Deliverable 2.1: Big Data Value Chain and Requirements

WP2 – System Requirements and Specifications

May 2021

www.bd4nrg.eu

Disclaimer

Copyright Message

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613

Grant Agreement Number: 872613 Acronym: BD4NRG

Full Title	BIG DATA FOR NEXT GENERATION ENERGY		
Topic	DT-ICT-11-2019 | Big data solutions for energy		
Funding scheme	H2020- IA: Innovation Action		
Start Date	January 2021	Duration	36
Project URL	www.bd4nrg.eu		
Project Coordinator	ENG		
Deliverable	D2.1 - Big Data Value Chain and Requirements		
Work Package	WP2 - System Requirements and Specifications		
Delivery Month (DoA)	M5	Version	V1.0
Actual Delivery Date	28/05/2021		
Nature	Report	Dissemination Level	Public
Lead Beneficiary	ASM		
Authors	Francesca Santori [ASM], Tommaso Bragatto [ASM], Marco Antonio Bucarelli [ASM], Jaime Gómez Tribiño [VEOLIA], Marina Bartolomé Lorenzo [VEOLIA], Francesco Bellesini [EMOT], Edoardo Mancinelli [EMOT], Uroš Kerin [ELES], Miha Bečan [ELES], Denis Kozjek [ELES], Borut Rajer [BORZEN], Aliene van der Veen [TNO], José Hernández [CAR], Gema Hernández [CAR], Sofía Mulero [CAR], Giuseppe Castro [ENELX], Ruben Bäumer [Centrica], Bert Claessens [Centrica], Marzia Mammina [ENG], Diego Arnone [ENG], Alessandro Rossi [ENG], Arsim Bytyqi [ENTSO-E], Gerasimos Anastasatos [NTUA]		
Quality Reviewer(s):	Vangelis Marinakis [NTUA], Stefan Dahling [RWTH]		
Keywords	Big Data, BD4NRG tools, methodology, requirements, Use Cases, Key Performance Indicators, scenarios, pilot.		

Preface

BD4NRG focuses on addressing emerging challenges in big data management for energy sector with an innovative open holistic solution for smart grid-tailored, near real time, energy-specific and AI-based open Big Data Analytics modular framework. The vision is to deliver holistic services for techno-economic optimal management of Electric Power and Energy Systems (EPES) value chain. Services range from optimal risk assessment for energy efficiency investments planning, to optimized management of grid and non-grid owned assets, improved efficiency, and reliability of electricity networks operation, while at the same time contributing to achieve fair energy prices to the consumers and laying the foundations for an EU-level energy-tailored data sharing economy. The BD4NRG Toolbox will be implemented and validated in real-life pilots in 12 large-scale demo sites across 10 countries for:

- Increasing the efficiency and reliability of the electricity network **BD-4-NET**
- Optimising the management of assets connected to the grid **BD-4-DER**
- De-risking investments in energy efficiency and increasing the efficiency and comfort of buildings **BD-4-ENEF**

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Who We Are

	Participant Name	Short Name	Country Code	Logo
1	ENGINEERING – INGEGNERIA INFORMATICA SPA	ENG	IT	
2	NATIONAL TECHNICAL UNIVERSITY OF ATHENS	NTUA	GR	
3	RHEINISCH-WESTFAELISCHE TECHNISCHE HOCHSCHULE AACHEN	RWTH	DE	
4	EUROPEAN DYNAMICS LUXEMBOURG SA	ED	LU	
5	INTERNATIONAL DATA SPACES EV	IDSA	DE	
6	EUROPEAN NETWORK OF TRANSMISSION SYSTEM OPERATORS FOR ELECTRICITY AISBL	ENTSO-E	BE	
7	PANEPISTIMIO DYTIKIS ATTIKIS	UNIWA	GR	
8	ATOS SPAIN SA	ATOS	ES	
9	FUNDACION CARTIF	CARTIF	ES	
10	UNIVERZA V LJUBLJANI	UNILJ	SL	
11	ENEL X SRL	ENELX	IT	
12	REN - REDE ELECTRICA NACIONAL SA	REN	PT	
13	CENTRO DE INVESTIGACAO EM ENERGIA REN - STATE GRID SA	RDN	PT	
14	UNINOVA-INSTITUTO DE DESENVOLVIMENTO DE NOVAS TECNOLOGIASASSOCIACAO	UNINOVA	PT	
15	ENERCOUTIM - ASSOCIACAO EMPRESARIALDE ENERGIA SOLAR DE ALCOUTIM	ENERC	PT	
16	FIWARE FOUNDATION EV	FIWARE	DE	
17	CENTRICA BUSINESS SOLUTIONS BELGIUM	CENTRICA	BE	
18	NEDERLANDSE ORGANISATIE VOOR	TNO	NL	

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

	Participant Name	Short Name	Country Code	Logo
	TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO			
19	ASM TERNI SPA	ASM	IT	
20	VIDES INVESTICIJU FONDS SIA	LEIF	LV	
21	COMSENSUS, KOMUNIKACIJE IN SENZORIKA, DOO	COMSENSUS	SL	
22	HOLISTIC IKE	HOLISTIC	GR	
23	INTERUNIVERSITAIR MICRO-ELECTRONICA CENTRUM	IMEC	BE	
24	TERRASIGNA SRL	TS	RO	
25	UBIMET GMBH	UBIMET	AT	
26	ELEKTRO LJUBLJANA PODJETJE ZADISTRIBUCIJO ELEKTRICNE ENERGIJE D.D.	EKL	SL	
27	BORZEN, OPERATER TRGA Z ELEKTRIKO, D.O.O.	BORZEN	SL	
28	AJUNTAMIENTO DE SANT CUGAT DEL VALLES	AJSCV	ES	
29	ELES DOO SISTEMSKI OPERATER PRENOSNEGA ELEKTROENERGETSKEGA OMREZJA	ELES	SL	
30	E-LEX - STUDIO LEGALE	ELEX	IT	
31	OSMANGAZI ELEKTRIK DAGITIM ANONIM SIRKETI	OEDAS	TR	
32	VEOLIA SERVICIOS LECAM SOCIEDAD ANONIMA UNIPERSONAL	VEOLIA	ES	
33	STICHTING EG	EGI	NL	
34	CINTECH SOLUTIONS LTD	CN	CY	
35	EMOTION SRL	EMOT	IT	

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Contents

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

List of Figures

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

List of Tables

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

List of Acronyms and Abbreviations

AHI	*Asset Health Index*
AI	*Artificial Intelligence*
API	*Application Programming Interface*
ARM	*Automated Revenue Metering*
AUI	*Asset Utilization Index*
B2B	*Business-to-business*
BD4NRG	*Big Data for Next Generation Energy Project*
BD	*Big Data*
BDV	*Big Data Value*
BDVA	*Big Data Value Association*
BESS	*Battery Energy Storage System*
BIM	*Building information modelling*
BTM	*Behind the meter*
CAPEX	*Capital Expenses*
CB	*Circuit Breakers*
CENCENELEC	*European Committee for Electrotechnical Standardization*
CEP	*Clear Energy for all Europeans package*
CHP	*Combined heat and power*
COMTRADE	*Common format for Transient Data Exchange for power systems*
COSMAG	*Comprehensive Architecture for Smart Grid*
CPV	*Concentrator photovoltaics*
CQRS	*Command Query Responsibility Segregation*
DAIRO	*Data, AI and Robotics Association*
DAPS	*Dynamic Attribute Provisioning Service*
DAS	*Data Acquisition System*
DR	*Demand Response*
DSO	*Distribution system operator*
EABB	*Architecture Building Blocks*
ECM	*Energy Conservation Measure*
EMFIP	*Electricity Market Fundamental Data Information Platform*
ENTSO-E	*European Network of Transmission System Operators for Electricity*
EPBD	*Energy Performance of Building Directives: 2010/31/EU and 2018/844/EU*
EPC	*Energy Performance Certificates*
EPES	*Electrical Power Energy Systems*
ESCO	*Energy Service Company*
ETSI	*European Telecommunications Standards Institute*
EV	*Electric vehicle*
EVSE	*Electric Vehicle Service Equipment*
FIM	*Federation Identity Management*
FTM	*Front of the meter*
GDPR	*General Data Protection Regulation*
HPC	*Cloud and High-Performance Computing*
HV	*High Voltage*
HVAC	*Heating Ventilation and Air Conditioning*
HW	*Hardware*
ICT	*Information and communications technology*
IDP	*Industrial Data Platforms*
IDS	*International Data Spaces*
IDSA	*International Data Spaces Association*
IEC	*International Electrotechnical Commission*
IED	*Intelligent Electronic Devices*

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

IEEE	*Institute of Electrical and Electronic Engineers*
IMB	*Identity Management Broker*
IoT	*Internet of Things*
IP	*Internet Protocol*
ISB	*Imbalance settlement period*
ISO	*International Organization for Standardization*
IT	*Information Technology*
ITU-T	*International Telecommunication Union – Telecommunication Standardization Bureau*
KPI	*Key Performance Indicator*
LIDAR	*Light detection and ranging*
LSP	*Large Scale Pilots*
LV	*Low Voltage*
MDP	*Markov Decision Process*
ML	*Machine Learning*
MQTT	*Message Queuing Telemetry Transport*
MV	*Medium Voltage*
O&M	*Operation and Maintenance*
OHL	*Overhead lines*
OPEX	*Operating expenses*
P2P	*Peer to Peer*
PDP	*Personal Data Platforms*
PMU	*Phasor Measurement Units*
PoC	*Proof of concept*
PPP	*Big Data Value Public Private Partnership*
PV	*Photovoltaic plant*
RDP	*Research Data Platforms*
RES	*Renewable Energy Sources*
RoW	*Right of Ways*
Rq	*Requirement*
SAIDI	*System Average Interruption Duration Index*
SAIFI	*System Average Interruption Frequency Index*
SCADA	*Supervisory Control And Data Acquisition*
SDO	*Standards Development Organisation*
SRI	*Smart Readiness Indicator*
SRIA	Strategic Research and Innovation Agenda developed by the BDVA
TP	*Transparency Platform*
TSO	*Transmission System Operators*
UC	*Use case*
UDP	*Urban/City Data Platforms*
W3C	*World Wide Web Consortium*
WP	*Work packages*

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Executive Summary

The use of Big Data (BD) for the energy system represents a great opportunity to develop a more resilient and reliable infrastructure that can manage all functions and needs of stakeholders effectively and with an integrated view. On the other hand, the implementation of BD systems internally fits into already proven and effective mechanisms and must be user-friendly, efficient and easy to use.

This document has the ambition to introduce the BD Value Chain providing with details and specifications the definitions and functionalities of BD and IoT tools and services, within the BD4NRG Framework. Projects partners, both pilot site owners and technology providers, were deeply analysed with the aim of exploiting the peculiarities of each of them, in order to develop a mutually beneficial cooperation. With respect to reference architectures, the document evaluates, and reports activities already carried out within previous European projects, with particular reference to the GAIA-X project.

The main objective of this document is to identify and present the characteristics of the twelve Large Scale Pilots (LSP), giving a description of each Use Case (UC), explaining reasons and benefits of the innovation in that area and clearly defining the requirements. The activity was carried out in tight collaboration with all the partners and produced the high-level description of 18 UCs and the related requirements, as a basis for service development and integration.

Besides, this document represents a pipeline between the Demonstration Pilots and the BD4NRG tools, which will be developed and implemented according to the functional and non-functional requirements of each UC and is therefore a key document for the implementation of the next WPs.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

1. Introduction

1.1. Purpose of the Document

D2.1 – Big Data Value Chain and Requirements is produced within the WP2 – System Requirements and Specifications and it gathers the activities carried out in T2.1 - User Stories' Definition and Requirements Analysis and T2.2 - Edge Energy Services Specifications for BD/IoT Data Framework infrastructure. The purpose of this document is to define BD value chain functionalities and specify the user requirements taking into account the input coming from the usage scenarios of the pilot cases.

The document sets the foundation of the project and serves as the key starting point for the innovation activities of the BD4NRG project. By means of a thorough analysis of the BD value chain functionalities of BD4NRG framework, the user stories, usage scenarios and user requirements have been defined and pilot partners and technologies providers have been actively involved in an iterative process aiming at setting user stories and usage scenarios directly related with the use of BD4NRG tools and services.

Moreover, the document describes and compiles all the results of the first and second tasks of WP2. Although during the project lifetime requirements may evolve based on the new knowledge, a commonly shared understanding of the goals during the first months is needed for building up a suitable architecture and the new functionalities foreseen in the BD4NRG project.

1.2. Relation to other documents

The present document plays an important role within the BD4NRG project, paving the way of other project tasks. Firstly, there are close dependencies with the other WP2 documents, as tasks T2.3 to T2.5 use the content of this document as a starting point for their analysis and task T2.6 starts from this document to have a deeper look inside the reference architectures and their relations with BD4NRG. D2.1 outcome is fundamental for WP3 - Data Governance and Management and WP4 - Data Modelling & Open Modular AI-based edge-level Analytics Toolbox, where technical partners will develop the tools required by the pilots and mentioned in this document. There is also an intensive collaboration with WP6 - Pilot Requirements and Assessment and with tasks T6.1 - Pilot Operation, Engagement and M&V Plan, T7.1 - Baseline Analysis Report and 'BD-4-NET' Large Scale Pilots Preparation, T8.1 - Baseline Analysis Report and 'BD-4-DER' Large Scale Pilots Preparation and T9.1 - Baseline Analysis Report and 'BD-4-ENEF' Large Scale Pilots Preparation.

1.3. Document Structure

This document is organized as follows: chapter 2 reports the methodology to collect requirements and use cases, chapter 3 provides an overview of the BD Value Chain, depicting existing architectures, services, and associations; chapter 4 outlines context scenarios of each of the twelve LSP; chapter 5 specifies UCs addressed to each pilot whilst chapter 6 gathers UC Requirements. Finally, chapter 7 concludes the document.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

2. Use Cases and Requirements Methodology

This section gives an overview of the methodological approach followed within BD4NRG, in order to specify the requirements needed for achieving the project objectives, to identify specific UCs and to define measurable indicators of the impact of the proposed innovative technological solutions. This methodology helped project partners to describe, formalize and track the project requirements in an explicit and unambiguous manner.

Basically, a template was prepared by Task2.1 coordinator, agreed with all the participants and shared with the partners involved in the task. This activity was managed and led through several video calls between the task leader and the partners. Contributions were provided in two rounds, supported by a review and a follow-up activity carried out by the task leader.

Partners gave contributions for a general description of the scenario, a description of the UC and a list of requirements. In the following sections templates and guidelines used to map UCs and requirements are presented.

Scenario

It provides a high-level description of the need enlarging the description of the rationale included in the proposal. A context scenario is a brief description of actor(s) using a product to achieve a goal; it describes an instance of use of the system in context. Context scenarios are usually narratives that tell a story describing one or more tasks in a specific environmental situation. Each scenario was produced as preparatory work to develop the project architecture in compliance with the IEEE guideline [1]. The scenarios can be considered as the starting point for the definition of the UCs, which represent the low-level specifications.

UC Template

A UC describes in a generic and replicable way how actors interact with the system to achieve a goal. UCs of BD4NRG have been assigned to each Large-Scale Pilot and presented in Chapter 5, in the form of tables.

A specific template has been adopted for the definition of the UCs and it was circulated among the project partners to identify the first set of UCs. The template was based on the "basic UC template" as in [2] and descriptions provided by partners have been included in this document.

Use Case ID	
Cluster – main beneficiary	*This section helps to categorize the UC as well as the LSP leader states its role in the UC*
Description	*This section is related to the storyline of the UC, highlighting* • *Location of the UC,* • *main issues to be solved by UC realization,* • *asset involved,* • *specific needs,* • *sequence of events*
Goals	*This row briefly describes what the UC is aiming to do.*
Actors	*Include all the relevant stakeholders, expanding the list of the main beneficiaries.*
Envisioned analytics and tools	*A set of enabling tools and analytics (e.g., weather forecast, fault prediction) is enlisted.*
Precondition	*This section highlights the main dependencies for*

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	
	deploying and testing the UC (e.g., HW availability, sensors deployment, data gathering for a certain time period). *Moreover, it highlights the EPES conditions that make this UC applicable (e.g., high penetration of RES or EV, age of the equipment).*
Triggering events (if applicable)	*This row is filled in if the UC related to any particular events to be fully demonstrated (e.g., faults on the network, DR campaign, asset replacement).*
Post conditions	*This section describes the status of the system under study after the UC finalization.*
Related KPIs	*This section should give an idea about the most suitable KPIs for demonstrating this UC*

Requirements Template

With respect to requirements, those were defined in order to define a set of characteristics to be provided by the services that enable each UC. Requirement collection was based on consolidated methodology as in [3]. UCs requirements are presented in Chapter 6 in the form of tables, which are presented with the characteristics described below. The different requirements are classified by type, which can be functional or non-functional, and a description, a rationale, a priority level and related UCs are presented.

ID: Indicate Requirement number (e.g., Rq_01)

Description: A statement which describes the intention of the requirement.

Type: The type of the requirement, some examples:

- Functional requirement
- Data requirement
- Legal requirement
- Security requirement
- Cultural and political requirement
- Interoperability requirement

Rationale: A justification of the requirement

Priority (1 – 5): The importance for the customer of successfully implementing the requirement.

- 5 - High: Requirements in this class are either realizing a key innovation of the project or they are needed to realize it. These requirements are necessary to achieve the goals of the project.
- 4-3 Medium: Requirements in this class are not necessary to realize a key innovation but they are necessary or very helpful to realize the application prototypes. These requirements are important to the application developer.
- 2-1 Low: Requirements in this class are neither realizing a key innovation nor necessary for the application prototypes. However, in a broader context possibly beyond the scope of the project they may be important.

Related UC: To be filled in according to the UC.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

3. Data-driven initiatives and Reference Architectures

BD can help improve the smart grid management to a higher level. For example, by enhancing the accessibility of a customer's electricity consumption data, the DR will be expanded, and energy efficiency will be improved. Similarly, analyzing data obtained from PMUs and IEDs will help to improve customer service, prevent outages, maximize safety, and ensure service reliability. Furthermore, electric utilities are using prediction data analytics for estimating several parameters which are helpful to operate the smart grid in an efficient, economical, and reliable way. For instance, whether the excess energy is available from renewable sources can be predicted through accessing the ability of the smart grid to transmit it. At the same time, calculating equipment downtime, accessing power system failures, and managing unit commitment can also be done effectively in high correlation with integrating distributed generation. Thus, the grid management state will be enhanced, the efficiency will be improved, and the robustness of generation and deployment will be ensured. With the integration of various distributed generating sources, refined forecasting, load planning, and unit commitment are more convenient to avoid inefficient energy transmitting or dispatching extra generation **Errore. L'origine riferimento non è stata trovata.** [13] shows some of the potential applications of BD in smart grid useful for various stakeholders.

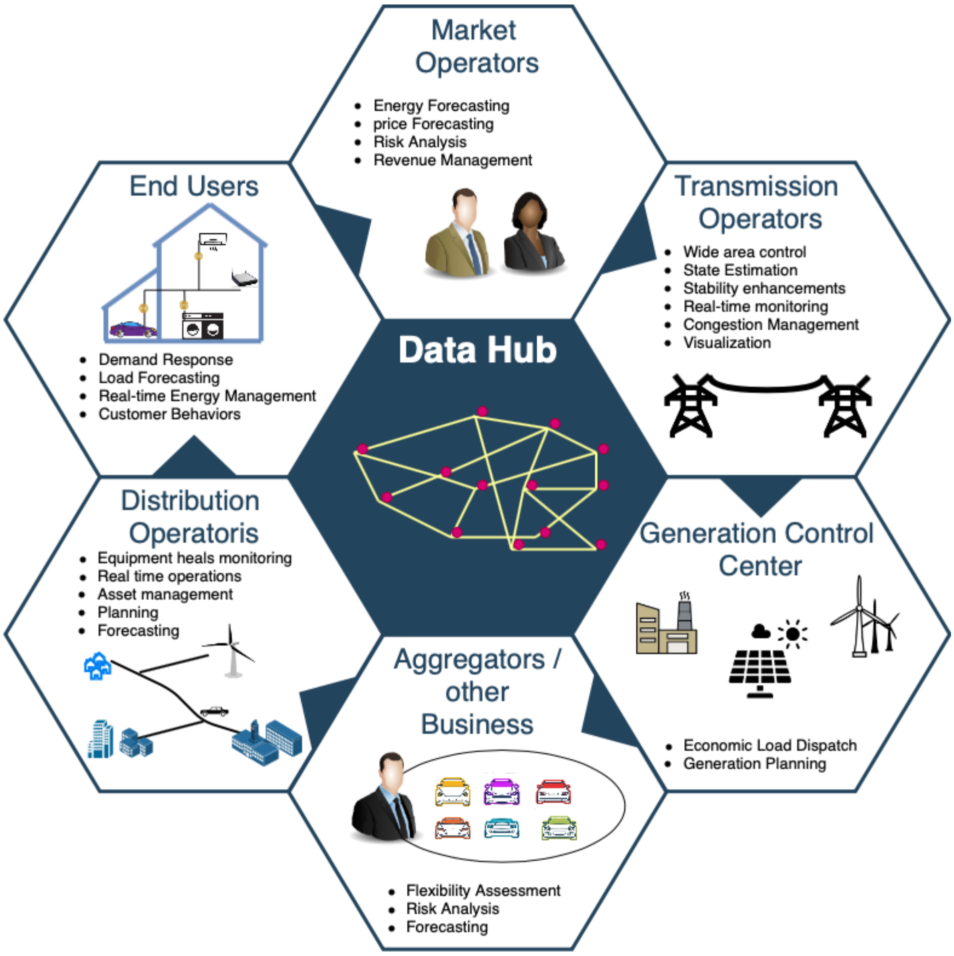


Figure 1 - Some of the potential applications of the big data analytics in smart grids.

Big Data have potential to unlock novel groundbreaking opportunities in power grid that enhances a multitude of technical, social, and economic gains. As power grid technologies evolve in conjunction with measurement and communication technologies, this results in unprecedented amount of heterogeneous BD. In particular, computational complexity, data security, and operational integration of BD into power system planning and operational frameworks are the key challenges to

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

transform the heterogeneous large dataset into actionable outcomes. In this context, suitable BD analytics combined with visualization can lead to better situational awareness and predictive decisions.

This chapter presents a comprehensive state of the art on BD applications, and BD analytics in power grids, and provides some detailed information and examples from ENTSO-E. Finally, detailed information about ENTSO-E transparency platform (TP) is presented which can be used as a useful example of the adoption of the BD in energy sector.

