This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 957047.

H2020-LC-SC3-EE-2020-1/LC-SC3-B4E-6-2020

Big data for buildings

Building Information aGGregation, harmonization and analytics platform

Project Nº 957047

D2.1 - Detailed description of Use cases and end-user services

Responsible:	CSTB
Document Reference:	D2.1
Dissemination Level:	Public
Version:	1.0
Date:	29/06/2021

Executive Summary

The BIGG project aims at demonstrating the application of big data technologies and data analytic techniques for the complete buildings life-cycle of more than 4000 buildings in 6 large-scale pilot test beds, achieved by: 1) The *Open Source BIGG Data Reference Architecture 4 Buildings* for collection/funnelling, processing and exchanging data from different sources (smart meters, sensors, BMS, existing data sets); 2) An interoperable buildings data specification*, BIGG Standard Data Model 4 Buildings*, based on the combination of elements from existing frameworks and EC directives, such as SAREF, INSPIRE, BIM, EPCHub that will be enhanced to reach full interoperability of building dates; 3) An extensible, open, cloud-based *BIGG Data Analytics Toolbox* of service modules for batch and real-time analytics that supports a wide range of services, new business models and support reliable and effective policy-Making. These solutions will be deployed and tested cross pilot and country validation of at least two business scenarios in Spain and Greece.

This deliverable presents the in-depth analysis of the use cases defined by pilots in the task 6.1. The deliverable D.6.1 had detailed the pilot goals and the technical assets that are pre-existing, especially the data sets in use and the ICT systems in place. They were presented as a textual description of a system's behavior corresponding to pilot's point of view.

These use cases are briefly introduced in the table below extracted from the deliverable D6.1.

Business Cases		Use Cases	
#1	**Benchmarking** and **Energy Efficiency** tracking in Public Building	#1	Benchmark and Monitoring of Energy Consumption
		#2	Energy Efficiency Measures, EEM: Registration and Evaluation
#2	**Energy Certification** (EPC) in Residential and Tertiary Buildings	#3	Integration of INSPIRE spatial data with Energy Performance Certification
		#4	Adoption of sustainability indicators of EU framework Level(s) in building Certification
#3	**Building Life-Cycle**: From Planning to Renovation	#5	Interoperability between BIM, BMS, CMMS and building simulation engines
		#6	Interoperability of BIGG with EEFIG-DEEP
		#7	Interoperability between EUBSO and national/regional EPC hubs
#4	**Energy Performance Contract** (EPC) based savings in commercial buildings: from Planning to Renovation	#8	Building **Assets management**
		#9	Actual savings tracking realized by the **ECM**s (monitors on daily/weekly/monthly basis)
		#10	Energy Performance Contract (**EPC**) **management**
#5	Buildings for occupants: **Comfort Case**	#11	Optimization using **weather forecast**
		#12	Optimization using **occupancy forecast**
		#13	Optimization using **price forecast**
#6	**Energy Certification** (EPC) in Residential and Tertiary Buildings	#14	On demand-response for **Electricity**
		#15	On demand-response for **Natural Gas**

The core objective of the work carried out to deliver this document is to go further in the previous analysis in order to propose extended details about each pilot processes and their potential relation with the BIGG platform.

The WP2 is dedicated to the specification (and integration) of the BIGG technical framework. As such, the objective of the first tasks of WP2 is to analyse the end-user oriented representation brought by the work done in WP6 and to produce more detailed requirements. The first step of this exercise is done in the T2.1 and the result is presented in the current document.

In order to do so, all the use cases have been detailed adopting a specific approach based on the decomposition of these use cases in processes. Then these processes have been analysed to identify commonalities (similar functions) making thus the need for specific services appearing.

Contributors Table

DOCUMENT SECTION	AUTHOR(S)	REVIEWER(S)
Section 1	CSTB (Bruno Fiès, Eric Pascual, Nicolas Pastorelly)	ECTP and some experts, Members of the ECTP (Juan Jacobo Peralta Escalante - CEMOSA, Noemi Jiménez Redondo – CEMOSA, Erling Onstein – NTNU, Lizhen Huang – NTNU, Sebastjan Meža - ZAG, Gema Hernández Moral – CARTIF, Sofia Mulero Palencia – CARTIF, Pieter Pauwels – TUE) INETUMBE (María Pérez)
Section 2	UCs owners From UC 1 to UC7: CIMNE (Stoyan Danov and Eloi Gabaldon) UC 8-10 : HELEXIA/ENERGIS (Pierre Lehanneur, Romain Hollanders, Frederic Wauters) UC 11-13 : HELEXIA/ENERGIS (Pierre Lehanneur, Riccardo De Vivo, Frederic Wauters) UC 14: HERON (Konstantina Mentesidi, Athanasios Papakonstantinou, Dimitris Chatzigiannis) UC 15: DOMX (Stratos Keranidis, Polychronis Symeonidis, Ellie Efstathiou)	ECTP and some experts, Members of the ECTP (Juan Jacobo Peralta Escalante - CEMOSA, Noemi Jiménez Redondo – CEMOSA, Erling Onstein – NTNU, Lizhen Huang – NTNU, Sebastjan Meža - ZAG, Gema Hernández Moral – CARTIF, Sofia Mulero Palencia – CARTIF, Pieter Pauwels – TUE) INETUMBE (María Pérez)
Section 3	CSTB, All	ECTP and some experts, Members of the ECTP (Juan Jacobo Peralta Escalante - CEMOSA, Noemi Jiménez Redondo – CEMOSA, Erling Onstein – NTNU, Lizhen Huang – NTNU, Sebastjan Meža - ZAG, Gema Hernández Moral – CARTIF, Sofia Mulero Palencia – CARTIF, Pieter Pauwels – TUE) INETUMBE (María Pérez)

Annexes	UCs owners From UC1 to UC 7: CIMNE (Stoyan Danov and Eloi Gabaldon) UC 8-10 : HELEXIA/ENERGIS (Pierre Lehanneur, Romain Hollanders, Frederic Wauters) UC 11-13 : HELEXIA/ENERGIS (Pierre Lehanneur, Riccardo De Vivo, Frederic Wauters) UC 14: HERON (Konstantina Mentesidi, Athanasios Papakonstantinou, Dimitris Chatzigiannis) UC 15: DOMX (Stratos Keranidis, Polychronis Symeonidis, Ellie Efstathiou)	

