards this direction, the Flexibility Settlement & Remuneration Engine will leverage innovative baseline algorithmic techniques - further complemented by adjustment and normalization methods – that are defined in the Analytics Workbench (Section 4.5.1) and executed in the Data Analytics Execution Service (Section 4.5.4) to enable analysis of real time big data streams coming from a wealth of data points (flexibility sources) on the way to support

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

continuous and "individualized" settlement and remuneration of the different market entities at different contexts and timings.

6.2.12.2 List of Features

The list of features offered by the Flexibility Settlement & Remuneration Engine include:

- **FSR_1 – Flexibility Settlement on the basis of an accurate Performance Measurement and Verification Process:** Following the contractual process and the active participation of flexibility sources in demand response campaigns, the Flexibility Settlement & Remuneration Engine shall ensure that the different flexible sources (and subsequently the business entities that own these sources) are fairly settled for their activation in these campaigns. The settlement should be performed on the basis of the actual metering data streams as gathered by the metering systems available in place.
- **FSR_2 – Flexibility Remuneration for the provided flexibility under a specific contract:** Along with the settlement for flexibility sources activation, appropriate methods should be available in the Flexibility Settlement & Remuneration Engine to facilitate the prompt remuneration of the different business entities on the basis of the contractual agreements. By taking into account the contractual terms and the actual participation of flexibility sources in demand response campaigns, the financial remuneration of them is taking place.

6.2.12.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 46: Mapping of the Flexibility Settlement & Remuneration Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
FSR _1	WP6_212, WP6_213, WP6_214, WP6_215, WP6_216, WP6_217, WP6_218
FSR _2	WP6_211, WP6_219, WP6_220, WP6_221, WP6_222, WP6_223, WP6_224, WP6_230, WP6_231

6.2.12.4 Exploited Technology Stack

In order to provide its intended functionalities, the Flexibility Settlement & Remuneration Engine will build on state-of-the art technologies, namely: (a) in the back-end layer, the Nest (NodeJS) web framework as a mature framework for delivering efficient, reliable and scalable server-side applications (b) in the front-end layer, VueJS for custom front-end design for the presentation of

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

settlement and remuneration results; (c) in the data storage layer, for performance measurement, depending on the applicable data assets' licenses (that determine whether local storage will be eventually allowed), PostgreSQL (as the relational database for the management of the energy and flexibility related properties), Elasticsearch (as the indexing engine to facilitate efficient clustering and segmentation of data). On the other hand, the settlement and remuneration results will be stored in the blockchain, but also in a relational database (PostgreSQL).

6.2.12.5 Updates from Draft Architecture

The scope and functionalities of the Flexibility Settlement & Remuneration Engine remain the same as presented in the SYNERGY Deliverable D2.6.

6.2.13 Blockchain Wallet

6.2.13.1 Component Overview

The blockchain-enabled DR Smart Contract Management Application aims to provide a market – based mechanism for aggregators and flexibility source owners to negotiate and transparently set contractual agreements. In order to achieve their business objectives and create a viable business case, aggregators and flexibility source owners need to get access in this marketplace in a secure and transparent way. The role of the Blockchain Wallet is to act as the user registry and management layer to enable end user's interaction with the blockchain-enabled DR smart contract monitoring, handling, settlement and remuneration platform.

The intended functionalities of the Blockchain Wallet described in this section are addressed by the Wallet Manager implementation (as described in Section 5.2.1) and thus the blockchain-enabled DR Smart Contract Management Application leverages the respective functionalities.

6.2.13.2 List of Features

The list of features offered by the Blockchain Wallet component include:

- **WM_1 – Authorisation based user's registration in the marketplace:** In order to enable the participation of the different business entities in the marketplace, a wallet presence (that represents the blockchain account) is required. A registration process should be supported by the platform to ensure that all actors of the marketplace can easily enrol to the functionalities offered by the platform. In order to ensure transparency on the marketplace operation, any

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

registration to the platform should be also assigned to the blockchain network of the marketplace.

6.2.13.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 47: Mapping of the Blockchain Wallet to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
WM_1	WP6_227, WP6_228, WP6_229

6.2.13.4 Exploited Technology Stack

In order to provide its intended functionalities, Wallet Manager will be based on Truffle[88] framework for managing the blockchain identity of the user (as the Ethereum based distributed ledger deployed in the project for the validation of the marketplace concept). The user configuration parameters will be also stored in a relational database (PostgreSQL).

6.2.13.5 Updates from Draft Architecture

The scope and functionalities of the Blockchain Manager remain essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the draft architecture, the only change is related to the implementation of the Blockchain Manager that leverages the respective functionalities of the SYNERGY Wallet Manager implementation.

6.3 Building/ District-level Analytics for Optimized Energy Performance Management

6.3.1 AI-boosted Renovation Decision Supporting Service

6.3.1.1 Component Overview

The aim of "AI-boosted renovation decision supporting service (AI-RDSS)" component is to perform pre-analysing and iterative pre-selection of potential renovation actions for ESCOs, building owners and other related business actors. This service utilizes occupants' behaviour and comfort profiles and

[88] https://www.trufflesuite.com/truffle

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

introduce iterative analytics loops when proposing alternative renovation scenarios of selected buildings. It will leverage and enhance digital twin for renovation analyses and module for machine learning the digital twin parameters by using the building energy consumption and related weather data.

6.3.1.2 List of Features

The list of feature offered by the component

- **AI-DRSS_1a**: Collect hourly history data of building energy consumption (electricity, heating, cooling) and related weather data. Collect the basic information of the studied building.
- **AI-RDSS_1b:** Collect the occupancy scheduling and occupants' thermal comfort parameters as critical parameters to be utilized in the simulation performed in the component for adapting to expected use of the to-be-renovated building.
- **AI-RDSS_2** - Build simplified and fast building energy consumption model (digital twin) by using the collected basic information of the building and machine learning the missing energy model parameters towards minimising the gap between measured energy consumption and model results.
- **AI-RDSS_3** - By using the learned building model, calculate selected renovation actions (e.g. improve thermal performance of windows, add solar collectors, etc.) one by one automatically in loop and store related results.
- **AI-RDSS_4** - Pre-select the most potential renovation actions for deep analysis using target criteria (e.g. payback time, investment costs, energy costs, carbon footprint and heating, hot water, cooling & electricity consumption, tenant comfort).
- **AI-RDSS_5** - Store data of building energy consumption and related weather data (needed for AI-RDSS_2). The machine learning algorithms for learning the studied building model needs background information of the existing building energy consumption (at least one year hourly energy and weather data – optionally also water data, excel file or online).
- **AI-RDSS_6** - Optional default: collect the country specific information such as investment costs, installation and energy costs and related carbon footprints.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.3.1.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 48: Mapping of the AI-boosted renovation decision supporting service to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
AI-RDSS _1	WP7_003; WP7_0012-WP7_0015; WP7_017-WP7_025; WP_071-WP7_087; WP7_091-WP7_098; WP7_099-WP7_113; WP7_154
AI-RDSS _2	WP7_012
AI-RDSS _3	WP7_132-WP7_134; WP7_138-WP7_142
AI-RDSS _4	WP7_132-WP7_134; WP7_138-WP7_142
AI-RDSS_5	WP7_132-WP7_134; WP7_138 - WP7_142
AI-RDSS_6	WP7_132-WP7_134; WP7_138-WP7_142

6.3.1.4 Exploited Technology Stack

In order to provide its intended functionalities, AI-RDSS will build on state-of-the art technologies and standards, namely: (a) in the back-end layer, the open source based Apache Tomcat for Java Servlet is required for running the studied renovation module and related obixStore database (access by REST interface) and NodeJS for running renovation module back-end GUI code, (b) in the front-end layer, e.g. Angular JavaScript front-end related GUI implementation; (c) in the data storage layer, PostgreSQL (as the relational database for building energy and weather data storage).

6.3.1.5 Updates from Draft Architecture

The scope and functionalities of the AI_RDSS component remains essentially the same as presented in the SYNERGY Deliverable D2.6, except that two new features AI-RDSS_5 and AI-RDSS_6 have been introduced to address the need to store building energy performance data and weather information locally (to be used by ML algorithms) and also to support the prioritization of renovation actions by cost in other EU countries, respectively.