In order to benefit from the potential of BD, proper technologies able to analyse huge quantities of data are necessary [5]. BD can be considered an evolution from traditional data analysis; therefore, a common structure explaining how BD complements and differs from existing analytics, business intelligence, databases, could be useful to have. This common structure is called a reference architecture and consists of a document or set of documents that provides best practices to interested parties. The objective of a reference architecture is to:

- providing a common language for the various stakeholders;
- encouraging adherence to common standards, specifications, and patterns;
- providing consistent methods for implementation of technology to solve similar problem sets;
- illustrating and improving the description of the various BD components, processes, and systems, in the context of a vendor- and technology-agnostic BD conceptual model;
- facilitating the analysis of candidate standards for interoperability, portability, reusability, and extendibility.

3.1. State of the art about Big Data for energy

Energy is one of the most important parts in human life. As a significant application of energy, smart grid is a complicated interconnected power grid that involves sensors, deployment strategies, smart meters, and real-time data processing. It continuously generates data with large volume, high velocity, and diverse variety.

Integrating telecommunication, automation and electric network control, smart grid requires reliable real-time data processing. To support this requirement and benefit future analysis and decision making, huge amount of historic data should be well fetched and stored in a reasonable time budget. In addition, there are various sources that huge amount of data can be generated through diverse measurements acquired by Intelligent Electronic Devices (IEDs) in the smart grid:

- Data from power utilization habits of users;
- Data from Phasor Measurement Units (PMUs) for situation awareness;
- Data from energy consumption measured by the widespread smart meters;
- Data from energy market pricing and bidding collected by Automated Revenue Metering (ARM) system;
- Data from management, control and maintenance of device and equipment in the electric power generation, transmission and distribution in the grid;
- Data from operating utilities, like financial data and large data sets which are not directly obtained through the network measurement.

All aforementioned sources increase the grid's data volume. According to the investigation [7], by 2009 the amount of data in electric utilities' system has already reached the level of TeraBytes (TBs). With every passing day, more TBs of data are emerging at infrastructure company data centers. This fact accelerates the pace of the so-called BD technology applying to the smart grid area.

BD technologies are a new generation of technologies and architectures designed to manage and to extract value from very large volumes of a wide variety of data by enabling high-velocity capture,

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

discovery, and/or analysis. BD bring a set of data management challenges that can be represented by the so called 5 Vs, that is:

- **Volume**, dealing with large scales of data within data processing;
- **Velocity**, dealing with streams of high frequency of incoming real-time data;
- **Variety**, dealing with data using differing syntactic formats (e.g., spreadsheets, XML, DBMS), schemas, and meanings.
- **Veracity**, dealing with quality and uncertainty;
- **Value**, dealing with ability to transform the large amount of data into business.

The BD Value Chain has been introduced to describe the flow of information within a BD system as a series of steps for generating value and useful insights from data. [6] describes the steps as follow:

- Data Acquisition is the process where data are collected, filtered, and cleaned before being stored. The infrastructure that supports the acquisition of BD must ensure low and predictable latency both in capturing and querying data, must be able to handle high volumes of transactions, and must support flexible data structures.
- Data Analysis makes raw data acquired suitable for domain-specific usage. Data analysis involves exploring, transforming, and modelling data, thus extracting useful hidden information relevant from a business point of view.
- Data Curation refers to the active management (content creation, selection, classification, transformation, validation, and preservation) of data over their life cycle, while ensuring that data are trustworthy, discoverable, accessible, reusable, and fit their purpose.
- Data Storage refers to the persistence and management of data in a scalable way satisfying the needs of applications that require fast access to the data. Relational Database Management Systems (RDBMS) have been the main, and almost unique, solution to the storage paradigm for nearly 40 years; the lack of flexibility with regard to schema changes and the performance and fault tolerance when data volumes, and complexity grow has pushed towards new technologies (like NoSQL database) that have been designed with the scalability goal in mind and present a wide range of solutions based on alternative data models
- Data Usage covers the data-driven business activities that need access to data or their analysis and the tools needed to integrate the data analysis within the business activity. In business decision-making, data usage can enhance competitiveness through reduction of costs, increased added value, or any other parameter that can be measured against existing performance criteria.

BD is a double-edged sword, indeed, it enables the development of new technologies and services that improve our lives, such as personalized medicine or targeted advertising. On the other hand, it also has potential to harm individuals by profiling their behaviour without any knowledge of it on their part.

In energy sector BD can be used to profile people, and this is something thoroughly thought. It does not matter if the data collected is encrypted or anonymous data; with BD, one can figure out a lot of things about a person simply by analysing their patterns.

One of the most important things we should think about is that BD poses a threat to privacy. With BD, one can profile an individual's behaviour without their knowledge or consent.

This allows us to build profiles on individuals without their knowledge or consent. In addition, BD can be analysed in a way that reveals patterns which are invisible to human observers.

The biggest challenge of BD for energy will be to come up with the right metrics. Without the right metric, and just using a wrong one, all conclusions are going to be off base.

Another big challenge will be to come up with a broad enough set of data.

In the following sections, the main data-driven initiatives, and Reference Architectures of interest to BD4NRG will be introduced; further information on BD initiatives and reference architectures can be found in the deliverable D2.5.

3.2. Big Data Value Association (BDVA)

The Big Data Value Association is an industry-led international non-profit association that has been established in 2014 [1]. BDVA is the private counterpart to the European Commission in the Big Data Value Public Private Partnership (PPP) aiming to position Europe as the world leader in the creation of Big Data Value. Members of the BDVA comprise small and large European organisation from public and private sector of business, industry, academia, and research that support the objectives of the BDVA as described in the Statutes:

- Develop technical solutions for the supply and the demand side and end user of the entire data business value chain.
- Implement innovations in industry, business, and public sector to strengthen the European competitiveness and to address the societal challenges, e.g., climate and health.
- Understand and improve the businesses and citizens' acceptance of BD value creating technology and application for a faster take up and wider integration in process of public sector service provisioning and process efficiency and product development and business model creation in industrial and business.
- Help creating ecosystems that allow for flourishing of a vast variety of viable business models small and large entities and entrepreneurs.

BDVA has the mission to develop the Innovation Ecosystem to enable the data and AI-driven digital transformation while delivering maximum economic and societal benefit to Europe and sustaining Europe's leadership on BD value creation and artificial intelligence (AI).
The Strategic Research and Innovation Agenda (SRIA), developed by the BDVA, defines the overall goals, main technical and non-technical priorities, and a research and innovation roadmap for the European PPP on BD value.
Staring from input received from technical experts and stakeholders along the whole Big Data Value (BDV) chain, the BDVA has developed the BDV Reference Model, a reference framework that addresses the main concerns and aspects to be considered for BDV systems and, thus, allowing to locate BD technologies on the overall IT stack.
As depicted in Figure 2, the BDV Reference Model is structured into horizontal and vertical concerns.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

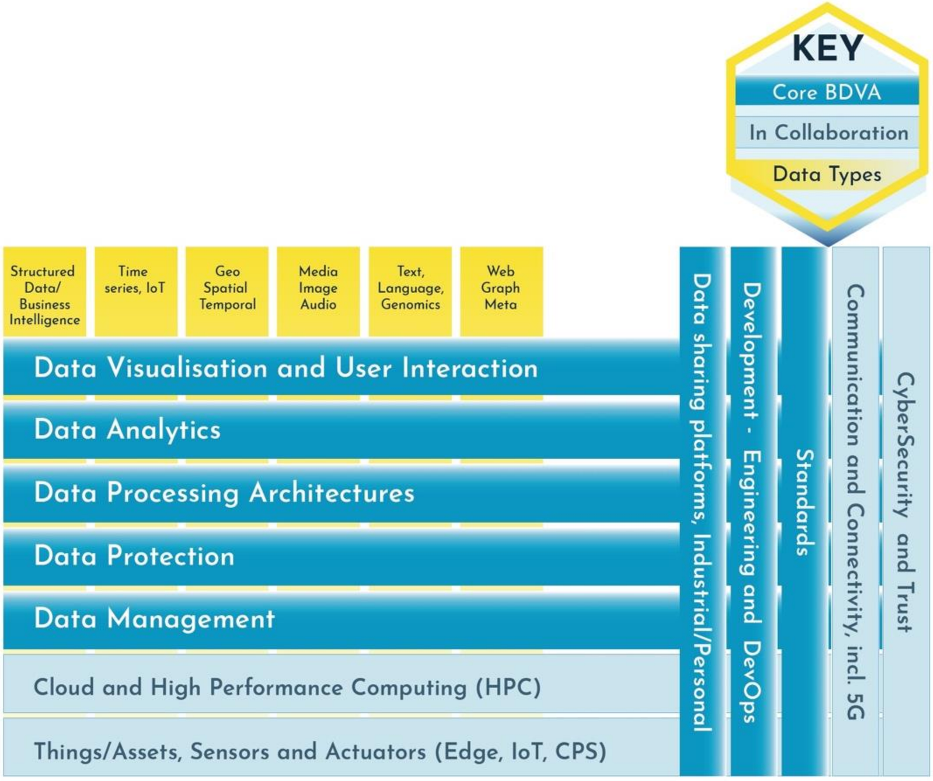


Figure 2 - The Big Data Value Reference Model developed by the BDVA

Horizontal concerns cover specific aspects along the data processing chain, that is data collection, ingestion, and visualization:

- Data Visualisation and User Interaction: advanced visualisation approaches for improved user experience.
- Data Analytics: data analytics to improve data understanding, deep learning and the meaningfulness of data.
- Data Processing Architectures: optimised and scalable architectures for analytics of both data-at-rest and data-in-motion, with low latency delivering real-time analytics.
- Data Protection: privacy and anonymisation mechanisms to facilitate data protection.
- Data Management: principles and techniques for data management.
- The Cloud and High-Performance Computing (HPC): effective BD processing and data management might imply the effective usage of HPC infrastructures.
- IoT, CPS, Edge, and Fog Computing: a main source of BD is sensor data from an IoT context and actuator interaction in cyber physical systems. In order to meet real-time needs it will often be necessary to handle BD aspects at the edge of the system.

The horizontal concerns do not imply a layered architecture; for example, data visualisation may be applied directly to collected data (data management) without the need for data processing and analytics.
Vertical concerns address cross-cutting issues, which may affect all the horizontal concerns and may also involve non-technical aspects:

- BD Types and Semantics: the following 6 BD types have been identified, based on the fact that they often lead to the use of different techniques and mechanisms in the horizontal concerns, which should be considered, for instance, for data analytics and data storage: (1) Structured data; (2) Time series data; (3) Geospatial data; (4) Media, Image, Video and Audio data; (5) Text data, including Natural Language Processing data and Genomics representations; and (6) Graph data, Network/Web data and

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Metadata. In addition, it is important to support both the syntactical and semantic aspects of data for all BD types.

- Standards: standardisation of BD technology areas to facilitate data integration, sharing and interoperability.
- Communication and Connectivity: effective communication and connectivity mechanisms are necessary in providing support for BD.
- Cybersecurity: BD often need support to maintain security and trust beyond privacy and anonymisation. The aspect of trust frequently has links to trust mechanisms such as blockchain technologies, smart contracts, and various forms of encryption.
- Engineering and DevOps for building Big Data Value systems: this topic will be elaborated in greater detail along with the NESSI Software and Service community.
- Marketplaces, Industrial Data Platforms and Personal Data Platforms (IDPs/PDPs), Ecosystems for Data Sharing and Innovation Support: data platforms for data sharing include, in particular, IDPs and PDPs, but also other data sharing platforms like Research Data Platforms (RDPs) and Urban/City Data Platforms (UDPs). These platforms facilitate the efficient usage of a number of the horizontal and vertical BD areas, most notably data management, data processing, data protection and cybersecurity.

Standardisation is essential to create a Digital Single Market and Data Economy, ensuring interconnectivity and interoperability in an ICT-centric economy. The PPP lead the way in the development of technology and data standards for BD by:

- leveraging existing common standards as the basis for an open and successful BD market;
- supporting Standards Development Organisations (SDOs), such as ETSI, CENCENELEC, ISO, IEC, W3C, ITU-T and IEEE, by making experts available for all aspects of BD in the standardisation process;
- aligning the BDVA Big Data Reference Model with existing and evolving compatible architectures;
- liaising and collaborating with international consortia and SDOs through the TF6SG6 Standards Group and Workshops;
- integrating national efforts on an international (European) level as early as possible;
- providing education and educational material to promote developing standards.

In 2020, BDVA members decided to strengthen the Association by giving it a new mandate and expanding its scope and breadth of activities. In 2021, BDVA becomes DAIRO, which stands for Data, A.I. and Robotics (DAIRO), as evidence of the ambition of the Association to closely collaborate with other communities in order to jointly engage at the intersection of the key disciplines of Data, AI and Robotics.

3.3. The GAIA-X Project

The GAIA-X project [7] was launched in autumn 2019 to push forward the development of a sustainable and innovative data economy in Europe. GAIA-X aims at creating a secure and federated data infrastructure in Germany and in Europe more broadly, strengthening data sovereignty for business, science, the state, and society, in the context of storing, exchanging and using data and services. The GAIA-X ecosystem extends domain knowledge among European players, reinforces the exchange and sharing of data within domains and between domains, and paves the way for innovative business models in the digital market.

GAIA-X connects federated centralised and decentralised infrastructures (in particular cloud services and edge services) into a homogeneous, user-friendly system. Those infrastructures make data and services available for AI applications while protecting rights, interests and intellectual property.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Each node of the federated infrastructure forms an autonomous unit, follows the reference architecture, and is unambiguously identifiable and reachable. A self-description of the node provides transparency specific aspects and capabilities, such as statements on the location of data storage and processing, on the technologies used, on calculation and storage performance, and on the functionality made available. Moreover, further characteristics are stated, such as real-time capability, data sovereignty based on certified degrees of protection, consumption models, etc.
The GAIA-X ecosystem is structured into the Infrastructure Ecosystem and the Data Ecosystem. Activities in the Infrastructure Ecosystem are focused on providing or consuming infrastructure services, which are represented in GAIA-X by the asset called Node. In Data Ecosystems, the main Asset is called Data. Participants, typically representing organizations engaged within these ecosystems, are differentiated into two major roles, that is Provider and Consumer. A Participant could be a Provider as well as a Consumer at the same time. Data and Infrastructure Ecosystems are not separable. Services are the binding element between Providers and Consumers, also tying Data and Nodes together.
The GAIA-X ecosystem relays on a common and solid foundation consisting of Policy Rules, an Architecture of Standards of interconnection. Figure 3 [7] gives a high-level overview of the GAIA-X reference architecture.

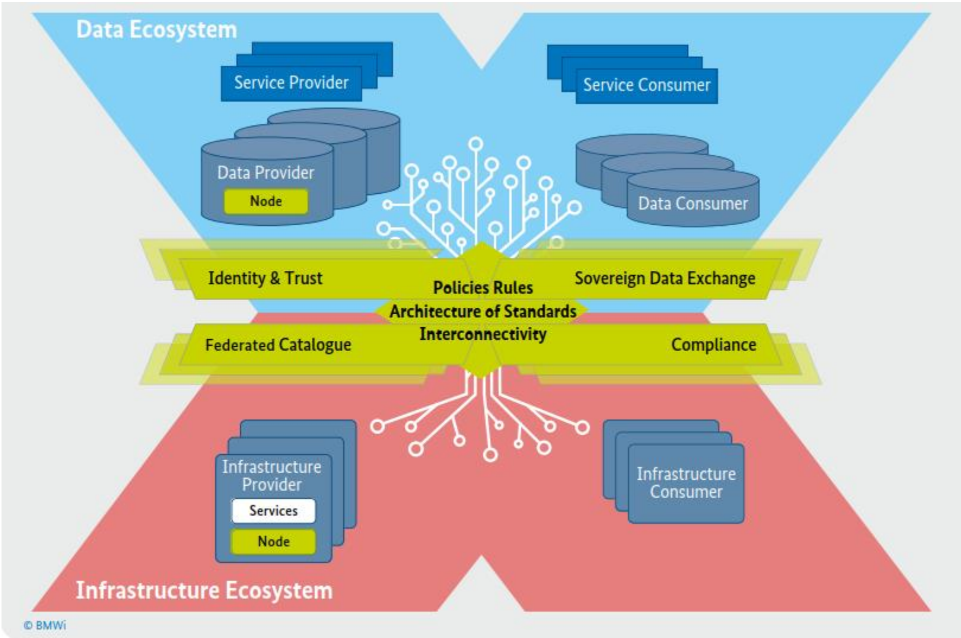


Figure 3 - High-level overview of the GAIA-X architecture

A Node is a computational resource or infrastructure element, such as an edge resource, a virtual machine, a container, a data centre or any other generic infrastructure building block where services can be deployed on. Nodes are placed in the Infrastructure Ecosystem.
A Service describes a cloud offering including all kinds of cloud services made available by a Service Provider. If a service relays on a GAIA-X Node, it is called a Service Instance. Services can be combined thus enabling the creation of service cascades. During all Service executions, GAIA-X enables interoperability, so that Services become portable from one Node to another.
A Data Asset is a data set that is available or hosted on GAIA-X Nodes and provided to consumers via GAIA-X Services. Data Assets can be combined with each other forming data spaces according to the federated character of the ecosystem.
A GAIA-X Participant, which is typically a business organization, a natural, or any form of a legal person participating in the ecosystem, obtains the role of a provider, consumer or both, depending on the respective business case. Each Asset has an associated GAIA-X Provider belonging to the GAIA-X ecosystem so that no provisions from outside without self-descriptions are allowed. Other roles

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

envisioned in the GAIA-X ecosystem are Service Provider, Service Instance Provider, or Node Provider.
Participants and Assets express their properties and characteristics through the obligatory GAIA-X Self-Descriptions, which contain structured, searchable metadata about concepts such as data owners and usage policies.
The core of GAIA-X is represented by the Federated Services, which interconnect the Infrastructure and Data Ecosystem. The Federation Services comprise both infrastructure services and organisational support functions. They include Federated Catalogues, Identity and Trust Services, Compliance Services, and Sovereign Data Exchange Services.
GAIA-X Federated Catalogue is an open mechanism allowing existing and future ecosystems to have the opportunity to include domain specific catalogues into the GAIA-X Federation of ecosystems as long as they comply with the Transparency, Fairness, Security & Trust requirements of a GAIA-X Catalogue. GAIA-X is responsible for the verification and certification process for these catalogues. The discovery of Assets will take place by means of the Federated Catalogue: proper Self-Description are found through an open and transparent query algorithm satisfying consumer needs. GAIA-X Data Connectors are part of the Federated Catalogue.
Identity and Trust Federated Services ensure that Participants' identities can be verified and trust towards their capabilities and Assets can be created. Moreover, the federated identity management makes identities operable between different domains by connecting several national and international identity providers.
Sovereign Data Exchange services are established by GAIA-X to enable and ensure a natural or legal person´s ability to decide exclusively how his data are to be used. A prerequisite is interoperability on the metadata level, as it enables self-determined decision-making regarding the use and processing of his data. A key component of data sovereignty is the enforcement of Usage Policies, so-called Usage Control.
Figure 4 shows main relations between GAIA-X Assets and GAIA-X Participants. Participants can take on multiple roles.

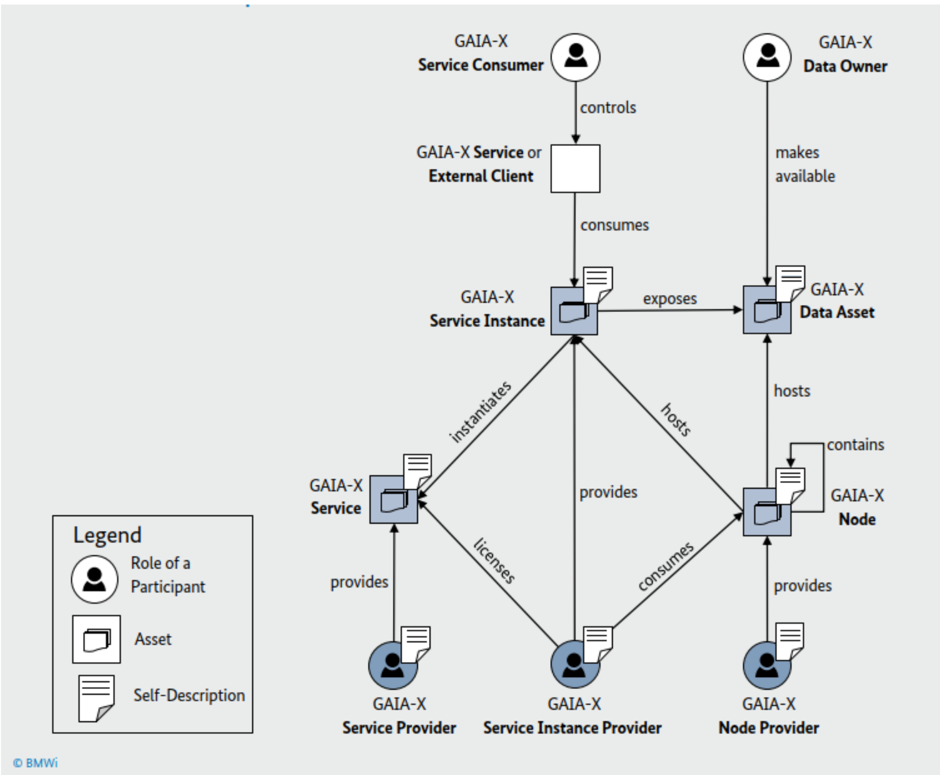


Figure 4 – Main relations between GAIA-X Assets and GAIA-X Participants

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

3.4. International Data Spaces Association (IDSA)

The Association (IDSA) is a non-profit organization founded jointly by business, politics, and research with the mission of establishing both the development and the use of a Reference Architecture Model for secure data spaces and sovereign data sharing on a European and International level. The IDSA defines a framework and governance principles for the Reference Architecture Model, as well as interfaces aiming at establishing an international standard.
International Data Spaces (IDS) Reference Architecture Model links different cloud platforms through policies and mechanisms for securing the exchange and sharing of trusted data (through the principle of data sovereignty). Indeed, industrial data clouds, individual enterprise clouds, on-premises applications and individual, connected devices can be connected to the IDS ecosystem through a dedicated software component called IDS Connector, as illustrated in Figure 5. The Connector works as an interface between the internal data sources and enterprise systems of the participating organization and the Industrial Data Space.