Table of Contents

Table of Figures

List of Tables

Table of Acronyms and Definitions

AHU	Air Handling Unit
AI	Artificial Intelligence
API	Application Programming Interface
BMS	buildings management systems (BMS), computerized maintenance systems (CMMS)
BPMN	BPMN is the chosen notation to represent the UCs. Business Process Model and Notation (BPMN) is the standard for business process modelling. It is provided by the Object Management Group (OMG).
CMMS	Computerized maintenance systems
DEEP	De-Risking Energy Efficiency Platform
DHF	BIGG Harmonized Format (BHF)
DHW	Domestic Hot Water
DR	Demand Response (DR)
DSF	Demand Side Flexibility
ECM	Energy Conservation Measure
EEM	Energy efficiency measures (EEM)
EFFIG	Energy Efficiency Financial Institution Group
EPC	Energy Performance Certificate .
EPCo	Energy Performance Contract
ES	Energy Conservation Measures (ECMs)
ESCO	Energy Service Company
EUBSO	EU Building Stock Observatory (EUBSO) and national/regional Energy Performance Certification (EPC)
HVAC	Heating Ventilation and Air Conditioning
INSPIRE	The INSPIRE Directive, establishing an infrastructure for spatial information in Europe to support Community environmental policies, and

	policies or activities which may have an impact on the environment entered into force in May 2007. INSPIRE is based on the infrastructures for spatial information established and operated by the Member States of the European Union. The Directive addresses 34 spatial data themes needed for environmental applications. See https://inspire.ec.europa.eu/
Process	A process is a logical grouping of operations manipulating data : retrieval, publication, transformation, derived data extraction,… Processes can be defined in a hierarchical way, for instance to provide a higher level view of a group of more elementary actions.
RAF	Reference Architecture Framework
RES	Renewable Energy Source
Service	Services are processes (whatever is their level) that are mainly intended to be available outside BIGG, and are used by external applications consuming them to implement their own actions. Note that nothing forbids some of the BIGG own processes to consume such services.
UC	Use Case. In this document, the various use cases mentioned are taken from the D6.1 and detailed according to a chosen formalism.

I. INTRODUCTION

I.1. Purpose and structure of the document

The global work in this workpacakge aims at defining the specifications and design of the reference architecture based on the needs of the enduser's requirements expressed in the early tasks of WP6.

This document corresponds to the first deliverable of the workpackage dedicated to the technical specification of the BiGG plateform. The role of the platform is enabling to collect and expose building related information or data so that they can be shared among the different services expected by the Pilots.The goal of this document is to present the in-depth analysis of the use cases defined by pilots in the task 6.1. The deliverable D.6.1 *"Detailed description of pilot's technical assets: ICT tools and accessibility to data sets"* had detailed the pilot goals and the technical assets that are pre-existing, especially the data sets in use and the ICT systems in place. They were presented as a textual description of a system's behavior corresponding to pilot's point of view.

The objective of task T2.1 *"Use case / End user services definition"* is to go further in the analysis in order to propose extended details about each pilot processes and their potential relation with the BIGG platform. The expectation of the T2.1 (and the corresponding D2.1) is to make the transfer from an end-user oriented representation to a developer oriented documentation.

In order to do so, all the uses cases have been detailed adopting a specific approach based on the decomposition of these use cases in processes. Then these processes have been analysed to identify commonalities (similar functions) making thus the need for specific services appearing.

The document is organized into 4 main parts:

- The present introduction
- The "Use Cases analysis based on BPMN": This part first introduces the reason why this language was chosen and the approach decided in order to stress the expected role of the BIGG platform and the interactions between the platform and the existing pilot's assets. In this section, a short introduction to the work performed on each use case in provided but in order to keep a reasonable length to the document, all the details have been placed in the annexes.
- The "BIGG platform services": This section is the analysis of the BPMN diagrams. The services identified at this stage are organized into four main categories. At the end of the section a matrix is provided making the link between the services and the concerned Use Cases (UCs).
- The Annexes: In this section, all the detailed description of the UCs is compiled. The descriptions are made based on the same template.

I.2. Scope and audience

This document is a first step that sorts and lists even at high level the interactions between the BIGG platform and the UCs owners. This document will be reused by the next tasks in the current work package (Technical framework: Specifications and integration) but also in other work packages that will focus more speciafically on dedicated technical topics (like : "End-users communication and security layers", "Data harmonization layer" and "AI toolbox".The first organization of services described here will constitute a starting point to go deeper in the definition of the technical requirements.

As such the audience for this document is mainly composed of the project partners that will be involved in the developments of the BIGG components. It may also interest more generally because of the methodology adopted to move from end-user requirements to technical requirements described through services.

II. USE CASES ANALYSIS

II.1. Approach and template

When this analysis task started, the main initial decision to make was to choose tools and methodologies that should be used within the consortium to detail pilot's current processes and go beyond the detailed descriptions of pilots technical assets delivered by previous tasks. It was needed to find a high-level way to explain how software systems (pilots' assets + BIGG platform) operate without being lost in potentially complex conceptual architectures at this early phase of the project.

It was decided to work with diagrams in order to effectively communicate the ideas for both technical and non-technical audiences. Indeed in the project's consortium we do have business-oriented individuals as well as product and technical managers.

It was important to demonstrate that the potential architecture of a BIGG platform brings solutions for different business use cases among pilots, that's why we needed to describe actors and chronological views of a BIGG workflow for each pilot. It has been decided to use a BPMN model [1]to further analyze our different use cases.

One of the requirements for the analyses was to identify the technical actors involved, in order to be able to define more precisely the BIGG ecosystem . Swim lanes of BPMN model clearly represent the different technical actors of a workflow, moreover we wanted to be sure to precisely formulate the assumptions of how the business processes can go from one step to another, both inside and outside of a BIGG platform. As far as methodology is concerned, then it was decided to provide a common BPMN representation pattern to each pilot so they could express their understanding and requirements for a technical workflow involving the BIGG platform.

[1]https://www.bptrends.com/publicationfiles/07-04%20WP%20Intro%20to%20BPMN%20-%20White.pdf

Here is the proposed BPMN pattern which is described below:

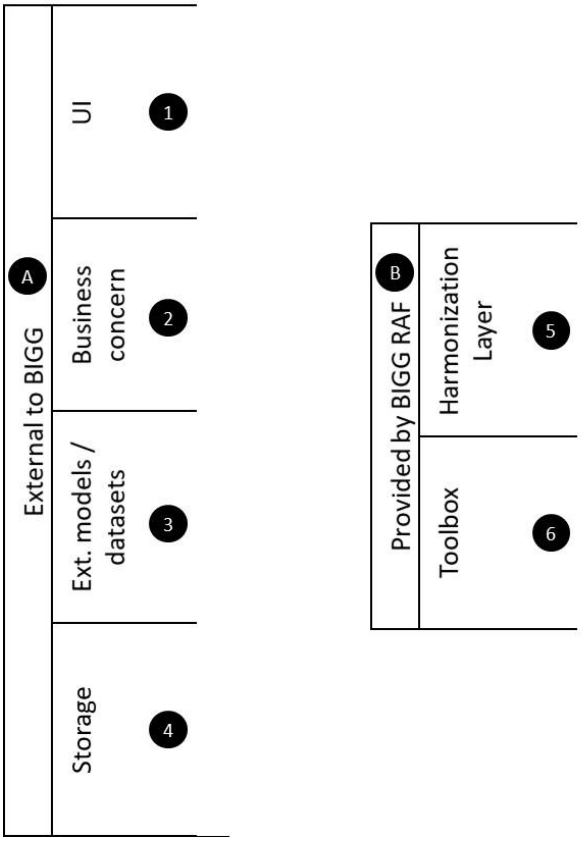


Figure 1: BPMN proposed template

The first step consisted in establishing a difference between business processes happening locally (in pre-existing pilots' infrastructures) and business processes happening inside the BIGG platform as depicted in Figure 1.

