RESEARCH IN ENTERPRISES 2023

“An innovative AI framework for supporting the resilient operation of electricity distribution grids”

D10 – GridAI Requirements and Detailed Technical Design

Work Package	WP3 – Requirements definition and GridAI Technical Design		
Authors	Suite5		
Status	Final	**Classification**	Public
Date	31/01/2025	**Version**	1.0

Funded by the European Union NextGenerationEU

Republic of Cyprus

GridAI Project profile

RPF PROPOSAL No:	ENTERPRISES/ENERGY/1123/0018
Acronym:	GridAI
Title	An innovative AI framework for supporting the resilient operation of electricity distribution grids
URL:	https://gridai.eu/
Start Date:	01/08/2024
Duration	15 months

Partners

Suite5 We Deliver Intelligence	SUITE5 DATA INTELLIGENCE SOLUTIONS LIMITED (SUITE5)	Cyprus

Document History

Version	Date	Status	Author
1.0	31/01/2025	Final review, document ready for submission	Suite5

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Executive Summary

This deliverable, D10, summarizes the progress of two key tasks, name T3.1: End-user and Business Requirements Elicitation and T3.2:Technical Architecture Design and Semantic Interoperability Specifications, as outlined in the contract [3]. These tasks contribute to the broader objectives of WP3 by defining and documenting the requirements that will guide the development of the GridAI solution, drawing from previous and ongoing activities. The deliverable also outlines the technical design for the platform's architecture, detailing the necessary specifications for technology development and integration efforts within the project.

In the business domain, several business scenarios have been identified, which address both horizontal aspects of the data-related operations of Distribution System Operators (DSOs), along with business aspects that need to be considered and can be effectively addressed through data analytics and Artificial Intelligence (AI). These scenarios form the basis for the overall GridAI functionality and will inform the design and development of the solution.

From a technical perspective, a set of use cases has been developed, aligned with the business scenarios. These use cases help translate high-level business needs into concrete platform functionalities, ensuring the system addresses all user needs and expectations.

This deliverable also includes a comprehensive set of business and user requirements, primarily focusing on functional aspects, data availability, quality, interoperability, security, and data analytics. These requirements provide a solid foundation for the development of various bundles, components, and functionalities, ensuring alignment with GridAI project objectives.

The deliverable also offers a comprehensive overview of the components within the platform's bundles, along with the technological infrastructure implemented to support each bundle, including:

- **Energy Data Collection Services Bundle**: This includes components responsible for data collection, semantic interoperability, data cleaning, and metadata management, ensuring reliable and consistent data acquisition for the system.
- **Energy Data Security Services Bundle**: This bundle encompasses components related to data storage (including backup and recovery mechanisms), data sovereignty, data governance, security, data model storage and user authentication all aimed at protecting data integrity and preventing unauthorized access.
- **Data Analytics Services Bundle**: This bundle includes an array of AI models that address the forecasting, profiling and outlier analysis needs of DSOs, along with the required components that enable AI-driven insights generation through a comprehensive dashboard that will be made available for analytics results' interpretation and decision support.
- **GridAI Platform Horizontal Services Bundle**: This horizontal service focuses on the orchestration, operation, and scalability of the platform. It includes key components such as Platform management & orchestration, data search, data retrieval, data model management.

The deliverable provides an in-depth overview of the requirements and specifications that will drive the development of the GridAI solution, as well focusing on security, interoperability, data management, and advanced analytics capabilities. This comprehensive approach will ensure the platform can effectively address the data and analytics needs of DSOs, driving efficiency, resilience, and data-driven decision-making across grid operations.

Table of Contents

List of Figures

List of Tables

Glossary

Acronym	Full name
AI	Artificial Intelligence
API	Application Programming Interface
BS	Business Scenario
CSV	Comma-separated values
D	Deliverable
DER	Distributed Energy Resources
DL	Deep Learning
DSO	Distribution System Operator
EMS	Energy Management System
EV	Electric Vehicle
GUI	Graphical User Interface
ICT	Information and Communication Technology
JSON	JavaScript Object Notation
ML	Machine Learning
PubSub	Publish-Subscribe
RES	Renewable Energy Sources
T	Task
TSO	Transmission System Operator
UC	Use Case

UI	User Interface
XML	Extensible Mark-up Language
WP	Work Package

1. Introduction

1.1 Purpose of the document

The purpose of this deliverable is to provide a comprehensive framework for the development and integration of the GridAI platform, focusing on both business and technical aspects. It establishes the necessary foundation by capturing key user and business requirements, translating them into actionable specifications for the GridAI technical bundles, components and high-level architecture. Through the identification of critical business scenarios and use cases, this deliverable ensures that the GridAI platform will meet the data-related and analytics-related needs of DSOs.

Additionally, it outlines the key bundles, components and functionalities that constitute the GridAI solution, and will be further documented, elaborated and illustrated in its prototype release in M10 of the project in the context of Deliverable D12 [5], addressing data management, security, analytics, and platform orchestration, with an emphasis on ensuring scalability, security, and interoperability.

1.2 Structure of the Document

The structure of this document is as follows:

- Section 2 provides and elaborates on the business scenarios driving the GridAI solution, covering the business aspects of the electricity data value chain and electricity market sector;
- Sections 3 and 4 detail the process and outcomes of defining the system use cases and requirements for the GridAI solution
- Section 5 depicts the architecture of the GridAI solution, along with an overview of the various bundles and components involved in the technical specifications.
- Section 6 concludes the document by summarizing the work completed and outlining the next steps for the development and delivery of the GridAI solution. It highlights the final stages of development, focusing on the planned releases of key components:
    - GridAI Prototype Solution (M10): This release will showcase the full functionality of the GridAI platform for DSOs, including data governance, AI analytics pipelines, and the GridAI dashboard. It will demonstrate how the platform aligns with the project's requirements in a real-world context.
    - GridAI Technical Verification Results (M12): This report will provide the results of technical testing, including assessments of system integration, performance, and scalability, ensuring that the platform meets all the defined requirements.

2. Business Scenarios in GridAI

The increasing digitalization of the electricity domain, coupled with the widespread deployment of smart meters and the growing penetration of Distributed Energy Resources (DERs), has fundamentally transformed the way electricity grids are monitored and managed. Distribution System Operators (DSOs) face key challenges related to the complexity of grid operations, the significant increase in data from the grid edge, and the need for real-time, data-driven decision-making. GridAI addresses these challenges by delivering an integrated, AI-driven solution designed to support DSOs in optimizing critical operations and enhancing grid resilience based on AI-driven insights and improved forecasting capabilities.

The business scenarios outlined in this chapter reflect the high-level, contextual business value that GridAI is expected to provide to DSOs. Each scenario is defined by its business context and challenge, as well as the business functions and anticipated benefits that will be realized through the implementation of GridAI.

2.1 Effective management of DSO data, ensuring high-quality, reliability and interoperability (BS1)

Business Context and Challenge

DSOs face the critical task of harnessing data from their diverse asset portfolio (smart meters, DERs) to optimize grid operations within a complex, DER-rich environment. Data heterogeneity—varied granularity, assumptions, and quality—compounded by siloed storage and disparate modeling standards, presents a significant integration challenge.

Key Stakeholder Needs

To effectively navigate the complexities of modern grid management, Distribution System Operators (DSOs) require robust data management and curation tools. These tools must facilitate timely data collection from diverse sources, enable seamless data harmonization and semantic understanding, ensure data quality through rigorous cleaning processes, guarantee compliance with privacy regulations like GDPR, and provide comprehensive control over data storage and access.

Business Functions and Anticipated Benefits

This approach empowers DSOs to gain deeper insights into their data, significantly reduce integration complexities, and inherently enhance data quality. Consequently, data collection and maintenance become more efficient over time, ensuring reliable availability for appropriate (re)use.

2.2 Maintenance of full control of DSO data, under enhanced security and sovereignty terms (BS2)

Business Context and Challenge

Digitalization and smart metering are generating an exponential growth of data assets across the electricity grid. However, effective data management (and in certain cases, exchange with external stakeholders that are in need of DSO data) is hampered by trust, confidentiality, and security concerns.

Key Stakeholder Needs

In this context, DSOs are in need of robust mechanisms that can effectively secure and safeguard their data assets, against any confidentiality and unauthorized access risk. Utilization of distributed data

assets in AI pipelines needs to be associated to solid mechanisms for access control and privacy preservation, so that they can effectively comply with relevant regulations (GDPR, Data Act) and ensure immutable and non-repudiable data exchange with other stakeholders are entitled to access specific data portions for realizing their operations (e.g. electricity retailers for billing purposes).

Business Functions and Anticipated Benefits

In this context, a bullet-proof framework can be created, increasing trust of DSOs in digital services and AI analytics operations, thus paving a secure and trustworthy pathway to their digital transformation and transition.

2.3 Advanced Demand Analysis and Forecasting for Grid Stability (BS3)

Business Context and Challenge

The electricity demand side is progressively transforming from a passive energy system element into an active player, affecting the majority of operations and respective decisions for the distribution grid operational management. Demand analysis and forecasting are considered valuable tools in the hands of DSOs towards promptly identifying anticipated imbalances, peak periods or congestions and proactively defining effective strategies that take advantage of cheap demand flexibility sources in order to ensure stability and resilience of electricity grids.

Key Stakeholder Needs

In fact, in order to achieve a secure and uninterrupted operation of the electricity grid, DSOs need to take informed decisions and define operational plans based on knowledge captured and extracted from demand side data (e.g. smart meters) regarding its anticipated short- and mid-term behaviour (consumption) across their whole portfolio of customers and in varying spatial (substation, whole portfolio) and temporal granularities (from 15 mins to day- and 72 hours-ahead).

Business Functions and Anticipated Benefits

Having this knowledge in hand, they will be able to promptly anticipate for possible congestions in specific segments of the grid (substation level), imbalances that may occur and affect voltage preservation within secure operation limits (substation-level, portfolio-level) and peak consumption periods that may drive further decisions and actions for the triggering of flexibility events to shed/ shift demand and optimize the relevant demand curves. To this end, accurate demand forecasts are required, starting from 72 hours-ahead forecasts (allowing network operators to prepare themselves for anticipated events and trigger market mechanisms) to day-ahead and intraday (hourly and 15-mins ahead), according to the timelines of the respective energy and flexibility markets.

2.4 Advanced Distributed Generation Analysis and Forecasting for Grid Stability (BS4)

Business Context and Challenge

RES Generation analysis and forecasting are considered valuable tools in the hands of DSOs towards promptly identifying anticipated imbalances (when analysing generation forecasts together with respective demand forecasts) or congestions and proactively defining effective strategies that take advantage of cheap demand flexibility sources or distributed storage in order to ensure stability and resilience of electricity grids.

Key Stakeholder Needs

In order to achieve a secure and uninterrupted operation of the electricity grid, DSOs need to take informed decisions and define operational plans based on knowledge captured and extracted from distributed Renewable Energy Sources (RES), regarding its anticipated output in the very short and short-term across their networks and in varying spatial (substation, whole portfolio) and temporal granularities (from 15 mins to day- and 72 hours-ahead).