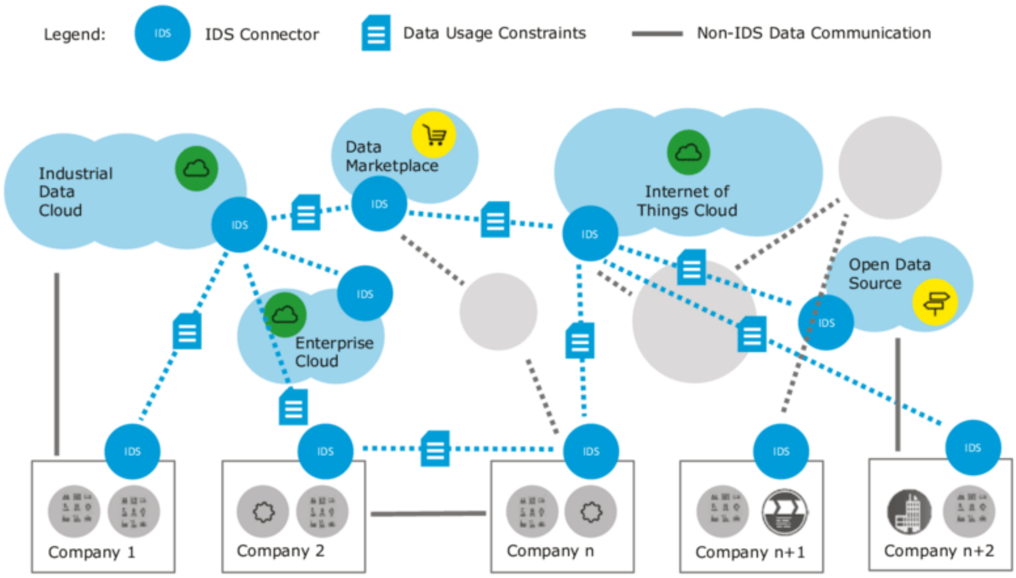


Figure 5 - Industrial Data Space and Cloud Platforms

As depicted in Figure 6, in conformity with common system architecture models and standards, the Reference Architecture Model uses a five-layer structure that express stakeholder concerns and viewpoints at different levels of granularity:

- Business Layer: specifies and categories roles, activities and interactions of the different stakeholders of the Industrial Data Space.
- Functional Layer: includes the functional requirements of the Industrial Data Space and abstract and technology-agnostic functionality of logical software components derived from the requirements.
- Process Layer: provides a view of the dynamic behaviour of the architecture by describing the interactions between the different components of the Industrial Data Space using the BPMN notation.
- Information Layer: uses the "linked data" principles to define a conceptual model describing both the static and the dynamic aspects of the Industrial Data Space constituents, for example Data Endpoints deployed, Data Apps advertised, or datasets exchanged.
- System Layer: relates to the decomposition of the logical software components, considering aspects such as integration, configuration, deployment, and extensibility of these components.

Furthermore, the three cross-sectional perspectives (Security, Certification, and Governance) of the Reference Architecture Model represents the implementation of core concepts of the Industrial Data

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Space regarding each of the layers. Establishing trust for data sharing and data exchange is essential in IDS. Two basic types of trust are defined in the Reference Architecture Model:

- Static Trust, based on the certification of participants and core technical components, and
- Dynamic Trust, based on active monitoring of participants and core technical components.

Some preliminary actions and interactions are necessary for every participant to share and exchange data in the IDS. These actions involve a Certification Body, Evaluation Facilities, and the Dynamic Attribute Provisioning Service (DAPS).

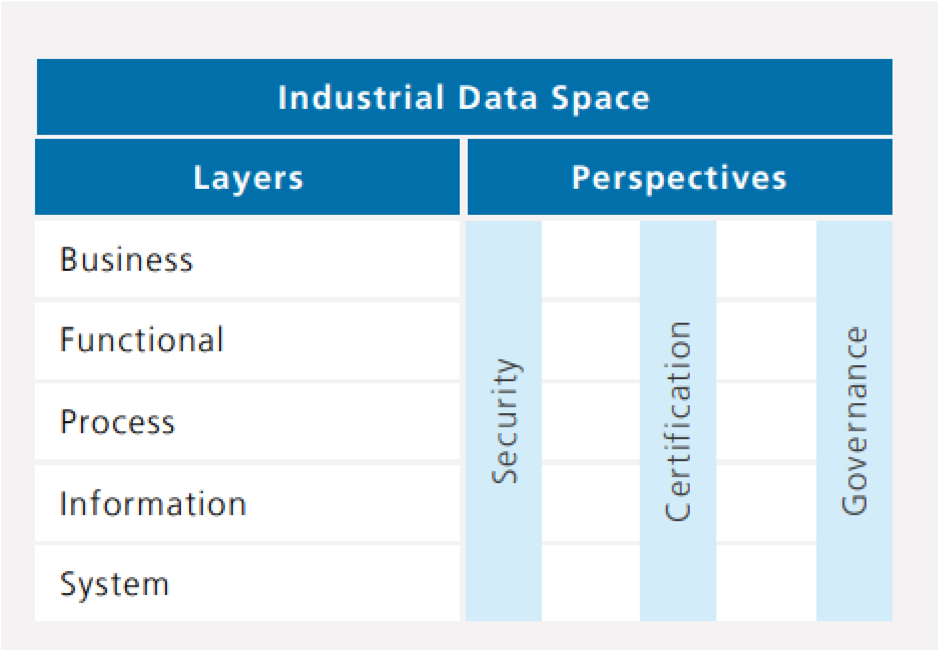


Figure 6 – General Structure of IDS Reference Architecture Model

Figure 7 represents the interactions of components on System Layer of the IDA Reference Architectural Model. Besides the Connector, which is the central technological building block, the IDS role model includes:

- Broker: broker service provider is an intermediary responsible for storing and managing information about the data sources available in IDS. It receives and provides metadata that allow provider and consumer connectors to exchange data.
- App Store: application provider develops applications that can be used by both data providers and consumers in the data space. Applications are typically downloaded from the remote app store and run inside the containerized connector. This role is optional in IDS.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

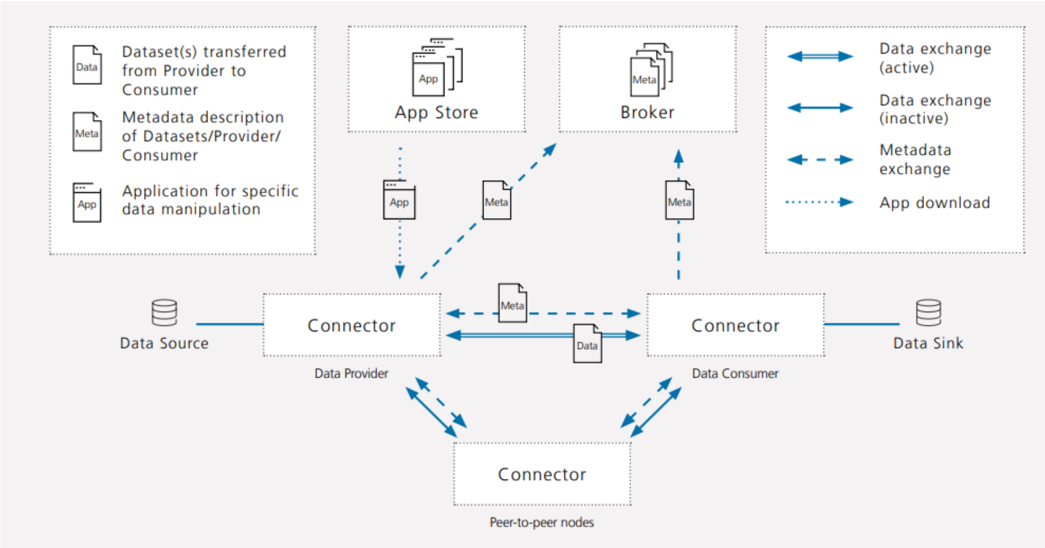


Figure 7 – interaction between technical components of IDS Reference Architecture Model

The Connector provides metadata to the Broker, which include description of the technical interface, authentication mechanism used, exposed data sources, and associated data usage policies. Only metadata is submitted to the Broker, while the actual data are transferred only between the Connectors of the Data Provider and the Data Consumer. Different types of Connector implementations are envisioned, based on different technologies and implementing different functions. Two basic examples are the Base Connector and the Trusted Connector. A Connector can be classified as external or internal. External Connectors execute the exchange of data between participants of the Industrial Data Space, providing data via the data endpoints it exposes. An External Connector is typically operated behind a firewall in a specially secured network segment of a participant (Demilitarized Zone), so that direct access to internal systems is not possible. An external Connector can be reached using the standard Internet Protocol (IP). A participant may operate multiple external Connectors, for example to meet load balancing or data partitioning requirements. External Connectors can be operated on-premises or in a cloud environment. Internal Connectors facilitate access to internal data sources in order to provide data for external Connectors. Internal Connectors are typically operated in an internal company network not accessible from outside.

Figure 8 [8] depicts the IDS Connector Reference Architecture. It uses Application Container Management technology to ensure an isolated and secured environment for individual Data services. Data Services are Data Apps that are deployed inside Connectors. Data Apps encapsule data processing and/or transformation functionality. In addition, Data Apps can be divided into:

- System Adapters, which exposes interfaces to external enterprise information systems and mainly add metadata to data (in addition to wrapping the enterprise information system).
- Smart Data Apps, which execute any kind of data processing. Normally, data provided from or to a Smart Data App are already annotated with metadata.

The Execution Core Container provides components for app orchestration and communication:

- Message Bus to a Connector, which stores data between services or Connectors, and
- Message Router, which executes multiple workflows, invokes Data Services according to defined workflow steps, and is also responsible for sending data to and receiving data from the Message Bus.

A Custom Container provides a self-developed Data Service and usually no certification is required. An App Store Container is a certified container downloaded from the App Store, providing a specific Data Service to the Connector.

A Data Service provides the executable activities for a workflow to be executed by a Message Router; it defines a public API that is invoked from a Message Router. Data Services can be implemented in

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

any programming language and target different runtime environments. Furthermore, existing components can be reused to simplify the migration from other integration platforms. The Runtime of a Data Service depends on the selected technology and programming language.
The Configuration Manager represents the administrative part of a Connector and it is responsible for managing and validating of the Configuration Model, followed by the execution of a deployment. The deployment task is delegated to a collection of Execution Configurators by the Configurator Management. The Configuration Model is an extendable domain model for describing the configuration of a Connector. The Configurator Management loads and manages an exchangeable set of Execution Configurators, which are exchangeable plug-ins that execute or translate single aspects of the Configuration Model to a specific technology.

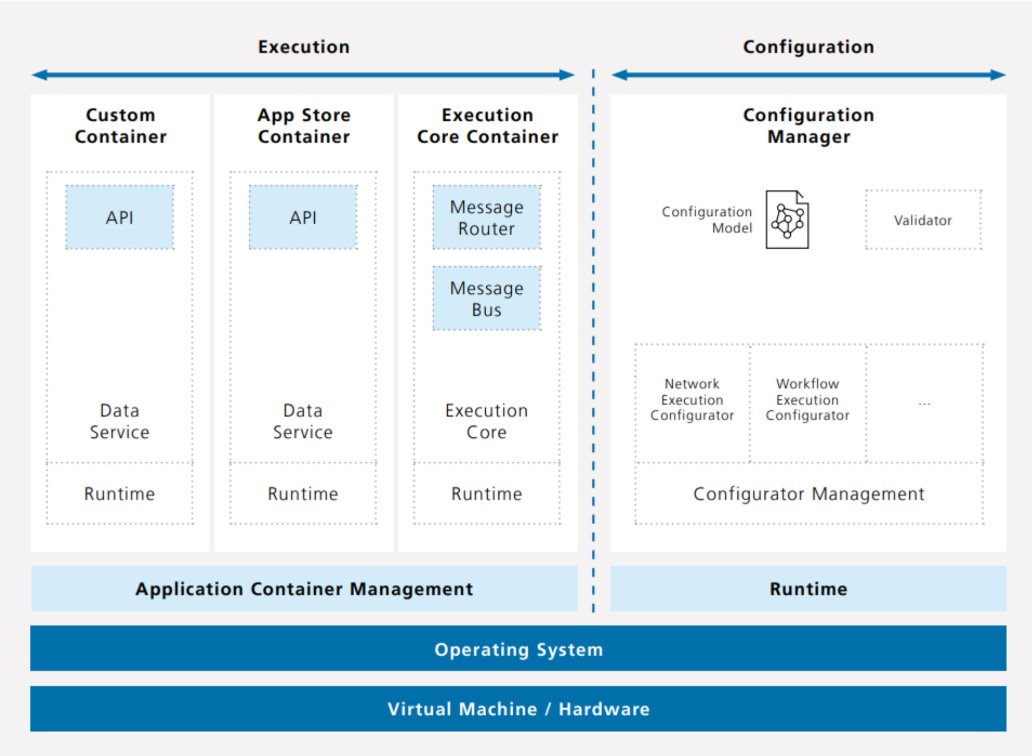


Figure 8 – IDS Connector Reference Architecture

3.5. Comprehensive Architecture for Smart Grid (COSMAG)

Comprehensive Architecture for Smart Grid refers to the collection of specifications defining any possible interaction for data exchanging among all the possible actors in the context of the smart grid, starting from the analysis current standards to check if they offer the proper interfaces to enable traditional and new business processes or if it is necessary to highlight the need of new standards. COSMAG has not the aim to introduce any new standards, but rather to exploits and collects results of previous projects or standardization activities.
The interactions must be defined to support the implementation of the vision of the European Commission as from the Winter Package.
The reference architecture must offer “open gates”, that is data interaction points that can be used for future expansion and futuristic UCs.
Figure 9 represents the starting point for the analysis to identify all the possible communication options; it depicts the structure of the modern electricity market and actor interactions. The idea is to analyse each possible flow to understand the maturity level per type of flow, that is the current status of protocol definition and data models. This activity depends on regulation, business models, technologies, etc.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

This list of actors derived from Figure 8 is not complete; other actors could be identified, such as suppliers, community manager or data manager and others.
Classical generation is not considered in this study since the operation of such a player is already well established, as well as all the interfaces to Wholesale Market give that they have been well established too.
A hypothetical new player called local market operator has been considered in COSMAG, even if this operator could potentially be the DSO but this identification is not yet accepted generally.
A preliminary overview of data exchange and data format in the Smart Energy sector is available at [14]

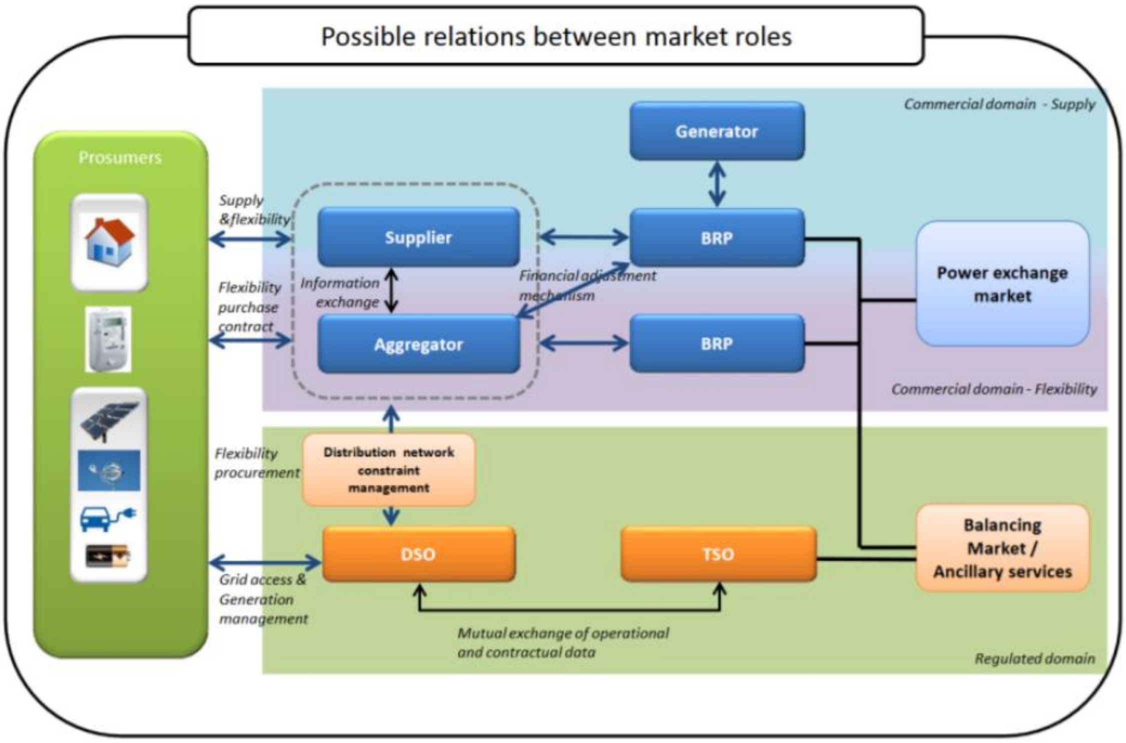


Figure 9 – The structure of the electricity market and actor interactions

Given the future role, the interfaces at level of customer/prosumers are extremely important. Indeed, large amount of data are related to the customer/prosumer level, therefore, data platform able to aggregate data are an important part of the picture and these data play a central role in the energy system. The digitalization of the energy sector demands Higher levels of operational excellence and new technologies are necessary to foster cross-domain data sharing and data-driven innovation that is not so simple to integrate into a single data management solution. The needed services cover a wide range of functionalities and their business opportunities affect various vertical energy domains. In this re-invented market value chain context, new energy actors are not the only newcomers, but they should also rely on incumbent actors that bring emerging technologies at ICT level, thus entailing more complexity and fragmentation to an already heterogeneous environment.
The goal of reference architectures is to contribute to the creation of controlled collaboration spaces where existing and emerging technologies could better exploit in secure and trustable ways the data provided by energy actors and the insights derived by data innovators.
The interoperability between the key elements of each platform and space is one of the more relevant aspects of the platform. Common standards agreed data models and ontologies, comprehensive documentation and Open APIs should harmonise the access to data and data processes, and facilitate the interactions among the different actors and platforms in a Federated Data Solution Space (Figure 10).

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

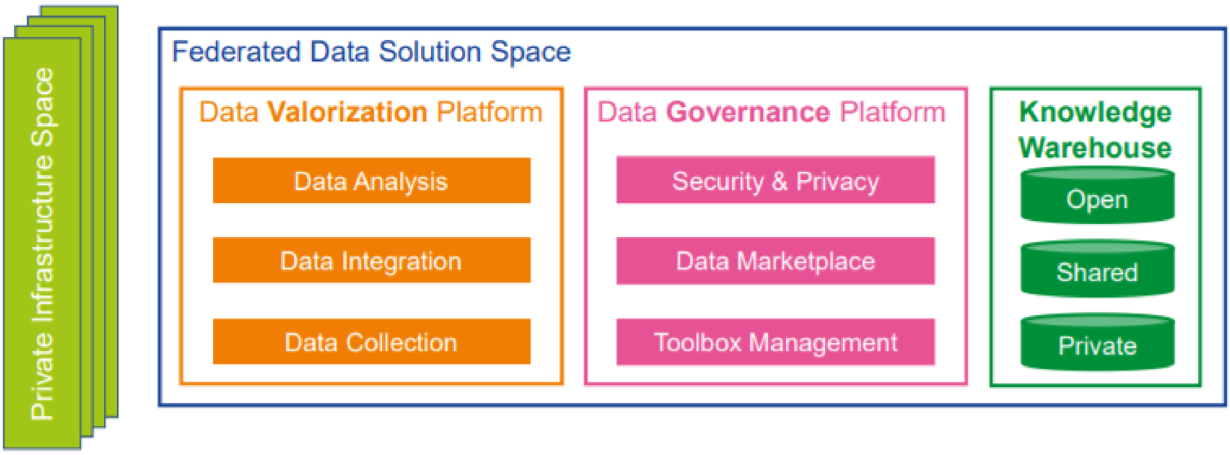


Figure 10 – Federated Data Solution Space

Key elements in data management to support a data economy can be summarized as follow:

- Data business model, defining appropriate data models to enable the interoperability;
- Data Commodity Monetisation, creating proper business models without violating privacy;
- Semantic, facilitating interoperability through clear and recognized data modeling;
- Context, defining context for bridging through different verticals;
- Sovereignty, defining who can manage the data and for which purpose to protect privacy and customer interest;
- Open API, offering the perfect bridge between private infrastructure spaces.

Another key element to take into consideration is the emerging role of sector coupling, avoiding data silos, not just between electricity, heat, gas, etc. within the energy sector, but also coupling of services with other sectors such as health, security, etcetera. In this respect new standards as emerged from the recent work of ETSI, e.g., the FIWARE Context Broker, are of paramount importance.

3.6. Data Streaming platform for ENTSOE Transparency Platform

The TP was implemented under the Electricity Market Fundamental Data Information Platform (EMFIP), the project with the primary goal to gather and publish data related to the energy market in a predefined form.
The ENTSO-E Transparency Platform provides free, continuous access to pan-European electricity market data for all users, across six main categories: Load, Generation, Transmission, Balancing, Outages and Congestion Management. The platform is used as a central collection and publication of the electricity generation, transportation and consumption data and can be considered as a good example how the big data can serve the pan-European market. This publication offers a one stop shop for TSO/ENTSO-E propriety data.
Data submission is made using IEC 62325-451-x series of standards and the principal data exchange mechanism is IEC 62325-503 (MADES).
From the legal perspective, the TP is the IT system designed to implement COMMISSION REGULATION (EU) No 543/2013, the transparency regulation. The regulator states the following:
“That a central information TP shall be established and operated in an efficient and cost-effective manner within the European Network of Transmission System Operators for Electricity (the

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

'ENTSO for Electricity'). The ENTSO for Electricity shall publish on the central information TP all data which TSOs are required to submit to the ENTSO for Electricity in accordance with this Regulation." [9]

"The central information TP shall be available to the public free of charge through the internet and shall be available at least in English. The data shall be up to date, easily accessible, downloadable, and available for at least five years. Data updates shall be time-stamped, archived, and made available to the public."

It is foreseen that the implementation of new requirements linked to the publication of the balancing energy bids (article EB GL 12.a) will significantly increase the volume of messages to be processed by TP. Some estimates showed that the system might need to cope with processing up to a million messages per hour. That new technical requirement makes the processing of energy bids message a suitable UC for testing the performance and benefits of a streaming platform.