Pre-existing pilots' infrastructures have been detailed in the documents "D6.1- Detailed description of pilots technical assets- ICT tools and accessibility to data sets" where all partners involved in trials -providing data and/or building infrastructure – produced detailed descriptions of the technical assets the the BIGG data platform and solutions need to interact with. Pilots' infrastructures support business requirements that are called use cases in their integration process with BIGG.

It was thus decided to create two different BPMN pools for each pilot:

- one pool for the processes happening outside of the BIGG platform (see A in Figure 1),
- one pool for any process that may happen inside the BIGG platform (see B in Figure 1).

It must be stated that BPMN notation has been used here as a common language in the working team for its "expressivity" (using swimlanes, boxes and arrows …) and its practicality, but it has been extended from its original intent to modelize building blocks more than business processes.

Here is the detail of the different swim lanes numbered from 1 to 6 in Figure 1 inside the two pools:

1. This swim lane represents any interaction that is happening with the user-interface of the pilot system. It has been decided that the BIGG platform shall be a toolbox and shall not provide any direct interface to the end-users. That's why modelling

local processes where users may interact with the BIGG platform through a local system is very important.

2. This swim lane shall capture requirements of any business process that could happen locally in the pilots' platforms. It is essential to understand which local business processes can be in relation with business processes and enablers hosted inside a BIGG platform.
3. This swim lane represents the pilot's locally managed data. It seems relevant to specify them in order to understand how local data can be set in relation with the harmonized data and processes inside the BIGG platform.
4. Finally, requirements for local storage are very important to be described as well. This swim lane represents pilots' local storages where outputs of a BIGG platform may be stored.
5. One key feature of a BIGG platform is to harmonize data among different use cases. At this stage of the analysis of the project, it is fundamental to start modelling the requirements for normalized data in each pilot use cases.
6. Another key feature of a platform is to propose a toolbox providing enablers for common computation matters among pilots. It is then necessary for each pilot to specify what kind of process is expected from the BIGG toolbox in this swim line.

The question is still pending about whether the BIGG platform shall propose a common storage solution or not. A partner can extend the BPMN template swim lines if required.

Each pilot has worked on defining its requirements and perception of BIGG-platform-related-assets through this BPMN representation. Use cases requirements will be summarized in the next chapter and will be detailed more extensively the annexes of the document. Chapter III will summarize what could be the technical intersections found in this detailed analysis of BIGG use cases. Since the BIGG platform is a service platform, the following chapters will summarize – based on the first steps described in section II.1, the emerging common requirements that seem to appear in the different use cases, especially in terms of the data standardization and toolbox enablers usage.

II.2. Use cases analysis based on BPMN

II.2.1. UC 1 – Benchmarking and monitoring energy consumption

Overview	
Comparison of energy consumption between similar buildings and evaluation of changes in the energy consumption trends.	
Scope	
Building energy management in large organizations with a significant number of buildings is not an easy task. This use case attempts to improve the overview of the energy performance status of the building stock. With the objective of: 1) facilitating continuous monitoring of energy consumption and performance trend of buildings. Evaluating the current energy consumption of buildings in comparison with their own historical evolution of energy consumption. 2) facilitating decision-making by identifying the energy performance of buildings in real and continuous use. Comparing the normalized energy consumption with similar buildings, not only the general energy consumption but also considering different situations of use (heating periods, cooling periods, base load, etc.) These results will be presented in a web visualization application and should allow public authorities, decision makers and energy managers to improve the understanding of how their buildings are consuming the energy and understand which ones need improvement actions, commissioning or maintenance. All this by reducing the dedicated resources and ensuring the correct location of investments to improve the energy performance of the building park.	
BPMN diagram	V.1.1.
Data collection, harmonization and preparation	This process contemplates the ingestion of raw data available in different external repositories, through different APIs or by uploading files (Excel, XML or CSV), these data will be stored in the local database in raw format.. This raw data will be stored in a local database and then mapped into the BIGG data model, thus ensuring harmonization of this data. The BIGG infrastructure will provide the mapping of the data in the BIGG model and these will be stored again in a harmonized database in the local application. Once the data has been harmonized, it will be processed to verify its quality and, if necessary, apply a cleaning process. For these pre-processes, the analysis toolbox developed in BIGG will be used. The execution of these tools may be carried out in BIGG's own infrastructure or in the "local" Infrastructure, as appropriate in each case. The harmonized and elaborated data will be accessible to end users so that they can assign the different data to each of the buildings, either globally or individually.
Performance trend	The final user can ask for some calculations and also set some parameters (for example: to fix the training period, validation period, model inputs, etc.)

evaluation and Benchmarking (Comparison of each building consumption energy with itself and with similar buildings)	in order to obtain a best baseline model to compare historical energy consumption period against the current ones. The analysis will use analysis tools developed in the BIGG environment, these will be executed on the local infrastructure or on the BIGG itself, depending on the needs. The expected results of these analyses are: Performance trend analysis - Segmentation of historical consumption according to different conditions (calendar characteristics, weather conditions, etc.) - Consumptions KPI's: kWh/m2ºC (annual, monthly, weekly, week's day...) also split on base load, Heating dependence, Cooling dependence, etc... - Selection of base line models for building and energy source - Detection of anomalies or significant change on trends on energy consumption for each building. Benchmarking analysis - Building characteristics Clustering, using consumption KPI's, building characteristics, etc… to select the similar buildings. - Classification Modelling, to rate the buildings energy consumption in comparison to similar buildings - Average KPI's, to extract general KPI's for Building typologies, organization, etc... The results of these analysis are expected to be stored in the harmonized database of the local infrastructure.
Presentation and exchanging results	The final users will have a web application (Dashboard) to view and explore the different results obtained from the processes described above. Admin users will have the possibility of exchanging buildings data or results with external applications through the web application.
Processes BPMN description	V.1.2.

II.2.2. UC 2 - Energy Efficiency Measures (EEM) registration and evaluation

Overview	
Continuous registration and evaluation of the implementation of energy efficiency measures in buildings.	
Scope	
The registration of the actual application of energy efficiency measures (EEM) in public administrations has not been carried out, so far, in a systematic and continuous way. The fact of not having EEM registered implies that it is impossible to jointly and systematically evaluate their impact on the energy efficiency of buildings, as well as the consequent improvement in decision-making in the application of EEM. This use case attempts to improve the recollection and evaluation of EEM applied in buildings. With the objective of: 1) Facilitating the continuous and systematic collection of EEM applied in the building stock in a harmonized way. 2) Evaluating the real impact of each of the energy improvement actions and extracting valuable information at the statistical level from the set of EEMs applied and registered. These results will be presented in a web visualization application and should allow public authorities, decision makers and energy managers to improve the understanding of the real impact of the actions they carry out on their buildings, and on the other hand to improve making decisions for actions to be carried out in the future.	
BPMN diagram	V.2.1.
Processes	
Data collection, harmonization and preparation	This process contemplates the ingestion of raw data available in different external repositories or provided directly form the User Interfaces. This raw data will be stored in a local database and then mapped into the BIGG data model, thus ensuring harmonization of this data. The BIGG infrastructure will provide the mapping of the data in the BIGG model and these will be stored again in a harmonized database in the local application. Once the data has been harmonized, it will be processed to verify its quality and, if necessary, apply a cleaning process. For these pre-processes, the analysis toolbox developed in BIGG will be used. The execution of these tools may be carried out in BIGG's own infrastructure or in the "local" infrastructure, as appropriate in each case. The harmonized and elaborated data will be accessible to end users so that they can assign the different data to each of the buildings, either globally or individually.
EEM Evaluation (EEM KPIs and	The final user can ask for some calculations and also set some parameters (for example: to fix the training period, validation period, model inputs, etc.)