6.3.2 IDA Indoor Climate and Energy (IDA-ICE) Renovation Analysis Service

6.3.2.1 Component Overview

The IDA-ICE renovation analysis service (IDA-ICE-RAS) component provides simulation services with aim to support more deep analysis of the pre-selected renovation actions resulted from AI boosted renovation decision support module. The necessary extensions of simulation libraries and analysis models will be developed in the IDA ICE environment.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.3.2.2 List of Features

The list of features offered by the component:

- **IDA-ICE-RAS_1** - Do detailed level modelling of the studied building and related pre-selected renovation actions. A detailed digital twin of the building will be created based on the available design parameters.
- **IDA-ICE_RAS_2** - Calculate the pre-selected renovation actions with chosen timestep (seconds, minutes, hours). Results will reveal heating, cooling and electricity saving potential of the actions.
- **IDA-ICE-RAS_3** - Prioritise the calculated renovation actions and select the best one, for example, it is possible, if needed, to prioritise the actions by economic optimization with MOBO optimization tool which can be linked to IDA ICE calculations.

6.3.2.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 49: Mapping of the IDA-ICE renovation analysis service to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
IDA-ICE-RAS_1	WP7_003; WP7_0012-WP7_0015; WP7_017-WP7_025; _071 - WP7_087; WP7_091 - WP7_098; WP7_099-WP7_113; WP7_154; WP7_132 - WP7_134; WP7_138 - WP7_142
IDA-ICE-RAS_2	WP7_132 - WP7_134; WP7_138 - WP7_142
IDA-ICE-RAS_3	WP7_132 - WP7_134; WP7_138 - WP7_142

6.3.2.4 Exploited Technology Stack

Simulations and libraries will be developed in the IDA-ICE simulation environment, which is widely used by ESCOs

6.3.2.5 Updates from Draft Architecture

The scope and functionalities of the IDA-ICE-RAS component remains the same as presented in the SYNERGY Deliverable D2.6.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.3.3 Near Real-time City Monitoring and Visualization Tool

6.3.3.1 Component Overview

The Near Real-time City Monitoring and Visualization tool (NRCMV) component collects and visualize (in near-real-time) the energy performance of group of buildings (i.e. district) and will offer advanced visual analytics on the energy performance of whole districts thus enabling the identification of weaknesses that need to be properly addressed by city decision maker.

6.3.3.2 List of Features

The list of feature offered by the component:

- **NRCMV_1** - Collect near-real time information about the energy performance of buildings and other information that may influence the energy performance of district
- **NRCMV_2** - Perform machine learning boosted data analytics and related forecasts based on measured heating, cooling and electricity consumption and production data. In addition calculate related KPI values
- **NRCMV_3** - Simulate selected city area buildings and related area energy demand estimates. The analytical methods will produce new indicators related to the energy performance of buildings in the studied area. Example of those area energy demand and production related KPIs are: area energy demand (heating, cooling, electricity separately); area energy production (heating, cooling, electricity separately) demand; energy demand and production per population; energy demand and production per weather (e.g. degree days); energy demand and production per area size; building average energy efficiency (kWh/m2) in studied area, area energy positive index in studied time frame, etc.
- **NRCMV_4** - Provide the list of new indicators to the SUPS component for their further integration into the simulation environment and respective visualization of targeted SECAP targets, hereby supporting city planner with the ability to design alternative urban transformation strategies (enhanced building level energy management capabilities, and district energy solutions such as large heat pumps, waste heat, seasonal storages, VPP emulations).

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.3.3.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 50: Mapping of the Near Real-time City Monitoring and Visualization tool to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
NRCMV_1	WP7_078 - WP7_088; WP7_114 - WP7_131; WP7_143 - WP7_151
NRCMV_2	WP7_078 - WP7_088; WP7_114 - WP7_131; WP7_143 - WP7_151
NRCMV_3	WP7_078 - WP7_088; WP7_114 - WP7_131; WP7_143 - WP7_151
NRCMV_4	WP7_155-WP7_161

6.3.3.4 Exploited Technology Stack

The City Monitoring and Visualization service (NRCMV) will be built on the following technologies to model and visualise city data: CityGML (City Geography Markup Language), EnergyADE, an extension of the CityGML standard for urban energy modelling and simulation. The Urban Analytics software will be built on the state-of-the art technologies and standards, namely: (a) in the back-end layer, the open source based Apache Tomcat for Java Servlet is required for running the studied solver and NodeJS for running renovation module back-end GUI code, (b) for updating the existing tool GUI, e.g. Angular JavaScript related GUI implementation; (c) in the data storage layer, PostgreSQL (as the relational database for building and area energy and weather data storage and related default values for Finland building stock).

6.3.3.5 Updates from Draft Architecture

The scope and functionalities of the NRCMV component remains essentially the same as presented in the SYNERGY Deliverable D2.6. The feature NRCMV_3 has been précised and combined with two features from SUPS component (SUPS_2, SUPS_3) with the progress of design and development activities of WP7. In addition, new NRCMV_4 function which is related to interaction with SUPS component and integration of energy performance indicators to the SUPS simulation environment has been added.

6.3.4 Strategic Urban Planning Supporter

6.3.4.1 Component Overview

The Near-Real Time City Monitoring and Visualization component will be complemented by “Strategic Urban Planning Supporter (SUPS)” component, which aims at supporting the city planner in the design of alternative urban transformation strategies (such as optimized energy management, building

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

renovation, interaction of buildings with energy markets, transportation interventions and strategies for EVs penetration), assess them and decide optimal routes to satisfy target KPIs defined in the city's SECAP plan.

6.3.4.2 List of Features

The list of features offered by the component:

- **SUPS_1** - Enable user to qualitatively model specific strategic use cases and scenarios including inserting new technologies into the context of existing settings (district, city).
- **SUPS_2** - Enable user to qualitatively simulate alternative scenarios for the selected cases such as for example alternative technology solutions and their effects on, for example, local businesses or attractiveness of neighbourhood or the effects of regulations, end user acceptance on the energy transition paths towards the reach of SECAP (Sustainable Energy and Climate Action Plan) targets.

6.3.4.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component

The list of the SYNERGY requirements addressed by the specific components is presented in the following table.

Table 51: Mapping of the Strategic Urban Planning Supporter to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
SUPS_1	WP7_155-WP7_161; WP7_026-WP7_028
SUPS_2	WP7_155-WP7_161; WP7_026-WP7_028

6.3.4.4 Exploited Technology Stack

In order to provide its intended functionalities, the Strategic Urban Planning Supporter (SUPS) will leverage the simulation libraries provided by Vensim.

6.3.4.5 Updates from Draft Architecture

The scope and functionalities of the SUPS component remains essentially the same as presented in the SYNERGY Deliverable D2.6. The two features from SUPS component (SUPS_2, SUPS_3) have been moved (combined) with NRCMV_3 feature, with the progress of design and development activities of WP7.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.3.5 Facility Management Monitoring Engine

6.3.5.1 Component Overview

The "Facility Management Monitoring Engine (FMME)" comprises a highly effective set of features allowing Facility Managers to get a deep and comprehensive understanding of the energy-related behaviour of the facilities under their responsibility, through the calculation and proper visualization of relevant metrics and KPIs.

6.3.5.2 List of Features

The list of features offered by the Facility Management Monitoring Engine includes:

- **FMME_1 – Visualization and drill-in of building available data**: all available data coming from the sensors of the building will be displayed in a comprehensive manner to help the assessment task of the building energy behaviour performed by the facility manager. The data may include measurements coming from smart meters (both at supply point or submetering), environmental sensors (e.g. temperature, humidity, presence, CO_2 or different particles concentration), production of RES units, operation of storage units, Electric Vehicle Charging Stations, etc. Collection and processing of the aforementioned data, together with results of personal and industrial/ energy analytics provided through the SYNERGY Data Analytics Service Bundles will enable the realization of rich dashboards towards continuously and simultaneously monitoring and assessing energy performance of the building, RES output and forecasts, demand metrics and forecasts, energy behaviours of occupants in buildings, status and capacity of storage systems and devices, flexibility capacity and forecasts, along with economic and environmental indices, at different levels/ zones of a building, as well as for a specific building in whole.
- **FMME_2 – Calculation and representation of relevant KPIs**: the component will also make use of the SYNERGY platform to calculate an appropriate set of KPIs relevant to the facility manager, in order to complement the available data and extract extra information from it that will help to determine the impact of the energy behaviour in other contexts, mainly economic and environmental. Examples of those KPIs include energy costs (potentially per type of load), RES-related KPIs (self-consumption, self-sufficiency), RES and storage-driven energy and cost savings, available flexibility, equivalent CO_2 emissions.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **FMME_3 – Tag and comparison of KPIs among different time periods**: the purpose of this feature is to allow facility managers to get a clear picture of the impact of specific measures (e.g. renovation, execution of optimized asset scheduling, etc) in the energy behaviour and related KPIs of the facility.
- **FMME_4 – Comparison of KPIs against baselines**: the purpose of this feature is to allow facility managers to understand how a particular facility behaves when compared to other reference facilities (e.g. similar facilities without optimized asset schedules).