Business Functions and Anticipated Benefits

The availability of such accurate forecasts can equip DSOs with the capability to promptly anticipate for possible congestions in specific segments of the grid (substation level), imbalances that may occur across whole portfolios and may affect severely the secure operation of their networks. To this end, accurate RES generation forecasts are required, starting from 72 hours-ahead (allowing network operators to prepare themselves for anticipated events and trigger market mechanisms) to day-ahead and intraday (hourly and 15-mins ahead), according to the timelines of the respective energy markets functioning.

2.5 Historical Portfolio Data Analytics and Correlation with external factors for supporting longer-term planning of distribution grids (BS5)

Business Context and Challenge

With the increasing penetration of DERs across distribution grids, DSOs are faced with the challenge of adapting to constantly changing consumption and generation patterns that need to be considered when defining longer-term operational plans (that could ultimately drive investment decisions or the contracting of flexibility assets). Such patters are not solely affected by the introduced DERs themselves, but are also dependent to external factors such as local weather conditions (and their alterations due to climate change), energy market prices and extraordinary events that may severely affect the demand side (e.g. similar to the ones experienced during the COVID lockdown period).

Key Stakeholder Needs

The continuously evolving energy system requires DSOs to obtain a better understanding of the behaviour of the various DERs and assets operating in their networks, so that they can proceed with informed decisions to the long-term planning of their networks, anticipate events that may occur under specific circumstances and evidently drive grid reinforcement investments or contracting of additional flexibility sources for the secure operation of their grids. To this end, more comprehensive analysis of historical demand and generation data is required to reveal valuable performance insights over their portfolio (e.g. patters referring to the timing of demand and generation peaks, demand behaviour across days, weeks, months and seasons), identify previously non-observed relationships between demand, generation and external factors (weather, energy prices, revealing operational insights that can be associated to extreme weather conditions or increased market prices) and understand the impact of extraordinary events upon the demand side, thus enhancing their preparedness in case of similar occurrences in the future.

Business Functions and Anticipated Benefits

Historical energy data analysis is considered as a powerful tool in the hands of DSOs, towards increasing their understanding on the operation of the various assets operating in their networks, revealing key insights and patterns that can effectively support longer-term planning of their operations, allowing them to analyse specific conditions associated with incidents that have occurred in their networks (asset-related or external ones) and increasing their awareness and preparedness

for extraordinary events that require tailored strategies to ensure the uninterrupted supply of electricity towards the demand side.

2.6 Comprehensive Energy Generation and Demand Visualization (BS6)

Business Context and Challenge

With the increasing number of DERs and fluctuating energy demand, continuous monitoring and visualization of both energy generation and consumption data have become critical. DSOs need a unified view of generation and demand data to make quick adjustments and ensure grid stability.

Key Stakeholder Needs

DSOs need access to continuous streams of generation data alongside demand data to make informed decisions that keep the grid balanced and stable. A unified dashboard that consolidates both sets of data will allow operators to view the entire grid's performance and take immediate action in case of unexpected fluctuations. This will help them avoid issues like over-generation, under-generation, or demand imbalances.

Business Functions and Anticipated Benefits

The business function here is to provide real-time, interactive visualizations of both generation and demand data. The benefits include improved real-time decision-making, enhanced grid stability, and the ability to react quickly to unexpected changes in demand or generation. By doing so, DSOs can maintain optimal grid operations and gain a comprehensive view of energy demand and generation trends, enabling better operational control.

3. GridAI Use Cases

The use case elicitation process follows the IEC 62559-2 standard (IEC, 2015) [1], offering a structured framework to define and document each use case. This methodology outlines not only the core objectives but also specifies the roles and responsibilities of the various stakeholders involved. It details the actions and interactions between these stakeholders, as well as the tools and technologies developed within the project to achieve the desired outcomes. Additionally, the process captures the preconditions and postconditions essential for each use case's successful implementation, ensuring clarity and precision in execution.

In total, 13 use cases have been identified, covering the full scope of Work Package 4. Each use case is prioritized to reflect its significance within the broader project context. By adhering to this standardized approach, it is ensured that every aspect — from stakeholder engagement to technical requirements — is well-defined, facilitating a clear and effective path toward achieving the project's objectives.

3.1 Use Cases Overview

The following table presents a comprehensive overview of GridAI's use cases. This structured summary provides a clear insight into the diverse functionalities and impact areas of GridAI.

Table 1 GridAI Use Cases

Use Case ID	Use Case Title
GridAI_UC_.1	*Efficient, multi-modal ingestion of real-time and batch data*
GridAI_UC_.2	*Seamless semantic interoperability of energy data*
GridAI_UC_.3	*Data Quality Enhancement through advanced data cleaning methods*
GridAI_UC_.4	*Data Security and Confidentiality preservation*
GridAI_UC_.5	*Performant search for data assets*
GridAI_UC_.6	*Efficient data exploration*
GridAI_UC_.7	*AI Analytics Pipelines Design and Configuration*
GridAI_UC_.8	*AI Analytics Pipelines Management and Execution*
GridAI_UC_.9	*AI-powered Grid-level energy demand and generation forecasting*
GridAI_UC_.10	*Interactive visualization and reporting from the raw data assets and/or derivative AI analytics results*

3.2 GridAI Use Cases Detailed Breakdown

This section provides an in-depth representation of the use cases defined throughout the project. The detailed documentation, presented in Table 2 to Table 11, follows the IEC 62559-2 standard, but further incorporates the correlation among the different UCs, as well the interdependencies with the Business Scenarios outlined in Section 2 of the deliverable, ensuring a structured and comprehensive overview of each use case.

Table 2 GridAI Use Case 1 documentation, following the IEC 62559-2 Standard

Title	Efficient, multi-modal ingestion of real-time and batch data
ID	GridAI_UC_1
Description	DSOs require a robust and adaptable system to collect and manage both real-time and batch data while ensuring transparency, trust, and interoperability. A secure and highly configurable data ingestion mechanism need to be provided, allowing stakeholders to efficiently gather and exchange energy data without modifying their internal systems. For batch data, ingestion shall occur through direct file uploads or via APIs exposed by DSOs' systems, enabling seamless integration with existing processes. Real-time data, on the other hand, need to be collected using advanced streaming mechanisms that ensure continuous data flow, supporting dynamic grid operations and timely decision-making. The entire ingestion process shall be designed to be flexible, allowing DSOs to choose the methods that best suit their operational needs while maintaining full control over their data. To ensure data integrity and security, all ingestion mechanisms shall rely on trusted and protected data containers, providing a secure environment for data storage
Priority	5
Triggering Event	N/A
Preconditions	1. The availability of data assets (in batch or real-time format) from the DSOs
Postconditions	1. The data assets are properly provisioned and are made available for further processing based on the needs of the DSO
Correlation to Business Scenarios	BS1

Table 3 GridAI Use Case 2 documentation, following the IEC 62559-2 Standard

Title	Seamless semantic interoperability of energy data
ID	GridAI_UC_2
Description	Data in the energy sector often originates from diverse sources, following various standards and formats. This inconsistency poses significant challenges to the integration and effective use of data. To overcome this, the GridAI solution shall implement a standardized data mapping process that transforms disparate data into a unified format based on a unified data model, ensuring smooth interoperability across different systems and stakeholders. The solution shall include a mechanism for annotating and aligning data with appropriate semantics, regardless of the original standards or organizational conventions. This will address internal inconsistencies such as differing naming conventions or data types (e.g., datetime formats in sensor measurements). By facilitating the mapping of ingested data to the GridAI Data Model, the solution shall ensure that data is automatically aligned with pre-defined guidelines, enabling consistent and seamless data management.
Priority	5
Triggering Event	Provision of a new data asset
Preconditions	1. The availability of data assets originating from multiple diverse data sources and in multiple formats. 2. A flexible and highly configurable mechanism that will enable the ingestion of real-time or batch energy data assets
Postconditions	1. Heterogeneous data assets are properly represented with the appropriate context and semantics and are available for further exploitation.
Correlation to Business Scenarios	BS1

Table 4 GridAI Use Case 3 documentation, following the IEC 62559-2 Standard

Title	Data Quality Enhancement through advanced data cleaning methods
ID	GridAI_UC_3
Description	Ensuring high-quality data is crucial for DSOs, as data quality impacts decision-making and operational effectiveness. A robust mechanism shall be provided to guarantee that ingested data meets necessary standards for consistency, accuracy, and relevance over time. This mechanism shall enable seamless data discovery, retrieval, and monitoring to ensure data remains valuable and usable. DSOs shall have the ability to define and customize data cleaning methods to fit their specific needs. The solution shall also offer actionable recommendations and guidelines to continuously improve data quality, ensuring best practices are followed. This shall include automated or user-defined corrective actions to address incomplete, inaccurate, or missing data. Data quality metrics shall enhance the perceived value of data assets. Continuous monitoring shall ensure data freshness, consistency, and compliance with predefined timeliness and accuracy standards, providing DSOs with reliable, high-quality data for their operations.
Priority	4
Triggering Event	Provision of a new data asset
Preconditions	1. The availability of data assets (in batch or real-time format) from the data value chain stakeholders. 2. A flexible and highly configurable mechanism that will enable the ingestion of real-time or batch energy data assets 3. The availability of a holistic mechanism that will facilitates the representation and interpretation of energy data assets from diverse sources
Postconditions	1. The data health is efficiently monitored.
Correlation to Business Scenarios	BS1

Table 5 GridAI Use Case 4 documentation, following the IEC 62559-2 Standard

Title	Data Security and Confidentiality preservation
ID	GridAI_UC_4
Description	Data security and privacy are critical concerns for DSOs. To ensure comprehensive protection, the solution shall enable DSOs to define security measures during the design phase, which will be enforced at runtime to prevent data breaches, mitigate exposure risks, and ensure regulatory compliance. The solution shall incorporate a variety of mechanisms to safeguard data confidentiality, privacy, and integrity. Privacy-preserving techniques shall address concerns related to sensitive or personal information, ensuring compliance with legal and regulatory standards. Additionally, DSOs shall be able to define access policies and licenses, to prevent unauthorized access and minimize the risk of malicious users or competitors accessing critical data.
Priority	5
Triggering Event	Provision of a new data asset
Preconditions	1. The availability of data assets (in batch or real-time format) from the data value chain stakeholders. 2. A flexible and highly configurable mechanism that will enable the ingestion of real-time or batch energy data assets. 3. The availability of a holistic mechanism that will facilitates the representation and interpretation of energy data assets from diverse sources.
Postconditions	1. All the data assets of a data provider are stored in a secure and trusted storage that is empowered with a set of data security and privacy-ensuring mechanisms in order to be further exploited and combined in an efficient manner.
Correlation to Business Scenarios	BS2

Table 6 GridAI Use Case 5 documentation, following the IEC 62559-2 Standard

Title	Performant search for data assets
ID	GridAI_UC_5
Description	Efficiently discovering and accessing relevant data assets is crucial for energy stakeholders looking to expand their data-driven services. A performant search for data assets enables stakeholders to identify valuable data, whether it comes from open sources, proprietary data from other industry players, or broader assets like trained models and accompanying reports or visualizations. This process requires robust mechanisms to search and identify data based on key attributes such as content, origin (owner), generation methods, and volume.
Priority	5
Triggering Event	A user wants to create a service without having at his/her disposal all the necessary data assets
Preconditions	1. Availability of data, analysis results, reports, visualizations allowing for automated consumption.
Postconditions	1. The user is able to consume the data asset.
Actors Involved	Distribution System Operator, Aggregators, Retailers
Correlation to Business Scenarios	BS3, BS4, BS5