savings evaluation)	in order to obtain the best model to predict the usage pattern based on specific time of the year and weather conditions. The analysis will use analysis tools developed in the BIGG environment, these will be executed on the local infrastructure or on the BIGG itself, depending on the needs. The expected results of these analyses are: - Best base line model applied to simulate the whole period consumption time series, considering that any other EEM was implemented. - Assessment of the energy savings of each EEM by $\sum_{k=0}^{n}(\text{real Q} - \text{base line Q})$ - Assessment of EEM impact in terms of variation in optimal parameters and energy savings.
Presentation and exchanging results	The final users will have a web application (Dashboard) to view and explore the different results obtained from the processes described above. Admin users will have the possibility of exchanging buildings data or results with external applications through the web application.
Processes BPMN description	V.2.2.

II.2.3. UC 3 - Integration of INSPIRE spatial data with Energy Performance Certification (EPC)

Overview	
Automating the integration of the INSPIRE data with the EPC input data for completing and cross-checking the information and improving the reliability of the services.	
Scope	
Automating the integration of the INSPIRE data with the EPC input data for completing and cross-checking the information and improving the reliability of the services	
BPMN diagram	V.3.1.
Processes	
Data collection, harmonization and preparation	This process contemplates the ingestion of raw data available in different external repositories or provided directly from the User Interfaces. This raw data will be stored in a local database and then mapped into the BIGG data model, thus ensuring harmonization of this data. The BIGG infrastructure will provide the mapping of the data in the BIGG model and these will be stored again in a harmonized database in the local application. Special attention will be placed on alignment with INSPIRE schema (https://inspire.ec.europa.eu/). Once the data has been harmonized, it will be processed to verify its quality and, if necessary, apply a cleaning process. For these pre-processes, the analysis toolbox developed in BIGG will be used. The execution of these tools may be carried out in BIGG's own infrastructure or in the "local" Infrastructure, as appropriate in each case. The harmonized and elaborated data will be accessible to end users so that they can assign the different data to each of the buildings, either globally or individually.
Modelling EPC (EPC indicators predictor based on weather and cadastral data)	The final user can ask for some calculations and also set some parameters (example: to fix the training period, validation period, model inputs, etc ...) in order to obtain a best model to predict the energy performance pattern based on specific time of the year, cadastral data and weather conditions. The analysis will use analysis tools developed in the BIGG environment, these will be executed on the local infrastructure or on the BIGG itself, depending on the needs. The expected results of these analyses are: - EPC indicators predictor based on weather and cadastral data. - Characterization of the theoretical energy consumption/demand of a certain subset of buildings - Characterization of the correlation between geographical areas and EPC indicators

	- Comparison to actual geographically-aggregated consumption data and socio-economic data - Estimation of the energy performance gap of the buildings.
Presentation and exchanging results	The final users will have a web application (Dashboard) to view and explore the different results obtained from the processes described above. Admin users will have the possibility of exchanging buildings data or results with external applications through the web application.
Processes BPMN description	V.3.2.

II.2.4. UC 4 - Adoption of the sustainability indicators of common EU framework Level(s) in building certification

Overview	
Adoption of Level(s) indicators that is possible with the currently available energy certification input data and specification of the extended input necessary for the calculation of the rest of indicators for the future evolving of the certification. Level(s) is a new European approach to assess and report on the sustainability performance of buildings, throughout the full life cycle of buildings. (https://ec.europa.eu/environment/levels_en)	
Scope	
This use case tries to improve the Energy Efficiency Certificates, as in UC3. Additionally, this Use Case 4 aims to facilitate the alignment of EPC results with the KPI Level(s) framework. In some cases, Level(s) KPIs will be calculated. In other cases, most of them, the way to calculate them will be studied to be able to add them in the future. Interoperating with other existing databases to obtain the necessary parameters or suggesting adding these additional fields in the Building Certification methodology. On the other hand, to improve making decisions for actions to be carried out in the future.	
BPMN diagram	V.4.1.
Processes	
Data collection, harmonization and preparation	This will be the same process presented in UC3.
EEM Evaluation (EEM KPIs and savings evaluation)	This will be the same process presented in UC3.
Evaluation of Level(s) KPIs:	Building energy certification results are explored to identify Level(s) KPIs that can be calculated and additional parameters required to calculate the other KPIs are identified.
Processes BPMN description	V.4.2.

II.2.5. UC 5 - Energy Efficiency Measures (EEM) registration and evaluation

Overview	
To guarantee the interoperability between the different data acquisition/generation systems that can be found during the life cycle of buildings.	
Scope	
This use case intends to guarantee the interoperability between different data acquisition/generation systems during the building cycle life. These systems can be monitoring systems, facilities control systems, buildings management systems (BMS), computerized maintenance systems (CMMS) and Building Information models. And it aims to achieve the following objectives: 1)To harmonize and map the different input data sets. 2) To allow the users to set, update and select the more trustworthy parameters in case of different values available from various sources. These results should allow public authorities to have control of their building data in a single point, thus facilitating its possible reuse in future applications.	
BPMN diagram	
Processes	
Data collection, harmonization and preparation	This process contemplates the ingestion of raw data available in different external repositories or provided directly form the User Interfaces. This raw data will be stored in a local database and then mapped into the BIGG data model, thus ensuring harmonization of this data. The BIGG infrastructure will provide the mapping of the data in the BIGG model and these will be stored again in a harmonized database in the local application. Once the data has been harmonized, it will be processed to verify its quality and, if necessary, apply a cleaning process. For these pre-processes, the analysis toolbox developed in BIGG will be used. The execution of these tools may be carried out in BIGG's own infrastructure or in the "local" Infrastructure, as appropriate in each case. The harmonized and elaborated data will be accessible to end users so that they can assign the different data to each of the buildings.
Cross-validation process	This process will be done in the local infrastructure without interaction with BIGG developments. The basic actions to be done are: - Assignation of each parameter to each building. - Validation of results provided by different sources

Processes BPMN description	V.5.2.