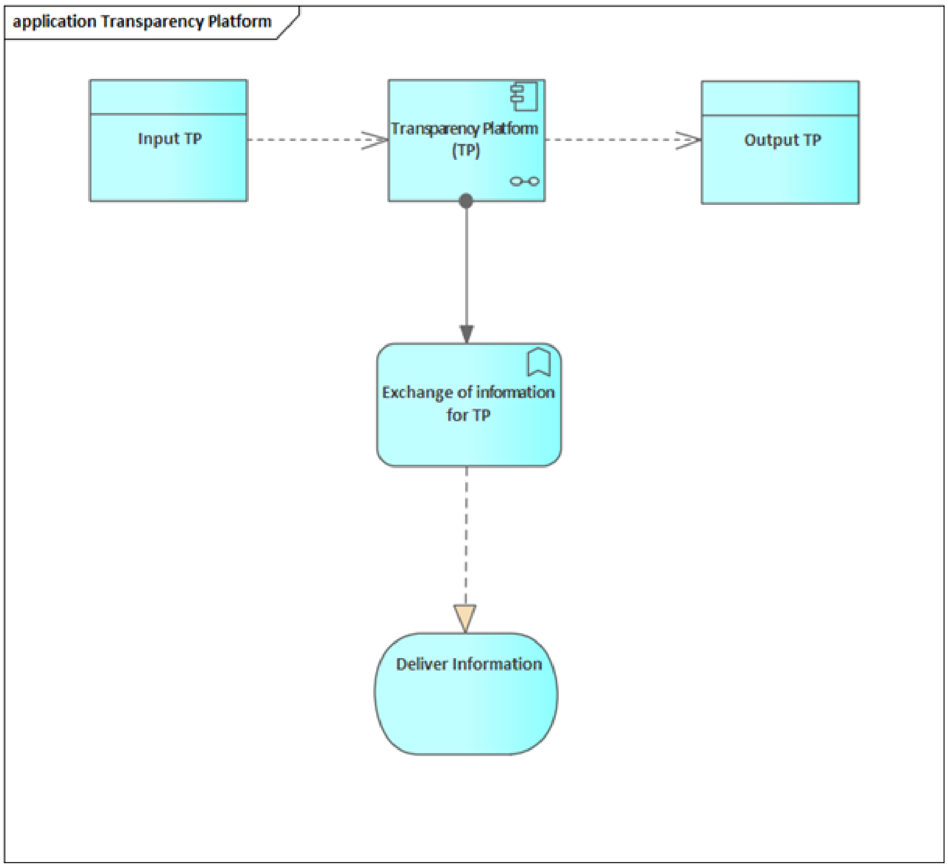


Figure 11 - Transparency platform Architecture Building Blocks

Architecture Building Blocks' (EABB) of a Streaming Platform

This section lists technologies and technical building blocks required for implementing of the streaming platform proof-of-concept for the TP.

Kafka cluster

Concerning the Kafka version, version 2.5.1 (Apache open-source project) was used with docker image for its deployment. Alternatively, Confluent also provides docker images where the recommended version is EAP 5.5.1. About the security of the cluster, the platform offers configuration files that enable the following features:

- Network data encryption using SSL/TLS.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

- A producer or a consumer authenticates using SASL protocol.
- Access control List (ACL). This feature allows a user to read/write to a specific topic.

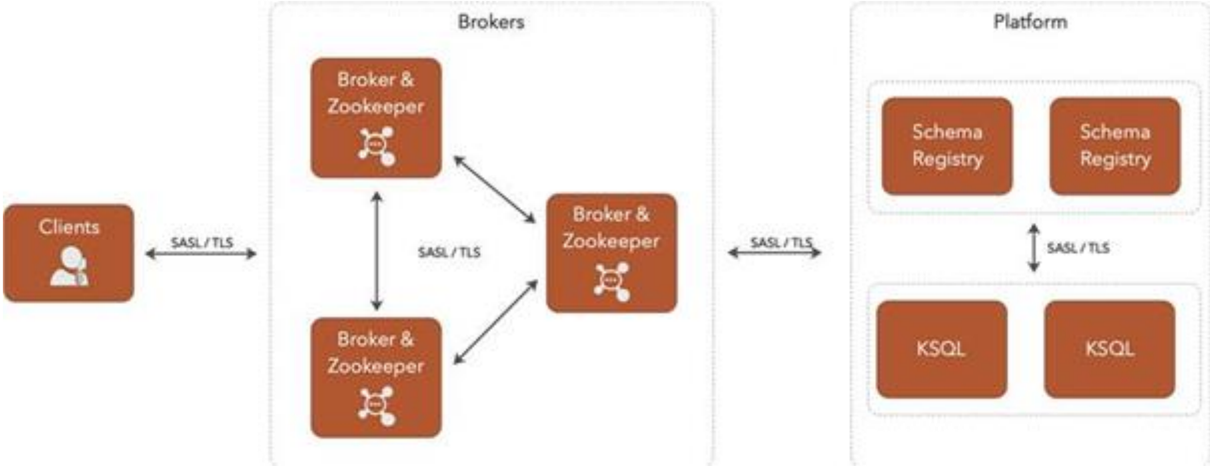


Figure 12 - Secured Kafka cluster

Single Sign-On & Federation Identity Management (FIM)

An Identity Management Broker (IMB) is an intermediate service that centralizes the management of identities across different areas or security domains. As such, IMB facilitates the development of applications and microservices as they remove the necessity from development teams to code the security logic. It is also essential to notice that IMB rely themselves on identity providers which are using specific protocols such:

- SAML 2.0
- OAuth 2.0
- OpenID Connect

Keycloak [10] is an open-source IMB that it is highly configurable. In this study, we had tested Keycloak with SAML 2.0, and OAuth2.0 & OpenID connect. Figure 15 illustrates the deployment of Keycloak broker with SAML 2.0.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

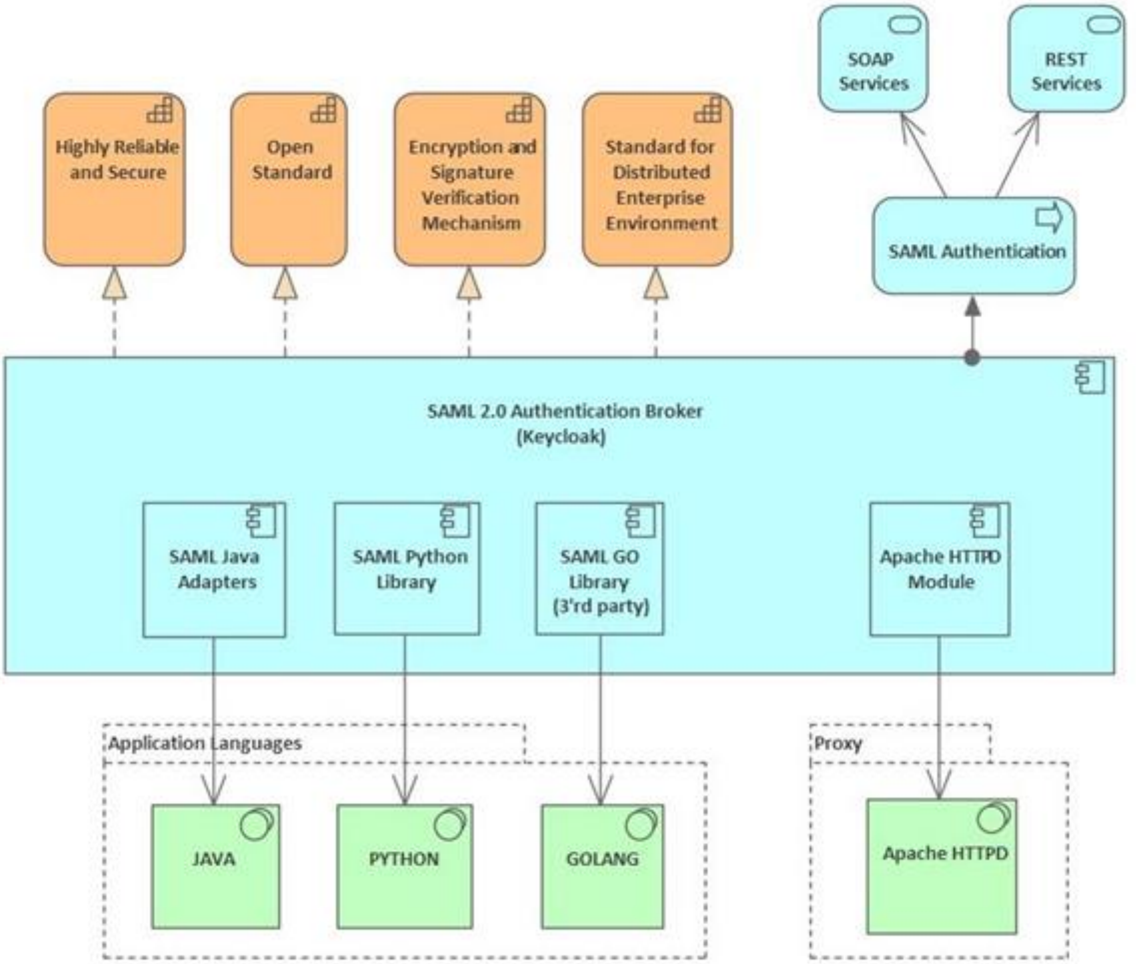


Figure 13 - Keycloak configured with SAML 2.0

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

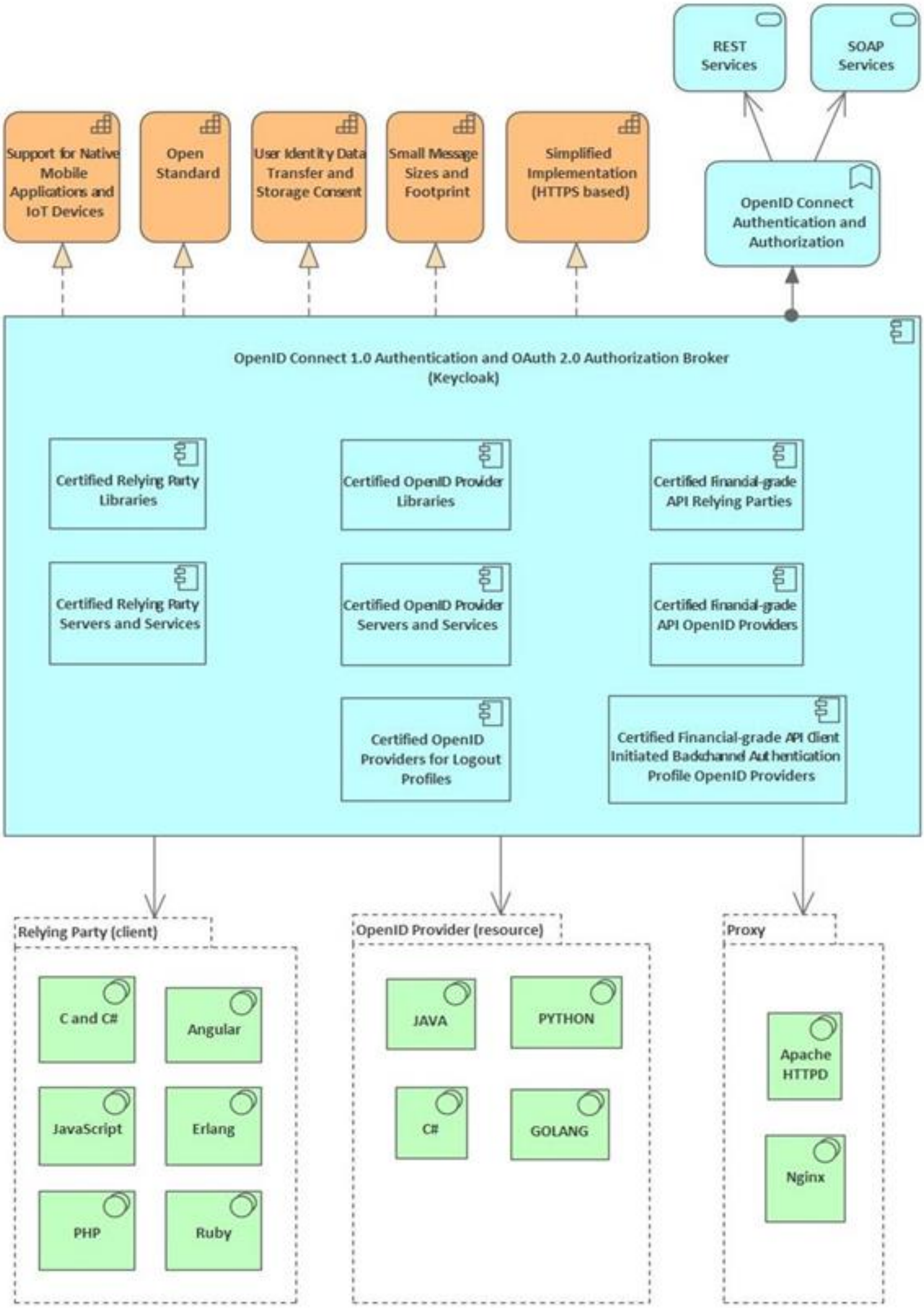


Figure 14 - Keycloak configured with OpenID

It is important to notice that the day-to-day management of IMB is requiring the adoption of new security policy, best practices and implementation guidelines. We found also that many companies are creating new roles of Site Reliability Engineer (SRE). These SRE Engineer are defining reliability goals which are continually measured and working to improve the quality of service as needed.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Transparency Platform – (POC)

This sub-section presents the proof of concept (POC) that was realized during this study. Main objectives of the POC were the following:

i. Create an application that can process up to a million of bids messages per hour.
ii. Elaborate an Architecture that implements Microservices.
iii. Demonstrate that proposed Architecture is scalable and fault tolerant.
iv. Demonstrate that the Architecture allows the development and the integration of independent components (no vendor lock-in).
v. Implement state-of-the-art security patterns.
vi. Implement a GUI that mimic the existing TP.

Context Diagram

Figure 15 presents the realized prototype. On the diagram, one could see that 6 microservices were developed. The following table represents the implemented microservices with their languages. In addition, one could observe that we implemented the Command Query Responsibility Segregation (CQRS) design pattern which decouples the inbound from outbound data flows.

Table 1: Implemented micro-services

Service Name	Language
Business Rule engine	Java
NTP Proxy	Python
State service	Python
Faust Consumer	Python
NTP Publisher	Python
NTP dashboard	Python

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

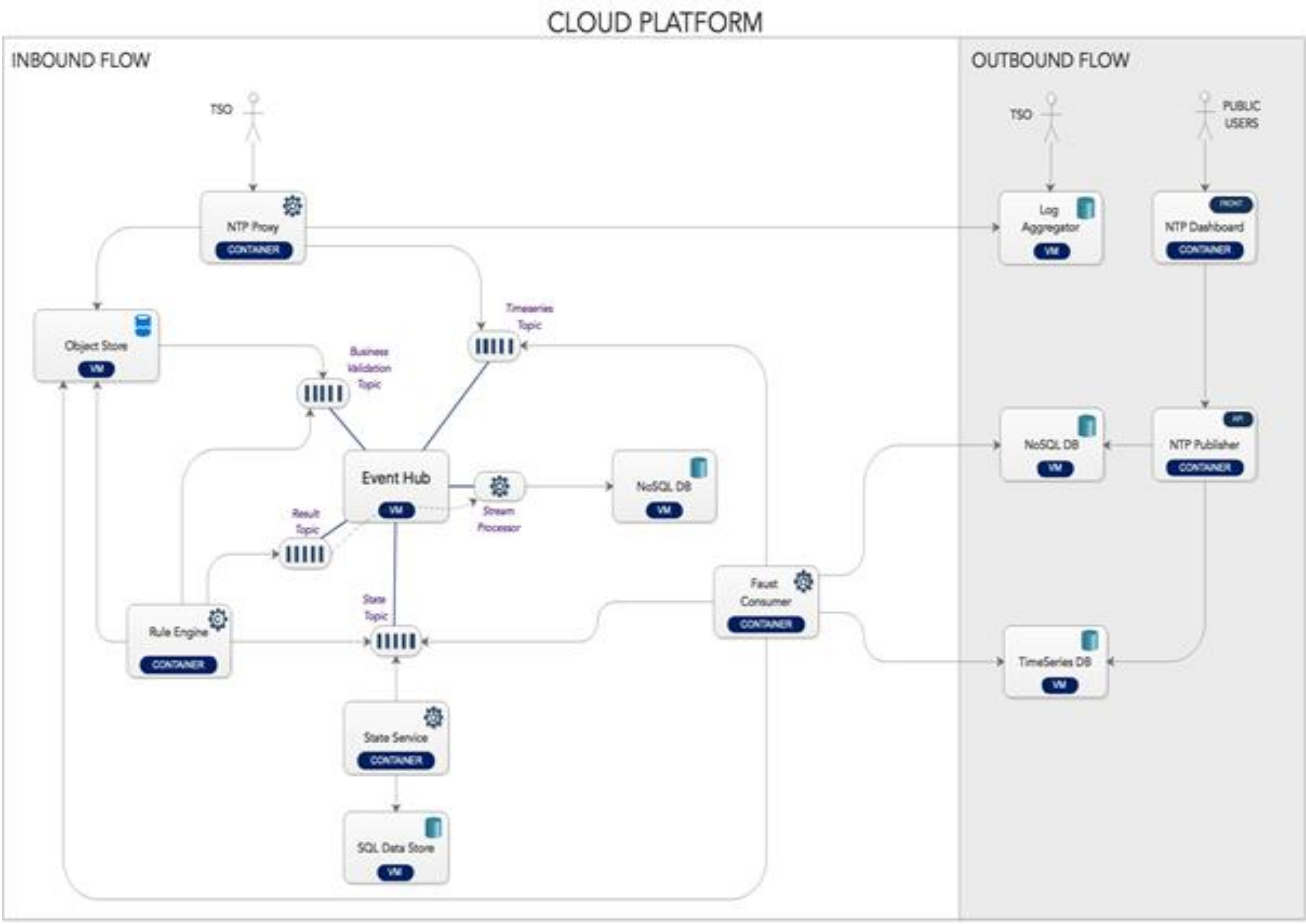


Figure 15 - Context diagram of the Transparency POC

The following table presents the number of lines of code that were implemented during this POC.

Table 2: Number of lines of code per language

Language	Component(s)	Number of lines of code (LOC)
Java	Business Rule	
Python	Backend	2408
Vue.js	GUI	537
Script.sh	Maintenance script	296
Xml	XSD validation data	15000
Perl	Maintenance script	64

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Technical Architecture

The Figure 16 presents the application components and used technology. Concerning the used technologies, we found out that Kassandra offers a greater integration, efficiency than its separate components.

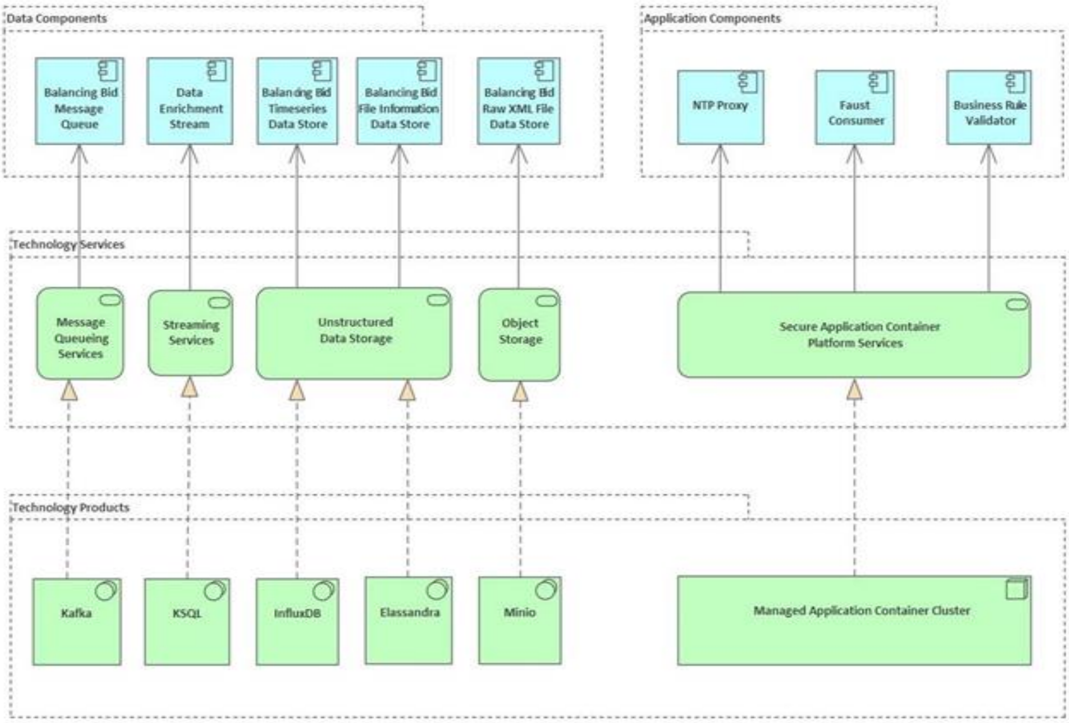


Figure 16 - Technology overview

3.6. Presentation Layer

This sub-section presents the presentation layer that was mimic the existing TP. The following figure represents the reception of the energy bids messages.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