Table 7 GridAI Use Case 6 documentation, following the IEC 62559-2 Standard

Title	Efficient data exploration
ID	GridAI_UC_6
Description	Interactive data exploration plays a critical role in unlocking the potential of existing data assets for analytics services. By enabling DSOs to gain deeper insights into data characteristics and uncover hidden patterns, it accelerates the development of new services while improving the overall user experience. However, the large volume and diverse nature of energy data can make exploration complex. Therefore, it is essential for users to have tailored mechanisms that support seamless exploration. Beyond basic aggregations and charts, effective data exploration allows users to filter and create relevant subsets of data, helping them assess its value and potential for use in various energy services.
Priority	5
Triggering Event	N/A
Preconditions	1. Availability of data required for domain specific analysis.
Postconditions	1. A clearer understanding of a data asset's content, including outliers, data distribution, key statistics, correlations, and conditional summaries.
Correlation to Business Scenarios	BS3, BS4, BS5

Table 8 GridAI Use Case 7 documentation, following the IEC 62559-2 Standard

Title	AI Analytics Pipelines Design and Configuration
ID	GridAI_UC_7
Description	The development of intelligent energy services relies on advanced analytics and insights that provide a knowledge base for optimizing functions performed by DSOs. To address this need, the goal is to offer an analytics pipeline designer that allows users to: (a) manipulate and aggregate data to prepare it for core analysis; (b) apply AI-based baseline algorithms to extract valuable insights from raw data streams; and (c) iteratively refine the pipeline until the results meet the desired criteria. The objective is to provide an end-to-end environment that simplifies the work of energy data scientists, enabling them to easily configure various steps of the analytics process, from basic data manipulation to the integration of machine learning or deep learning techniques in a plug-and-play setup for seamless execution.
Priority	5
Triggering Event	Data manipulation functionalities (including filtering, feature extraction, analytics) in the form of re-usable services have been developed and can be used, Baseline algorithms in the form of re-usable services have been developed and can be used
Preconditions	1. An AI analytics pipeline is successfully designed, configured, and ready for execution. 2. Data has been properly pre-processed and is ready for analysis.
Postconditions	1. An analytics pipeline configuration made available for further exploitation during the execution phase.
Correlation to Business Scenarios	BS3, BS4, BS5

Table 9 GridAI Use Case 8 documentation, following the IEC 62559-2 Standard

Title	AI Analytics Pipelines Management and Execution
ID	GridAI_UC_8
Description	Once the data analytics pipelines are designed and configured, their execution must still align with the specific needs and preferences of DSOs. The system should provide a unified execution environment where DSOs can run analytics models either on-demand or at scheduled intervals. The focus is on using raw data or performing computations based on available datasets. The execution environment should support users by ensuring that data usage and analysis can occur in a way that respects the type of data available, adheres to security/ privacy policies, and generates valuable insights. Advanced monitoring is essential to ensure the health and reliability of the analytics pipelines, with automated alerts and troubleshooting tools that provide complete visibility across all stages of the data pipeline process. This approach allows users to track the performance and integrity of the pipeline, ensuring efficient execution and addressing any issues that may arise in real time.
Priority	5
Triggering Event	Design/Configuration of an analytics pipeline, data assets made available
Preconditions	1. The pipeline has been successfully designed and configured. 2. The required data for the pipeline execution is available (either real-time or batch data). 3. Security and data-sharing policies are in place, ensuring compliant data usage during the analysis.
Postconditions	1. The result of the analytics pipeline is available to the location of preference of a stakeholder
Correlation to Business Scenarios	BS3, BS4, BS5

Table 10 GridAI Use Case 9 documentation, following the IEC 62559-2 Standard

Title	AI-powered Grid-level energy demand and generation forecasting
ID	GridAI_UC_9
Description	To establish a sustainable and resilient operational environment across the electricity grid, accurate and detailed energy demand and generation forecasting is essential. Advanced AI analytics techniques shall be employed to derive demand and generation forecasts at grid and substation level. The focus shall be on developing flexible methods that allow forecasting models to adapt to different business scenarios and contexts. These models shall accommodate various time scales, ranging from 15-minute intervals to forecasts for days ahead. They shall also cover a broad scope, from individual assets to substations and entire network areas, ensuring that the needs and priorities of DSOs are met.
Priority	5
Triggering Event	The need for forecasting arises based on real-time operational requirements, upcoming energy demand, or the specific time horizon (e.g., future energy predictions for substations or communities).
Preconditions	1. Forecasting models are designed and configured using AI techniques (e.g., time-series analysis, machine learning). 2. Grid operators have defined the required time scale for the forecast (e.g., 15-min intervals, daily, days ahead).
Postconditions	1. Accurate and actionable forecasts for energy demand and generation are produced for the specified time period.
Correlation to Business Scenarios	BS3, BS4, BS5

Table 11 GridAI Use Case 10 documentation, following the IEC 62559-2 Standard

Title	Interactive visualization and reporting from the raw data assets and/or derivative AI analytics results
ID	GridAI_UC_10
Description	To support DSOs in managing and interpreting large volumes of complex, real-time data, intuitive and interactive user interfaces (UI) that enhance decision-making and operational efficiency are needed. DSOs shall be able to visualize key metrics like energy demand and generation, along with their forecasted timeseries, effectively filtering them and correlating them with external factors to extract valuable insights and advance evident decision making for the resilient operation of their assets. The GridAI dashboard shall offer users a wealth of visualizations, using a range of built-in charts, and consolidate multiple charts for each analytics pipeline in a user-friendly manner. The dashboard shall support real-time interaction, enabling users to drill down into data, apply filters, and view trends for deeper insights.
Priority	5
Triggering Event	A dataset and/or the results of an analytics pipeline execution are available.
Preconditions	1. Visualizations are generated for energy metrics. 2. Reports and Charts consolidated for each analytics pipeline in a user-friendly manner
Postconditions	1. A visualization dashboard has been created and can be accessed.
Correlation to Business Scenarios	BS6

4. Definition of GridAI Requirements

The definition of GridAI platform's technical and business requirements follows the ISO/IEC/IEEE 29148:2018 [2] international standard, which outlines best practices for systems and software engineering, life cycle processes, and requirements engineering. The objective was to transform the needs expressed in the Business Scenarios and Use Cases into well-defined requirements that include relevant constraints and conditions. These requirements serve as a bridge between stakeholder expectations and the technical specifications needed to develop a system capable of delivering the desired services, as outlined in ISO/IEC/IEEE 29148:2018.

To ensure effectiveness, GridAI's requirements adhere to the SMART criteria:

- Specific – Clearly defined and focused on particular functionalities.
- Measurable – Quantifiable to track progress and success.
- Attainable – Realistic and achievable given available resources.
- Relevant – Aligned with stakeholder needs and project goals.
- Timely – Defined within a feasible timeframe for implementation.

The requirements elicitation process aimed at establishing a well-structured set of requirements that:

- ✓ Promote a shared understanding among diverse stakeholders.
- ✓ Address real-world needs across the entire electricity data value chain.
- ✓ Provide a strong foundation to guide the design and development of the **GridAI** solution.

To ensure alignment with the ISO/IEC/IEEE 29148:2018 standard, specific guidelines were established regarding both syntax and linguistic parameters. Adhering to these guidelines was essential to maintain uniformity, clarity, and effectiveness throughout the requirement definition process.

Each requirement had to meet the following fundamental principles:

- Be concise and focused, addressing a single concept while remaining implementation-independent.
- Maintain clarity and precision, avoiding ambiguity or vague language.
- Be traceable, allowing easy identification of its origin and relationship with other requirements.
- Ensure consistency, avoiding contradictions within the system or across different requirements.
- Be comprehensive, providing all necessary details without leaving gaps.
- Remain easily understandable, ensuring stakeholders from different backgrounds can interpret it correctly.
- Be verifiable, meaning it can be tested or validated against objective criteria.
- Be realistic and achievable, considering technological and resource constraints.
- Stay relevant and necessary, maintaining long-term applicability, or, if subject to change over time, be accompanied by provisions for modification or replacement.

Through this process, a total number of 120 requirements were defined and are analyzed in the following paragraphs.

4.1 Requirements analysis overview

The table below, titled "**UCs - Requirements (**Table 12**)"** outlines the direct relationship between each Use Case (UC) and its associated requirements. This table is designed to provide a clear and organized overview of which specific requirements are linked to each UC, offering a comprehensive understanding of the underlying needs and constraints for each use case. By analyzing this table, one can gain valuable insights into how the requirements align with the various use cases, aiding in the deeper analysis and classification of GridAI's requirements.

Table 12 UC - Requirements relationships

UCs - Requirements
GridAI_UC_1 (15 Requirements)
Req_001, Req_002, Req_003, Req_004, Req_005, Req_006, Req_007, Req_008, Req_009, Req_010, Req_018, Req_020, Req_027, Req_028, Req_038
GridAI_UC_2 (13 Requirements)
Req_003**,** Req_014, Req_015, Req_016, Req_017, Req_018, Req_019, Req_021, Req_022, Req_023, Req_024, Req_025, Req_026
GridAI_UC_3 (11 Requirements)
Req_003**,** Req_006**,** Req_018**,** Req_019**,** Req_020**,** Req_025**,** Req_027, Req_028, Req_029, Req_030, Req_031
GridAI_UC_4 (11 Requirements)
Req_011, Req_012, Req_013, Req_032, Req_033, Req_034, Req_035, Req_036, Req_037, Req_038, Req_043
GridAI_UC_5 (11 Requirements)
Req_004, Req_011, Req_012, Req_015, Req_039, Req_040, Req_041, Req_042, Req_043, Req_044, Req_045
GridAI_UC_6 (3 Requirements)
Req_037, Req_044, Req_045
GridAI_UC_7 (26 Requirements)
Req_047, Req_048, Req_049, Req_050, Req_051, Req_052, Req_053, Req_054, Req_055, Req_056, Req_068, Req_069, , Req_070, Req_071, Req_072, Req_073, Req_074, Req_075, Req_079, Req_080, Req_081, Req_082, Req_083, Req_084, Req_085, Req_086
GridAI_UC_8 (37 Requirements)
Req_020, Req_027, Req_028, Req_047, Req_053, Req_054, Req_055, Req_057, Req_058, Req_059, Req_060, Req_061, Req_062, Req_063, Req_064, Req_065, Req_066, Req_067, Req_068, Req_069,

Req_070, Req_071, Req_072, Req_073, Req_074, Req_075, Req_076, Req_077, Req_078 , Req_079, Req_080, Req_081, Req_082, Req_083, Req_084, Req_085, Req_086,
GridAI_UC_9 (19 Requirements)
Req_068, Req_069, Req_070, Req_071, Req_072, Req_073, Req_074, Req_075, Req_076, Req_077, Req_078, Req_079, Req_080, Req_081, Req_082, Req_083, Req_084, Req_085, Req_086,
GridAI_UC_10 (34 Requirements)
Req_087, Req_088, Req_089, Req_090, Req_091, Req_092, Req_093, Req_094, Req_095, Req_096, Req_097, Req_098, Req_099, Req_100, Req_101, Req_102, Req_103, Req_104, Req_105, Req_106, Req_107, Req_108, Req_109, Req_110, Req_111, Req_112, Req_113, Req_114, Req_115, Req_116, Req_117, Req_118, Req_119, Req_120

4.2 Analysis of GridAI Requirements

The elaboration of the use cases in Section 3 formed the basis for defining the GridAI solution's requirements. During the requirements definition phase, 120 requirements were identified. Table 13 outlines both functional and non-functional requirements, specifying the necessary system capabilities. To ensure clarity, a consistent syntax is used to indicate priority levels: "The system should..." signifies desired features, while "The system shall..." denotes mandatory requirements.