II.2.6. UC 6 - Interoperability of BIGG with EEFIG-DEEP

Overview	
Adoption of the specifications of De-Risking Energy Efficiency Platform (DEEP) https://deep.eefig.eu/ , its data model and EEM definitions to ensure compatibility; development of Application Programming Interfaces (API) for data exchange with DEEP and its latest version (DEEP2); development of a standard and objective procedure for monitoring and validation of resultant energy and cost savings based on continuous data collection and evaluation and results export to DEEP.	
Scope	
This use case attempts to improve the recollection and evaluation of EEM applied in buildings, as it happens in UC2. Additionally, this Use Case 6 pretends to facilitate the exchange of information collected and processed in UC with EFFIG- DEEP Platform. With the specific objective of: 1) Standardizing the EEM information with DEEP platform specifications. 2) Facilitating the exchange of information with DEEP platform.	
BPMN diagram	V.6.1.
Processes	
Data collection, harmonization and preparation	This will be the same process presented in UC2.
EEM Evaluation (EEM KPIs and savings evaluation)	This will be the same process presented in UC2.
Data anonymization	To facilitate the exchange of information with the BIGG platform. The buildings and EEM data will be anonymized so that it is not possible to identify exactly which building the data belongs to.
Exchanging results	Admin users will have the possibility of exchanging buildings data or results with external applications through the web application.

Processes BPMN description	V.6.2.

II.2.7. UC 7 - Interoperability between EU Building Stock Observatory (EUBSO) and national/regional Energy Performance Certification (EPC) hubs through BIGG

Overview	
The BIGG data model will be mapped to the European Building Stock Observatory (EUBSO) https://ec.europa.eu/energy/topics/energy-efficiency/energy-efficient-buildings/eu-bso_en and the necessary transformations developed in order to ensure interoperability with the Catalonia EPC hub.	
Scope	
This use case attempts to improve the Energy Performance Certificates, as it happens in UC3. Additionally, this Use Case 7 intends to facilitate the exchange of information collected and processed in UC7 with EUBSO Platform. With the specific objective of: 1) Standardizing the EPC information with EUBSO platform specifications. 2) Facilitating the exchange of information with EUBSO platform.	
BPMN diagram	V.7.1.
Processes	
Data collection, harmonization and preparation	This will be the same process presented in UC3.
EEM Evaluation (EEM KPIs and savings evaluation)	This will be the same process presented in UC3.
Data anonymization	To facilitate the exchange of information with the BIGG platform. The EPC data will be anonymized so that it is not possible to identify exactly which building the data belongs to.
Exchanging results	Admin users will have the possibility of exchanging buildings data or results with external applications through the web application.

Processes BPMN description	V.7.2.

II.2.8. UC 8 - Assets management to store, view, update all relevant assets such as buildings, contracts, invoices, meters, sub-meters, sensors, equipment, ...

Overview	
Enable the user to store, access, view and manage all relevant data regarding the management of an EPCo (Energy Performance Contract). The data includes the relevant information to describe the building, the contract, the invoices and consumption data, the monitoring hardware (meters, sensors) and the equipment.	
Scope	
The process concerns the management of an EPCo, which involves the management of the building equipment operation, the impact of occupant's behaviour and the conditions of the building maintenance contract's implementation. This use case is strictly focused on building operation on either new buildings or existing ones). It can be implemented for any building which is operational (built and commissioned). The process involves the building owner, the building occupants and the building manager or energy service company that has been missioned to operate the building in its behalf. To achieve the scope presented in the previous section, several processes must be carried out in the existing "local" application, Energis Cloud owned by Energis and rendered accessible to the members of the consortium through Helexia , improved with the interaction with the BIGG system. These processes will be in the scope of this use case.	
BPMN diagram	V.8.1.
Processes	
Data collection, harmonization and preparation	- Data collection, This process contemplates the ingestion of raw data available in different external repositories. This raw data will be stored in a local database. This data sources can be various (existing data bases, Excel spreadsheets, digitized energy bills, existing maintenance contractors information, energy performance contracts terms…) - Data Harmonization The data will then be mapped into the BIGG data model, thus ensuring harmonization of this data. The BIGG infrastructure will provide the mapping of the data in the BIGG model and these will be stored again in a harmonized database in the local application

	- Data Preparation Once the data has been harmonized, it will be processed to verify its quality and, if necessary, apply a cleaning process. For these pre-processes, the analysis toolbox developed in BIGG will be used. The execution of these tools may be carried out in BIGG's own infrastructure or in the "local" Infrastructure, as appropriate in each case. The harmonized and elaborated data will be accessible to end users so that they can assign the different data to each of the buildings, either globally or individually.
Processes BPMN description	V.8.2.

II.2.9. UC 9 - Actual savings tracking realised by the Energy Conservation Measures (ECMs) undertaken by the ESCO are monitored on a daily/weekly/monthly basis, ...

Overview	
Ease the process of quantifying the impact of an Energy Conservation Measure on a given building through an accurate modelling of the building consumption. Enable the user to track the implemented ECMs in time and their impact on the managed asset.	
Scope	
The process concerns the management of an EPC which involves the management of the building equipment operation, the impact of occupant's behaviour and the conditions of the building maintenance contract's implementation. This use case is strictly focused on building operation on either new buildings or existing ones. It can be implemented for any building which is operational (built and commissioned). The process involves the building owner, the building occupants and the building manager or energy service company that has been missioned to operate the building in its behalf.	
BPMN diagram	V.9.1.
Processes	
General Description	Energy Conservation Measures (ECMs) cover a broad range of measures so as to conserve energy while at the same time guaranteeing building occupant comfort. ECMs can be related to the design of a building (e.g., openings orientation) or internal condition monitoring and optimization (e.g., adjustment of CO_2 levels to a desired / law compliant value). Such ECMs can be undertaken both for new constructions, in which case they occur either during the building design phase or for renovations of old buildings, and existing buildings. The present use case –at least as far as the pilot scenario is taken into consideration– focuses on pre-existing buildings so measures are rather linked to different consumptions

	adjustment (e.g., electricity) and optimization of internal conditions (e.g., CO_2 concentrations lighting levels) rather than design plans (orientation of windows/openings).
Processes BPMN description	V.9.2.

II.2.10. UC10 - Energy Performance Contract Management to manage the EPCo life cycle and perform actions (e.g. reporting) according to contractual milestones

Overview	
Energy Performance Contract Management to manage the EPCo life cycle and perform actions (e.g. reporting) according to contractual milestones.	
Scope	
The process concerns the management of an EPCo which involves the management of the building equipment operation, the impact of occupant's behaviour and the conditions of the building maintenance contract's implementation. This use case is strictly focused on building operation on either new buildings or existing ones. It can be implemented for any building which is operational (built and commissioned) and when an Energy Performance Contract is signed. The process involves the building owner, the building occupants and the building manager or energy service company that has been missioned to operate the building in its behalf.	
BPMN diagram	V.10.1.
Processes	
General Description	An Energy Performance Contract (EPCo) is a contract between an Energy Service Company (ESCO) and a building owner to improve the building performances through the implementation of one or more Energy Conservation Measures (ECMs) with the particularity that the ESCO's service fee is function of the actual savings achieved. The main aspects of an EPCo are the following: • Existing Baseline: The current performances of the building in terms of energy efficiency and energy costs; • Existing factors of influence: The factors that have an impact on the existing baseline which are not under the control of the service provider (weather conditions, production level, occupancy...); • Projected savings: A simulated model that quantifies the expected results of the contract;

	• Actual savings: actual results of the projects (taking into account the factors of influence) which will be a dynamic data stream collected on site on an on-going basis. This use case will focus on facilitating the management of such a contract by providing the stakeholder with functionalities to simplify the on-going management process, the assessment of the impact of the implemented ECMs and to make sure the current building performances are in line with the contractual objectives The BIGG platform will provide data analytics services to enable the EPCo manager to verify the performances of the building against the projected savings that are contractually defined. The platform will enable to model the building consumption prior to ECM implementation and verify that the level of savings reached is in line with the milestones of the EPCo.
Processes BPMN description	V.10.2.