6.3.5.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 52: Mapping of the Facility Management Monitoring Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
FMME_1	WP7_001, WP7_003, WP7_004, WP7_005, WP7_016, WP7_022, WP7_024, WP7_025, WP7_042, WP7_043, WP7_044, WP7_045, WP7_046, WP7_047, WP7_048, WP7_068, WP7_069, WP7_074, WP7_075, WP7_076, WP7_077, WP7_225, WP7_226, WP7_227, WP7_228, WP7_262, WP7_270
FMME_2	026, WP7_049, WP7_050, WP7_051, WP7_052, WP7_055, WP7_056, WP7_059, WP7_060, WP7_061, WP7_062, WP7_063, WP7_064, WP7_203, WP7_204, WP7_205, WP7_206, WP7_207, WP7_208
FMME_3	N.A.
FMME_4	WP7_053, WP7_054, WP7_057, WP7_058, WP7_065

6.3.5.4 Exploited Technology Stack

The main software libraries and platforms foreseen for the development, include NodeJS for backend services, Meteor for front-end/back-end communication and reactive data delivery to clients, and React for reusable and highly modular front-end development. In order to host component-specific data, MongoDB database is used.

The component has been packaged in a Docker image, which facilitates the deployment both on-premise and in the most extended cloud platforms.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

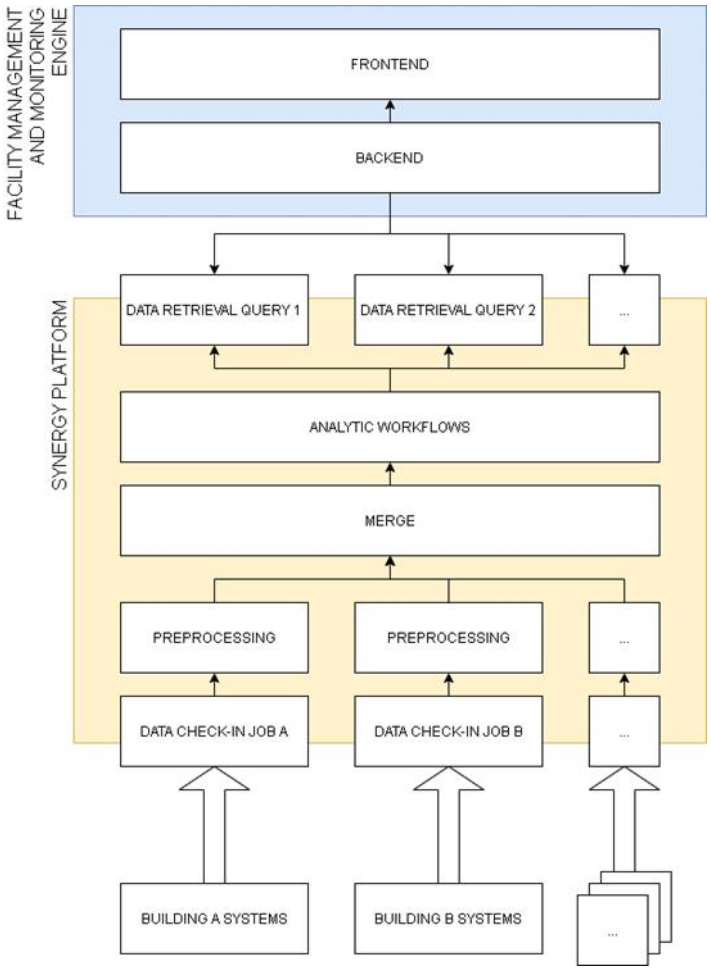
All data analytics used by this component have been developed in the form of Data Analytic Workflows that are executed by the SYNERGY Platform. Storage and data ingestion are outsourced to the SYNERGY Platform as well.

The resulting application will be accessible to end-users by means of a website, following a SaaS approach (i.e. the website will be multitenant and multi-user as required by the design specifications).

6.3.5.5 Updates from Draft Architecture

The scope of the FMME is the same as presented in deliverable D2.6.

Main updates on the architecture are related with the integration and usage of the SYNERGY Platform, and in particular on the configuration of the different Data Analytic Workflows in order to be able to achieve the design requirement of SaaS approach. This integration has been documented in D7.2, and an overview is presented hereby for the sake of completion.



After having access to the working version of the SYNERGY Platform, and after checking different options, the integration in general terms was defined as follows:

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- Data ingestion is outsourced to the SYNERGY Platform. In this sense, it is expected that building managers will upload data from their systems and devices into the SYNERGY Platform (via *Data Check-in Jobs*), and signed the appropriate contracts so the partner in charge of FMME setup has access to this data
- As part of the integration work of FMME with a particular customer, specific pre-processing steps need to be performed in order to transform the uploaded data into a common format (meaning common set of datasets, all of them with the same columns). As part of this pre-processing jobs, a unique id is attached to all data related to the same building
- The results of this pre-processing step get merged into common datasets that include data from all different customers. This step allows that all processing happening afterwards is stable and unaffected by the inclusion of new customers (I.e. allows multitenancy and scalability in terms of number of customers that can make use of the application)
- Core functionality of the component is configured as a set of *Analytic Workflows* which are executed within the SYNERGY Platform. Results are accessible via *Data Retrieval Queries*
- Results of analytics are accessible by end users from a website

6.3.6 HVAC Predictive Maintenance Service

6.3.6.1 Component Overview

The "HVAC Predictive Maintenance Service (HVAC-PMS)" component supports predictive maintenance tasks for large energy HVAC systems installed in buildings. It aims at recognizing building technical system malfunctions, inefficiencies and optimization possibilities.

6.3.6.2 List of Features

The list of features offered by the component:

- **HVAC-PMS_1** - Continuous monitoring and history data collection (measurement values, controllers' set point values) from studied HVAC system(s) and related indoor environment conditions including building and HVAC system level energy consumption (and production if available) by Building Automation and Control System (BACS) and weather data by open data API.
- **HVAC-PMS_2** - Providing integrated analytics for monitoring the technical performance of AHU and Heating Network functionalities. The analytics will include various data sets analyzed

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

against the needs and requirements of the occupants and visitors or historical performance data of specific facility systems in order to detect performance discrepancies or discrepancy related occupant discomfort.

- **HVAC-PMS_3** - Complementing and enhancing of HVAC-PMS_2 with Digital Twin based malfunctions, inefficiencies and optimization possibilities detection will be done for building level heating/cooling energy systems and human thermal comfort based approach for space heating/cooling system. The aim is to learn related digital twin parameters and compare learned digital twin results with real time measurement values. If the difference will start roaming or change rapidly then inform about possible malfunction.
- **HVAC-PMS_4** - Providing neural network based malfunctions, inefficiencies and optimization possibilities detection will be done using an autoencoder (optionally supported by simulations models)
- **HVAC-PMS_5** - Visualization of analysis on GUI to end-customer

6.3.6.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 53: Mapping of the HVAC Predictive Maintenance Service to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
HVAC-PMS_1	WP7_068, WP7_069
HVAC-PMS_2	WP7_068, WP7_069
HVAC-PMS_3	WP7_068, WP7_069
HVAC-PMS_4	WP7_068, WP7_069
HVAC-PMS_5	WP7_068, WP7_069

6.3.6.4 Exploited Technology Stack

The HVAC Predictive Maintenance Service (HVAC-PMS) will build on state-of-the- art technologies and standards, namely: (a) in the back-end layer, the open source based Apache Tomcat for Java Servlet is required for running the studied predictive maintenance module and related obixStore database (accessed by REST interface) and NodeJS for running predictive maintenance module back-end GUI code, (b) in the front-end layer, e.g. Angular JavaScript front-end related GUI implementation; (c) in the data storage layer, PostgreSQL (as the relational database for weather, building and HVAC system measurement data storage).

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.3.6.5 Updates from Draft Architecture

The scope and functionalities of the HVAC-PMS component remains essentially the same as presented in the SYNERGY Deliverable D2.6. However two featured presented in D2.6 have been further split into 5 related features as resulted of the progress made in WP7 design and development. More specifically, previous HVAC-PMS_2 feature, that have been containing all functionalities such as occupants comfort profiles analysis, aggregated data analysis, HVAC units malfunction detection algorithms and visualization of results to the end user, is fine-grained by the detention of respective 4 features separately as presented in the section earlier.