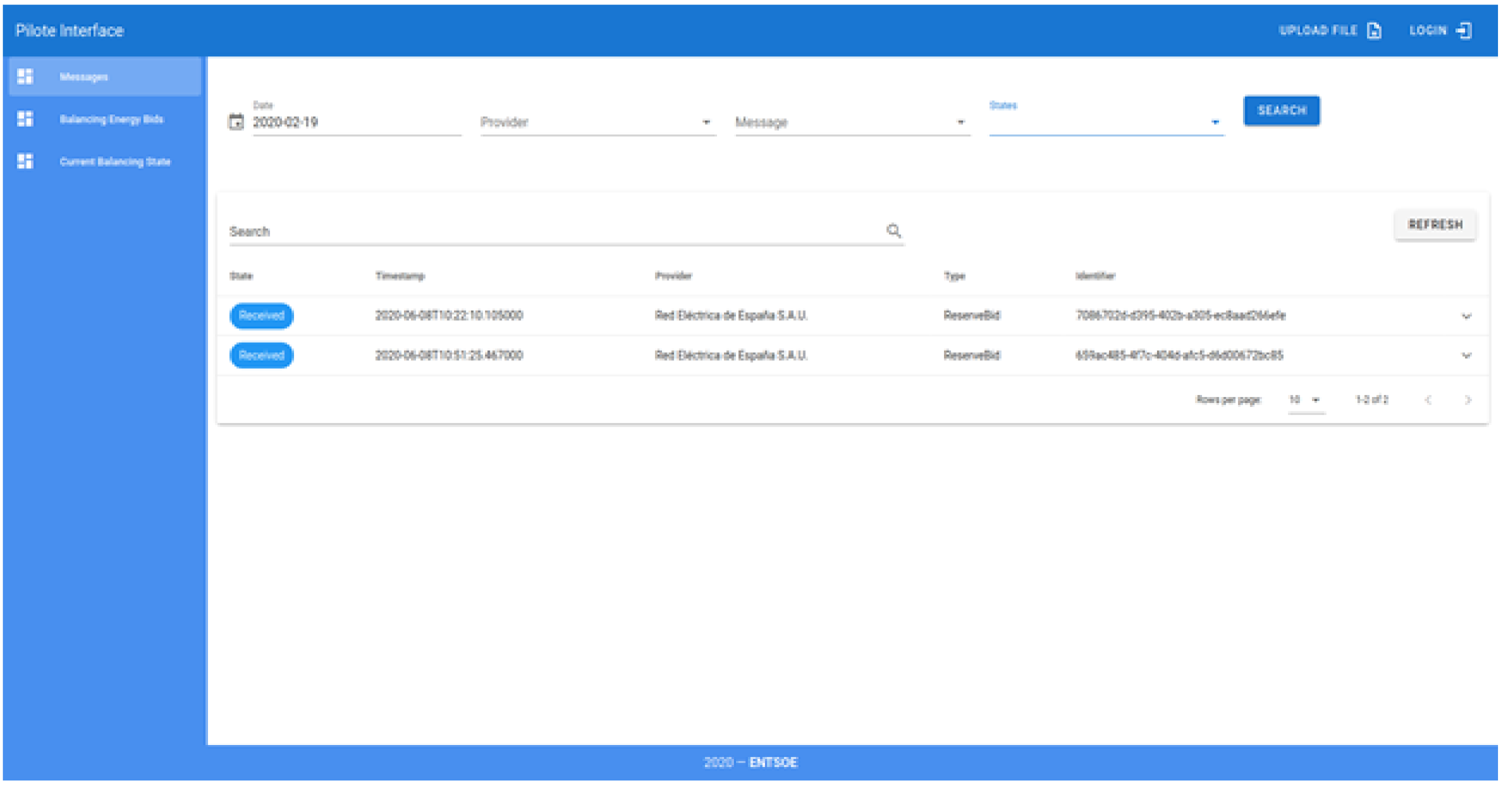
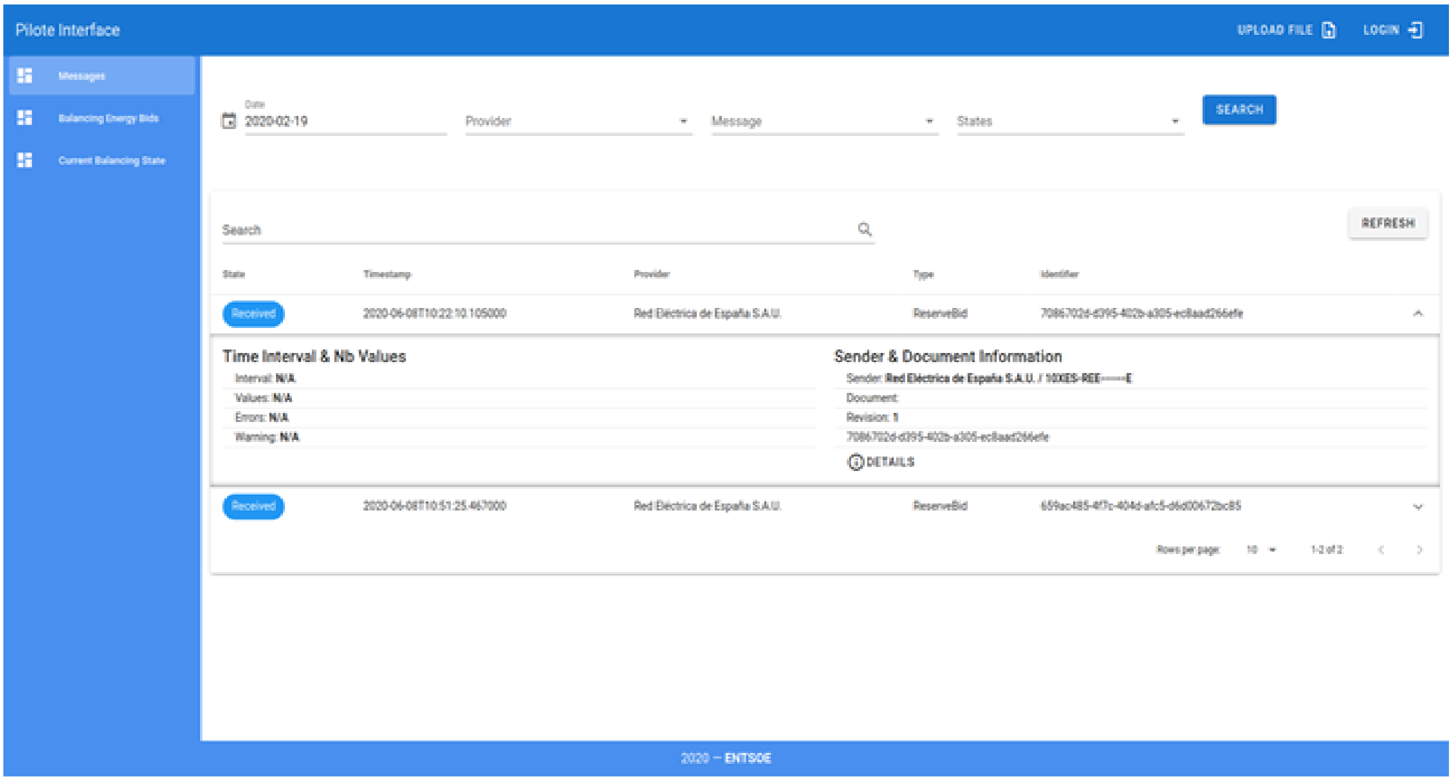


Figure 17 - Energy bid message reception

The following screen represents details of a Bid message:



BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Figure 18 - Detailed view of a Bid message

Tableau report

The adoption of Microservices and the CQRS design pattern gave us the possibility to integrate very easily with legacy system. The following figure illustrates the integration of the backend with Tableau reporting tool.

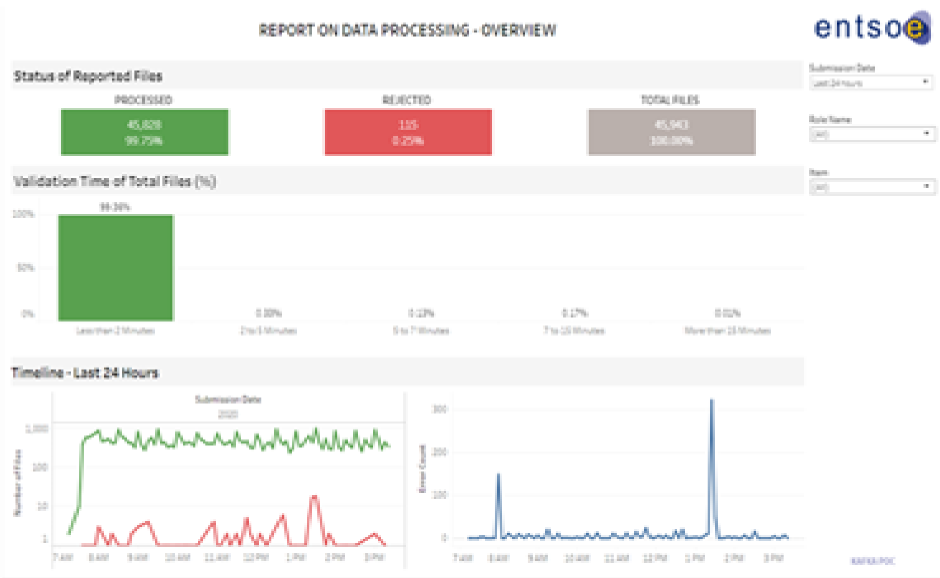


Figure 19 - Tableau dashboard integration

Performance

Current system can handle the processing of 432k messages. In order to validate the performance and scalability of the proposed architecture, the following performance test was executed.

Table 3: Benchmark result

Concurrent User	Message size	Performance
200	XML file with 200 bid messages	147 million per hour

The result of the performance test is very good and validate the scalability and resilience of the proposed Architecture design. In addition, the new Architecture should allow to reduce significantly the operational cost (OPEX) of future application(s).

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

4. Value Chain and BD4NRG Scenarios

BD4NRG aims to address the emerging challenges in BD management (introduced in Chapter 3) in the energy sector by providing an open holistic solution that proposes a European approach to the business to business (B2B) platforms able to provide competitive edge services to the stakeholders involved in the energy value chain and to create new market opportunities.

As matter of fact, regulatory, security, organizational, privacy and domain-specific technological / communication barriers still hamper the full-scale realization of the data economy in the energy value chain, even if the opportunities are evident. BD4NRG intends to tackle those barriers and to create a real data economy for energy by proposing an incentivizing conceptual and technological framework facilitating and motivating energy stakeholders to share data. Indeed, despite a number of architectures and underlying technological implementations of BD solutions for smart energy grids has emerged, the lack of a standardized approach for BD architecture interoperability is clearly lagging behind. While most value chains are cross-sector, existing BD applications in B2B energy sectors tend to be "siloed" and do not provide easy ways for sharing, reusing and repurposing data assets, cognitive models and AI services across domains/stakeholders. The lack of semantic and business interoperability across data stream providers is hindering the high potential of BD in the energy domain. Combined data from different domains (energy, climate, economics and financing, user behaviour and societal trends, regulations and many more) could pave the way for a wide spectrum of novel services. Dealing with cross-domain data requires a precise understanding of their nature and the support of new digital technologies, like AI and BD services, to process large quantities of data efficiently, within tolerable times and taking into consideration best practices to ensure data security. The lack of energy-tailored data ownership and sovereignty management policies has actually preventing in many cases to make available and share data. Moreover, limited data provision for near real time proactive analytics-based decision-making and control has implied a limited harnessing the capability of capturing new value streams from BD.

Lastly, the traditional approach to power system data management focuses only on the most crucial data that are immediately needed to supervise and oversee predesigned operations and applications. For example, in most protection relays and related sensors, the data collected is discarded shortly after internal use. For many power system sensors, including many emerging and state-of-the-art sensors, the majority of data is either not logged, or they are overwritten very quickly (lack of support for BD management and persistence capabilities).

The overall vision and main objective of BD4NRG is to deliver an innovative framework which leverages on AI-based cross-sector analytics for integrated and optimised smart energy grid management (including operation and planning), based on seamless data-information-knowledge exchange under respective sovereignty and regulatory principles.

BD4NRG intends to align and integrate the existing specifications for data-driven initiatives and underlying B2B Reference Architectures at the interplay among smart energy grids, AI, IoT, BD, smart home/building, industry/manufacturing, to provide a living reference architecture specifically tailored to the energy value chain, which will lay the foundation for facilitating seamless B2B multi-lateral interoperability.

In this chapter context scenarios of each Large-Scale Pilot (LSP) have been collected, sorted from LSP 1 to LSP 12. First of all, the context where BD4NRG tools and solutions will be demonstrated was depicted, followed by the description of the actors involved and the path to be followed within the project to achieve the objectives.

The overall BD4NRG service analytics reference framework will be demonstrated in these 12 LSP, implemented in 10 countries. Rationale of pilot applications is to address in a combined way two of the major challenges actually hindering big data analytics value capturing smart energy grids:

1) Nurturing the shifting towards predictive / prescriptive analytics;
2) Enabling multiple data sources analytics for multiple applications.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

The following Figure 20 [15] shows the geographical location of the pilots and the type of application to be demonstrated.

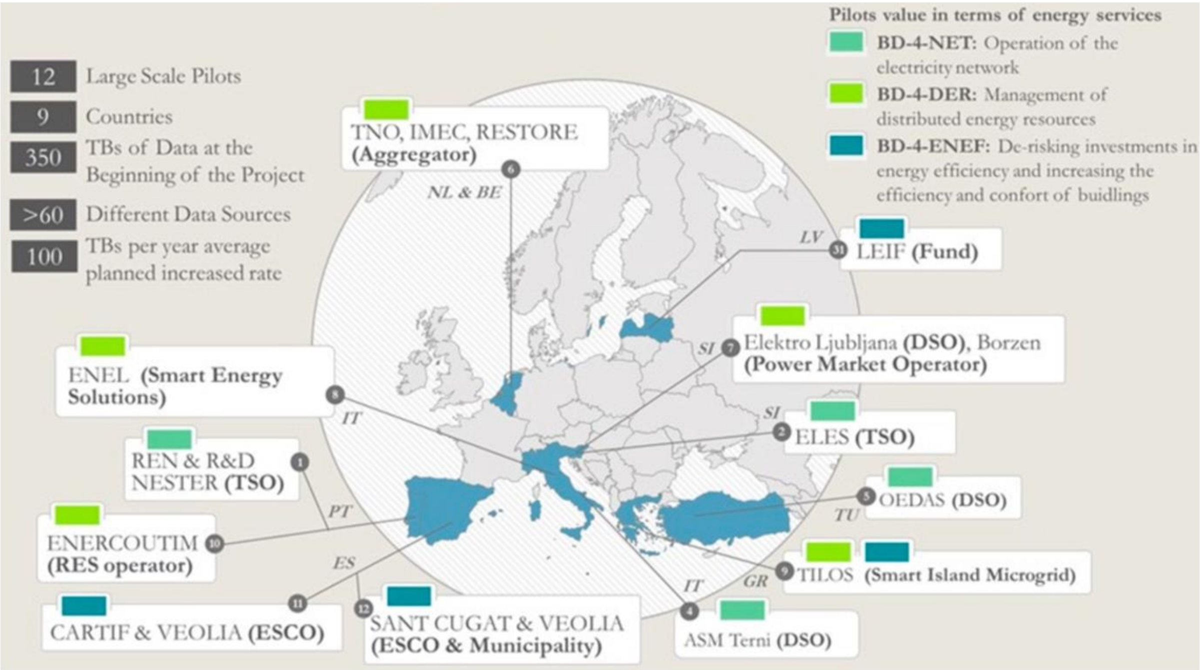


Figure 20 - Pilots' location

The BD4NRG pilots aim to analyse different components of the energy system, in particular 3 types of LSP are distinguished.

- BD-4-NET LSPs Increasing the efficiency and reliability of the electricity network;
- BD-4-DER LSPs Optimizing the management of assets (Distributed Energy Resources – DER) connected to the grid;
- BD-4-ENEF LSPs De-risking investments in energy efficiency and increasing the efficiency and comfort of buildings

The following tables indicates the different LSPs with their title, location and leading partners, distinguished with their placement within the BD4NRG project.

Table 4 - BD-4-NET LSPs

Pilot	LSP Description	Leader Partner	Location
LSP 1	Integrating off-grid Data with Condition based Monitoring for Enhanced Predictive Asset Maintenance	REN & RDN	Portugal
LSP 2	Cross-functional Integration of Grid Operation with Predictive Asset Management	ELES	Slovenia
LSP 3	Cross-Domain Cross-functional Integration of Predictive Asset Management and Grid Operation while coordinating with EVs smart charging	ASM	Italy
LSP 4	Cross-functional Predictive Asset Management for MV Grid Planning	OEDAS	Turkey

Table 5 - BD-4-DER LSPs

Pilot	LSP Description	Leader Partners	Location

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Pilot	LSP Description	Leader Partners	Location
LSP 5	Cross-stakeholder Transfer Learning for Flexibility Assets Forecasting	CENTRICA	Belgium
LSP 6	Coordinating grid-owned and behind-the-meter assets management prediction for improved grid operation and rear real time power market operation settlement	EKL, BORZEN & CS	Slovenia
LSP 7	Predictive Analytics Forecasting of storage optimized operation for improved market reserve participation	ENELX	Italy
LSP 8	Predictive Demand and Generation Forecasting for Optimized Local Energy Community Management	UNIWA	Greece
LSP 9	Collaborative aggregated energy generation prediction	ENERC	Portugal

Table 6 - BD-4-ENEF LSPs

Pilot	LSP Description	Leader Partners	Location
LSP 10	Predictive and Prescriptive Analytics for improved Energy Performance Certificates Reliability	CARTIF & VEOLIA	Spain
LSP 11	Weather-enhanced building thermal comfort prediction to enable energy efficiency reliable forecasting	AJSCV & VEOLIA	Spain
LSP 12	Predictive Analytics for Energy Efficiency Investments de-risking	LEIF	Latvia

4.1. Integrating off-grid data with Condition based Monitoring for Enhanced Asset Maintenance

Asset management departments of Transmission System Operators (TSO) are traditionally data intensive. Over the last years, with the ongoing digitalization of power systems, the amount of data streams and data sources have significantly increased, showing the need for improved data analytics tools, in order to derive predictive models for assets maintenance. These models are intended to extend the lifetime of assets and optimize O&M assets costs, hence enabling sustainable risk management and, in a later stage, the integration of asset management with grid improved network operation.

The proposed UCs, coordinated by REN and R&D Nester, aims to assess the condition of Circuit Breakers (CB) and Overhead Lines (OHL) in order to provide answers to key challenges in assets management: implementation of optimal condition-based maintenance plans, incident characterization, predictive maintenance and implementation of maintenance plans. Both are core elements of the electricity network and the challenges represented are related to data processing, due to the different nature of the associated datasets.

4.2. Cross-functional Integration of Grid Maintenance Planning with Predictive Asset Management

Nowadays increasing stochastic power generation behaviour with high penetration of renewable energy sources is compounding the effects of the traditionally stochastic consumption on the power system operation. This requires high flexibility of transmission system, system operation and closely connected activities such as maintenance activities and the associated equipment outages. Also

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

considering the asset management and the transition from time-based maintenance to condition base or predictive maintenance, the transmission system operation and planning are increasingly demanding.
Traditionally outage planning is typically done manually with minimal integration of various internal and external actors.
The primary goal of the UC is to develop a decision support tool to help the experienced human planners in maintenance scheduling process. Consequently, the outage planner will need less workhours to produce an outage plan.
This decision support tool will help TSO in its ambition towards fewer outages, reduced outage duration, better outage utilization for maintenance purpose, increased system reliability, easier and faster outage rescheduling, better coordination with outside partners, possibility of dealing with more complex outages and therefore operation of the power system closer to the operational limits.
Reversible analytics and better process understanding can be used to trigger planning for purchasing equipment with higher capabilities than typically needed with a view to determine how asset predictive maintenance timing may linearly affect TSO grid operational capability and power market transaction prediction, and how predictive models for grid assets maintenance may be affected by grid planning decisions.

4.3. Cross-Domain Cross-functional Integration of Predictive Asset Management for Grid Operation and coordination with flexible EVs charging

DSOs face already severe network stability problems to the power distribution grid due to the increasing share of intermittent generation from renewable energy and to the new loads as energy storage and electric mobility, which are modifying the operating conditions of the assets, originally designed for almost steady and unidirectional power flow. Moreover, the envisioned penetration of e-mobility is going to significantly affect residual hosting capacity of the power distribution networks; in this respect, a DSO aims to maintain properly the existing infrastructure.
On the one hand, asset maintenance is deeply linked to BD analytics of the sensors deployed in the network; on the other hand, it could be enhanced by leveraging active flexibility sources in the context of DR programs. In both cases, the amount of data, that a DSO will collect and treat, will be significantly higher than the current datasets managed and stored by utilities and their IT infrastructures. BD analytics will definitely boost the DSO as one of the main actors for supporting energy transition and customer-centric energy markets.

4.4. Cross-functional Predictive Asset Management for MV Grid Planning

Especially for the distribution companies which are serving in large areas, number of customers and demand for energy are increasing day by day. Increasing numbers of distributed energy resources such as PVs and battery storages are also creating some problems like voltage quality problems, capacity problems etc. Because of these reasons, distribution network operators are facing with more problems than before. Thus, there is a need for optimal network operations and also a need for professional maintenance & asset management in order to reduce faults due to reasons mentioned before.
As the digitalization in distribution grids increases, the monitorability of the system also increases, which creates a large data pool. This is a situation that can be beneficial in preventing or predicting some problems in the distribution system.
Reliability indexes such as SAIDI and SAIFI may be one of the most important indexes for DSO companies and these indexes directly reveal the performance of the distribution company. On this

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

point, evaluation of the existing and historical grid data and using them as an input for predictive algorithms will increase the performance of distribution network with preventing possible faults.
For example, having a data driven predictive maintenance process is more beneficial than having traditional inefficient methods. With these methods, maintenance processes become more professional. Preventing of faults will increase customer satisfaction and also reduce SAIDI, SAIFI indexes. This will have also huge benefits from an economic and financial point of view reducing costs associated with maintenance. Through analytic approaches, investment planning will be more accurate.