II.2.11. UC11-12-13 - Optimization using weather, occupancy and price forecasts

Overview
Weather and occupancy conditions have a direct impact on the energy demand of buildings (e.g. the necessity of heating/cooling) and RES production (e.g. solar PV). In use case 11, being able to forecast weather conditions will allow to proactively match energy demand and supply (e.g. heat less if large solar gains are expected later in the day). In use case 12, occupancy forecasts will allow, for example, to make office spaces or meeting spaces comfortable before they will be occupied (pre-cooling/pre-heating) or to turn off HVAC devices a few hours before they will be unoccupied to leverage the building thermal inertia and reduce the consumptions. Moreover, it is expected that energy prices will become more dynamic in future years and their influence of related energy costs with respect to usage profiles will be more significant. In use case 13, a control algorithm will be developed that uses energy prices forecast in addition to the occupancy forecast and weather forecast to allow using always the most sustainable and/or cheapest energy available. The facility management company will perform a site survey of the on-site equipment (HVAC, chiller, AHUs, …), collect information from occupants to identify comfort issues in the site, building or in zones of the building.
Scope
The process concerns the energy management of building which involves the optimization of the building equipment operation, the impact of occupant's behaviour and the conditions of the building maintenance contract's implementation. From a building life cycle perspective, this use case intervenes in the operational phase of the building. It can be implemented in any building that built and commissioned.

The use case is geared toward the building owner and building manager or energy service company that has been missioned to operate the building.	
BPMN diagram	V.11.
Processes	
General Description	The facility management company needs to collect, transform and store data of the assets, namely: ● Technical information about the equipment to control to allow actuating them from the BIGG platform ● Consumption data of the controlled equipment ● Environmental data about the zones for which the comfort must be optimized, such as indoor temperature, CO2, relative humidity ● Weather, occupancy and energy price data, including forecasts, to be exploited by the optimization When required, extra sensors will be installed on site. Commands will be executed via the Engie Connect platform, the BMS or directly via devices like Modbus modules. When required, extra actuators will be installed on site. The facility manager will design and implement the ruleset to optimize comfort and energy usage. Assets and metrics will be used in the AI ruleset. The BIGG platform will allow us to build the models which will be used to make predictions of consumption & comfort. An example of such a rule is to adapt the current temperature setpoints when high solar gains are being forecasted for a building zone and considering its orientation. Different rulesets will be identified for different time periods (e.g., summer and winter) and for the characteristics of the zone (e.g., orientation, activity type, zone type). Possibilities to optimize the rulesets using AI will be explored to further improve the comfort level, decrease the energy consumption, and react to unforeseen changes in the rooms (e.g., reinforcement learning techniques are able to identify). The facility manager will have a dashboard for comfort & energy monitoring of the different buildings and zones. This will help him or her maintain occupants' comfort in buildings by decreasing the time to take corrective actions and to reduce occupants' complaints. The occupants of the building will have a local dashboard to see the decision logic of the optimization algorithms provided by the BIGG platform and to follow the comfort related KPIs.
Processes BPMN description	V.11.2.

II.2.12. UC14- Electricity Demand Response (DR)

Overview	
This use case describes the process where residential electricity end-users can monitor their power/energy consumption of their connected devices. On top of that users could perform manual actuation (ON/OFF) for controllable devices or automatic actuation based on specific time-schedule events. Comparison of energy consumption between similar buildings and evaluation of changes in the energy consumption trends.	
Scope	
This use case describes how end-users can monitor in real-time their power/energy consumption, both total and at phase/relay level for their connected devices using a mobile App or a web-based platform. On top of energy monitoring users can perform manual actuation for their connected devices at relay or phase level such as water heaters, IoT sensors or electric space heaters. In addition to manual management users can benefit from automated actuation based on rules/events both set by themselves or by allowed/agreed upon to be performed by third parties, e.g in the context of demand side flexibility (DSF) requests. The above data along with the end user preferences will provide valuable input for the BIGG platform to learn the behaviour of consumers and their trends. To achieve the above goal, consumers' energy consumption and environmental data should be monitored and collected by various smart meters and sensors that will be eventually installed in residential set-ups during the pilot execution. In addition, the following objectives are foreseen to be achieved: Typical consumption patterns will be identified considering electricity consumption per flexible device (e.g., water heater) and for the aggregate portfolio on a daily and/or monthly basis to make customers aware of their energy wastages. Evaluation of the current energy consumption of residential consumers in comparison with similar trends of other households or their own consumption history. Classification of individual residential consumers across their monthly energy-cost analysis. Identification of load profile (electricity use and PV production) at prosumer level through data monitoring and analysis for evaluating the available flexibility to be harnessed. Identification of the flexibility potential by enabling users providing their flexibility preferences through a user interface (mobile app or web-based). Facilitation of continuous monitoring of user engagement through dashboards, while reporting the successful status change actuations.	
BPMN diagram	V.12.1.
Processes	
Data entry, collection, storage, harmonization and data processing/validation	This process contemplates the collection of data available in different external repositories as well as the ingestion of user preferences (e.g. time-schedule for the operation of the water heaters). This data will be stored in a local database and then mapped into the BIGG data model, thus ensuring harmonization of this data. The BIGG infrastructure will provide the mapping of the

	data in the BIGG model and these will be stored again in a harmonized database in the local system. Once the data has been harmonized, it will be checked against several quality metrics such as outliers, missing values and, if necessary, apply a cleaning process. For these pre-processes, the BIGG analytics toolbox will be used. The harmonized and elaborated data will be accessible to project partners and especially the technology providers for the performance of various statistical analyses.
Flexibility capacity estimation	This process involves the real-time monitoring of electricity consumption at household and/or relay level as well as of PV generation for prosumers. In addition, sensor data (such as temperature, humidity or movement) will be monitored to evaluate consumers' flexibility potential. Key points of the process will be: a. Formulate clusters based on dynamic consumption data (e.g. heavy vs. light consumers, day vs night consumers etc.) b. Provide input to optimization (e.g. find the optimal combination of equipment/ smart appliances to fulfil the request). Analytics performed within the BIGG toolbox will incorporate user preferences to identify user behaviour & consumption patterns which will comprise input for benchmarking (consumption segmentation, baseline consumption). The user can also monitor the state of its energy consumption (compared to similar households or their own historical consumption). **KPIs involved:** Several metrics can be used such as: recordings of energy consumption(kWh),corresponding CO_2 emissions and electricity cost (€), inventory of controllable devices based on their specs (Watt), and number of engaged users or number of participants that use the mobile app.
Optimization of energy savings and costs	This process involves consumption management through the activation of DR events for minimizing the costs (based on market observed and forecasted data) and/or maximizing savings. Vertical modelling for analysing energy consumption trends per household. Analytics performed within the BIGG toolbox may include reinforcement learning, regression and computational/numerical methods (optimization). **KPIs involved:** Evaluate the DR offer (before and after DR) using a set of metrics (CO_2, kWh, costs).
Evaluation of energy efficiency	End-user decision related to opt-in/opt-out from the DR scheme. Dynamic consumption clusters based on real-time consumption or sensor data. Peak shavings vs. base load shifting (in pursuit of monetary gains or environmental signals).