6.3.7 Building-level Energy Performance Optimisation Manager

6.3.7.1 Component Overview

As indicated in D7.1, the Building-Level Energy Performance Optimisation Manager (BL-EPOM) module supports facility managers, in presence of RES, storage and other energy flexibility sources (EVs), to design appropriate flexibility control strategies and interacting mechanisms with the urban energy network in order to be able to maximize self-consumption and reduce energy costs without compromising occupants' requirements for comfort at building level. This module addresses the objectives raised by use case 7.5 and can be summarized as the calculation of the operation profile of manageable systems that maximize self-consumption and reduce energy costs according to the expected building energy demand and generation and its systems state.

6.3.7.2 List of Features

The list of features offered by the Building-Level Energy Performance Optimisation Manager (BL-EPOM) include:

- **BL-EPOM_1**: Calculate solar PV and wind generation for every building. The BL-EPOM will optimize local generation use, this will allow to choose the better use for the renewable energy potential: use it, store it or even program/propose curtailment in extreme situations (available generation but no possibility of using, storing or selling energy).
- **BL-EPOM_2**: Calculate batteries use for every building. An output of the optimization process of the BL-EPOM will be the local batteries usage plan: charging and discharging power for the next optimization period.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **BL-EPOM_3**: Calculate EV batteries charge for every building. An output of the optimization process of the BL-EPOM will be the EV batteries usage plan: charging and discharging (if V2G use is available) power for the next optimization period.
- **BL-EPOM_4**: Calculate energy exchanges with the grid. Optimize the energy exchanges with the grid (purchases and sales) by considering as input the available flexibility at building asset level as defined and scheduled through comprehensive flexibility analytics that will be made available as input to the component.
- **BL-EPOM_5**: Calculate manageable demands use for every building without compromising occupant requirements for comfort. An output of the optimization process of the BL-EPOM will be the HVAC system and manageable demands usage plan.

6.3.7.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 54: Mapping of the Building-Level Energy Performance Optimisation Manager to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
BL-EPOM_1	WP7_273 to WP7_276 WP7_281 to WP7_284
BL-EPOM_2	WP7_289 to WP7_295
BL-EPOM_3	WP7_303 to WP7_309
BL-EPOM_4	WP7_317 to WP7_321
BL-EPOM_5	WP7_330 to WP7_332

6.3.7.4 Exploited Technology Stack

Related to the optimization processes of the BL-EPOM, the tool will work in two stages:

1. Optimization of the electric energy demand derived from HVAC systems. Artificial Neural Network (ANN) will be used to model and characterise HVAC effect on building average temperature (creating a “transfer function”). Periodically, the ANN model will be updated to reflect any change in the building or in the user’s behaviour. A Genetic Algorithm (GA) will be used to obtain the HVAC profile that minimizes energy costs and maintains user comfort, by minimizing the difference between the temperature setpoint and the building temperature induced by the HVAC system use. The ANN will be implemented with Keras, based on Tensorflow, and codified in Python. The GA will be a CIRCE development, also developed in Python.

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

2. Optimization of other electric systems. In this second stage, the optimal use of HVAC systems (calculated in the first stage) will be considered as an input and part of non-manageable demands. A MILP model/problem will be used to emulate building electric behaviour and to optimize the use of flexible or manageable devices to minimize energy cost and maximizing self-consumption. It is going to be used GUROBI as optimization solver.

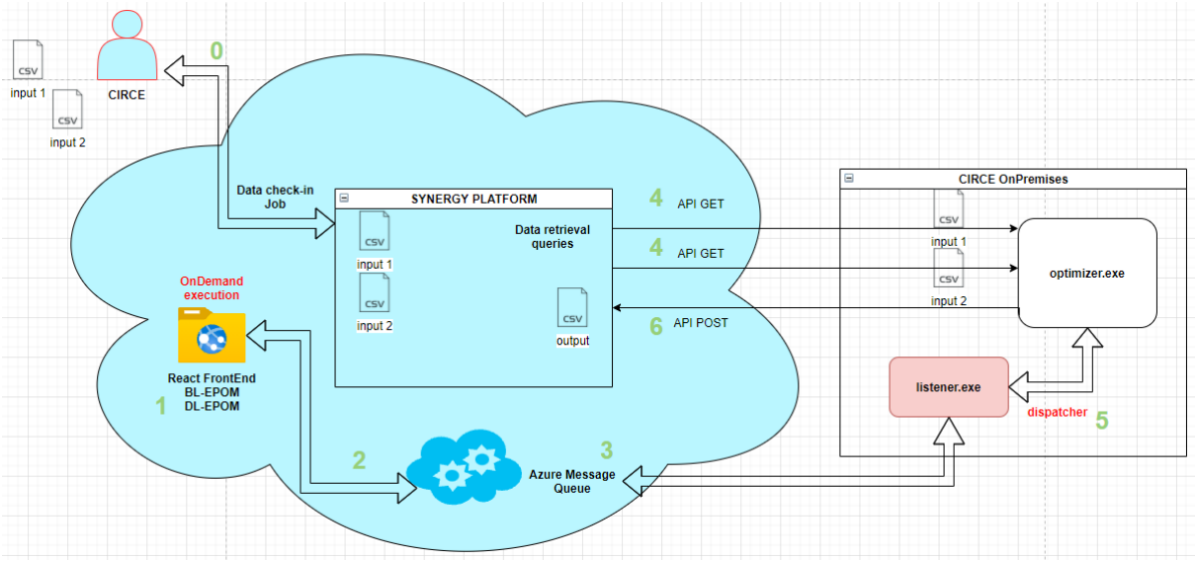
One of the objectives of these two optimization steps is skipping the first one if there is no manageable HVAC system in the building. Besides these tools different libraries are used:

- Pandas (for PYTHON)
- React (FrontEnd)
- React-dom
- React-bootstrap
- Listener (.NET Core)
- ServiceBusClient
- ServiceBusMessage
- ServiceBusMessageBatch
- Axios API Connect

6.3.7.5 Updates from Draft Architecture

The scope of the BL-EPOM is the same as presented in the SYNERGY document D2.6.

This is the architecture diagram with all related technology that is used for this preliminary OnDemand execution process.



This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Step 0: CIRCE uploads input files (.csv) to the SYNERGY Platform through a Data Check-in Job

Step 1: CIRCE, through the user interface which is being developed in React FrontEnd, launches an OnDemand execution, which internally triggers a message to Azure Message Bus

Step 2: CIRCE OnPremises has a program that act as a listener for that message queue.

Step 3: Once the message has been received ("OnDemand execution needed"), it triggers an execution of the implemented optimizer.exe for this stage.

Step 4: Files uploaded into the SYNERGY Platform must be downloaded to the CIRCE OnPremises in order for this .exe to run properly.

Step 5: Listener triggers the execution of optimizer.exe, which has all the required information to run and generate the result

Step 6: output file (.csv) is generated as a result of this OnDemand execution locally on CIRCE OnPremises, and could be also uploaded into the SYNERGY Platform through the use of an API POST method.

6.3.8 District-level Energy Performance Optimisation Manager

6.3.8.1 Component Overview

As indicated in D7.1, the District-Level Energy Performance Optimisation Manager (DL-EPOM) module supports facility managers, in presence of RES, storage and other energy flexibility sources (EVs), to design appropriate flexibility control strategies and interacting mechanisms with the urban energy network in order to be able to maximize self-consumption and reduce energy costs of groups of buildings (districts). This component addresses the objectives raised by use case 7.6 and can be summarized as the calculation of the operation profile of manageable systems that maximize self-consumption and reduce energy costs according to the expected district energy demand and generation and its systems state.

6.3.8.2 List of Features

The list of features offered by the Building-Level Energy Performance Optimisation Manager (BL-EPOM) include:

- **DL-EPOM_1**: Calculate solar PV and wind generation for every building. The DL-EPOM will optimize local generation and program/propose curtailment if needed.
- **DL-EPOM_2**: Calculate batteries use for every building. An output of the optimization process of the DL-EPOM will be the local batteries usage plan: charging and discharging power for the next optimization period.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- **DL-EPOM_3**: Calculate EV batteries charge use for every building. An output of the optimization process of the DL-EPOM will be the EV batteries usage plan: charging and discharging (if V2G use is available) power for the next optimization period.
- **DL-EPOM_4**: Calculate energy exchanges with the grid. Optimize the energy exchanges with the grid (purchases and sales) by considering as input the available flexibility at building asset level as defined and scheduled through comprehensive flexibility analytics that will be made available as input to the component.
- **DL-EPOM_5**: Calculate energy exchanges between buildings through the grid. Optimize the energy exchanges among buildings through public the grid.
- **DL-EPOM_6**: Calculate manageable demands use for every building without compromising occupant requirements for comfort. An output of the optimization process of the DL-EPOM will be the manageable demands usage plan.