4.5. Cross-stakeholder Transfer Learning for Flexibility Assets Forecasting and aggregation

Residential flexibility that is valorised on multiple energy services is becoming increasingly interesting. Residential assets such as electric vehicles (EV), electric boilers, electric storage heaters, batteries, micro-CHP's or heat pumps are an untapped potential of flexible energy to integrate renewable energy by providing balancing services and mitigate grid congestion. Driven by increasing connectivity of these devices unlocking monitoring and control and users that are looking for energy insights in their household appliances and acting with their devices on their energy bill. Automation and user experience will be important success factors to seamlessly address upcoming customer needs.
The UC is divided in 2 subtopics:

- Transfer learning, transferring knowledge characterized by a control policy across similar assets in different geographical locations, device types or customer segments: In order to control the assets efficiently, flexibility providers are building models of the assets to optimize their control actions. This can be a labor-intensive and time-consuming process that limits scalability. Centrica and IMEC are therefore approaching this challenge by leveraging on data-driven algorithms in a Markov Decision Process (MDP) setting, such as reinforcement learning algorithms. The challenge that arises here is the fact that training data-driven models needs a considerable amount of data and can take time to converge to an optimal policy. Bad data quality, unavailability or failure of sensors could make it even more challenging to derive a data-driven policy. Cross-context transfer learning is explored as a solution and will be demonstrated in a real-life pilot. Transfer learning is a technique to exchange knowledge between actors, for example by sharing state information, tuned hyperparameters for regression architectures or even a learned policy or advantage function. Besides tackling the cold start challenge, exchanging knowledge between agents could improve the performance significantly by sharing exploration experiences of the state-action space from other devices. An important contribution of this project is the usage of collected field data and the aim to come to solutions that can be implemented at large scale within operational boundaries such as memory usage, sensor failures, etc. The algorithms will be developed on historical datasets.
- Cross-context grid friendly DR: Flexibility providers, such as aggregators have access to very granular data over each individual asset such as power and voltage at fine-grained timescales (in the minutes to seconds range) through proprietary metering solutions. This topic contributes to the observability of the grid, which will be the foundation of grid management that becomes more and more complex due to deployment of variable generation, electrification of heating and transport and consumers actively acting in the energy transition. The data logged by a flexibility provider can be leveraged in a cross-domain context to perform DR in a grid friendly way, by performing DR in a way that supports the grid, mitigating voltage and congestion issues. The exchange of grid representations derived from different datasets allows to strive for generalisation and offset limited availability or poor quality data.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

4.6. Coordinating grid-owned and behind-the-meter assets management prediction for improved grid operation and near real time power market operation settlement

The escalating grid complexity can no longer be accommodated by slight infrastructure over-dimensioning to withstand worst case energy provision requirements but necessitates data-driven transformation to instantiate situational awareness, allow the operation closer to the margin, and react on network events/disturbances. DSOs are at the core of utility industry transformation and local energy systems are becoming key assets that need to be properly managed and coupled with the infrastructure. This technology transformation also calls for an economic transformation in which DSOs must recognize new roles and opportunities and adapt long-established business models. Services-oriented economy, as already exemplified by other business segments, paves the way towards the idea of an asset-free utility. The pilot application relates to the optimized management of grid connected assets (i.e., small-scale RES and demand-response flexible loads), while the rationale is in the design of data and analytic services marketplace allowing various stakeholders of the energy supply value chain to collaborate and facilitate business relations on a P2P.

The near-real time data using advanced analytics will enable the scenarios in relation to the independent aggregator. There are two main demonstrative fields, the first one and the main one is the real time data driven. In this scenario, the main role will be given to the DSO, the second one goes into more detail about the legislative framework changes foremost stemming from the "Clean Energy for all Europeans" EU legislative package (CEP), provide for new business models relating primarily to DSO-level connected assets and their flexibility. The key difference is the new role of "independent aggregator", meaning a market participant engaged in aggregation who is not affiliated to the customer's supplier. Another important difference is the emergence of so-called "submetering", allowing metering data stemming from behind the DSO-grid threshold to be used in settlement of various market services, including for example organised electricity trading on day-ahead or intra-day markets, as well as ancillary services procurement, primarily balancing. The aim of this Pilot is to address some of the key challenges, that arise because of the aforementioned legislative changes.

CEP provides for general guidance but does not imply specific solutions. More precisely: "Member States should be free to choose the appropriate implementation model and approach to governance for independent aggregation while respecting the general principles set out in this Directive. Such a model or approach could include choosing market-based or regulatory principles which provide solutions to comply with this Directive, such as models where imbalances are settled or where perimeter corrections are introduced. The chosen model should contain transparent and fair rules to allow independent aggregators to fulfil their roles as intermediaries and to ensure that the final customer adequately benefits from their activities. Products should be defined on all electricity markets, including ancillary services and capacity markets, to encourage the participation of DR."

The aim of this pilot is thus to enhance the coordinated management of grid and third-party-owned grid connected assets and optimize their flexibility potential with a view to support different and new business model and subsequent anticipated settlement. The DSO's reactive (SCADA) and preventive (protection scheme) operational applications will be complemented with proactive (schedules) and corrective (mode of operation, set point modulation, etc.) operation schemes based on data analytics. Internally within the DSO (intra-stakeholder) this demands for the integration of growing data stores of disparate systems (e.g., meter data management, customer information from smart meters, outage management, etc.) with the aim of obtaining critical insights to enhance operations as well as strategic decision-making. The final expected outcome is to:

- increase efficiency and reliability of network operation via coordinated optimization and management of connected assets,

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

- include different possible models of consumer participation, design of rules, settlement process and assessment criteria and
- design infrastructure with reliable and near real-time metering data aggregation to support analytical tools.

4.7. Predictive Analytics Forecasting of storage optimized operation for improved market reserve participation

Nowadays, with the diffusion and use of renewable resources combined with storage systems as key-value of the de-carbonization process, it is very important to be able to predict any critical conditions that can negatively affect performance, causing disservices on FTM (front-of-the-meter) and BTM (behind-the-meter) applications.
The use case is focused on Battery Energy Storage System (BESS) plants integrated in the power system in order to provide large-scale energy storage system testing activities for grid applications.
The primary aim of the pilot is to create a BESS predictive maintenance tool and related procedures, starting from condition-based monitoring, in order to identify under-performance cases and predict critical events thus to improve BESS reliability and, consequently, improve plant performance while reducing operating maintenance costs.
The first step to achieve the goal is to integrate physical equipment into the testing environment, and to collect and store the data produced into a data acquisition system.

4.8. Predictive Demand and Generation Forecasting for Optimized Local Energy Community Management

Non-interconnected island electricity systems are largely dependent on the use of imported oil for electricity generation purposes. This increases production costs considerably, while also introducing technical limitations to the large-scale exploitation of RES power generation. At the same time, significant load variation noted over the course of the year further challenges increased RES penetration, with many island systems presenting increased shares of RES curtailments, especially during the winter (non-touristic) period. This reflects upon the need for the introduction of flexibility means including the components of storage, DR, vector coupling and also advanced energy management relying on BD analytics and modern forecasting techniques. Similar portfolios and bundles of solutions may unlock the potential for further RES exploitation, offering at the same time readily available tools for local community scale actors.

4.9. Collaborative aggregated energy generation prediction

Large scale solar power plants are driven by a large number of components that regularly need maintenance, observation, troubleshooting and monitoring. Maintenance like the cleaning of the CPV panels (especially important in this kind of PV technology) is done on a periodical basis instead of being done according to the production conditions. Within systems observation as well as external observation capacities could be combined to increase overall system efficiencies and performance. Drone observation technologies are evolving, which would allow for large fields scanning in addition to the on-site generated data. This new technology can produce better results if various data types are analysed.
Several processes can be singled out for overall improvement and a further level of analytics. For example, electrical troubleshooting is only done after faults occur with no reliable fault prediction analytics or data models.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Another example is related to the prediction of production, which is currently based on weather models and forecasts, but their accuracy can be improved, weather data could be enriched on the local level and equipment performance could be improved.
The use of BD and its application with AI in such a pilot can help improve all of the above existing and future operations Both for the operator of the large-scale production sites, self-consumption operators and deliver services to the grid operator increasing grid management capacity and coordination.
Maintenance can be done on a need basis and fault prediction algorithms can be created for some of the components. Production forecasting can also be improved by coupling weather forecast data with production models. Better production forecasts also enable more grid services that can be provided.

4.10. Predictive and Prescriptive Analytics for improved Energy Performance Certificates Reliability

UE Member States were asked to develop and implement Energy Performance Assessment and Certification schemes to improve the energy performance of their building stock, based on the entry into force of the Energy Performance of Buildings Directive (EPBD, 2010/31/EU and 2018/844/EU). Nowadays, EPCs and their accompanying assessment methodologies are well-established in many EU countries. In this context, and even when the potential of EPCs to act as catalyst to support the energy transition and boost energy refurbishments is really high, a review of these schemes has shown different interpretations, approaches and levels of compliance. As a result, this ultimately leads to a climate of distrust by the general public in the whole process and in the actors involved in it.
It is clear that improvements to these schemes should be implemented in a coherent and cost-effective manner, even though it will require a great amount of time and resources for the whole approach to Energy Performance Assessment and Certification schemes to be altered based on the currently identified shortcomings.
However, the current increase of data availability that can support EPCs generation, as well as BD techniques bring new opportunities to overcome these barriers, enabling scalable integration and management of large heterogeneous cross-domain and cross-stakeholders amounts of data in a harmonized way, from real-time smart meters to relevant contextual information. Thus, now is a good opportunity for reviewing the implementation of EPC schemes; and, with a more consistent, reliable, user-friendly, cost-effective and validated methodology, what EPCs should be used for in the future, while comparable good quality and compliance with EU legislation are granted. To this end, predictive and prescriptive analytics will be developed to support the implementation of next generation EPC.

4.11. Weather-enhanced building thermal comfort prediction to enable energy efficiency reliable forecasting

The Energy Performance Directive 2010 (2010/31/EU) established the starting point for the use of the EPCs. Their full potential has not been exploited yet, mainly due to the lack of interoperability between data sources and the complexity for its management. This leads to current situation, where its implementation remains unreliable and unfeasible from a practical point of view.
BD techniques bring new opportunities to overcome these barriers, enabling scalable integration and management of large heterogeneous cross-domain and cross-stakeholders amounts of data in a harmonised way, from real-time smart meters to relevant contextual information.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

4.12. Predictive Analytics for Energy Efficiency Investments de-risking

In the current situation a large number of decision makers, head of departments and project managers, have only limited information on the actual energy consumption of buildings. Energy Performance Certification (EPC) has been in place since the implementation of the Energy Performance Directive 2010 (2010/31/EU) and more and more public and private buildings receive EPC, but the actual data shows that EPC is not very reliable in assessment of energy performance, especially when it comes to deciding on the effectiveness of financial investments in these buildings.

Many public institutions and project developers also lack information and knowledge about the implementation of many energy efficient projects and their economic justification. This is especially evident in situations where the project could be implemented with the support of an ESCO company. It is also important for ESCO companies and their financiers to make sure of the historical energy data consumption of buildings when implementing projects.

The pilot will reduce the uncertainty linked to energy efficiency investments, which can be attributed to the lack of relevant skills and ability to assess investments. It will strengthen debt and equity financing of energy efficiency projects. More specifically, the platform will provide investors / financiers and project developers the opportunity to: evaluate quickly and easily key performance indicators for projects (analysis and assessment techniques for projects and companies, benchmark tools, financial performance indicators for investors, special indicators "Green barometer", written analysis); Explore past and future projects (project history, synopsis, projects/companies more attractive to investors); Make a profound risk measurement and evaluation (implement key indicators, such as debt service coverage ratio, integrate a cross-check for legal requirements of environmental laws and regulations).

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5. BD4NGRG Use Cases

The BD4NRG project has twelve LSPs, each of which has one or more UCs, that analyze the use of BD in different sectors of the energy system. The following chapter shows all UCs sorted by LSP numbering for easy placement. Each UC is presented in the form of a table, according to the template proposed Chapter 2.

5.1. Integrating off-grid data with Condition based Monitoring for Enhanced Asset Maintenance

Table 7: AI methods applied to Condition-based monitoring of Circuit Breakers

Use Case ID	AI methods applied to Condition-based monitoring of Circuit Breakers
Cluster – main beneficiary	TSO
Description	This UC analyzes different CBs operating at several very high voltage (HV) levels, controlled and managed by the TSO. These CBs are associated with several recorded operations (mainly trips due to faults and switching operations), which may variate in order of importance according to the operational conditions of the asset. The operation records contain a lot of unused information about the faults, stored in COMTRADE data, mainly as oscillography data from current and voltage channels. Therefore, this UC aims to apply data processing and analytics tools to incident report data from CBs in order to typify incidents, track actuation times and identify incident causes (categorization of faults), by integrating incident report data together with off-grid data. These results will be used to both automatically extract key insights from CB incidents on the fly and predict the probability/likelihood of future failures and performance reduction - feeding enhanced asset maintenance processes.
Goals	• Categorization of cause of faults using unsupervised clustering methods, integrating off-grid data • Estimate likelihood index of failure (actual failure or reduction of performance) for CBs in order to support predictive maintenance
Actors	• Asset Management Division of TSO • Protection Engineers
Envisioned analytics and tools (BD4NRG components involved)	• Failure probability likelihood index estimation • Categorization of the cause of incidents/faults from reports
Precondition	In order to implement this UC, CB activation incident historic data, including oscillography and manufacturer's manuals should be available, together with data retrieved from SCADA.
Triggering events (if applicable)	Activation incidents of CBs (trips and switching operations)
Post conditions	Better insights at potential causes of CB activation events. Optimal operation of CBs under maintenance plan, by integrating failure/malfunction probabilities/like hood indexes and extracted knowledge from incident events.
Related KPIs	• Number of real switching operations

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	AI methods applied to Condition-based monitoring of Circuit Breakers
	• Circuit breaker fault clearance time • Circuit breaker closing/opening time % of successful auto reclosing prediction • Percentage of unknown fault cause categorization • Secondary arc extinguishing time (for OHL).

Table 8: Semi-Automatic Maintenance Plan for Right-of-Ways (RoW) in Overhead Lines (OHL)

Use Case ID	Semi-Automatic Maintenance Plan for Right-of-Ways (RoW) in Overhead Lines (OHL)
Cluster – main beneficiary	TSO
Description	This UC analyzes different OHL operating at several very HV levels. From transmission line towers LiDAR cylindrical scan data, use point cloud data to correctly classify each tower, in accordance to the specification of TSO catalogue. Furthermore, distinguish structure and identify potential abnormalities (e.g. in insulators), stork nests, fiber optic boxes, using data analytics and computer vision techniques. Use these results to derive condition based monitoring based on state of individual components. Complement LiDAR data with data from other external sources (e.g. vegetation index data and climate data) to better assess RoW state and derive optimal maintenance techniques and classify state of RoW.
Goals	• Usage of unstructured scan data to identify and evaluate OHL Towers condition • Derive maintenance plan based on state of RoW of OHL assessed with data mining techniques
Actors	Asset Management Division of TSO
Envisioned analytics and tools (BD4NRG components involved)	Classification of OHL transmission towers and their state • Point cloud feature extraction • Point cloud data classification • Computer Vision
Precondition	In order to implement this UC, LAS file data from LiDAR inspections for OHL towers and samples of Right-of-Way should be available.
Triggering events (if applicable)	-
Post conditions	Semi-automatic maintenance plan that allows for enhanced operation of OHL, based on data mining insights extracted from scan data.
Related KPIs	• Accuracy identifying the transmission tower type • Accuracy identifying the number of structural elements in the tower • Number of different type of tower and RoW features and entities identified • Meshed ground profile along the RoW • Vegetation index mass

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5.2. Cross-functional Integration of Grid Maintenance Planning with Predictive Asset Management

Table 9: Maintenance Plan Optimization

Use Case ID	Maintenance Plan Optimization
Cluster – main beneficiary	TSO
Description	Location: The UC takes place in 400 kV, 220 kV, 110 kV transmission system grid. Current status: Today the grid operation, asset management, maintenance etc. are not fully exploiting the advantages of historic and predictive data for generating cross-sector optimized maintenance plans. Primarily planning considers well established schedules, proven timelines, past maintenance interval with reference to equipment type and operational security (N-1 criteria). Time-based optimization is used only to a limited extent. Using term maintenance planning with reference to system operation the equipment outage planning for maintenance purposes is normally considered. Issue to be solved: Improvement of established ways of maintenance planning by relying on data from various data sources and using multi-criteria optimization procedures. Advanced maintenance Planning will take into account internal and external signals and find a coordinated outage plan for all transmission system elements in focus. Assets involve: Power transformers or/and transmission lines and connected assets (e.g. circuit breakers and measurement equipment,) Sequence of events: • Importance of specific available data for UC is assessed. • Identified relevant data for UC are pre-processed. • Developed algorithm(s) are thoroughly tested. • Needed algorithm(s) modifications are harmonized. • Final outage plan proposal is presented as a possible support to traditional outage planning process.
Goals	Optimization of the maintenance plan: • Automatic generation of Gantts chart for outage plan proposal • Analytical planning approach • Flexibility of planning • Digitalization of planning • Maintenance planning considering a wider range of influencing parameters • Use of various criteria functions • Reduction of effort for producing the plan • Increased role of data in maintenance planning • Use of analytical reversibility approach to determine the influence of introduction of equipment with higher capabilities
Actors	• Internal TSO departments: system operation, system maintenance, asset management • distribution system operators (DSOs), • HV-grid connected consumers and producers of electricity

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	Maintenance Plan Optimization
Envisioned analytics and tools (BD4NRG components involved)	• Generation forecast • Load flow forecast for the regulation zone HV grid • Operational security limit (N-1 criterion) validation • Asset condition estimation: Asset Health Index (AHI); for energy transformers • Asset utilization estimation: Asset Utilization Index (AUI); for energy transformers • Asset relevance index • Probability of failure estimation of assets
Precondition	• Detailed technical and maintenance historical data for all targeted assets should be available • Equipment condition data should be known • Condition based maintenance should have been adopted • Basic criteria for decision making should be set • Use of advanced analytics for maintenance plan should be generally accepted as reliable and useful
Triggering events (if applicable)	• Periodically with respect to present time interval • On-demand triggering by internal and external signals for maintenance plan update • On-demand triggering by TSO on asset health change
Post conditions	Ambitiously aiming for more agile, efficient and integrated maintenance planning process delivering proposed optimized plan based on criterion function with respect to internal and external signals. Thus, setting foundations for enabling better asset utilization at the same system reliability level.
Related KPIs	• Power system reliability indices (e.g. energy not served) • AUI increase (per asset and per fleet) • AHI level (per asset and per fleet) • Increased social welfare (additional energy transfer enabled due to optimized maintenance plan; additional TSO revenue for transmission capacity due to optimized maintenance plan) • CAPEX and OPEX reduction • Faster decision making

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5.3. Cross-Domain Cross-functional Integration of Predictive Asset Management for Grid Operation and coordination with flexible EVs charging

Table 10: Efficient management of flexibility resources for supporting grid operation

Use Case ID	Efficient management of flexibility resources for supporting grid operation
Cluster – main beneficiary	DSO
Description	The UC takes place in an urban district where RES and EV charging stations are present. The lack of coordination between EV and RES is going to overload assets of the LV network (i.e., cables and transformers). Therefore, the BD4NRG tools will firstly perform analysis of big data produced by Smart Meters and EV, EVSE monitoring devices and technical provider, such as load forecast and EV charging stations consumption forecast. secondly, forecasting algorithm will be tested. Finally, a set of DR campaigns will be carried out for reducing the loads of the assets. The reduction of assets loading is directly linked to their ageing as well as to the whole efficiency of the distribution network.
Goals	• Definition and implementation of optimal scheduling and usage of RES and EVSE for reducing equipment loading during normal service. • In case of degraded service, RES and EVSE will support DSO whilst it implements topology changes and repairing activities.
Actors	• EV manager, • technical provider
Envisioned analytics and tools (BD4NRG components involved)	Forecast of energy consumption of the district. Forecast of the usage of EV and charging stations. Asset status according to energy and no-energy measurements.
Precondition	In order to implement this UC, RES and EV are available. Moreover, energy measurements should be integrated both as historical data and real – time data. Components have to be deployed in LSP IT infrastructure
Triggering events (if applicable)	The UC demonstration is linked to the definition and implementation of a set of DR campaigns which involve EV users.
Post conditions	The EV charging infrastructure modulates its power output to provide flexibility to the grid and to maximise the RES integration. Overall asset loading is reduced. An e-Mobility platform hosting charging station managers and EV users shall be implemented and deployed
Related KPIs	SAIDI SAIFI Reduction of losses Reduction of peaks Reduction of Reverse Power Flows Number of cable and transformers replacements

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Table 11: Anticipated prediction of asset failures

Use Case ID	Anticipated prediction of asset failures
Cluster – main beneficiary	DSO
Description	The UC takes place in a power distribution network. The main issue to be solved is the lack of awareness about Low Voltage Therefore, the BD4NRG tools will firstly collect data produced by Smart Meters and suitable sensors supplied by specialist manufacturers deployed in the secondary substation. Then big data are analyzed by technical provider in order to find a correlation between measured quantities and faults. The prediction of assets failures is directly linked to BD4NRG outcomes that will identify anomalous paths in the data shared by the DSO.
Goals	Definition and implementation of new plan for asset replacement and maintenance based on predictive analytics
Actors	• technical provider, • manufacturers
Envisioned analytics and tools (BD4NRG components involved)	Analysis of audio-based unstructured data from digital sensors Weather forecast and historical weather data (through UBIMET service) Geographical imagery (through TS partner service) Analysis of energy data
Precondition	This UC needs sensors deployed in the network assets as well as at customer premises
Triggering events (if applicable)	-
Post conditions	Asset failures rates are reduced
Related KPIs	SAIDI SAIFI Reduction of losses Reduction of peaks Reduction of Reverse Power Flows Number of cable and transformers replacements

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5.4. Cross-functional Predictive Asset Management for MV Grid Planning

Table 12: Predictive maintenance and fault prediction for MV grid and cross functional predictive asset management

Use Case ID	Predictive maintenance and fault prediction for MV grid and cross functional predictive asset management
Cluster – main beneficiary	DSO
Description	The UC takes place in a medium voltage network distribution network. For the pilot locations selection, an old network section with lots of MV faults in an urban district will be considered. Since DSO does not have a professional maintenance plan and data evaluation processes, some unexpected faults on MV grid happen. In order to reduce faults and have more accurate predictive maintenance processes, the BD4NRG AI tools will initially work on long term Historical data to analyze the accuracy of its short-term historical data. This fine tuning process will serve as a continuous feedback loop for better prediction performance. Once the AI achieves a certain level of accuracy then the algorithm, knowing the weather forecasts and the future demand scenario, will be used to predict and analyze short term future operation, maintenance and failure of the MV network assets, and hence failure along the grid. According to plan, the created model will be trained on one asset (preferably for transformer or any other asset which has historical data) and then the same pre-trained model is further trained on different assets
Goals	• Definition of the data sets and using of data as input to AI algorithm in order to have predictive maintenance plan in MV level. • Definition of optimal scheduling of maintenance operations to reduce unnecessary maintenance. • Definition and implementation of predictive fault monitoring in order to perform troubleshooting before faults occur.
Actors	-
Envisioned analytics and tools (BD4NRG components involved)	AI based predictive maintenance tools Fault prediction tools Weather forecasting tools Demand forecasting tools
Precondition	Necessary data sets will be defined previously and according to data sets, historical and near real-time data gathering processes will be done with the MV equipment on field. Components have to be deployed in LSP IT infrastructure
Triggering events (if applicable)	UC will be demonstrated through the implementation of maintenance and fault prediction operations.
Post conditions	With predictive maintenance results, maintenance processes will become more professional and efficient and also with optimized maintenance and fault predictions, it is expected to have less MV faults on grid. Reduction of faults

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	Predictive maintenance and fault prediction for MV grid and cross functional predictive asset management
	Reduction of losses Number of components replacements
Related KPIs	SAIDI SAIFI Reduction of faults Productive maintenance/preventing unnecessary maintenance Number of components replacements

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5.5. Cross-stakeholder Transfer Learning for Flexibility Assets Forecasting and aggregation

Table 13: The agents responsible for controlling residential demand response assets can benefit from the transferred knowledge from other agents in an MDP setting.