	Analytics performed within the BIGG toolbox may involve horizontal modelling for comparing households with similar trends in order to: a. Detect similar households. b. Classification modelling for rating the buildings' energy consumption in relation to similar buildings. c. Identify clusters of end users based on their consumption behaviour. **KPIs involved:** Measure active users and successful DR activations (cases where consumers do not override DR recommendation). Calculate a % of kWh of fulfilled DR as part of the historical data and device profiles and the corresponding savings in € and CO_2 emissions. Analytics for the different user clusters e.g. number of peak shaving users vs. load shifters, volumes of energy savings per cluster. The results of these analyses are expected to be stored in the harmonized database of the local infrastructure.
Presentation and exchanging results:	The users will have a mobile app as well as a web-based platform to visualize the different results obtained from the processes described above.
Processes BPMN description	V.12.2.

II.2.13. UC15 - Natural Gas Demand Response (DR)

Overview	
This use case provides all the processes required for enabling Gas Suppliers to deliver Energy Efficiency and Flexibility Management services for legacy space heating gas boilers of residential consumers. The core aim of the Natural Gas supplier to correct the daily imbalances deriving from non-optimal estimation of their consumer portfolio daily demand, while offering incentives for the end consumers to participate in flexibility services, while minimizing the impact of the achieved climate comfort. The overall process progress is being monitored both by the supplier at the portfolio level and the participating consumers at the individual household level.	
Scope	
The proposed concept focuses on the Management of Natural Gas consumption in buildings, by actively controlling and optimizing the indoor environment, with the aim of (a) improving energy efficiency through load reduction and (b) contributing to energy system flexibility providing real-time gas balancing services. The core innovation of the proposed concept builds on the interconnection of major consuming heating devices (boilers, DHW preparation, radiators, etc.) with the gas network, through the seamless integration of the domX heating controller with legacy devices. The system is interconnected with a cloud-based energy management system that constantly collects, stores and analyses the detailed data collected from connected heating devices. The process involves as the main actors: a) the Natural Gas supplier that interacts through the dedicated dashboard, b) the Technology provider offering the Energy Management system and c) the gas consumers that interact with their residential heating system through a smartphone application. Within BIGG, several more important data sources will be connected in a harmonized system, including building characteristics, weather forecasts, actual natural gas consumption and billing data and natural gas balancing prices. The full list of heterogeneous data types that will be harmonized within this UC in BIGG include: - heating data collected from the boiler (water temperature, modulation, etc.), the thermostat (room temperature and target) and through climate sensors (outdoor temperature, etc.) - user requirements (temperature comfort limits, heating schedules) collected through the smartphone application; - building characteristics (size, orientation, insulation, etc.) and boiler specs, as captured by the building Energy Performance Certificate (EPC) data - forecast weather data collected from external service - natural gas consumption and billing data collected form the Energy Supplier - Natural gas balancing prices from the TSO	
BPMN diagram	V.13.1.
Processes	
Data entry, collection, storage	Data collection considers a) the static datasets collected by different sources (consumption and billing data from the Supplier, Building EPC from the National EPC registry) b) the real-time datastream coming from domX heating controllers, sensors and smartphone

	application are stored on the domX timeseries DB c) the public datasets collected from different public data sources (Weather data service, Gas Balancing Prices from the TSO) to the system. Different types of storage are considered, such as SQL, Time-series and Big Data DBs.
Data preparation-harmonization	The collected datasets (static dataset, real-time datastream, public dataset) are processed for removing outliers, missing items and inconsistent data and harmonized to match the BIGG data format.
Energy Efficiency evaluation	This process focuses on improving the energy efficiency of legacy natural gas boilers, by dynamically adapting the space heating in order to constantly meet the prevailing: - user comfort limits and heating schedule - building and boiler performance - outdoor weather variations The process aims to quantify the attained energy savings, by comparing the performance of the adaptive heating mode with the baseline mode.
Flexibility Potential evaluation	This process focuses on dynamically adapting the demand of connected gas consumers for adjusting the daily imbalances of the Gas Supplier. The considered steps include: - Real-time monitoring and consumption analysis - Identification of the flexibility potential of connected buildings and gas boilers - Calculation of gas imbalances to be corrected by the Supplier - Dynamic adaptation of the heating process per subscribed consumer - Quantification of the attained flexibility
Presentation and exchanging results	The overall process is being monitored by: - End users through the smartphone application - Natural Gas supplier through a dedicated dashboard Both user groups receive detailed reports on the achieved performance, covering both individual consumption point level and aggregate supplier portfolio level.
Processes BPMN description	V.13.2.

III. BIGG PLATFORM SERVICES

The various functionalities identified by the use cases description will be exposed by BIGG as services. This section summarizes their definition based on the analysis of the detailed BPMN models produced by the partners involved either as consumers or as providers of these functionalities.

It must be noted that the definition of services aims at sharing common functionalities expressed in distinct use cases and merging the specificities, if any, so that the technical specification of the software development can be derived directly from them. It is important to keep in mind that features that are too much use-case-specific are not included, and that only common interest features are taken in account.

III.1. BIGG services

Services are made available by the BIGG plaform as deployable units. These units will be provided in a technical form that allows being deployed on a central shared platform as well as on the service consumer own infrastructure if more relevant with respect to its operational context.

III.1.1. Data ingestion (S1)

This set of services provides entry-points for data entering the BIGG ecosystem from external sources. Two operating modes are to be addressed:

- Push: data are sent by external sources, such as instrumentation systems, either in real time or in periodic batches;
- Pull: data are retrieved from external sources on-demand by processing in BIGG services.

The role of the services is to handle the protocol and decode the format used by the communication with the external sources, so that the data can be passed to the harmonization mechanism afterwards in the homogenous form.

This relates to the communication layer represented in the global architecture diagram.