6.3.8.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 55: Mapping of the District-Level Energy Performance Optimisation Manager to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
DL-EPOM_1	WP7_277 to WP7_281 WP7_285 to WP7_288
DL-EPOM_2	WP7_296 to WP7_302
DL-EPOM_3	WP7_310 to WP7_316
DL-EPOM_4	WP7_318 to WP7_321
DL-EPOM_5	WP7_322 to WP7_329
DL-EPOM_6	WP7_333 to WP7_335

6.3.8.4 Exploited Technology Stack

The DL-EPOM optimization process works in wot stages as previously explained for the BL-EPOM. The libraries used are:

- Gurobipy (for PYTHON)
- Pandas (for PYTHON)

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

- React (FrontEnd)
- React-dom
- React-bootstrap
- Listener (.NET Core)
- ServiceBusClient
- ServiceBusMessage
- ServiceBusMessageBatch
- Axios API Connect

6.3.8.5 Updates from Draft Architecture

The scope of the DL-EPOM is the same as presented in the SYNERGY document D2.6.

The architecture changes of the DL-EPOM are same as the previously presented for the BL-EPOM

6.3.9 eDECs (enhanced Display Energy Certificates) Calculation Engine

6.3.9.1 Component Overview

As described in D7.2 the Facility Management Energy Analytics, Self-Consumption Optimization & Predictive Maintenance toolbox will integrate the two additional components dealing with the assessment of the buildings, the enhanced Display Energy Certificates (eDECs) and the Smart Readiness Indicator (SRI).

The eDECs module will deliver dynamic certificates of building energy performance in variant resolutions (e.g. annual, monthly, daily), for the building as a whole or per designated zone. It will utilize real-time field data of the building (consumption, production, etc.), static building data and benchmark values (drawn from the SYNERGY platform) and will dynamically calculate the energy performance indicators associated with the Display Energy Performance certificate as per the standard that will be adopted. The eDECs application will retrieve the necessary input data uploaded by the facility managers to the SYNERGY Platform (e.g. the time series of meter readings). In any case, facility managers will be able to fill in any missing piece of data (mainly static properties of the facilities under study) directly to the eDECs application, by using the corresponding forms on the user interface. The data collected through the UI will also be stored in SYNERGY platform.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.3.9.2 List of Features

- **eDEC-CE_1**: Close to real-time calculation of building's energy performance operational rating
- **eDEC-CE_2**: Close to real-time calculation of CO2 (Carbon Dioxide) emissions due to electricity consumption

6.3.9.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 56: Mapping of the eDECs (enhanced Display Energy Certificates) Calculation Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
eDEC-CE_1	WP7_039, WP7_040, WP7_041, WP7_042, WP7_043, WP7_046, WP7_053, WP7_055, WP7_057, WP7_066, WP7_067
eDEC-CE_2	WP7_051, WP7_061, WP7_063, WP7_065

6.3.9.4 Exploited Technology Stack

"Go" programming language has been used for the development of the eDECs web API service while the following library was used for the same purpose.

Library	Version	License
Echo	4.4.0	Open source

6.3.9.5 Updates from Draft Architecture

The scope of the eDECs remains essentially the same as presented in the SYNERGY Deliverable D2.6. In respect to the functionalities though, one feature has been removed (eDEC-CE_3 - Close to real-time calculation of human preferences indicators (e.g. visual and thermal comfort, occupancy)) as it has been considered outside the final scope of the tool.

With regards to the architecture, since the delivery of D2.6 and as also described in deliverables D7.1 and D7.2, it has been agreed that this module will be integrated in the user interface of the FMME module described in section 6.3.5. The user will login to the FMME application and through a dedicated module (option), they will be able to perform the eDEC assessment of the building of interest.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

6.3.10 SRI (Smart Readiness Indicator) Calculation Engine

6.3.10.1 Component Overview

The SRI module will deliver the "Smart-Readiness" assessment of a building, by calculating its capability to a) support energy savings techniques b) respond to user needs and c) offer services to the grid. The tool will utilize static building data (drawn from the SYNERGY platform) and benchmark values and will calculate various indicators (disaggregated and building total) as per the SRI methodology, also described within D7.1.

6.3.10.2 List of Features

The list of features offered by the SRI Calculation Engine include:

- **SRI_CE_1**: Calculate the total Smart-readiness score of the building (One single score classifies the building's Smart Readiness)

6.3.10.3 Mapping with SYNERGY Requirements

The list of the SYNERGY requirements addressed by the specific component is presented in the following table.

Table 57: Mapping of the SRI Calculation Engine to the SYNERGY Requirements

Feature	Related Requirements (D2.2)
SRI_CE_1	WP7_351, WP7_380, WP7_381, WP7_382, WP7_383, WP7_384, WP7_385, WP7_386, WP7_387, WP7_388, WP7_389, WP7_390, WP7_391, WP7_392, WP7_393, WP7_394, WP7_395, WP7_396, WP7_397, WP7_398, WP7_399, WP7_400, WP7_401

6.3.10.4 Exploited Technology Stack

"Go" programming language has been used for the development of the eDECs web API service while the following library was used for the same purpose.

Library	Version	License
Echo	4.4.0	Open source

6.3.10.5 Updates from Draft Architecture

The scope and functionalities of the SRI remain the same as presented in the SYNERGY Deliverable D2.6. In respect to the architecture, similarly to the eDECs module, since the delivery of D2.6 and as also described in deliverables D7.1 and D7.2, it has been agreed that the SRI module will be integrated

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

in the user interface of the FMME module described in section 6.3.5. The user will login the FMME platform and will be able to select to perform the SRI assessment of the building of interest.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

7 Mapping of SYNERGY Framework Architecture to Reference Architectures

In this section, the mapping between the SYNERGY Framework Architecture and the BDVA Reference Model and the SGAM is described.

It needs to be noted that there are no changes in respect to the draft architecture mapping described in D2.6.

7.1 Alignment to BDVA Reference Architecture

As depicted in Figure 20, the overall SYNERGY Big Data Platform & AI Marketplace, along with its different Data Services Bundles is well aligned to the BDVA Reference Model defined in the European Big Data Value Strategic Research & Innovation Agenda (BDVA, 2017). From the horizontal axis perspective, topics around Data Management are appropriately addressed through the SYNERGY Data Collection and Data Governance Service Bundles. Data Protection is considered from an all-around perspective in the SYNERGY Data Security Service Bundle. Data Processing Architectures, Data Analytics and Data Visualization and User Interaction aspects have a similar context and orientation as in the SYNERGY Data Analysis Services Bundle. With regard to the vertical axes, the Data Sharing Platforms are indeed tackled through the SYNERGY Data Sharing Services Bundle that is innovative in introducing the concept of multi-party sharing. Development, Engineering and DevOps aspects are well embedded in the SYNERGY Platform Management Services Bundle. Finally, the Standards dimension is addressed within the SYNERGY Common Information Model that builds upon different energy data standards, ontologies and vocabularies.

It needs to be noted that with its perception of the data pipelines for the data check-in processes and the data analysis processes, SYNERGY is also well positioned and aligned with the ISO 20547-3 “Big Data Reference Architecture” (2020) and the under-development ISO 23053 “Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)” to which the BDVA Reference Model has also contributed.