Use Case ID	The agents responsible for controlling residential demand response assets can benefit from the transferred knowledge from other agents in an MDP setting
Cluster – main beneficiary	Aggregator
Description	Location: asset in LV/MV-grid Assets involved: To be specified during the project, for example EV, electric boiler, etc. Main issue: In order to control the assets efficiently, flexibility providers are building models of the assets to optimize their control actions. The challenge that arises here is the fact that training data-driven models needs a considerable amount of data and can take time to converge to an optimal policy: e.g. for new assets that join the portfolio there is no historical dataset available to train models on. The cost of acquiring interaction samples can be significant. Sequence of events: • A model is trained on one asset (where historical data is available) • The same pre-trained model is further trained on another asset (transfer learning)
Goals	• Developing knowledge transfer algorithms for residential DR • Demonstrating transfer learning on residential DR in the field, with a full end-to-end software chain
Actors	• residential customer, • DSO/TSO
Envisioned analytics and tools (BD4NRG components involved)	Transfer learning techniques, starting from the work on transfer learning in supervised learning, apply that to the MDP setting. Visualization of Transfer Learning techniques. The developed weather forecast tools might be leveraged.
Precondition	In order to implement this UC: • Historical data should be available to test the knowledge transfer approaches • a part of Centrica's portfolio should be available for the pilot • simulation and lab testing performance should be sufficient to ensure customer comfort and conform delivery of flexibility services • A software integration of the TL agents in Centrica's Flexpond DR platform has to be implemented • Software agents have to be deployed to production in order to start the trials • Specifications have to be written to formalize the transfer learning problem and categorize which knowledge has to be transferred. • Algorithm must be developed in simulation mode, exploring different TL

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	The agents responsible for controlling residential demand response assets can benefit from the transferred knowledge from other agents in an MDP setting
	techniques in an MDP setting
Triggering events (if applicable)	-
Post conditions	More efficient learning of new controllers leading to higher comfort and faster cost/revenue benefits of controllers
Related KPIs	• The final performance of the agent with and without knowledge transfer. • The time required for the target agent to reach certain performance thresholds, with and without knowledge transfer. • Number of historical samples required for training to fulfil the minimum cost saving target or related threshold • Number of transfer learning agents deployed in staging/production environment • The amount of computing power required when scaling up • The communication and data storage requirements

Table 14: Using residential DR assets to retrieve local grid information and use this information in the aggregation activations

Use Case ID	Using residential DR assets to retrieve local grid information and use this information in the aggregation activations
Cluster – main beneficiary	Aggregator
Description	Location: assets to be specified during the project, for example EV, electric boiler, etc Main issue: The data collected by the flexibility provider can be leveraged in a cross-domain context to perform DR in a grid friendly way, e.g. by performing DR in a way that supports the grid, mitigating voltage and congestion issues.
Goals	Mitigate voltage and congestion issues
Actors	• residential customer • DSO/TSO
Envisioned analytics and tools (BD4NRG components involved)	• Power flow simulations • Aggregation-disaggregation techniques
Precondition	• Relevant DSO topology and measurements should be available • Historical dataset of measurements from Centrica's DR portfolio • Data or a simulation model or assumption about congestion / voltage issue is required • Software agents have to be deployed to production in order to start the trials • The UC and methodology for grid friendly DR has to be defined
Triggering events (if applicable)	Before the activation of DR

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	Using residential DR assets to retrieve local grid information and use this information in the aggregation activations
Post conditions	The flexibility provider performs DR in a grid-friendly way
Related KPIs	% of congestion or voltage issues mitigated (in terms of time and volume/capacity) compared to conventional DR activation

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5.6. Coordinating grid-owned and behind-the-meter assets management prediction for improved grid operation and near real time power market operation settlement

Table 15: Identification of flexibility potential, based on near-real time data

Use Case ID	Identification of flexibility potential, based on near-real time data
Cluster – main beneficiary	DSO, grid users
Description	Location / assets involved: Preselected substations (DSO's assets) are equipped with new, additionally installed sensors (e.g. transformer oil temperature, housing temperature…). Beside sensors, the substations will be also equipped with smart meters with applied advanced functionalities (streaming data) Sequence of events: All data coming from the assets will be processed in real time as a data analytics stream method Main issues to be solved: • installation of the technically approved (based on the results of the pre-made analyses) metering and sensor equipment • integration of equipment with existing Big Data platform; enabling the data coming from the field to be collected. • data processing: data validation and data modelling • AI modelling and forecasted values testing (energy, power, temperature,) • A list of substations based on AI results, together with the identified flexibility needs • DSO will inform the independent aggregator about the desired flexibility (location – substation; power needs; time slot)
Goals	• increase efficiency and reliability of network operation • infrastructure design with reliable and near real-time metering data • exploiting the flexibility potential identified at the substations to defer investments or maintenances
Actors	• residential customers • RES producer • Independent aggregator • Market operator
Envisioned analytics and tools (BD4NRG components involved)	Power forecast Prediction with data analytics in flow Open-source AI tools
Precondition	In order to implement this UC: • Selection of the pilot area; it will base on the most frequently installed power

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	Identification of flexibility potential, based on near-real time data
	of the transformers; second criteria will be the share of RES: at least 30% of installed power (nominal power) must represent the renewables. • Fulfilled technical and technological requirements for the equipment will allow its field installation • Integration of the equipment with the Bigdata platform
Triggering events (if applicable)	Continuous process as stream data provision and results.
Post conditions	To approve the near real time data streaming and analytics; To send the information about the flexibility needs to the external, aggregator's platform (DSO, TSO, coordination)
Related KPIs	• To be specified more in details: number of installed digital technologies and pilots, • Improved reliability and efficiency of the power network (% improvement of the electricity quality indicators) • Cost reduction for management and maintenance of the power network assets, • Increased number of the flexible prosumers, • Contracted/activated flexible power provided only to cover the DSO needs, • Improved power network management, • Number of innovative, data driven business modes, in this case the focus is given to the role of the DSO and his needs, • Amount of the data provided, variety of the data,

Table 16: Independent aggregator flexibility activation, quantification, and settlement

Use Case ID	Independent aggregator flexibility activation, quantification, and settlement
Cluster – main beneficiary	DSO – Market operator
Description	A flexibility request is triggered by an actor (e.g., DSO). The request is handled by the independent aggregator, who has a flexibility services contract with the asset (RES producer or consumer). The provision of flexibility impacts the supplier of the asset via imbalances on the market (e.g., if consumption is reduced, then in principle the supplier's position switches to long). Location / assets involved: asset in DSO-grid (consumer and/or producer) Sequence of events: • Receiving the request for the Flexibility service form DSO • Activation (+ type of service) -> Metering / data management -> Market settlement / adjustment (primary relation supplier-independent aggregator) Main issues to be solved: • quantification of service in energy terms (e.g., baselines) • settlement models (including different options regarding sub-metering usage for market settlement) • potential impact on data flows and data usage /& adjustment

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	Independent aggregator flexibility activation, quantification, and settlement
	• appropriate simulation models, when / if real data not available
Goals	• Determine possible / optimal quantification methods for independent aggregator activation • Determine possible / optimal settlement models • Determine needed data structures and timing for implementation in larger scale • Determine possible models, adequate for customers with metering resolution above ISP (imbalance settlement period)
Actors	• Independent aggregator • (residential) customer • RES producer • Supplier
Envisioned analytics and tools (BD4NRG components involved)	Different baseline models with error assessment -> separation of single metered data stream into different market settlement components (independent aggregator, supplier) In terms of analytical tools, the point is the quantification of the flexibility service actually provided – i.e. usually the difference between metered data and ex-ante baselines (short term forecasts). A key issue is the impact of different approaches and the assessment of errors. The second part is more akin to accounting – different options on how to handle quantities (and at what prices, if any).
Precondition	In order to implement this UC: • Historical data should be available or otherwise - simulated • Possible models should be put in place before simulations, as they may affect it
Triggering events (if applicable)	-
Post conditions	Extended knowledge of different models' pros and cons.
Related KPIs	• Impact on final asset (consumer / producer) for different possible models. • Scalability of models (incl. equipment needs) • The communication and data storage requirements for different models

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5.7. Predictive Analytics Forecasting of storage optimized operation for improved market reserve participation

Table 17: Predictive Analytics Forecasting of storage optimized operation for improved market reserve participation

Use Case ID	Predictive Analytics Forecasting of storage optimized operation for improved market reserve participation
Cluster – main beneficiary	Aggregators
Description	Location: Testing system will be composed by a stand-alone energy storage system (250kW/500kWh) equipped with additional sensors able to validate the data collected by BMS and ESS, that can be coupled with 300kW Load/PV simulator. Issue to be solved: Improvement of plant performance with operating maintenance costs reduction by ML algorithms able to identify BESS under-performance cases and predict critical events through specific predictive maintenance procedures and analysis. Assets involved: BESS plant. Sequence of events: • Collect and store the diverse data from the multiple data sources (storage system, meters…) and analyse them. • Select the market scenario as UC context. • Define a BESS model using real-time and historical data. • Identify under-performance areas. • Develop and test predictive maintenance procedures to apply to the identified under-performance areas.
Goals	• Physical equipment installation and Data Acquisition System (DAS) upgrade. • Market scenario and grid services selection. • Test protocol design and execution. • DAS integration with BD4NRG platform. • BESS Condition Based Monitoring procedures implementation. • BESS Condition Based Monitoring procedures testing (under-performance areas identification).
Actors	ESCOs, Universities
Envisioned analytics and tools (BD4NRG components involved)	C&I Load Forecast. Weather forecast. PV production forecast. Battery Degradation forecast.
Precondition	• BESS system installed and connected to the Lab environment. • Lab DAS upgrade in order to collect the data produced by the selected testing system. • Market scenario definition and grid services selection. • Test protocol design and test execution procedures definition. • Availability of other data sources (such as grid services participation reports).
Triggering events (if applicable)	• BESS optimization accordingly to the market scenario. • BESS Grid Services participation

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	Predictive Analytics Forecasting of storage optimized operation for improved market reserve participation
Post conditions	BESS optimization through the identification of under-performance areas and predictive maintenance procedures (accordingly to the market scenario).
Related KPIs	• Number (percentage) of DAS measures collected by physical equipment. • BESS under performance areas identification. • Number of predictive maintenance procedures identified.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5.8. Predictive Demand and Generation Forecasting for Optimized Local Energy Community Management

Table 18: Optimal management strategies for cross-infrastructure flexibility resources to support higher RES shares at the island community level.

Use Case ID	Optimal management strategies for cross-infrastructure flexibility resources to support higher RES shares at the island community level
Cluster – main beneficiary	DSO
Description	Through the use and upgrade of existing smart metering infrastructure of an island, the pilot aims to explore strategies of increased flexibility that may support both higher RES shares and reduced costs at multiple ends, supported by the upscale of similar, cross-sectoral DR pools in island systems. It will do so through the monitoring of a representative and sufficiently diverse pool of loads and end-uses (residential, commercial, public buildings, EV charger, prosumers and municipal water pumps), eventually aggregated through energy communities or Aggregators, together with distributed PV assets and/or wind parks in operation in the broader electricity systems connected to the island. Data available (weather forecast, RES production forecast and status of assets) will inform the development of production and demand forecasting models aiming at the advancement of energy management strategies at the island community-scale, which may in turn generate relevant day-ahead schedules for the different components involved.
Goals	Definition of optimal scheduling and usage of deferrable loads of different elasticity to achieve higher RES shares at local (island) community level Development of a portfolio of services for economically optimal management of island municipalities' deferrable loads Definition of optimal scheduling of deferrable loads for offering distribution grid services at the island level (e.g. peak shaving)
Actors	• Island municipality, • end-customers, • local energy communities and aggregators (in general), • private RES actors / prosumers.
Envisioned analytics and tools (BD4NRG components involved)	Forecast of different energy consumption streams for sub-groups of loads and aggregate load demand Forecast of weather data Forecast of RES power generation Asset status according to energy and no-energy measurements.
Precondition	In order to implement this UC, smart metering infrastructure in dispersed end-loads on the island is available. Communication equipment is currently restored in certain instances, while for local end-customers' electricity consumption to be monitored and for relevant data to be processed, informed consent is first required, that will allow us to engage a group of residential end-customers. Additionally, and to broaden the scope of the UC, it is planned that detailed historical data will be requested from the local DSO for the Kos and Kalymnos electricity system, in order to involve additional RES assets.
Triggering events (if applicable)	-

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	Optimal management strategies for cross-infrastructure flexibility resources to support higher RES shares at the island community level
Post conditions	The pilot will indicate optimal, cross-sectoral asset management, with the aim being to trigger the generation of relevant incentives for the identified flexibility services towards the local grid and for the development of local energy community schemes.
Related KPIs	Asset and system related: Electricity costs (€/kWh) Annual RES shares (%/year) RES curtailments recovered (%/year)

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5.9. Collaborative aggregated energy generation prediction

Table 19: Optimizing and improving management and prediction of DER production

Use Case ID	Optimizing and improving management and prediction of DER production
Cluster – main beneficiary	RES operator
Description	The Solar Demonstration Platform is located in an infrastructure area. Maintenance of the assets is done on a regular basis and electrical troubleshooting is done after faults occur. There is also no way to forecast future production of the system in an accurate way. By developing specific tools and value-added services based on BD4NRG platform and tools to analyze the Big Data created at the inverters (power, voltage, current and frequency) and also local weather stations, maintenance, fault prediction and production algorithms will be created and combined with weather data models of services providers and other external sources. The algorithms will then be tested and compared.
Goals	• Definition and implementation of optimal scheduling of maintenance operations to reduce unnecessary or due maintenance • Definition and implementation of predictive fault monitoring in order to perform troubleshooting before faults occur • Definition and implementation of production forecast using weather data and production models to add value to the services provided
Actors	• National Weather institute - IPMA • Grid Operator • Weather services providers (UBIMET) • Operation and Maintenance services provider
Envisioned analytics and tools (BD4NRG components involved)	Weather data – both historical and forecasts in the region Fault prediction Production forecast models
Precondition	This UC focuses on the RES available, a large-scale solar plant and a free standing self-consumption system with EV charging and storage components. Data gathering will be done with sensors deployed at the inverter level and gateways/dataloggers with edge computing capabilities. Energy measurements should be both historical and real-time. The weather data will be provided from on-site weather station and the related gamut of sensors.
Triggering events (if applicable)	The UC will be demonstrated through the implementation of maintenance and fault prediction operations by the RES operators coupled with reliable production forecasts.
Post conditions	Maintenance operations are optimized, fault prediction is used, and the RES operators are able to use production forecasts in their services.
Related KPIs	• Reduction of faults • Reduction of losses • Number of components replacements

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Table 20: Regional energy sector aggregation digital twin

Use Case ID	Regional energy sector aggregation digital twin
Cluster – main beneficiary	Grid operators, aggregators, energy suppliers
Description	The UC is located in a region with about 270 MW of DER installed capacity consisting mainly of wind farms and solar plants. This installed capacity is distributed all over the region and there's currently no aggregation of the regional production, it is only done at the national level. By aggregating at a lower level, the regional data economy is created and energy services are made possible. After aggregating regional capacity and near-real time generation values, regional weather forecasts can also be implemented in order to produce short-term energy generation forecasts. CO2 emissions can also be calculated. Even though some analytics can and will be provided, Data-as-a-Service is the main offer of this use case.
Goals	• Aggregate regional energy production (both capacity and actual generation) • Real-time monitorization of production and short-term production forecasts • Implementation of energy services enabled by this data value chain • Creation of regional production electricity map and energy related carbon footprint
Actors	• EPES stakeholders: Suppliers, aggregators, DSO, TSO, ESCOs • Weather Service Providers • Electricity Map Open-Source provider (IT company *Tomorrow*)
Envisioned analytics and tools (BD4NRG components involved)	• Energy generation forecast of solar and wind (predictive analytics) • Data views (both historical and real-time) of regional installed capacity, actual generation and generation forecast (all of these by production type, in this case Solar and Wind) – Descriptive and visual Analytics • Data-as-a-Service
Precondition	Access to installed capacity and generation data from the DER operators and access to regional weather data
Triggering events (if applicable)	-
Post conditions	• Regional production capacity is known • Regional actual generation values by production type are aggregated • Short-term generation forecasts for the region are made • This Data Value Chain is exploited by regional energy sector
Related KPIs	Number of generation sources collected Number of data streams offered per type (installed capacity, generation type, actual generation and generation forecasts)

Table 21: FlexATwin, Regional Digital Twin with multiple interfaces for flexibility across sectors

Use Case ID	FlexATwin, Regional Digital Twin with multiple interfaces for flexibility across sectors
Cluster – main beneficiary	DSO, aggregator, resource managers of the region
Description	The UC is located in a region with an opportunity to aggregate and couple energy

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	FlexATwin, Regional Digital Twin with multiple interfaces for flexibility across sectors
	sectors. Assets like DER, storage, electric mobility and municipal buildings loads can be clustered to enable energy services like flexibility services and demand-side response services. Regions of interest near the solar demonstration platform will be characterized and relevant energy assets will be aggregated.
Goals	• Create Nodal VPPs • Create data-driven energy services, such as flexibility services and demand response for the regional energy sector • Improve regional resource management
Actors	• EPES stakeholders: Generators, suppliers, aggregators, Municipalities, DSO, TSO, Mobility services, ESCO • Weather Service providers
Envisioned analytics and tools (BD4NRG components involved)	• Create descriptive, predictive, prescriptive analytics for the aggregated assets, including for generation, consumption, grid stability • Data views – Visual Analytics • Data-as-a-Service
Precondition	Access to data sources from all the relevant stakeholders
Triggering events (if applicable)	-
Post conditions	Creation of a functional digital twin of the energy sectors in the area capable of providing data-driven energy services
Related KPIs	Number of assets aggregated Maturity Level of digital twin created Number of services enabled by FlexATwin

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5.10. Predictive and Prescriptive Analytics for improved Energy Performance Certificates Reliability

Table 22: Predictive and Prescriptive Analytics for improved Energy Performance Certificates Reliability

Use Case ID	Predictive and Prescriptive Analytics for improved Energy Performance Certificates Reliability
Cluster – main beneficiary	ESCOs
Description	The UC is applied in a variety of locations, with diverse characteristics (year of construction, renovation statuses, building typology…) and different ownership. Current EPCs are static, not considering dynamic energy data (i.e. real consumption) and contextual information. This LSP makes use of varied data to apply Big-Data techniques to come up with EPC v2.0 (or namely dynamic EPCs). The UC involves buildings themselves, the energy systems involved in the heating/cooling, as well as contextual information. The UC will follow the next sequence of events: 1. Ingest the diverse data from the multiple data sources and analyse them. 2. Calculate the EPC-related KPIs (i.e. PMV, Energy performance and SRI), including forecast of energy consumption. 3. Determine the labelling by applying the EPC calculation algorithm. 4. Feedback to continuously evaluate EPC.
Goals	• Overcome “reliability” barriers that hinder the high potential of EPCs • Contribute to increasing the use of EPCs in the EU and among end-users • Calculate next generation of EPC v2.0 (or dynamic EPC). • Provide recommendations for energy conservation measures (ECMs) to improve the EPC label • Dynamically update the EPC label • Contribute to establish connections between EPCs and standards/ other relevant approaches
Actors	• data provider, • public entities, • technology developer
Envisioned analytics and tools (BD4NRG components involved)	Forecast of energy consumption. PMV (comfort) analytics. SRI calculator. Weather forecast. ML algorithms. EPC calculation algorithm. ECM advisor.
Precondition	Smart meters must be available and their data must be gathered. Energy data must be integrated (historical and real time information) Availability of other data sources (such as Cadastre, BIM, current EPC databases…) to provide contextual information.
Triggering events (if applicable)	-
Post conditions	Modelling of the EPC v2.0 to be scalable, extensible and replicable to other building assets.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Related KPIs	Number of calculated EPCs Number of building typologies included in the next generation EPCs Number of different data sources integrated for EPC calculation Number of open data sources integrated for EPC calculation Number of maximum refresh rate reached (depending on the data used)

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5.11. Weather-enhanced building thermal comfort prediction to enable energy efficiency reliable forecasting

Table 23:Weather-enhanced Building Thermal Comfort Prediction to Enable Energy Efficiency Reliable Forecasting

<table>
<tr><th>Use Case ID</th><th>Weather-enhanced Building Thermal Comfort Prediction to Enable Energy Efficiency Reliable Forecasting</th></tr>
<tr><td>Cluster – main beneficiary</td><td>Building owners</td></tr>
<tr><td>Description</td><td>The UC is applied in a variety of locations, with diverse characteristics (year of construction, renovation statuses, building typology…) and different ownership.
Current EPCs are static, not considering dynamic energy data (i.e. consumption) and contextual information.
The UC involves buildings themselves, the energy systems involve in the heating/cooling, as well as contextual information.
Cross-domain integration of building/district energy consumption data with off-grid weather information will take place to enable thermal comfort analytics-based prediction applications at building level, and related trade-offs with energy efficiency, with a view to support near real time comfort-enhanced decision-making by building or district operator. We will use ML algorithms to bridge the gap between controllable building parameters and thermal comfort, by conducting an extensive study on the efficacy of different ML techniques for modelling comfort levels, while trading off with energy efficiency. Data of demand forecast, evaluation weather forecast and comfort are necessary to perform the analysis.
Main gained benefit will be consisting of finer grained and more accurate thermal comfort prediction, which is expected to be linked linearly with HVAC set points and actions on controllable loads, while taking into due consideration potential changes in energy consumption and how this could be reflected into building energy efficiency. We will be creating new business models to increase the efficiency and comfort of buildings and big data analytics to provide information to the private sector.
The UC will follow the next sequence of events:
1. Ingest the diverse data from the multiple data sources and analyse them.
2. Cross-domain integration of building/district energy consumption data with off-grid weather information will take place to enable thermal comfort analytics-based prediction applications at building level, Use ML algorithms to bridge the gap between controllable building parameters and thermal comfort
3. Conduct an extensive study on the efficacy of different ML techniques for modelling comfort levels, while trading off with energy efficiency.</td></tr>
<tr><td>Goals</td><td>• thermal comfort prediction applications at building level, and related trade-offs with energy efficiency.
• more accurate thermal comfort prediction,
• Creation of new business models to increase the efficiency and comfort of buildings.</td></tr>
<tr><td>Actors</td><td>• ESCOs,
• data provider,</td></tr>
</table>

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	Weather-enhanced Building Thermal Comfort Prediction to Enable Energy Efficiency Reliable Forecasting
	• public entities, • technology developer, • building users
Envisioned analytics and tools (BD4NRG components involved)	Forecast of energy consumption. PMV (comfort) analytics. SRI calculator. Weather forecast. ML algorithms. EPC calculation algorithm. ECM advisor.
Precondition	Smart meters must be available and their data must be gathered. Energy data must be integrated (historical and real time information) Availability of other data sources (such as Cadastre, BIM, current EPC databases…) to provide contextual information.
Triggering events (if applicable)	-
Post conditions	Modelling of the ML algorithm to be scalable, extensible and replicable to other building assets.
Related KPIs	• Increased energy efficiency measured as energy saving based on predictive consumption forecasting and comfort forecast. • Number of pilots where demonstration of data-driven energy management systems take place. • Big data availability for real life research • Potential number of platforms, applications, business processes and innovative business models validated via large-scale piloting • Maximum refresh rate reached (depending of data used).