Table 13 GridAI Requirements and UC relations

Requirement_ID	Requirement Description	Use Cases Relation
Req_001	GridAI shall allow the upload of data assets through APIs.	GridAI_UC_1
Req_002	GridAI shall allow the ingestion of real-time data assets published by the user's systems.	GridAI_UC_1
Req_003	GridAI shall transform and store all datetime data in UTC.	GridAI_UC_1, GridAI_UC_2, GridAI_UC_3
Req_004	GridAI shall enable the data asset retrieval once from Open Data APIs.	GridAI_UC_1, GridAI_UC_5
Req_005	GridAI shall allow the upload of data files (i.e., CSV, JSON, XML, etc.) to the platform.	GridAI_UC_1
Req_006	GridAI shall enable the ingestion of real-time data assets by subscribing to an external messaging system.	GridAI_UC_1, GridAI_UC_3
Req_007	GridAI should allow the periodic upload of data assets through APIs.	GridAI_UC_1
Req_008	GridAI should allow the scheduling of future upload of data assets through APIs.	GridAI_UC_1

Req_009	GridAI should enable the periodic data asset retrieval from Open Data APIs.	GridAI_UC_1
Req_010	GridAI should enable the scheduling of future data asset retrieval from Open Data APIs.	GridAI_UC_1
Req_011	GridAI should allow the update of data assets which have been already successfully checked in.	GridAI_UC_4, GridAI_UC_5
Req_012	GridAI should allow updates of data assets that have been ingested through Open Data APIs.	GridAI_UC_4, GridAI_UC_5
Req_013	GridAI should allow the deletion of data assets which have been already successfully checked in.	GridAI_UC_4
Req_014	GridAI shall allow the definition of the metadata that are necessary and that are not automatically calculated.	GridAI_UC_2
Req_015	GridAI shall provide predictions for the mappings of each concept of the data asset to the data model with different confidence levels.	GridAI_UC_2, GridAI_UC_5
Req_016	GridAI shall ensure that data transformations for datetime format and time zones are performed in the concepts of a data asset in accordance with the data model.	GridAI_UC_2
Req_017	GridAI shall ensure that data transformations for geographical coordinates are performed in the concepts of a data asset in accordance with the data model.	GridAI_UC_2
Req_018	GridAI shall allow the provision of samples for ingested data assets.	GridAI_UC_1, GridAI_UC_2, GridAI_UC_3
Req_019	GridAI shall process the samples of the data assets according to the mapping, cleaning, and anonymization rules that are applied to the data assets.	GridAI_UC_2, GridAI_UC_3
Req_020	GridAI shall allow the monitoring of the progress and status of ongoing data ingestion jobs.	GridAI_UC_1, GridAI_UC_3, GridAI_UC_8
Req_021	GridAI shall enable the selection of the GridAI Data Model for the data asset to be ingested.	GridAI_UC_2
Req_022	GridAI shall enable the semantic annotation of data assets based on the GridAI Data Model	GridAI_UC_2
Req_023	GridAI shall enable the manual editing of the mappings of each concept of the data asset to the equivalent concepts of the GridAI data model during the data ingestion configuration	GridAI_UC_2
Req_024	GridAI shall provide suggestions and guidelines during the mapping configuration process for each concept of the data asset.	GridAI_UC_2

Req_025	GridAI shall allow the definition of data harmonization and transformation rules that are to be performed in each field of a data asset in accordance with its mapping to the GridAI Data Model	GridAI_UC_2, GridAI_UC_3
Req_026	GridAI shall allow the running of the mapping configuration on sample data to show how the data are transformed to comply with the GridAI Data Model	GridAI_UC_2
Req_027	GridAI shall provide notifications when no data arrive after data ingestion execution.	GridAI_UC_1, GridAI_UC_3, GridAI_UC_8
Req_028	GridAI shall provide notifications when data with missing values arrive after data ingestion execution	GridAI_UC_1, GridAI_UC_3, GridAI_UC_8
Req_029	GridAI shall allow the configuration of cleaning rules for handling incorrect, incomplete, inaccurate, irrelevant, or missing parts of the data.	GridAI_UC_3
Req_030	GridAI shall enforce the application of manual rules for handling incorrect, incomplete, inaccurate, irrelevant, or missing parts of the data during data ingestion.	GridAI_UC_3
Req_031	GridAI shall monitor different health metrics for the data arriving after data ingestion	GridAI_UC_3
Req_032	GridAI shall enable the user to choose whether or not to anonymize the data assets to be ingested when configuring data ingestion	GridAI_UC_4
Req_033	GridAI shall allow the configuration of different data anonymization techniques to be applied on data	GridAI_UC_4
Req_034	GridAI shall perform (semi-)automatic data anonymization checks on the data sample during the configuration of data ingestion	GridAI_UC_4
Req_035	GridAI shall enable the configuration of fine-grained access policies for data assets	GridAI_UC_4
Req_036	GridAI shall enable the configuration of applicable licenses for data assets when a data asset is made available.	GridAI_UC_4
Req_037	GridAI shall enable the update of the applicable licenses for data assets	GridAI_UC_4, GridAI_UC_6
Req_038	GridAI shall provide a configurable data storage mechanism that allows the secure storage and retrieval of data assets	GridAI_UC_1, GridAI_UC_4
Req_039	GridAI shall provide views of the metadata of available assets.	GridAI_UC_5
Req_040	GridAI should provide views of the data samples for available data assets.	GridAI_UC_5

Req_041	GridAI shall allow searching for available assets based on the assets' contents.	GridAI_UC_5
Req_042	GridAI shall allow searching for available assets based on the assets' metadata.	GridAI_UC_5
Req_043	GridAI shall allow a user to review the license characteristics of assets owned by or obtained by his/her organization.	GridAI_UC_4, GridAI_UC_5
Req_044	GridAI should allow the export/ retrieval of a dataset to be downloaded locally by a user.	GridAI_UC_5, GridAI_UC_6
Req_045	GridAI should allow the export/ retrieval of a dataset through APIs	GridAI_UC_5, GridAI_UC_6
Req_046	GridAI should enable the configuration of a data analytics pipeline, selecting the analytics framework for its execution.	GridAI_UC_7
Req_047	GridAI should enable the editing of the data analytics pipeline details (name, description, configuration).	GridAI_UC_7, GridAI_UC_8
Req_048	GridAI should provide an interactive, graph interface in order to configure the data analytics pipeline.	GridAI_UC_7
Req_049	GridAI should enable the selection of the specific data assets that will set the input data of the analytics pipeline and configure the necessary parameters.	GridAI_UC_7
Req_050	GridAI shall enable the inclusion of different data manipulation methods available on the platform, in an analytics pipeline and configure the necessary parameters.	GridAI_UC_7
Req_051	GridAI shall enable the inclusion of different data analysis methods, including applying and evaluating machine learning and deep learning models available in the platform, in an analytics pipeline and configure the necessary parameters.	GridAI_UC_7
Req_052	GridAI shall provide a search functionality to enable search over the available data manipulation and analytics methods.	GridAI_UC_7
Req_053	GridAI shall enable the storage of the results of a data analytics pipeline and configure the necessary details.	GridAI_UC_7, GridAI_UC_8
Req_054	GridAI shall enable the export of the results of a data analytics pipeline.	GridAI_UC_7, GridAI_UC_8
Req_055	GridAI shall enable the test run of the data analytics pipeline over a sample of the input dataset.	GridAI_UC_7, GridAI_UC_8
Req_056	GridAI shall display the results of the test run of the data analytics pipeline.	GridAI_UC_7

Req_057	GridAI shall enable the finalization of a data analytics pipeline and running of the execution over the full (input) dataset.	GridAI_UC_8
Req_058	GridAI should enable the filtering of the available data analytics pipelines based on different criteria.	GridAI_UC_8
Req_059	GridAI should enable the search for a specific data analytics pipeline based on different criteria.	GridAI_UC_8
Req_060	GridAI should allow the configuration of the execution time of a data analytics pipeline.	GridAI_UC_8
Req_061	GridAI shall enable the monitoring of the execution incidents, performance, and health of a specific data analytics pipeline.	GridAI_UC_8
Req_062	GridAI shall enable tracking the history of past executions and check the status, incidents, and performance of each execution.	GridAI_UC_8
Req_063	GridAI shall provide a list of metrics to check the health status of the results derived from each analytics pipeline execution.	GridAI_UC_8
Req_064	GridAI shall enable the creation of custom visualizations on the results of a data analytics pipeline, utilizing a variety of built-in charts.	GridAI_UC_8
Req_065	GridAI shall enable the configuration and management of multiple charts per data analytics pipeline.	GridAI_UC_8
Req_066	GridAI shall enable saving visualizations, linking the visualization with the respective analytics pipeline.	GridAI_UC_8
Req_067	GridAI shall enable the export of the visualizations configured for a data analytics pipeline.	GridAI_UC_8
Req_068	GridAI shall support demand forecasting at portfolio level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_069	GridAI shall support 72 hours-ahead demand forecasts at portfolio and substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_070	GridAI shall support day-ahead demand forecasts at portfolio and substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_071	GridAI shall support hour-ahead demand forecasts at portfolio and substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_072	GridAI shall support 15 minute-ahead demand forecasts at portfolio and substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9

Req_073	GridAI shall support day-ahead forecasts of peak demand portfolio and substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_074	GridAI shall support day-ahead forecasts of peak demand at substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_075	GridAI shall support extraction of insights regarding demand outliers over portfolio's historical consumption data.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_076	GridAI shall support correlation of historical data and outliers with relevant events (mainly weather, national holidays, lockdown periods)	GridAI_UC_8, GridAI_UC_9
Req_077	GridAI shall support comparative insights between forecasted and actual data	GridAI_UC_8, GridAI_UC_9
Req_078	GridAI shall support correlation of how extreme weather or extraordinary events (outliers) have affected portoflio's consumption	GridAI_UC_8, GridAI_UC_9
Req_079	GridAI shall support demand forecasting at substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_080	GridAI shall support generation forecasting at grid level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_081	GridAI shall support 72-hour-ahead generation forecasting at grid and substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_082	GridAI shall support day-ahead generation forecasting at grid and substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_083	GridAI shall support hour-ahead generation forecasting at grid and substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_084	GridAI shall support 15 minutes-ahead generation forecasting at grid and substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_085	GridAI shall support 72-hour-ahead generation forecasting at grid and substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9
Req_086	GridAI shall support generation forecasting extraction at substation level.	GridAI_UC_7, GridAI_UC_8, GridAI_UC_9