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

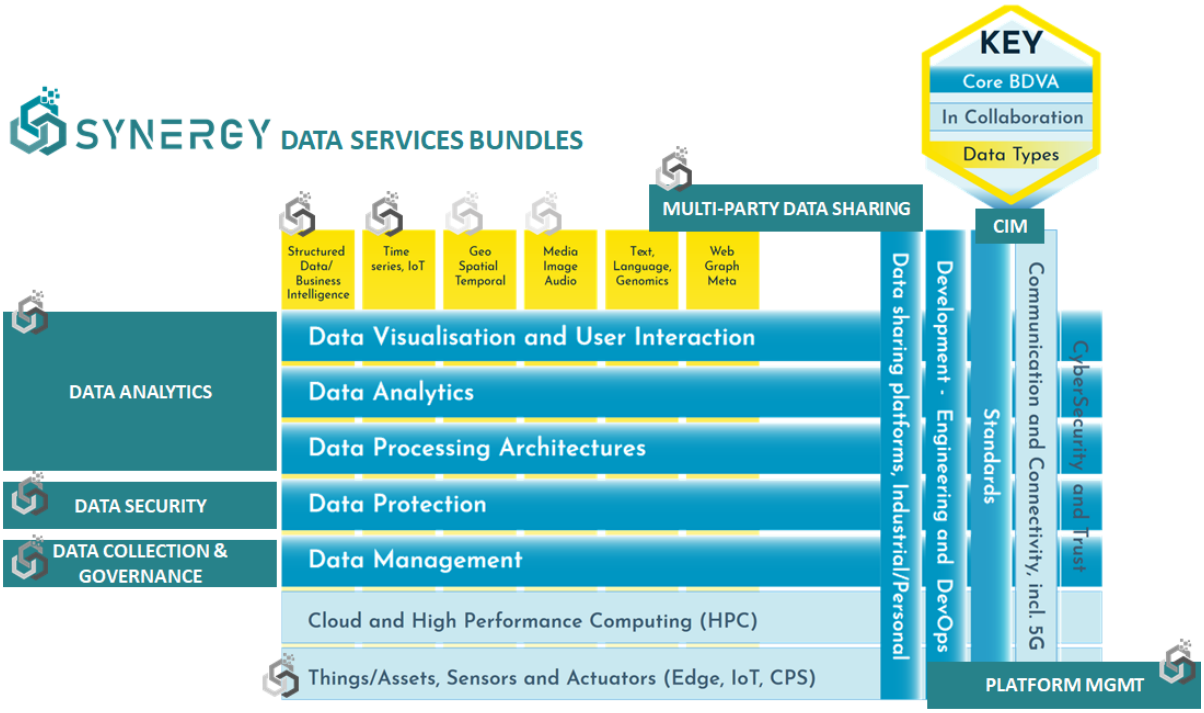


Figure 20: SYNERGY Platform Alignment to the BDVA Reference Model

7.2 Alignment to SGAM Reference Architecture

In the direction of Enterprise Architecture Frameworks, the Smart Grid Architecture Model (SGAM) (CEN-CENELEC-ETSI, 2012) aims at offering a broad framework for the design of smart grids use cases with an architectural approach allowing for a representation of interoperability viewpoints in a technology neutral manner. In accordance with the scope of the M/490 program (EC, 2011), the SGAM framework is practically a three-dimensional model merging the dimension of interoperability layers with the two dimensions of the Smart Grid Plane as depicted in the following figure.

The x-axis of the SGAM (Domains) divides the problem domain "electrical power supply" into the individual sections: Generation, Transmission, Distribution, Distributed Energy Resource, and Customer Premises. The y-axis of SGAM ("Zones") reflects the automation pyramid (Process, Field, Station, Operation), supplemented by the two zones "Enterprise" and "Market". The z-axis of an SGAM (Interoperability) layers the five interoperability levels, namely the Business Layer, the Function Layer, the Information Layer, the Communication Layer, and the Component Layer.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

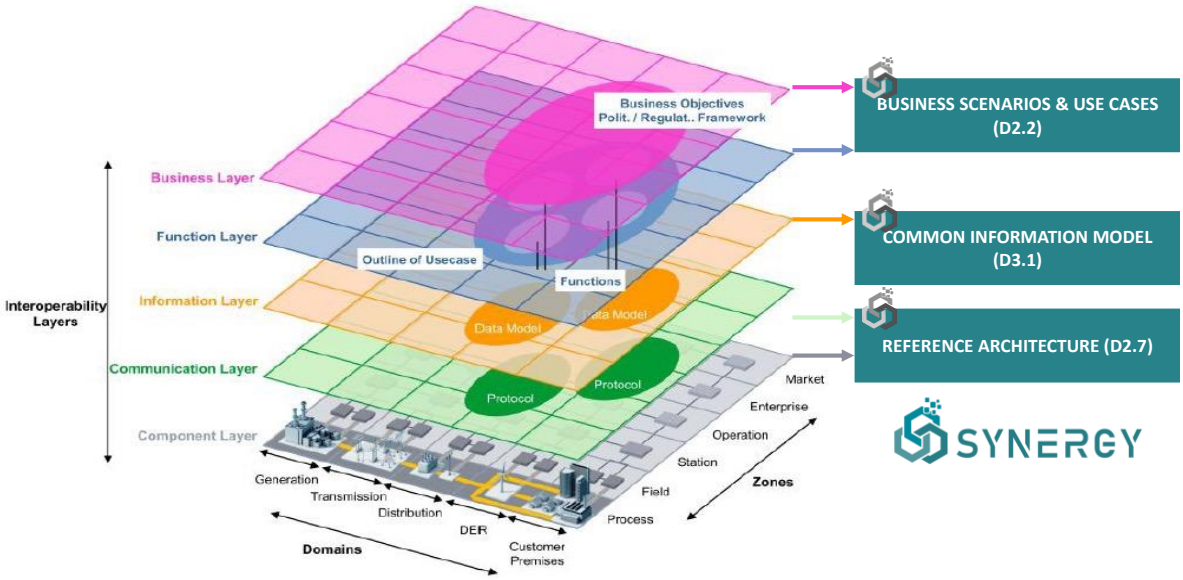


Figure 21: SGAM Architecture Model and SYNERGY

As it can be realized from sections 3-6, the SYNERGY architecture was designed following the SGAM philosophy and design patterns in a more loosely coupled manner. The SYNERGY Reference Architecture that was presented in section 2 can be conceptually mapped to the SGAM Component Layer as depicted in Figure 22. The details about the Business and Function layers of the SGAM Architecture Model have been elaborated through the business scenarios and use cases as presented in the SYNERGY Deliverable D2.2 while the specificities of the Information layer have been presented in the SYNERGY Deliverable D3.1 along with the definition of the SYNERGY Common Information Model. To this end, the focus in the present architecture deliverable has been laid on the definition of the Component Layer to define the physical distribution of all participating components in the smart grid context.

As depicted in Figure 22, the SYNERGY Cloud Infrastructure (including the Core Big Data Platform & AI Analytics Marketplace as well as the Secure Experimentation Playground) is placed at the core in the operational and enterprise zone spanning through the different domains of the energy value chain. Complementary to the SYNERGY Cloud infrastructure, the SYNERGY On-Premise Environments are positioned at the station layer to collect and run analytics over energy data while their edge edition also acts at aggregation level for field devices.

On top of the SYNERGY Core Data Platform in the Cloud Infrastructure (at the enterprise/market zone), the SYNERGY energy applications (and the respective architecture components) appear, namely: the Advanced Grid-level Analytics for Optimized Network and Asset Management Services and

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

Applications (focusing on the generation, transmission and distribution domain), the Portfolio-level Analytics for Energy-as-a-Service (EaaS) Applications (focusing mainly on the DER and Customer Premise domains) and the Building/ District-level Analytics for Optimized Energy Performance Management (focusing primarily on the Customer Premise domain but also on the DER domain).

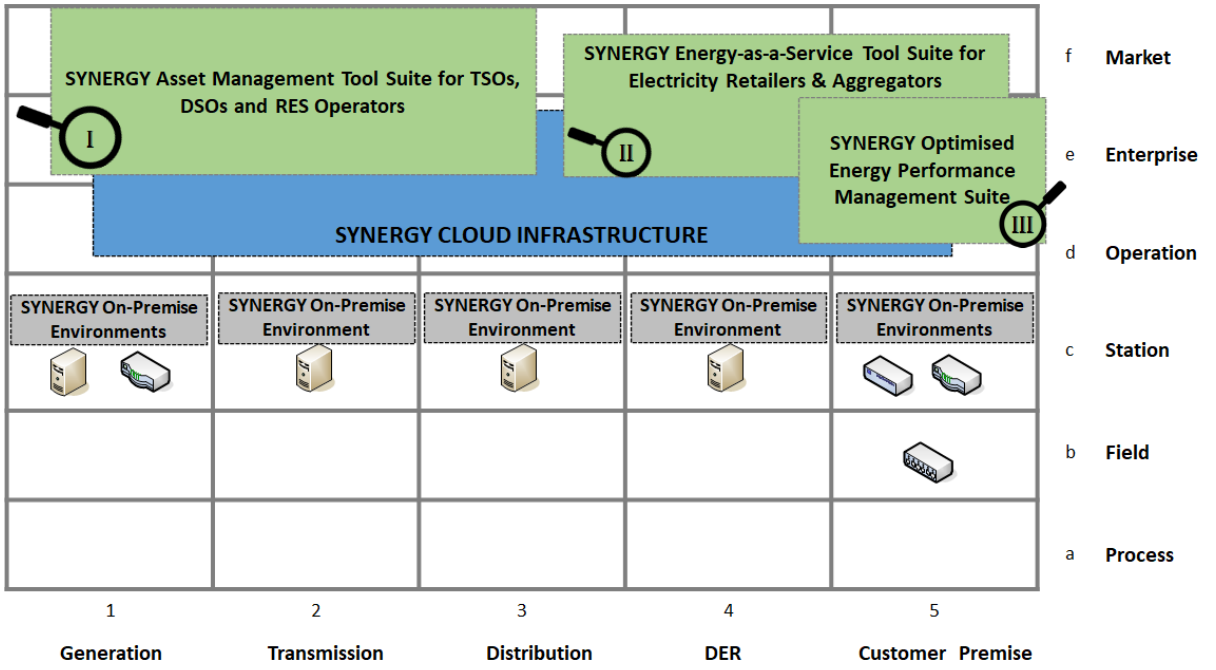


Figure 22: Positioning of the SYNERGY Reference Architecture Layers in the SGAM Architecture Model