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

5.12. Predictive Analytics for Energy Efficiency Investments de-risking

Table 24: De-risking energy efficient investments

Use Case ID	De-risking energy efficient investments
Cluster – main beneficiary	Investors
Description	In the current situation a large number of decision makers have only limited information on the actual energy consumption of buildings. Indeed, actual data shows that EPC is not very reliable in assessment of energy performance, especially when it comes to deciding on the effectiveness of financial investments in these buildings. Many public institutions and project developers, in particular when there is the support of an ESCO company, lack information and knowledge about the implementation of many energy efficient projects, their economic justification and historical energy data consumption. The pilot will reduce the uncertainty linked to energy efficiency investments, which can be attributed to the lack of relevant skills and ability to assess investments. It will strengthen debt and equity financing of energy efficiency projects. More specifically, the platform will provide investors / financiers and project developers the opportunity to: • evaluate quickly and easily key performance indicators (KPIs) for projects (analysis and assessment techniques for projects and companies, benchmark tools, financial performance indicators for investors, special indicators “Green barometer”, written analysis); • Explore past and future projects (project history, synopsis, projects/companies more attractive to investors); • Make a profound risk measurement and evaluation (implement key indicators, such as debt service coverage ratio, integrate a cross-check for legal requirements of environmental laws and regulations). In the UC assets involved are the buildings and building related energy systems (heating, cooling, lightning, ventilation, electricity), smart meters. Sequence of events: 1. Collect data on the energy consumption of the building and then analyze it, 2. Define energy efficiency measures that are going to be deployed on a building, 3. Define costs of potential energy efficiency measures, 4. Calculate the EPC-related KPIs and cross-domain financial and energy consumption,
Goals	• Analysing best and financially low risk refurbishment option in energy efficiency • Evaluate quickly and easily KPIs for projects • Explore past and future projects

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Use Case ID	De-risking energy efficient investments
	• Make a profound risk measurement and evaluation
Actors	• ESCO, • governmental institutions, • financiers, • project developers, • building owners (public and private sectors).
Envisioned analytics and tools (BD4NRG components involved)	• Forecast of energy consumption • Report on potential money savings, • ECM advisor, • Benchmark tools, • Financial performance indicators for investors, Special indicators "Green barometer", • Debt service coverage ratio, • Integrate a cross-check for legal requirements of environmental laws and regulations.
Precondition	Existing monitoring devices to measure the most significant parameters of the EPC are going to be installed.
Triggering events (if applicable)	-
Post conditions	The pilot will reduce the uncertainty linked to energy efficiency investments, which can be attributed to the lack of relevant skills and ability to assess investments. It will strengthen debt and equity financing of energy efficiency projects.
Related KPIs	Investment in energy efficiency [kWh/€] Investment CO2 efficiency: [tones /€] Renewable energy efficiency: [kWh/€] House heating/cooling efficiency: [kWh/m2]

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

6. BD4NRG Requirements

In the following Table 25 and Table 26 show the requirements of each UC divided into functional and non-functional, following the methodology in Chapter 2.

6.1. Functional Requirements

Table 25: Functional Requirements for all LSPs

ID	Description	Type	Rationale	Priority (1 – 5)	Related UC
Rq_01	Availability of system able to identify the type of tower from the data retrieved from LIDAR	Functional	Identify tower catalogue model and distinguish insulators, active conductors, shield wires, etc.	4	LSP01_02
Rq_02	System able to identify the condition of the transmission tower, based on LiDAR scan of each individual tower	Functional	Assess state of identified elements and if some element is missing, to feed maintenance processes	4	LSP01_02
Rq_03	The tools should support an experienced human outage planner in maintenance scheduling process	Functional	Producing the outage plan is a complex and multidimensional challenge. The tool should assist the human planner to deliver optimized plan based on a data available today	5	LSP02_01
Rq_04	Output should be presented with Gantt's chart	Functional	Organization of time and work	4	LSP02_01
Rq_05	Outage plan proposal should be automated taking into consideration as many of available key operational and maintenance data as possible	Functional	Automated processes speed up and optimize processes	5	LSP02_01
Rq_06	Tool must allow adding future additional optimization constrains and criteria	Functional	Optimization processes	3	LSP02_01
Rq_07	Typification and classification of circuit breaker recorded incidents	Functional	Automatic incident type (e.g. single line to ground fault)	5	LSP02_01

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

ID	Description	Type	Rationale	Priority (1 – 5)	Related UC
			detection and association with external causes (e.g lightning, fires)		
Rq_08	e-Mobility platform shall be enabled for user registrations	Functional	Charging stations and vehicles subscription	4	LSP03_01
Rq_09	e-Mobility platform shall be enabled to provide charging station/EV monitoring service	Functional	Charging stations and EVs monitoring	4	LSP03_01
Rq_10	e-Mobility platform shall be enabled to provide charging station remote management service	Functional	Charging stations remote management	4	LSP03_01
Rq_11	DSO must be able to select specific charging stations to be involved in incentives campaign to optimize power management	Functional	e-Mobility/DSO coordination for power management optimization	5	LSP03_01
Rq_12	e-Mobility platform must be able to execute smart micro-contracts and micro-payments	Functional	e-Mobility Smart Contracts and micro-payments	4	LSP03_01
Rq_13	Components have edge computing capabilities	Functional	A first processing of data saves space and alleviates computing stress in the cloud/server	5	LSP09_01
Rq_14	It should be possible to calculate the Smart Readiness Indicator (SRI), as well as analytics focused on energy performance and comfort.	Functional	Analytics are pivotal for the next generation EPC calculation	4	LSP10_01 LSP11_01
Rq_15	It should be implemented an algorithm to calculate EPC labeling	Functional	This is the objective to be obtained, therefore, the algorithm should be implemented	3	LSP10_01
Rq_16	It should be possible to feed the system with the new results obtained for EPCs back.	Functional	EPC should be continuously updated and feed backed	3	LSP10_01

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

ID	Description	Type	Rationale	Priority (1 – 5)	Related UC
Rq_17	KPIs, energy performance, investment risk, cross-domain financial and energy consumption analytics should be calculated	Functional	Analytics are most important aspect for the cross-domain financial and energy consumption analytics	5	LSP12_01
Rq_18	Real-time grid meter data should be integrated in the system	Functional	Capabilities to collect real-time data	5	LSP3_01 LSP7_01
Rq_19	BESS model should identify under-performance areas based on real time and historical data	Functional	Use of BESS model to optimize the system	5	LSP7_01
Rq_20	Existing platform should be reconfigurated to host the final pool of involved loads (municipal, residential, commercial loads, EV charger and distributed RES)	Functional	Existing platform should be adapted to buildings involved	4	LSP08_01
Rq_21	Data collected form acoustic sensors should be analyzed in order to map possible failures	Functional	Mapping of acoustic sources makes it possible to identify where partial discharges have occurred	5	LSP03_02

6.2. Non-Functional Requirements

Table 26: Non-Functional Requirements for all LSPs

ID	Description	Type	Rationale	Priority (1 – 5)	Related UC
Rq_22	Owners must be engaged for the application of EPC	Cultural and political	Users are the key users to be addressed	2	LSP10_01
Rq_23	A new EPC-based business models should be created	Cultural and political	Business models should drive the next generation of energy	2	LSP10_01
Rq_24	Data set should have historical operation data	Data	Useful for clustering and correlation	5	LSP01_01 LSP05_01 LSP05_02 LSP06_01 LSP06_02 LSP09_01 LSP12_01
Rq_25	Oscillography available in COMTRADE format	Data	Parsing COMTRADE files allows	5	LSP01_01

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

ID	Description	Type	Rationale	Priority (1 – 5)	Related UC
			retrieving relevant KPIs		
Rq_26	Incident report data should integrate information from different sources (e.g. lightning, fires)	Data	To infer correlations between incidents and causes	4	LSP01_01
Rq_27	Data stored in COMTRADE files should allow a representation of different types of incidents	Data	Useful for clustering	4	LSP01_01
Rq_28	LiDAR files include point cloud cilinder data centered around transmission tower	Data	To focus differentiation near or at the OHL tower and allow for feature extraction	5	LSP01_02
Rq_29	Very high-resolution geo images	Data	For precise cross-reference with LIDAR samples of RoW	4	LSP01_02
Rq_30	Asset monitoring data that are needed for calculation of AHI and AUI need to be available	Data	AHI and AUI can only be calculated for the assets for which asset monitoring data are available.	5	LSP02_01
Rq_31	Each individual charging station must have its own unique identifier	Data	Charging Station Unique Identifier	3	LSP03_01
Rq_32	Each individual EV must have its own unique identifier	Data	EV unique identifier	3	LSP03_01
Rq_33	Charging station must be connected to internet to provide data to e-Mobility platform	Data	Charging Station Connectivity	3	LSP03_01
Rq_34	EV must be connected to internet to provide data to e-Mobility platform	Data	EV Connectivity	3	LSP03_01
Rq_35	Charging station data shall be stored for result evaluation	Data	Charging Station Data Storage	4	LSP03_01
Rq_36	EV data shall be stored	Data	EV Data Storage	4	LSP03_01

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

ID	Description	Type	Rationale	Priority (1 – 5)	Related UC
	for result evaluation				
Rq_37	Historic data of maintenance must be collected and evaluated for being implemented in predictive maintenance	Data	Assessments of related data for interpretation before predictions	5	LSP04_01
Rq_38	Historic data of faults must be collected and evaluated for being implemented in predictive maintenance	Data	Assessments of related data for interpretation before predictions	5	LSP04_01
Rq_39	The specific assets for the demonstration must be identified and consent should be in place to perform the deployment of the specific agent	Data	Identification and consent of trial customer	5	LSP05_01 LSP05_02
Rq_40	Procedures for determining the energy quantity of the service (activation) should be identified and developed.	Data	As separate metering might not be available at all or not in the needed time resolution, procedures for quantity determination need to be established	4	LSP06_02
Rq_41	The specific assets for the demonstration must be identified and consent should be in place, if the asset owner is not a BD4NRG consortium member. Simulation is an alternative.	Data	Identification and consent of trial assets, including data provision	2	LSP06_01
Rq_42	Historic data of faults occurred	Data	Understand frequency of faults before fault prediction monitorization	3	LSP09_01
Rq_43	Weather data provided in near real-time and forecasts should have more than one time range (e.g. 6 hours, 24 hours, etc.)	Data	Near real-time data is used for verification of normal operations and weather forecast data for production prediction	4	LSP09_01

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

ID	Description	Type	Rationale	Priority (1 – 5)	Related UC
Rq_44	Systems need availability of contextual data (Cadastre, weather services, LiDAR, COPERNICUS, EPC databases	Data	Publicly availability and access to contextual data via open APIs and open data	4	LSP08_01 LSP10_01 LSP11_01 LSP12_01
Rq_45	Systems need availability of building information (BIM, audits…)	Data	Building status characterization is crucial for determining the building status	5	LSP10_01 LSP11_01 LSP12_01
Rq_46	Data must be pre-processed before integration (if needed, i.e. LiDAR)	Data	Some available data sources need to be previously processed	4	LSP3_02 LSP10_01 LSP11_01
Rq_47	A list of the Energy Efficiency Measures must be defined	Data	To help users to decide which energy efficiency measures to deploy on a building	4	LSP12_01
Rq_48	Average costs of energy efficiency measures must be defined	Data	To help users to decide which energy efficiency measures to deploy on a building regarding their budget	5	LSP12_01
Rq_49	Real-time data must be available.	Data	Data come from multiple sources	5	LSP7_01
Rq_50	Data harmonization, orchestration and synchronization should be performed	Data	Data come from multiple sources, in different formats and timestamps.	4	LSP03_01 LSP04_01 LSP08_01 LSP10_01 LSP11_01 LSP12_01
Rq_51	Weather data should be provided for weather forecasting.	Data	For the estimation of RES power generation and for correlation purposes with load data, and to inform data analytics and forecasting	5	LSP08_01
Rq_52	Historical RES data must be provided	Data	Collect power generation data to inform data analytics and forecasting	5	LSP08_01
Rq_53	Unique identifier for each of the 150 inverters of the Solar demonstration platform	Data	Component unique identifier	3	LSP09_01

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

ID	Description	Type	Rationale	Priority (1 – 5)	Related UC
Rq_54	Systems must receive and analyze smart meters data for IoT	Data	IoT capabilities to collect dynamic and real-time data	5	LSP08_01 LSP10_01 LSP11_01 LSP12_01
Rq_55	The system must collect the data provided by the storage	Data	Capabilities to collect real-time data	5	LSP7_01
Rq_56	Gateways/dataloggers of the sensors must be connected to the internet	Data	DER production data collection and storage	5	LSP09_01
Rq_57	e-Mobility platform must ensure interoperability between charging station and EV	Interoperability	EV/Charging Station Interoperability	5	LSP03_01
Rq_58	Technical equipment must be integrated into Big data platform	Interoperability	Fundamental needs	5	LSP06_01
Rq_59	Connection with the external aggregator's platform must be allowed	Interoperability	Models have to be specificated	4	LSP06_01, LSP06_02
Rq_60	API should be developed for integration of data services to middleware and BD4NRG platforms	Interoperability	Different APIs should be developed in order to provision of data sharing	4	LSP04_01
Rq_61	Use of communication protocol (MQTT – Mosquito?) and communication transport (NB-IoT, ZigBee, LTE, LoRa?)	Interoperability	Use of a protocol and transport layer already used by many different sources	5	LSP09_01
Rq_62	Data must be anonymized according to GDPR	Legal	Data processed externally must observe privacy requirements (GDPR)	5	LSP02_01 LSP03_01 LSP04_01 LSP08_01 LSP10_01 LSP11_01
Rq_63	As independent aggregators must be able to access all electricity markets – settlement models must therefore be identified and codified, based on market laws and rules	Legal	Fulfilment of the requirements of the Clean Energy Package	5	LSP06_02
Rq_64	EPC v2.0 must be adopted	Legal	Legislation changes for the application of next generation	2	LSP10_01

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

ID	Description	Type	Rationale	Priority (1 – 5)	Related UC
			EPCs		
Rq_65	Access to residential electricity data must be allowed	Legal	Informed consent of end-users	5	LSP08_01
Rq_66	Lab trial performance must be verified in a safe lab environment of IMEC HomeLab	Safety	The performance will be validated in a safe environment before field tests are starting	4	LSP05_01 LSP05_02
Rq_67	Smart meters data collection should apply security aspects for data accessing	Security	These data are available on databases, which should be protected	3	LSP08_01 LSP10_01 LSP11_01
Rq_68	All data (energy and financial) collection should apply security aspects for data accessing	Security	These data are available on databases, which should be protected	3	LSP12_01

6.3. Constraints

Table 27: Constraints of all LSPs

ID	Description	Type	Rationale	Priority (1 – 5)	Related UC
Rq_69	Operational security limits	Constraint	Operational security limits (e.g., N-1 criterion) needs to be considered	5	LSP02_01

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

7. Conclusions

In the current deliverable UCs and requirements of each pilot were identified, a necessary preparatory activity to those foreseen in WP 3 – Data Governance and Management and WP 4 – Data Modelling & Open Modular AI-based edge-level Analytics Toolbox for the realisation of the architectures, services and tools for the demonstration pilots.
For each UC the expected services, objectives and KPIs were depicted and they will be used as a basis for the demonstration activities to be performed in the 12 LSPs. Table 25 gathers together the 18 UCs analysed in this document, classified per pilot.

Table 28: UCs summary

Pilot	UC code	UC description
LSP 1	LSP01_01	AI methods applied to Condition-based monitoring of Circuit Breakers
	LSP01_02	Semi-Automatic Maintenance Plan for Right-of-Ways (RoW) in OHL
LSP 2	LSP02_01	Maintenance Plan Optimization
LSP 3	LSP03_01	Efficient management of flexibility resources for supporting grid operation
	LSP03_02	Anticipated prediction of asset failures
LSP 4	LSP04_01	Predictive maintenance and fault prediction for MV grid and cross functional predictive asset management
LSP 5	LSP05_01	The agents responsible for controlling residential DR assets can benefit from the transferred knowledge from other agents in an MDP setting
	LSP05_02	Using residential DR assets to retrieve local grid information and use this information in the aggregation activations
LSP 6	LSP06_01	Identification of flexibility potential, based on near-real time data
	LSP06_02	Independent aggregator flexibility activation, quantification, and settlement
LSP 7	LSP07_01	Predictive Analytics Forecasting of storage optimized operation for improved market reserve participation
LSP 8	LSP08_01	Optimal management strategies for cross-infrastructure flexibility resources to support higher RES shares at the island community level
LSP 9	LSP09_01	Optimizing and improving management and prediction of DER production
	LSP09_02	Regional energy sector aggregation digital twin
	LSP09_03	FlexATwin (Regional Digital Twin with multiple interfaces for flexibility across sectors)
LSP 10	LSP10_01	Predictive and Prescriptive Analytics for improved EPCs Reliability
LSP 11	LSP11_01	Weather-enhanced Building Thermal Comfort Prediction to Enable Energy Efficiency Reliable Forecasting
LSP 12	LSP12_01	De-risking energy efficient investments

Besides, the requirements were provided by the partners and categorised. A total of 20 functional requirements, 47 non-functional requirements were distinguished and reported in Table 25 and Table 26. Non-functional requirements are divided in categories, such as data, interoperability, legacy or safety requirements. A summary of Requirements is available in Table 26.

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

Table 29: Requirements' summary

Use Case ID	Functional Requirements	Non-Functional Requirements
LSP01_01	-	Rq_24, Rq_25, Rq_26, Rq_27
LSP01_02	Rq_01, Rq_02	Rq_28, Rq_29
LSP02_01	Rq_03, Rq_04, Rq_05, Rq_06, Rq_07	Rq_30, Rq_62
LSP03_01	Rq_08, Rq_09, Rq_10, Rq_11, Rq_12	Rq_31, Rq_32, Rq_33, Rq_34, Rq_35, Rq_36, Rq_57, Rq_62
LSP03_02	Rq_18, Rq_21	Rq_46, Rq_50
LSP04_01	-	Rq_37, Rq_38, Rq_50, Rq_60, Rq_62
LSP05_01	-	Rq_24, Rq_39, Rq_66
LSP05_02	-	Rq_24, Rq_39, Rq_66
LSP06_01	-	Rq_24, Rq_41, Rq_58, Rq_59
LSP06_02	-	Rq_24, Rq_40, Rq_59, Rq_63
LSP07_01	Rq_18, Rq_19	Rq_49, Rq_55
LSP08_01	Rq_20	Rq_44, Rq_50, Rq_51, Rq_52, Rq_54, Rq_62, Rq_65, Rq_67
LSP09_01	Rq_13	Rq_24, Rq_42, Rq_43, Rq_53, Rq_56, Rq_61
LSP09_02	-	-
LSP09_03	-	-
LSP10_01	Rq_14, Rq_15, Rq_16,	Rq_22, Rq_23, Rq_44, Rq_45, Rq_46, Rq_50, Rq_54, Rq_62, Rq_64, Rq_67
LSP11_01	Rq_14	Rq_44, Rq_45, Rq_46, Rq_50, Rq_54, Rq_62, Rq_67
LSP12_01	Rq_17	Rq_24, Rq_44, Rq_45, Rq_47, Rq_48, Rq_50, Rq_54, Rq_68

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.

References

[1] ISO/IEC/IEEE 42010:2011

[2] Alistair Cockburn, Basic UC Template, http://cis.bentley.edu/lwaguespack/CS360_Site/Downloads_files/Use%20Case%20Template%20%28Cockburn%29.pdf

[3] IEC 62559-2:2015: Use case methodology – Part 2: Definition of the Templates for Use Cases, Actor List, and Requirements List, 2015.

[4] https://bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf

[5] https://www.bigdataframework.org/big-data-architecture/#:~:text=The%20NIST%20Big%20Data%20Reference%20Architecture%20is%20a%20vendor%2Dneutral,develop%20a%20Big%20Data%20architecture.&text=Along%20the%20Information%20Value%20axis,results%20following%20the%20value%20chain.

[6] Curry E. (2016) The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches. In: Cavanillas J., Curry E., Wahlster W. (eds) New Horizons for a Data-Driven Economy. Springer, Cham. https://doi.org/10.1007/978-3-319-21569-3_3

[7] https://www.data-infrastructure.eu/GAIAX/Redaktion/EN/Publications/gaia-x-technical-architecture.html

[8] https://www.fit.fraunhofer.de/content/dam/fit/en/documents/Industrial-Data-Space_Reference-Architecture-Model-2017.pdf

[9] Commission Regulation (EU) No 543/2013 of 14 June 2013 on submission and publication of data in electricity markets and amending Annex I to Regulation (EC) No 714/2009 of the European Parliament and of the Council Text with EEA relevance

[10] https://www.keycloak.org

[11] ISO/IEC 20547 standards

[12] NIST Big Data interoperability Framework (NBDIF)

[13] Bishnu P. Bhattarai et al. Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions. IET Smart Grid, 2019, Vol. 2, pp. 141-154.

[14] http://ec.europa.eu/research/participants/portal/doc/call/h2020/dt-ict-11-2019/1850117-background_document_on_big_data_and_energy_en.pdf

[15] Grant Agreement 872613 – BD4NRG – Part B, page 20

BD4NRG project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 872613.