Req_087	GridAI shall support the visualization of the historical overview of portfolio consumption	GridAI_UC_10
Req_088	GridAI shall support the visualization of daily energy consumption data (per hour)	GridAI_UC_10
Req_089	GridAI shall support the visualization of weekly energy consumption data	GridAI_UC_10
Req_090	GridAI shall support the visualization of monthly energy consumption data	GridAI_UC_10
Req_091	GridAI shall support the visualization of annual energy consumption data	GridAI_UC_10
Req_092	GridAI shall support zoom in and out in timeseries data visualisations	GridAI_UC_10
Req_093	GridAI shall support the visualization of the energy demand forecast for the next 72 hours as a timeseries	GridAI_UC_10
Req_094	GridAI shall support the visualization of comparative timeseries (actual demand vs forecasted demand) for the same period that the 72 hours-ahead demand forecast refers to	GridAI_UC_10
Req_095	GridAI shall support the visualization of the energy demand forecast for the next 24 hours as a timeseries	GridAI_UC_10
Req_096	GridAI shall support the visualization of comparative timeseries (actual demand vs forecasted demand) for the same period that the day-ahead demand forecast refers to	GridAI_UC_10
Req_097	GridAI shall support the visualization of the energy demand forecast for the next hour as a timeseries	GridAI_UC_10
Req_098	GridAI shall support the visualization of comparative timeseries (actual demand vs forecasted demand) for the same period that the hour-ahead demand forecast refers to	GridAI_UC_10
Req_099	GridAI shall support the visualization of the energy demand forecast for the next 15 minutes	GridAI_UC_10
Req_100	GridAI shall support the visualization of comparative timeseries (actual demand vs forecasted demand) for the same period that the 15 min-ahead demand forecast refers to	GridAI_UC_10
Req_101	GridAI shall support the visualization of the forecasted peak demand for the next day (at 15-min level, hour-level according to data granularity)	GridAI_UC_10
Req_102	GridAI shall support the visualization of the detected outliers over historical data	GridAI_UC_10

Req_103	GridAI shall support the visualization of a comparative analysis of actual and forecasted data for the same period that an outlier was detected	GridAI_UC_10
Req_104	GridAI shall support the visualization of an analysis of portfolio's demand between weekdays and weekends (temporal filters to be applied)	GridAI_UC_10
Req_105	GridAI shall support the visualization of an analysis of portfolio's demand across different areas (temporal and spatial filters to be applied)	GridAI_UC_10
Req_106	GridAI shall provide historical visual overview of portfolio generation	GridAI_UC_10
Req_107	GridAI shall support the visualization of daily generation data (per hour)	GridAI_UC_10
Req_108	GridAI shall support the visualization of weekly generation data	GridAI_UC_10
Req_109	GridAI shall support the visualization of monthly generation data	GridAI_UC_10
Req_110	GridAI shall support the visualization of annual generation data	GridAI_UC_10
Req_111	GridAI shall support the visualization of a combined view of portfolio energy generation and demand data in different temporal granularities	GridAI_UC_10
Req_112	GridAI shall support the visualization of the generation forecast for the next 72 hours as a timeseries	GridAI_UC_10
Req_113	GridAI shall support the visualization of comparative timeseries (actual generation vs forecasted generation) for the same period that the 72 hours-ahead generation forecast refers to	GridAI_UC_10
Req_114	GridAI shall support the visualization of the generation forecast for the next 24 hours as a timeseries	GridAI_UC_10
Req_115	GridAI shall support the visualization of comparative timeseries (actual generation vs forecasted generation) for the same period that the day-ahead generation forecast refers to	GridAI_UC_10
Req_116	GridAI shall support the visualization of the generation forecast for the next hour as a timeseries	GridAI_UC_10
Req_117	GridAI shall support the visualization of comparative timeseries (actual generation vs forecasted generation) for the same period that the hour-ahead generation forecast refers to	GridAI_UC_10
Req_118	GridAI shall support the visualization of the generation forecast for the next 15 minutes as a timeseries	GridAI_UC_10

Req_119	GridAI shall support the visualization of comparative timeseries (actual generation vs forecasted generation) for the same period that the 15 minutes-ahead generation forecast refers to	GridAI_UC_10
Req_120	GridAI shall support the visualization of an analysis of portfolio's generation across different areas/ substations (temporal and spatial filters to be applied)	GridAI_UC_10

5. GridAI Architecture

This section provides an overview of the high-level architecture, outlining the various bundles within the architecture, the components contained in each bundle, and the purpose of each component.

5.1 GridAI Bundles Overview

As described in the Contract [3], GridAI introduces a proposed end-to-end solution for managing and analyzing big data, primarily or secondarily related to the electricity domain and DSO Operations. This data will come from diverse sources such as historical energy demand and generation data from open sources, weather data, energy market data, and other relevant datasets. The solution is designed to assist Distribution System Operators (DSOs) in effectively and interoperably managing the data they own. It will also enable the integration of external data required for driving operational decisions (e.g., weather data, energy market data) and combine this data to enhance internal intelligence. Ultimately, this will support data-driven decision-making for the resilient operation of the distribution grid by leveraging accurate AI-based demand and generation forecasts across varying spatial and temporal resolutions.

In GridAI, extracting valuable intelligence from the distribution grid data assets, carrying important electricity-related information requires a novel approach. This approach is built around three Core Services Bundles and their respective components, as depicted in Figure 1, tailored to the electricity domain:

- **Energy Data Collection Services Bundle**, handling Data Ingestion, Mapping to a comprehensive Energy Data Model, developed already in the frame of D11 [4], and Cleaning. It includes also Metadata Management, enriching data assets with supplementary details to enhance discoverability, traceability, and contextual understanding.
- **Energy Data Security Services Bundle**, responsible for safeguarding and securing any data asset that becomes available within the GridAI solution. This bundle ensures that data remains protected throughout its lifecycle, from ingestion to processing and storage, while also enforcing governance policies and access controls.
- **Data Analytics Services Bundle** enabling exploratory data analysis, design and execution of analytics workflows (including alternative demand and generation forecasting pipelines addressing different spatio-temporal resolutions) for AI-based demand and generation forecasting.

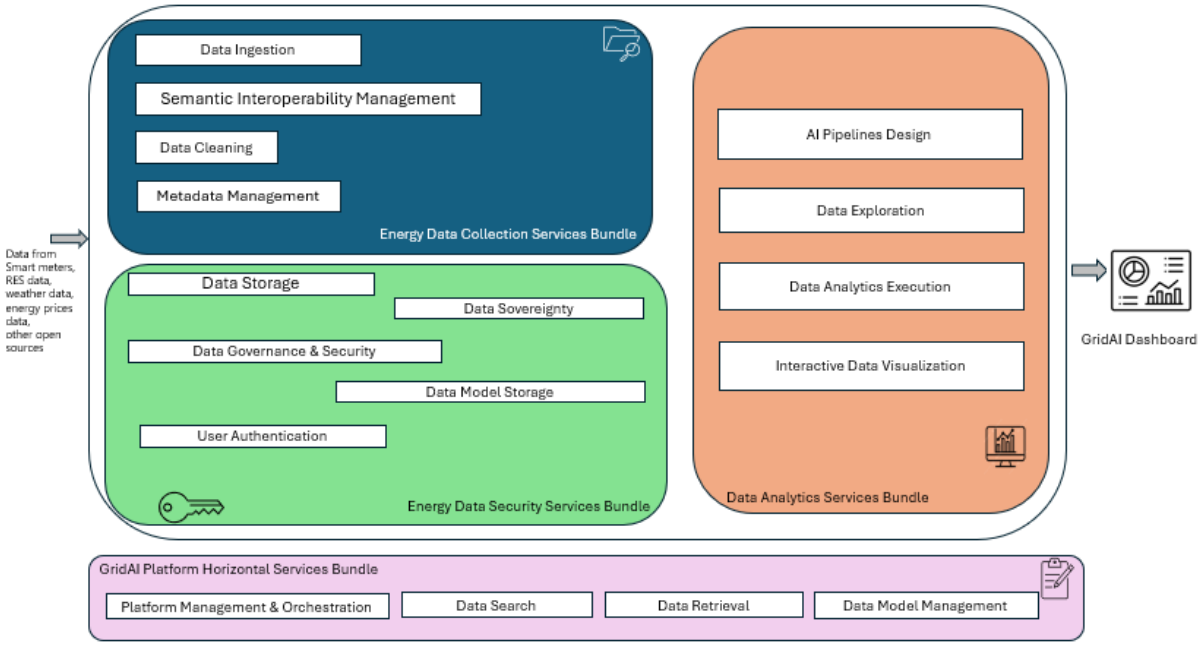

Figure 1 GridAI High-level Architecture

Additionally, GridAI includes the **GridAI Platform Horizontal Services Bundle**, a horizontal bundle that ensures efficient orchestration, operation, and scalability of the GridAI solution. This service supports user management, system performance optimization, and overall platform administration to maintain seamless functionality across all bundles. Moreover, it guarantees consistency, controlled modifications, and seamless alignment with the evolving data requirements of the GridAI Data Model.

5.2 Energy Data Collection Services Bundle

Energy Data Collection Services Bundle is a core element of the data collection process within the GridAI Platform and plays a pivotal role in the solution's overall functionality. This bundle supports a wide range of essential capabilities, including the ingestion of data from multiple sources, the application of semantic mapping to ensure semantic data interoperability, and the execution of data cleaning processes to guarantee the accuracy and quality of the data. Additionally, it encompasses metadata management to enhance data assets by adding supplementary details. The following sections provide a comprehensive overview of the components that constitute the Energy Data Collection Services Bundle.

5.2.1 Data Ingestion

The Data Ingestion Component is responsible for collecting and integrating data into the GridAI solution. It accommodates various ingestion methods to support stakeholder needs, including:

- Uploading datasets as files
- Retrieving data through APIs
- Streaming data using PubSub mechanisms

Each ingestion method comes with configurable options, such as scheduling API-based data collection, managing authentication (e.g., API keys and tokens), handling connection errors, and supporting

pagination. Given the sensitivity of API and PubSub parameters, they are securely stored in an encrypted form within the Energy Data Security Services Bundle.

For reliability, all ingested data is temporarily stored in an object storage, enabling rollback in case of failures in subsequent processing steps such as Semantic Interoperability Management, Data Cleaning, and Data Sovereignty.

5.2.1.1 Component Features

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

DIC1: Configurable ingestion settings - Supports flexible configuration of data collection, including scheduling, security, and error-handling.

DIC2: File-based data collection - Enables ingestion of structured files in formats like CSV, TSV, JSON, and XML.

DIC3: API-based data retrieval - Facilitates secure collection of data from third-party APIs and Open Data sources.

DIC4: Data Streaming via PubSub mechanisms - Allows real-time data ingestion from both GridAI-hosted and external PubSub systems.

DIC5: Pre-processing and storage - Filters and stores the required data subset in an object store, ensuring seamless handoff to downstream processing.

5.2.1.2 Features and Requirements linkage

Feature	Linked Requirements
DIC1	Req_001, Req_002, Req_004, , Req_005, Req_007, Req_008, Req_009, Req_010, Req_020, Req_027, Req_028, Req_035, Req_036, Req_037
DIC2	Req_005, Req_006, Req_007, Req_008
DIC3	Req_001, Req_004, Req_005
DIC4	Req_002, Req_005, Req_006
DIC5	Req_003, Req_016, Req_017, Req_018, Req_027, Req_028, Req_029, Req_030, Req_031

5.2.2 Semantic Interoperability Management

The Semantic Interoperability Management Component ensures that ingested datasets are transformed according to the GridAI Data Model by applying predefined mapping templates. These transformations standardize data structures, enforce format consistency, and align measurement units with expected specifications.