The positioning of the components related to Advanced Grid-level Analytics for Optimized Network and Asset Management Services and Applications (that is referred to as Asset Management Tool Suite for TSOs, DSOs and RES Operators in the figures, for brevity) is depicted in more detail in Figure 23 where it is confirmed that generation, transmission and distribution are the focus of nine (9) SYNERGY applications for TSOs, DSOs and RES operators.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

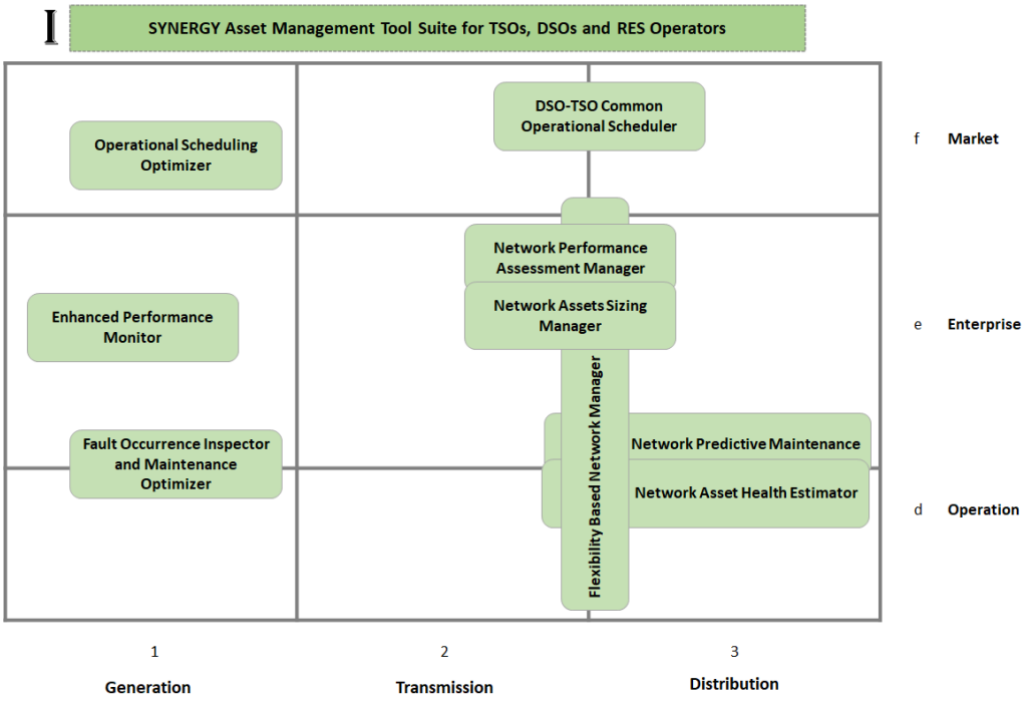


Figure 23: Positioning of the SYNERGY Asset Management Tool Suite for TSOs, DSOs and RES Operators in the SGAM Architecture Model

The positioning of the components related to Portfolio-level Analytics for Energy-as-a-Service (EaaS) Applications (that is referred to as SYNERGY Energy-as-a-Service Tool Suite for Electricity Retailers & Aggregators in the figures, for brevity) is depicted in more detail in Figure 24 where it is confirmed that DER and customer premise as well as slightly distribution are the focus of eleven (11) SYNERGY applications for electricity retailers and aggregators.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

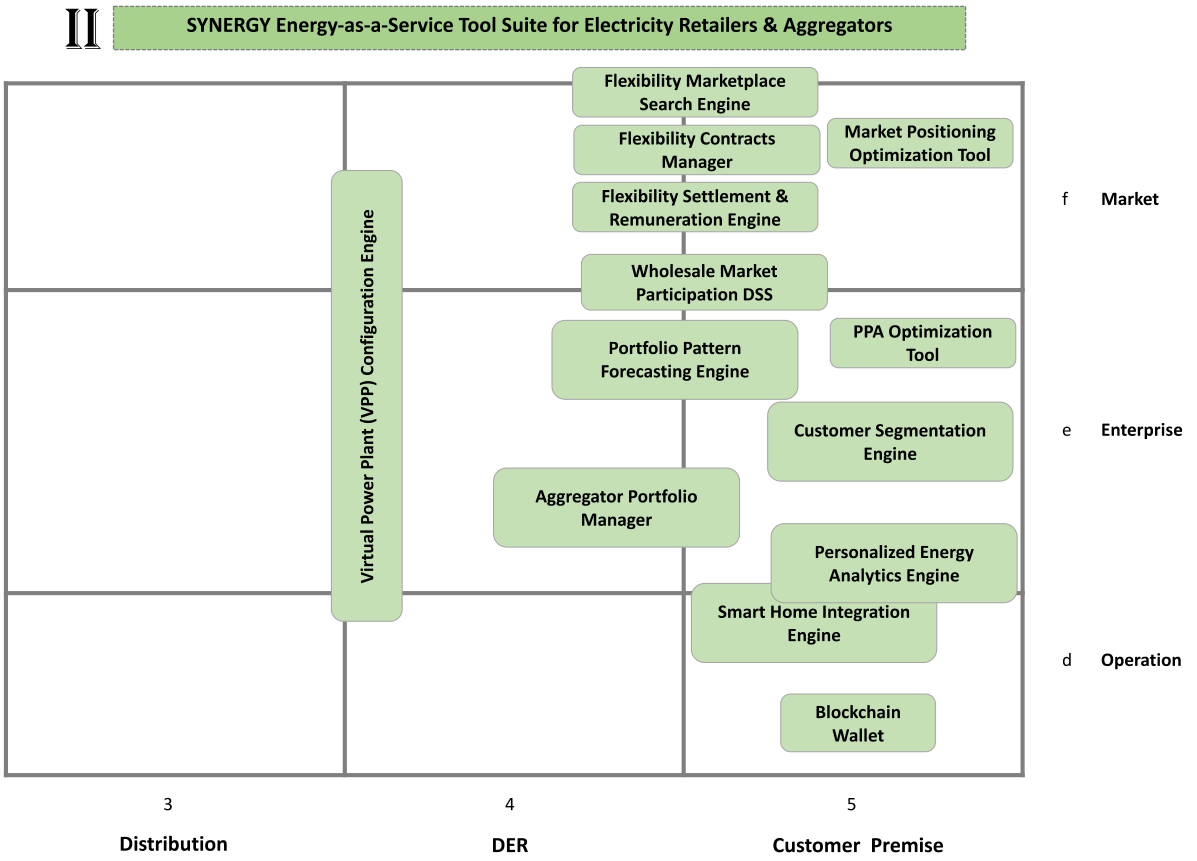


Figure 24: Positioning of the SYNERGY Energy-as-a-Service Tool Suite for Electricity Retailers & Aggregators in the SGAM Architecture Model

The positioning of the components related to Building/ District-level Analytics for Optimized Energy Performance Management (that is referred to as SYNERGY Optimised Energy Performance Management Suite in the figures, for brevity) is depicted in more detail in Figure 25 where it is confirmed that DER and mostly customer premise are the focus of ten (10) applications targeting Facility Managers and ESCOs.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