As the second step in the data collection process (following Data Ingestion), this component is crucial for enabling seamless integration across GridAI components. It ensures that datasets conform to standard types and formats by applying transformations such as attribute renaming, unit conversions, date formatting, and type casting.

To ensure data reliability, transformed datasets are temporarily stored before further processing, allowing rollback in case of failures in subsequent steps like Data Cleaning and Data Sovereignty.

5.2.2.1 Component Features

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

SIM1: Data Mapping to the GridAI Data Model – Responsible for the alignment of ingested data with the GridAI Data Model by renaming attributes (e.g., column names in tabular data, attribute names in JSON) based on predefined mapping templates.

SIM2: Data Transformation & Standardization - Converts data types, adjusts measurement units, reformats date/time attributes, and applies necessary transformations to ensure consistency with GridAI concept names.

SIM3: Transformation Insights & Reporting - Tracks and logs applied transformations, providing detailed reports on modifications for data providers to review and audit.

SIM4: Error Handling in Transformation Rules - Identifies and flags transformation errors, highlights failed rules, and prevents critical mapping issues.

5.2.2.2 Features and Requirements linkage

Feature	Linked Requirements
SIM1	Req_015, Req_021, Req_022, Req_023, Req_024, Req_026
SIM2	Req_003, Req_016, Req_017, Req_025
SIM3	Req_024, Req_026
SIM4	Req_020, Req_027, Req_028, Req_031

5.2.3 Data Cleaning

The Data Cleaning component is essential for ensuring that the data ingested into GridAI is accurate, consistent, and complete. This is achieved by adhering to custom rules defined during the data collection configuration process. The primary objective is to clean the data by removing or correcting incomplete, inconsistent, improperly formatted, or otherwise incorrect data. By doing so, the overall data quality is improved, which in turn ensures effortless reusability and more trustworthy insights extraction.

The Data Cleaning component is responsible for implementing functionalities that clean the data. These range from simpler tasks such as value substitution, reformatting, and duplicate removal to more advanced tasks like outlier detection and substitution.

Once the data is cleansed, it is stored in a temporary object storage for enhanced reliability and rollback options in the event of a failure during the subsequent pre-processing stages (Data Sovereignty).

5.2.3.1 Component Features

The Data Cleaning Component provides a set of features designed to support comprehensive and flexible data cleaning. These features allow users to configure, execute, and monitor data cleaning rules, ensuring high-quality data for further processing and analysis. The features offered by the Data Cleaning Component are as follows:

DC1: Flexible Configuration of Cleaning Rules - Allows users to define cleaning rules by specifying constraints and corrective actions for data quality assurance.

DC2: Data Cleaning Rules Execution - Executes data cleaning methods based on defined rules, such as validation options and corrective actions (drop/replace values).

DC3: Feedback from Executed Cleaning Rules - Provides feedback on the execution of cleaning rules, offering insights into how many times a rule was applied and if any issues occurred.

DC4: Easy Testing of Defined Cleaning Rules on Sample Data - Provides the functionalities to safely evaluate cleaning rules on sample datasets before applying them on full datasets.

5.2.3.2 Features and Requirements linkage

Feature	Linked Requirements
DC1	Req_025, Req_029, Req_030
DC2	Req_025, Req_029, Req_031
DC3	Req_020, Req_027, Req_031
DC4	Req_018, Req_026, Req_031

5.2.4 Metadata Management

The Metadata Management component of the GridAI solution ensures that all data assets generated during the data ingestion process are systematically organized and classified. Upon importing data, a comprehensive profile for each data asset must be created. This profile includes defining essential details such as the title, data type, format, language, temporal coverage, and temporal resolution units. These attributes are captured through an intuitive interface, utilizing drop-down menus for easy selection.

This structured approach to metadata enables efficient identification and discovery of data assets within the platform. The metadata index is regularly updated to reflect any modifications or additions to the data asset profile made by the data providers.

5.2.4.1 Component Features

The following feature is offered by the component:

MM1: Data Assets Metadata Storage - Along with the processed data assets, the Metadata Management component stores the accompanying metadata, creating the appropriate links between the stored data assets and their metadata information. This feature ensures that metadata is persistently stored and easily accessible for use across various components and services in the GridAI solution.

5.2.4.2 Features and Requirements linkage

Feature	Linked Requirements
MM1	Req_014, Req_038, Req_039, Req_042, Req_043

5.3 Energy Data Security Services Bundle

The Energy Data Security Services Bundle is central to the protection of data within the GridAI solution. It ensures that all data assets, including metadata, are securely stored and protected from unauthorized access. The bundle incorporates user authentication, which serves as the platform's

identity management system, enabling secure login and access control. Additionally, the Data Sovereignty feature allows users to define anonymization and privacy rules for the data they share, including options to exclude sensitive information and prevent unintended disclosures. This bundle also includes secure Data Storage mechanisms that ensure the safe retention of data and log information, supporting both operational data and metadata management. The Data Model Storage Component is also a part of this bundle, providing version control and security for the GridAI Data Model, ensuring that users work with the most up-to-date and accurate model for their data operations while maintaining its integrity. The following sections provide a comprehensive overview of the components that make up the Energy Data Security Services Bundle.

5.3.1 Data Storage

The Data Storage component provides comprehensive storage solutions for various data types within the GridAI solution. It ensures secure and scalable storage for processed datasets, which are stored in trusted data containers after ingestion. Additionally, metadata associated with these datasets is stored, establishing clear links between data and its metadata.

The component also accommodates the storage of AI models within dedicated containers for use in analytics pipelines. Temporary storage is available for intermediate files generated during data processing, enabling the pausing and continuation of the relevant processes in case of failure.

Furthermore, the Data Storage component secures sensitive data and credentials, such as API keys and passwords, in a dedicated database to ensure controlled access. Finally, it serves as a centralized repository for operational data, including platform data, user information, and operational metadata.

5.3.1.1 Component Features

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

DS1: Storage of Datasets in Trusted Data Containers - Scalable and secure storage for processed datasets after they are ingested. These datasets are stored in trusted data containers, ensuring optimized management, especially for big data.

DS2: Storage of AI Models – Storage of AI models used in analytics pipelines. It ensures that these models are saved for future use in data processing and analytics tasks

DS3: Storage of all Data Collection and Data Analytics Configurations – Storage of configurations related to data collection and data analytics. This includes configurations related to various components such as the Data Ingestion Component, Semantic Interoperability Management Component, etc.

DS4: Storage of Sensitive Data and Credentials - Provides a secure database for storing sensitive data like API keys, tokens, usernames, and passwords.

DS5: Storage for All Platform's Operational Data - Handles storage for all operational data, including platform data, user and organizational information.

5.3.1.2 Features and Requirements linkage

Feature	Linked Requirements
DS1	Req_003, Req_011, Req_012, Req_013, Req_018, Req_026, Req_038, Req_053, Req_049, Req_050, Req_063, Req_075
DS2	Req_053, Req_054

DS3	Req_025, Req_029, Req_035, Req_046, Req_047, Req_049 , Req_050, Req_051, Req_060, Req_061, Req_062
DS4	Req_032, Req_033, Req_034, Req_038
DS5	Req_011, Req_012, Req_013, Req_020, Req_031, Req_038, Req_039, Req_041, Req_042

5.3.2 Data Sovereignty

The Data Sovereignty component within GridAI plays a crucial role in protecting data assets from the unintended exposure of personal or corporate information. This component is responsible for defining the anonymization actions to be applied to data, specifically identifying which attributes need anonymization before making the data accessible within GridAI. As part of the Energy Data Security Services Bundle, the Data Sovereignty component helps data providers by:

- Identifying sensitive or personally identifiable information within data assets.
- Offering tools to anonymize this information, including any attributes considered potentially identifying when combined with others.

Since the Data Sovereignty component needs the complete dataset to function properly - ensuring consistency in data anonymization - it imposes certain configuration restrictions, particularly for streaming data or small datasets gradually acquired through APIs.

5.3.2.1 Component Features

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

DSo1: Identification of Attributes for Anonymization - The Data Sovereignty component assists in identifying dataset attributes containing sensitive or personally identifiable information. This can include identifying attributes or quasi-identifier attributes that, when combined, could identify individuals or organizations.

DSo2: Flexible Configuration of Anonymization Rules - Although data anonymization is performed during the data ingestion process, the Data Sovereignty component provides flexibility in configuring anonymization rules. This feature allows to define and adjust anonymization settings according to their needs, facilitated by backend services supporting these configurations.

DSo3: Data Anonymization Rules Execution - Based on the type of attribute (e.g., sensitive, identifying, or quasi-identifying) and its data type, the Data Sovereignty component enables data providers to select an anonymization method. Examples of supported methods include generalization (for creating arithmetic intervals or categories for numeric attributes) and masking (for string fields). The component does not offer a fully-fledged anonymization tool but provides essential anonymization functionalities that can be configured by the user.

DSo4: Insights Extracted from Executed Anonymization Rules - Data anonymization can significantly alter the original dataset, so it is essential for understanding the impact of anonymization rules. The Data Sovereignty component provides a utility to calculate the data loss resulting from anonymization, helping to assess the balance between protecting sensitive information and maintaining usable data. If the anonymization process leads to data loss exceeding the defined thresholds, the Data Sovereignty component will intentionally fail, allowing the user to review and adjust the anonymization rules.

5.3.2.2 Features and Requirements linkage

Feature	Linked Requirements
DSo1	Req_032, Req_033, Req_034
DSo2	Req_033, Req_034
DSo3	Req_016, Req_030, Req_033, Req_034
DSo4	Req_020, Req_031, Req_032, Req_033, Req_034

5.3.3 Data Governance & Security

The Data Governance & Security component in GridAI is essential for ensuring the security, privacy, and governance of data assets throughout their lifecycle. It orchestrates various components related to data ingestion, semantic interoperability management, cleaning, anonymization and access control. The component is designed to guarantee that data is properly secured and that only authorized users can access it, based on policies that are configured by the user.

The Data Governance & Security component in GridAI is responsible for defining and managing the following key tasks:

- Data Security & Privacy: Ensuring that sensitive data is anonymized or processed according to security and privacy requirements.
- Access Control & Policy Enforcement: Enabling flexible, fine-grained access policies for controlling which users or organizations can access specific data, ensuring that only authorized parties can view or modify it.
- Data Orchestration: Managing and automating data pipelines for both data collection and analytics processes, ensuring that the required data transformation and processing steps occur securely and efficiently.

5.3.3.1 Component Features

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

DGS1: Orchestration of Data Processes (Collection and Analysis) - The Data Governance & Security component orchestrates and automates data processing workflows for both data collection and data analysis. This includes managing data ingestion, semantic interoperability management, cleaning, anonymization, during the ingestion process, as well as orchestration of manipulation and analytics services during data analysis. The component ensures that these components are performing based on user-defined configurations, either once or on a recurring schedule, with a focus on maintaining data privacy and security at every stage of the process. By automating this orchestration, the component ensures that user needs are met while keeping data secure, ensuring privacy, and streamlining workflows across GridAI.

DGS2: Access Control and Enforcement of Policies - The Data Governance & Security component provides a flexible mechanism for defining, managing, and enforcing access policies for data assets within GridAI. Users can create and update rules that control who can access their data, with policies based on attributes of the data asset, the requester, or contextual factors such as the time and

location of the request. These access policies are automatically enforced during data ingestion, storage, and retrieval, ensuring that only authorized users or organizations can access specific data assets. The component maintains a clear separation between policy definition and enforcement, allowing for quick policy updates, which are immediately applied to ensure consistent and secure data access. Additionally, the component ensures that any access control decisions are performed efficiently, providing smooth user experience while maintaining strict security standards.

DGS3: Monitoring and Managing Service Execution - The Data Governance & Security component monitors the execution status of functionalities across various components. It ensures that components run efficiently, manages errors, and allows for process re-execution in case of failure.

5.3.3.2 Features and Requirements linkage

Feature	Linked Requirements
DGS1	Req_007, Req_002, Req_004, Req_006, Req_007, Req_008, Req_009, Req_010, Req_020, Req_025, Req_029, Req_033, Req_046, Req_047
DGS2	Req_035, Req_036, Req_037, Req_038, Req_043
DGS3	Req_020, Req_027, Req_028, Req_031, Req_061, Req_062, Req_063

5.3.4 Data Model Storage

The Data Model Storage Component plays a crucial role in managing the structured and standardized representation of concepts, fields, and relationships within the GridAI solution. It oversees the entire lifecycle of the Data Model, including its creation, specification, updates, while providing the necessary services to support these processes. By ensuring the integrity, consistency, and sustainability of the Data Model, this component enables seamless data management across the platform.

Designed for continuous evolution, the Data Model Storage Component adapts to the changing needs of users, supporting backward compatibility where possible and ensuring that updates propagate correctly across all dependent GridAI services and components. Additionally, it guarantees secure storage, proper versioning, and seamless accessibility, aligning the Data Model with all data activities.

5.3.4.1 Component Features

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

DMS1: Storage of Data Model Schemas and Vocabularies - Manages the structured schemas, standardized vocabularies, and hierarchical relationships that define the GridAI Data Model. Ensures consistency, semantic interoperability, and proper alignment of data concepts across the platform.

DMS2: Data Model Evolution and Versioning - Manages the lifecycle of the GridAI Data Model, including versioning, updates, and deprecation. Ensures backward compatibility, tracks schema modifications, and propagates changes across all dependent services while maintaining data integrity and consistency

5.3.4.2 Features and Requirements linkage

Feature	Linked Requirements
DMS1	Req_014, Req_015, Req_016, Req_017, Req_021, Req_022, Req_023, Req_024, Req_025, Req_026

DMS2	Req_014, Req_015, Req_016, Req_017, Req_021, Req_022, Req_024

5.3.5 User Authentication

User Authentication component is responsible for ensuring secure access to the solution by managing the identities of users. It acts as the core identity provider, handling the registration and authentication of users, verifying credentials, and enabling secure access to GridAI resources.

This component offers fine-grained authentication and authorization services to ensure that only authorized users can interact with the platform's components and bundles. It provides security measures to maintain the integrity of data and ensure trustworthiness across the solution's bundles, enabling secure access to data assets and analytics services.

5.3.5.1 Component Features

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

UA1: Identity information management - Manages the creation, registration, and verification of users to ensure secure access to GridAI.

UA2: Authentication and authorization - Delivers fine-grained access control for platform functions, ensuring that only authorized users can interact with the platform.

5.3.5.2 Features and Requirements linkage

Feature	Linked Requirements
UA1	Req_035, Req_036, Req_037, Req_043
UA2	Req_036, Req_037, Req_043

5.4 Data Analytics Services Bundle

The Data Analytics Services Bundle in the GridAI platform provides powerful tools for advanced data analysis, predictive insights, and AI-driven workflows. This bundle allows users to design custom analytics pipelines, generate accurate forecasts and visualize data insights. Its search and retrieval features enable quick and efficient access to the required data through advanced filtering and metadata indexing.

5.4.1 AI Pipelines Design

The AI Pipelines Design component is a comprehensive component designed to help users build, configure, and execute data analysis workflows tailored specifically for energy data. It offers a user-friendly interface that enables data scientists to easily engage with complex data processing tasks.

This component allows users to design end-to-end data analysis pipelines by combining various data manipulation functions, machine learning (ML) or deep learning (DL) models, and analytics techniques in a seamless manner. The component supports both predefined, energy-specific models and generic ML/DL algorithms, ensuring flexibility to meet diverse analytical needs. Users can configure these pipelines to process energy data, apply transformations, and generate valuable insights.

Additionally, users can specify execution parameters for the pipelines and the format in which the results should be exported. These results can be presented as raw data, visualizations, or reports, depending on the user's requirements.

5.4.1.1 Component Features

The key features of the AI Pipelines Design component are:

AIPD1: Configurable Data Processing and Analytics Models - Users can easily design pipelines by selecting and configuring reusable data manipulation functions and energy-specific machine learning or deep learning models. This feature supports both industry-specific algorithms and more generic models, allowing users to flexibly process data and apply analytics.

AIPD2: Customizable Data Input and Output Integration - This feature enables users to seamlessly configure data sources (both internal and external) as inputs to their analysis pipelines. Output can be customized to generate visualizations, reports, or data exports, giving users full control over the data flow and presentation.

AIPD3: Reusable Templates for Data Manipulation Functions - Users can create, save, and reuse function templates to streamline data preparation tasks. Configurable templates ensure that common data transformation operations can be consistently applied across multiple pipelines, improving efficiency and accuracy.

AIPD4: Data Transformation Insights - As users manipulate data, the component provides feedback on how each transformation impacts the dataset, with summary statistics and sample previews. This helps users better understand their data at every stage of the pipeline, making it easier to track changes and fine-tune the analysis.

5.4.1.2 Features and Requirements linkage

Feature	Linked Requirements
AIPD1	Req_046, Req_047, Req_048, Req_049, Req_050, Req_051, Req_052
AIPD2	Req_049, Req_053, Req_054, Req_064, Req_067
AIPD3	Req_047, Req_050
AIPD4	Req_055, Req_056, Req_063

5.4.2 Data Exploration

The Data Exploration component offers an intuitive and interactive platform for users to visually analyze selected data assets, enhancing their ability to uncover valuable insights. By engaging in this service, users can apply filters and focus on specific data subsets. The service provides a wide range of visualization tools that clearly display data relationships, trends, and patterns. These visualizations play a key role in identifying correlations, anomalies, and trends. Additionally, the Data Exploration component enables dynamic, interactive exploration, allowing users to investigate various facets of the dataset in real-time.

5.4.2.1 Component Features

The key features of the Data Exploration component are:

DE1: Interactive Data Exploration - Interactive Data Visualization provides users with dynamic, customizable visual representations of data, enabling them to explore datasets intuitively. Users can understand relationships, trends, and anomalies within the data. This feature empowers users to identify key insights by interacting with the data, filtering based on specific criteria, and examining detailed patterns.

5.4.2.2 Features and Requirements linkage

Feature	Linked Requirements
DE1	Req_040, Req_044, Req_045, Req_064, Req_066, Req_067

5.4.3 Data Analytics Execution

The Data Analytics Execution component handles the execution of data analysis jobs based on configurations created by the AI Pipelines Design. It enables users to run diverse data processing workflows by combining various blocks, which serve as wrappers for analytics models defined in the pipeline. The component applies machine learning (ML) and deep learning (DL) models to the specified input data based on user-defined parameters and generates results according to the configuration.

The Data Analytics Execution component is activated by the Data Governance & Security component, following the schedule set in the job configuration. It is allocated the necessary computational resources by the Platform Management & Orchestration Component.

5.4.3.1 Component Features

The key features of Data Analytics Execution component are:

DAE1: Use of Predefined ML/DL Models for Energy Problems - The component supports predefined algorithms tailored for energy-specific tasks (e.g., demand forecasting). These models are packaged as configurable functions, which can be applied to user datasets that match the required formats. Each model exposes configurable parameters, allowing users to adjust settings for executing the corresponding model on their data.

DAE2: Use of Generic ML/DL Algorithms - In addition to pre-trained models, the component provides flexibility for users to apply generic ML/DL algorithms such as classification, clustering, and regression. It supports the full data science pipeline, including training, testing, and evaluation of these models using various popular ML/DL libraries.

DAE3: Unified API for Analytics - The component offers a unified, simple API that ensures consistent interaction with all analytics blocks. This API hides the complexity of transforming configuration files into executable code, allowing for the use of different ML/DL libraries and execution engines under a common configuration template.

DAE4: Job Execution Feedback - During the execution of data analytics pipelines, the component generates and displays summary statistics on the analytics process. If any part of the pipeline fails, users receive clear feedback to identify the issue, allowing them to correct and re-run the affected part of the pipeline.

DAE5: Pipeline Configuration, Validation, and Execution Monitoring - Users can define the execution schedule for their data pipelines, validate pipeline configurations for errors, and view real-time insights into pipeline performance. In the event of failures, the system highlights problematic steps, ensuring quick identification and resolution of issues.

DAE6: Pipeline Update and Version Control - Once a pipeline has been executed, users can make targeted updates to its configuration without disrupting the results from previous runs. This ensures flexibility in adapting the pipeline to new requirements while maintaining the integrity of historical outputs.

5.4.3.2 Features and Requirements linkage

Feature	Linked Requirements
DAE1	Req_050, Req_051, Req_068, Req_069, Req_070, Req_071, Req_072, Req_080, Req_061, Req_082, Req_083, Req_084
DAE2	Req_051, Req_055, Req_056
DAE3	Req_046, Req_047, Req_050, Req_051
DAE4	Req_056, Req_061, Req_063
DAE5	Req_057, Req_060, Req_061, Req_062
DAE6	Req_047, Req_058, Req_059, Req_062

5.4.4 Interactive Data Visualization

The Interactive Data Visualization component plays a crucial role in the Energy Data Analytics Services Bundle by enabling users to visually explore insights from specific analytics results. This component allows users to create custom visualizations that highlight interesting data of the analytics pipelines they have designed and executed.

Working closely with the Data Analytics Execution component, the Interactive Data Visualization component ensures that the results of data analysis are effectively presented in the customized diagrams selected by the user.

5.4.4.1 Component Features

The key features of the Interactive Data Visualization component are:

IDV1: Built-in Chart Support – The component provides a wide range of predefined charts that can be easily configured and populated with data/ analytics results, allowing users to generate meaningful and visually appealing visualizations with minimal effort. These visualizations can display not only the results of data analysis pipelines but also raw dataset information.

IDV2: Custom Report Visualizations – Users can create reports accompanied by user comments to help viewers understand the insights and key takeaways.

IDV3: Saving and Exporting Visualizations and Reports – Visualizations and reports created with the Interactive Data Visualization component can be saved in both static (e.g., image or PDF) and dynamic formats. In the case of dynamic objects, a configuration template is generated, enabling the updated results to be rendered each time the visualization is accessed.