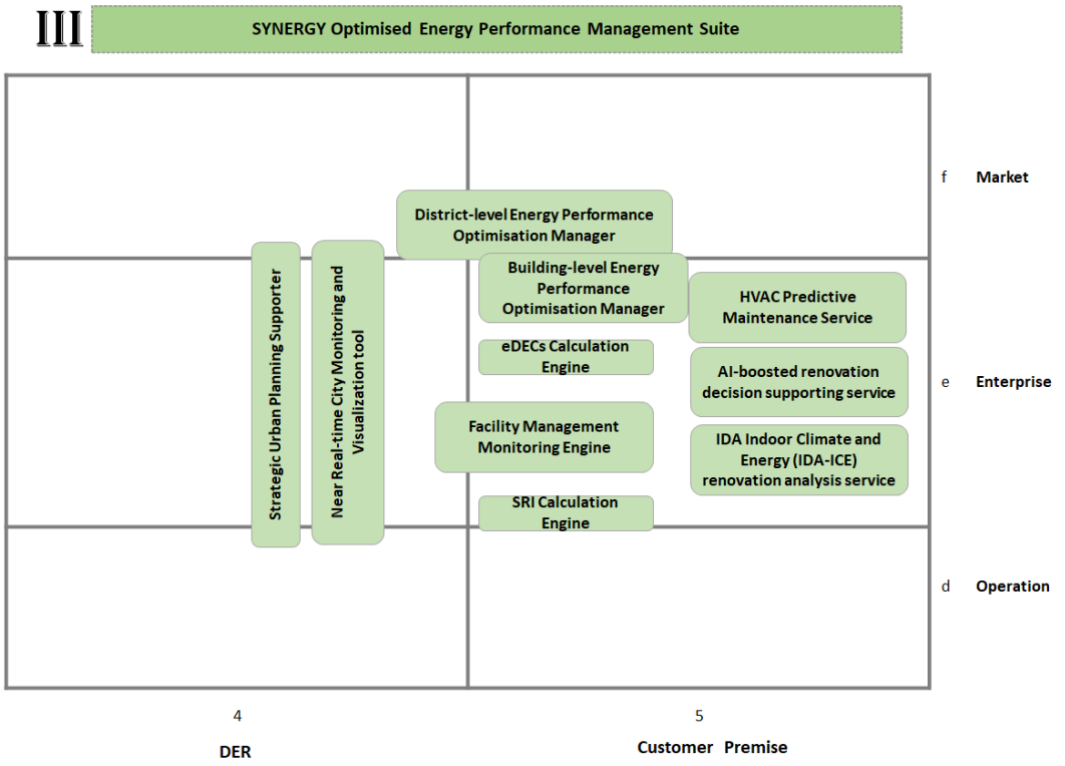


Figure 25: Positioning of the SYNERGY Optimised Energy Performance Management Suite in the SGAM Architecture Model

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

8 Conclusions

The purpose of the deliverable D2.7 entitled "SYNERGY Framework Architecture including functional, technical and communication specifications v2" was to deliver the final version of the overall conceptual architecture of the SYNERGY platform accompanied by the detailed specifications of the platform's components and their respective functionalities.

To this end, the deliverable presented the final detailed overall conceptual architecture of the SYNERGY platform that was optimised and enhanced in various ways from three different perspectives. At first, the deliverable presented the overall conceptual architecture from the layer's perspective, where the three core layers of the platform's architecture, namely the SYNERGY Cloud Infrastructure, the On-Premise Environments and the SYNERGY Energy Apps Portfolio, were documented highlighting their key elements as well as their scope in the overall architecture. Secondly, the deliverable presented the architecture from the data-driven services bundles and energy apps perspective, documenting the details of the eight (8) in total service bundles and the Energy App portfolio with twelve (12) apps. In particular, for each service bundle or energy app its main role and its overall offerings were presented. Finally, the architecture was presented from the components perspective, where the twenty eight (28) in total components that compose these service bundles and thirty one in total (31) that compose the Energy Apps were documented describing their context and positioning in the overall architecture. Finally, the description of the overall conceptual architecture of the SYNERGY platform is supplemented with the documentation of the concrete roles and users that will utilise the SYNERGY platform.

Following the documentation of the overall conceptual architecture of the SYNERGY platform, the updated SYNERGY platform's workflows were presented. The eleven (11) in total workflows received the required enhancements and updates and were presented in detail with the help of BPMN diagrams which depict the final functionalities that are offered by the SYNERGY platform. The different workflows were organised in three main workflow categories, namely the data collection, the data search and sharing and the data analytics workflows. Each workflow depicts a set of functionalities by documenting how the various users with different roles are interacting with the SYNERGY platform and how the various components are combined.

The deliverable presented the updated documentation of the design specifications of the different components of the SYNERGY platform. SYNERGY platform's architecture is composed of twenty eight (28) components in total which were presented based on the data-driven services bundle, as well as

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

thirty one (31) components which were presented based on the energy app that they belong, and the respective layer of the platform's architecture where the bundle or the energy app is positioned. The deliverable presented the updated documentation of their scope, the offered features and the requirements addressed by each component as well as the changes introduced in respect to the draft architecture presented in D2.6. In addition to this, the technologies which will be utilised for their implementation and the updates from the previous version are highlighted. Finally, the deliverable documented the mapping the SYNERGY platform's architecture aspects with the two core reference architectures, namely the SGAM and BDVA Reference Architectures, presenting the alignment of the SYNERGY platform's architecture aspects with their various aspects.

The deliverable at hand delivered the final version of the overall conceptual architecture of the SYNERGY platform, as well as of the final version of the detailed specifications of the platform's components. It constitutes the final report of the outcomes of Task 2.4 and provides the supplementary documentation of the initial documentation that was provided with deliverable D2.6 with refinements and optimisations from M13 till M24. The deliverable concludes the activities of the specific task per the SYNERGY Description of Action.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

9 References

BDVA (2017). European Big Data Value - Strategic Research and Innovation Agenda, Version 4.0. Retrieved from: https://bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf [Accessed 30 December 2020]

CEN-CENELEC-ETSI Smart Grid Coordination Group. (2012). Smart Grid Reference Architecture. Retrieved from: https://ec.europa.eu/energy/sites/ener/files/documents/xpert_group1_reference_architecture.pdf [Accessed 30 December 2020].

EC. (2011) Smart Grid Mandate Standardization Mandate to European Standardisation Organisations (ESOs) to support European Smart Grid deployment (M/490 EN). Retrieved from: https://ec.europa.eu/energy/sites/ener/files/documents/2011_03_01_mandate_m490_en.pdf [Accessed 30 December 2020].

ISO. (2020). ISO/IEC 20547-3:2020 Information technology — Big data reference architecture — Part 3: Reference architecture. Retrieved from: https://www.iso.org/standard/71277.html [Accessed 30 December 2020].

ISO. (2020). ISO/IEC CD 23053.2 Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML). Retrieved from: https://www.iso.org/standard/74438.html [Accessed 30 December 2020].

OMG. (2011). BPMN (Business Process Model And Notation) Specification Version 2.0. Retrieved from: https://www.omg.org/spec/BPMN/2.0/PDF [Accessed 30 December 2020].

SYNERGY Consortium. (2019). *SYNERGY GA - ANNEX 1.*

SYNERGY Consortium. (2020a). SYNERGY D2.1 "End-user and Business requirements analysis for big data-driven innovative energy services and ecosystems v1"

SYNERGY Consortium. (2020b). SYNERGY D2.3 "Socio-economic and regulatory analysis of obstacles to innovation v1"

SYNERGY Consortium. (2020c). SYNERGY D2.5 "Ex-Ante Pilot Audits and Pilot Deployment Plan"

SYNERGY Consortium. (2020d). SYNERGY D2.6 "SYNERGY Framework Architecture including functional, technical and communication specifications v1"

SYNERGY Consortium. (2021a). SYNERGY D3.1 "SYNERGY Common Information Model"

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.

SYNERGY Consortium. (2021b). SYNERGY D3.4 “SYNERGY Integrated Platform & Open APIs – Beta Release”

SYNERGY Consortium. (2021c). SYNERGY D3.5 "SYNERGY Data Collection, Security, Storage, Governance & Management Services Bundles – Release 1.00"

SYNERGY Consortium. (2021d). SYNERGY D4.3 "SYNERGY Data Analytics, Sharing & Matchmaking Services Bundles – Release 1.00"

SYNERGY Consortium. (2021e). SYNERGY D4.2 "SYNERGY Baseline Data Analytics – Draft Release"

SYNERGY Consortium. (2021f). SYNERGY D5.2 "1st Version of the Asset Management Tool Suite for Network and RES Plant Operators"

SYNERGY Consortium. (2021g). SYNERGY D6.2 "1st Version of the EaaS Tool Suite for Energy Retailers and Aggregators"

SYNERGY Consortium. (2021h). SYNERGY D7.2 "1st Version of the Optimized Energy Performance Management Suite"

SYNERGY Consortium. (2021i). SYNERGY D2.2 "End-user and Business requirements analysis for big data-driven innovative energy services and ecosystems v2"

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under Grant Agreement No 872734.