5.4.4.2 Features and Requirements linkage

Feature	Linked Requirements
IDV1	Req_040, Req_056, Req_064, Req_065
IDV2	Req_064, Req_065
IDV3	Req_053, Req_066, Req_067

5.5 GridAI Platform Horizontal Services Bundle

The GridAI Horizontal Services Bundle consists of the following components:

5.5.1 Platform Management & Orchestration

The Platform Management & Orchestration component ensures the seamless operation of the GridAI solution by efficiently provisioning and managing the necessary resources. Its role is to guarantee that the platform can consistently perform data collection and analysis tasks in a secure and isolated environment, without interruption. This involves allocating resources such as memory, compute power, and storage, based on the specific requirements of different data analysis jobs that need to be executed.

In close collaboration with the Data Governance & Security component, Platform Management & Orchestration ensures the dynamic deployment and scaling of resources for various components and jobs across the available virtualized infrastructure. It efficiently allocates and adjusts resources depending on the dynamic needs of each job, ensuring that the platform delivers high-quality service to all stakeholders in the electricity data value chain.

The Platform Management & Orchestration component is designed to ensure that data collection and analysis tasks are executed without failure due to resource limitations, while also ensuring that no resources are wasted on jobs that require lower memory or compute power.

5.5.1.1 Component Features

The key features of Platform Management & Orchestration component are:

PMO1: Resource Management for GridAI - The Platform Management & Orchestration is responsible for automating the deployment and monitoring of the overall platform components. It uses templates to configure cloud resources and their dependencies. This feature includes automatic load balancing, scaling, and failure management (e.g., automatically restarting failed containers without data loss). It also includes pre-built mechanisms for rollouts and rollbacks, ensuring near-zero downtime during these operations.

PMO2: Dynamic Resource Allocation for Data Collection - The Platform Management & Orchestration dynamically allocates resources to support the data collection processes within GridAI. This allocation is determined based on the volume of data being processed and the needs of associated components, such as Data Ingestion, Semantic Interoperability Management, Data Cleaning, and Data Sovereignty.

PMO3: Dynamic Scaling for Data Analysis Executions - The Platform Management & Orchestration is responsible for dynamically scaling memory and compute resources for running data analysis executions. It adjusts these resources based on the specific requirements of the data manipulation

and analytics services needed for each job, ensuring that data analysis tasks can be executed efficiently and securely.

5.5.2 Data Search

The Data Search component allows GridAI users to efficiently discover relevant data assets. This tool provides users with the ability to search across a wide range of GridAI assets, including not only data but also AI models and analytical results. The search capabilities go beyond traditional search functionalities by integrating advanced search mechanisms based on semantics and relationships derived from the GridAI data model.

5.5.2.1 Component Features

DSE1: Flexible Asset Search with Keyword and Facet Filters - The Data Search offers users a flexible approach to discovering assets by supporting both free-text keyword queries and filtering options. This allows users to search based on asset metadata (aligned with the GridAI metadata schema), asset structure (i.e., data model concepts mapped to asset attributes), data contents, and related meta-information. Faceted search is also available, enabling users to narrow down results based on frequently used facets for easier asset discovery.

DSE2: Comprehensive Search Across All Asset Types - This feature extends the search capabilities beyond data assets, supporting the discovery of other types of GridAI assets such as AI models and analytics reports. This ensures that users can conduct searches that are not limited to just data but encompass all available resources, offering a holistic view of GridAI's assets.

DSE3: Query Execution and Result Delivery - After configuring a search query, Data Search translates the query into a request that is processed by the Energy Data Collection Services Bundle (specifically the metadata management component). Results matching the search criteria are retrieved, processed, and presented to the user in an easy-to-understand format. Along with the search results, relevant details about each asset are provided to assist users in quickly identifying the most relevant assets for their needs.

DSE4: Search History and Query Re-execution - Users can save and revisit previous search queries for future use. This feature is particularly useful for queries involving advanced filters or data content retrieval. By storing and updating their queries, users can efficiently re-execute them to obtain updated or slightly modified data subsets, avoiding the need to repeat complex search steps.

5.5.2.2 Features and Requirements linkage

Feature	Linked Requirements
DSE1	Req_039, Req_041, Req_042, Req_052
DSE2	Req_039, Req_041, Req_042, Req_052
DSE3	Req_014, Req_038, Req_039, Req_042, Req_043
DSE4	Req_046, Req_047, Req_058, Req_059

5.5.3 Data Retrieval

The Data Retrieval component is activated to prepare and export data assets, along with results of data analysis pipelines, making them accessible for different uses. This feature enables the extraction of selected data and analytics results in the appropriate format and subsequently makes the exportable data available for transfer to the GridAI Dashboard.

5.5.3.1 Component Features

The key features of Data Retrieval component are:

DR1: Secure Export of Data and Analytics Results as a File - The Data Retrieval component allows the secure export of data and analytics results, as a downloadable file (e.g., CSV format) through a temporary link, ensuring secure data transfer.

DR2: Secure Export of Data and Analytics Results via API - Users can also securely export data or stored results of a data analysis pipeline through an API endpoint. This feature ensures that the analytics results are made available for integration to the GridAI Dashboard.

5.5.3.2 Features and Requirements linkage

Feature	Linked Requirements
DR1	Req_038, Req_044, Req_054
DR2	Req_038, Req_045, Req_054

5.5.4 Data Model Management

The Data Model Management component enables platform administrators to effectively manage and update the GridAI Data Model. It provides a comprehensive set of tools to review and modify the structure and content of the data model, while tracking all changes and updates. Administrators are alerted when new concepts or attributes are suggested, allowing for timely updates to keep the model current and relevant.

5.5.4.1 Component Features

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

DMM1: Data Model Handling - The component provides CRUD (Create, Read, Update, Delete) operations for concepts and attributes within the GridAI Data Model. All modifications follow strict validation rules to maintain the overall consistency of the model.

The user interface ensures that:

- New concepts and attributes cannot exist disconnected from the rest of the model.
- Concepts must have complete information (e.g., no concepts without a name, no attributes without parent concepts).
- Validation checks run during updates or deletions to prevent breaking knowledge representation or negatively affecting existing data assets mapped to the model.

DMM2: Data Model Navigation - The component offers an intuitive interface for model administrators to search, explore, and review the GridAI Data Model's structure, relationships, and specific concepts, aiding in informed decision-making for model evolution.

5.5.4.2 Features and Requirements linkage

Feature	Linked Requirements
DMM1	Req_014, Req_015, Req_016, Req_017, Req_021, Req_022, Req_023, Req_024, Req_025, Req_026, Req_038
DMM2	Req_014, Req_015, Req_016, Req_017, Req_021, Req_022, Req_024

5.6 GridAI Dashboard

The GridAI Dashboard is a separate component that centralizes monitoring and visualization for real-time analytics management, ensuring users can make informed decisions based on actionable insights. It provides an intuitive user interface that facilitates the visualization and interpretation of distribution grid portfolio data and the results of associated AI pipelines. It consolidates critical information, presenting it in a clear and accessible format, enabling users to quickly grasp key insights and trends.

5.6.1.1 Component Features

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

D1: Demand Forecasts Visualization – The Dashboard provides comprehensive visualization of demand forecasting analytics pipelines results in various temporal and spatial granularities

D2: Generation Forecasts Visualization – The Dashboard provides comprehensive visualization of generation forecasting analytics pipelines results in various temporal and spatial granularities.

D3: Comparative Analysis Visualization – The Dashboard provides comparative visualizations (in common charts) of the executed demand and generation forecasts against the respective actual data to help users assess the accuracy of the models.

D4: Historical Data Visualization – The Dashboard provides comprehensive visualizations of historical data and correlations of them to support the extraction of valuable insights regarding historical performance of demand and generation assets across distribution grids.

D5: Interactive Visualization Charts – The Dashboard supports zoom in and out functions across timeseries visualization to help the users easily navigate across the whole period address in their selection. Charts are also customly populated with the relevant data based on users' selections on the period applied or other filters.

5.6.1.2 Features and Requirements linkage

Feature	Linked Requirements
D1	Req_092, Req_093, Req_095, Req_097, Req_099, Req_101
D2	Req_092, Req_112, Req_114, Req_116, Req_118
D3	Req_094, Req_096, Req_098, Req_100, Req_103, Req_113, Req_115, Req_117, Req_119
D4	Req_087, Req_088, Req_089, Req_090, Req_091, Req_102, Req_104, Req_105, Req_106, Req_107, Req_108, Req_109, Req_110, Req_111, Req_120

D5	Req_087, Req_088, Req_089, Req_090, Req_091, Req_092, Req_104, Req_105, Req_106, Req_107, Req_108, Req_109, Req_110, Req_120

6. Conclusions

In conclusion, Deliverable D10, provides a critical foundation for the successful development and implementation of the GridAI solution. By comprehensively capturing and detailing the defined requirements, it establishes a clear framework for the solution's design and integration. The deliverable aligns business needs with the corresponding technical functionalities, ensuring that the system will effectively address the challenges faced by DSOs. Additionally, the specifications outlined for each solution component—ranging from data collection and security to advanced analytics and platform management—underscore the emphasis on scalability, interoperability, and security.

The GridAI solution is structured around key bundles that address distinct aspects of the solution's functionality. These bundles are designed to meet the needs of DSOs and ensure comprehensive service delivery.

- **Energy Data Collection Services Bundle**: This bundle focuses on the reliable and consistent acquisition of data, including components for data collection, semantic interoperability, data cleaning, and metadata management. It ensures that data is structured and accessible for further analysis.
- **Energy Data Security Services Bundle**: This bundle emphasizes the protection of data, including mechanisms for data storage, data sovereignty, data governance and security, and data model storage. It safeguards data integrity and prevents unauthorized access, ensuring compliance with security standards.
- **Data Analytics Services Bundle**: This bundle supports advanced analytics capabilities, including exploratory data analysis, the design and execution of AI-driven workflows, and the generation of predictive insights.
- **GridAI Platform Horizontal Services Bundle** : This service ensures efficient orchestration, operation, and scalability of the platform. It comprises critical components such as Platform Management & Orchestration, Data Search, Data Retrieval and Data Model Management.
- **GridAI Dashboard**: It consists of the GUI provided to DSOs to support improved decision making based on data insights and accurate demand and generation forecasts delivered by the GridAI analytics models.

With this comprehensive approach, the GridAI solution is positioned to deliver significant value in enhancing operational efficiency, data-driven decision-making, and grid resilience across the electricity value chain.

Looking ahead, the development and delivery of the GridAI solution are entering a crucial phase. Key milestones in the coming months include the release of the GridAI Prototype Solution, scheduled for M10, which will incorporate the full suite of functionalities, including data governance and security mechanisms, AI-driven analytics pipelines, and the GridAI dashboard. This prototype will allow the extensive testing of the solution against the requirements set, as well as, assessment of the accuracy of the GridAI models. Following this, the GridAI Technical Verification Results will be reported in M12. This deliverable will outline the outcomes of technical verification and testing activities, confirming the platform's performance, integration, and scalability to ensure it meets all the required technical standards for broader deployment.

References

1. International Electrotechnical Commission. 2015. *Use case methodology - Part 2: Definition of the templates for use cases, actor list and requirements list* (IEC 62559-2). Retrieved from https://webstore.iec.ch/publication/22349
2. https://www.iso.org/standard/72089.html
3. DoA Part B, GridAI Grant Agreement SMART GROWTH/RESEARCH IN ENTERPRISES ENTERPRISES/ENERGY/1123/0018
4. GridAI D11 "GridAI Data Model"
5. GridAI D12 "GridAI Prototype Solution"