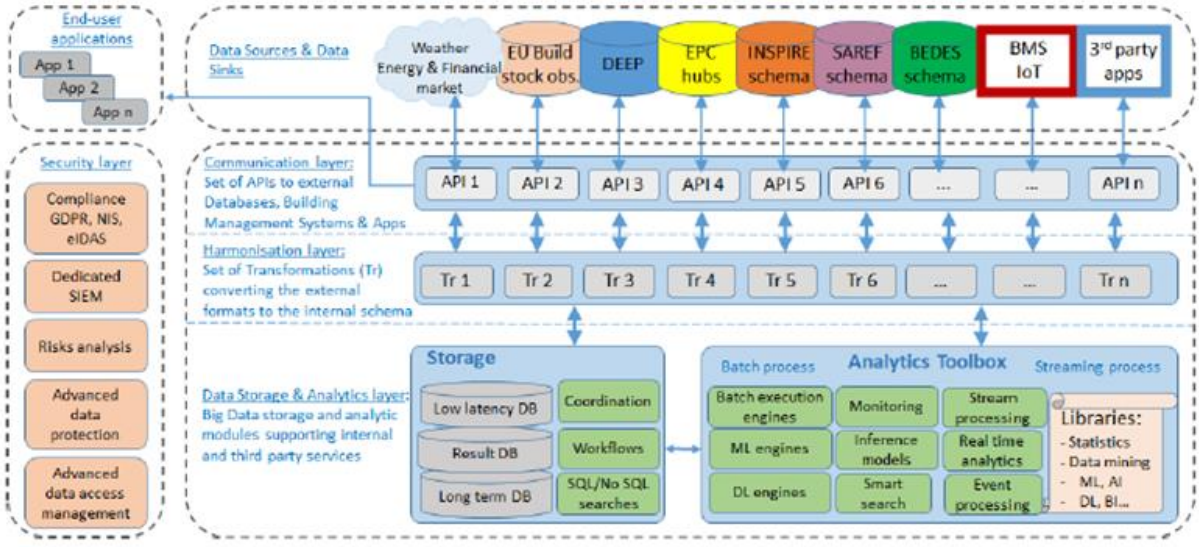


Figure 2: Global architecture diagram

All use cases operating on data obtained from external sources (instrumentation systems, weather data providers...) are candidate to use the services belonging to this family.

How many services to be created depends on how many communication channel types will be required to be supported by the use cases. This will be refined in the next design step.

III.1.2. Data harmonization (S2)

S2.1. External Format To BIGG Harmonized Format (EXT2BHF) data conversion

This service converts data provided in external formats to the relevant BIGG Harmonized Format (BHF). It allows consumers to “speak the same language” as long as data are related to the same topic. For instance, let’s say that weather data are available from various sources and using different formats, the conversion process will project them to the same BHF.

As mentioned as a general consideration above, the data conversion service can be deployed and executed on the use case consumer’s side to optimize data transfers and for security concerns.

S2.2. BIGG Harmonized Format to External (BHF2EXT) data conversion

This service is the counterpart of the EXT2BHF one. Its principle of operation and interface specification share the same approach.

III.1.3. Data preparation (S3)

S3.1. Gaps detection

This service operates on time series and detects gaps in periodic time series. It expects time series previoulsy harmonized in the BIGG format.

S3.2. Outliers detection

This service operates on time series and detects points which value is considered out of validity domain. The validity checking is based on the specification of bounds and other checking criteria.

Like the other analysis tools provided by BIGG, it expects time series previoulsy harmonized in the BIGG format.

S3.3. Cleaning

The cleaning process involves modifying the value to time series, in order to recover some time series with gaps or low impact anomalies that can be easily recovered without distorting the physical sense. If the impact of holes or anomalies in the data is low, they can be easily replaced by average or interpolated values. This could be applied, for example, in meteorological data time series, consumption time series.

Used by: UC1, UC2, UC3, UC4, UC5, UC6, UC7, UC9, UC10, UC11, UC12, UC13, UC14 UC15

III.1.4. Analytical process (S4)

S4.1. Data clustering

- Consumption curve: Extraction of the different usage profiles implicit in the time series of energy consumption. Process to use when no real usage/occupancy profile data is available. Depending on the granularity of the input data, they may provide different information.

  - Hourly data: hours of use / occupation, days with a similar operation, etc.
  - Monthly data: months of similar consumption behavior.

- Building characteristics. Similar buildings are identified in this process. Depending on the input data available, it may not be a trivial process (if we do not have the actual employment profile, supply inventory, hours of operation, etc.). The building characteristics clustering uses consumption KPIs, consumption profile features and these are processed them along with descriptive information about the building to create the best possible groups.

S4.2. Baseline extraction

A baseline model is extracted for each source of consumption for each of the buildings. This base model aims to be able to compare different periods of energy consumption of the building under the same operating conditions (weather, occupation, etc.).

S4.3. Modelling Used

- Baseline extraction
- Classification modelling. Get the statistical parameters of each item within the similar group. Global classification of the individual element within the group, average value of the group, deviation of the group, classification in base load, classification in heating load, classification in cooling load, classification of occupancy, etc.

  Used by: UC1, UC2, UC3, UC4, UC12, UC14
- Regression modelling: model identified from historical data variables predicting the consumption from temperature, occupancy, ...
- Data Modelling
  - EPCo
  - Obtain characteristics of the building based on time series of consumption
- Knowledge base modelling: Rule based controller optimising comfort and consumption based on rules checking conditions on actual and forecasted data and executing automated actions on building equipments (HVAC, …)

Used by : UC11, UC12, UC13

- Optimization of energy consumption and related costs based on market observed and forecasted data and vertical modelling for analysing energy consumption trends per household.

S4.4. Comparative analysis

- Comparative analysis between periods of consumption of a building.

This service refers to the detection of changes in the trend of the energy consumption profile of an element and in a specific period of time and to the evaluation of this change in the profiles. energy consumption. This analysis could be applied to different time series of consumption (global or partial in the building. The results can be used to assess changes in general consumption patterns of construction or to assess the impact of the application of EEM in buildings.

  - Trend analysis
  - Savings evaluation (EEM /EPC or buildings)
  - Anomaly detection
  - Forcasting

- Comparative analysis of energy consumption between buildings (Benchmarking)

Benchmarking serves as a mechanism to measure the energy performance of a single building over time, relative to other similar buildings.

  - Buildings Rankings
  - Building energy code

S4.5. Consumption KPI

Compute selected KPI over based on consumption time series and contextual information (building characteristics, weather, ...). These KPIs must describe the individual building consumption characteristics.

IV. CONCLUSION

In the previous chapters, the different Use Cases have been describe (and detailed in annexes) and a corresponding set of services have been defined.

S#	Ingest. (S1)	Data Harmonization (S2)		Data preparation (S3)			Analytical process (S4)				
UC		S2.1	S2.2	S3.1	S3.2	S3.3	S4.1	S4.2	S4.3	S4.4	S4.5
UC1	X	X		X	X	X	X	X	X	X	X
UC2	X	X		X	X	X	X	X	X	X	X
UC3	X	X		X	X	X	X	X	X	X	X
UC4	X	X		X	X	X		X	X	X	X
UC5	X	X		X	X	X					
UC6	X	X	X	X	X	X					
UC7	X	X	X	X	X	X					
UC8	X	X		X	X	X					
UC9	X	X			X	X		X	X		
UC10	X	X		X	X	X		X	X		
UC11	X	X		X	X	X			X		
UC12	X	X		X	X	X			X		
UC13	X	X		X	X	X			X		
UC14	X	X		X	X	X	X	X	X		X
UC15	X	X		X	X	X	X	X	X	X	X

Table 1: Correspondance among Services and Use Cases

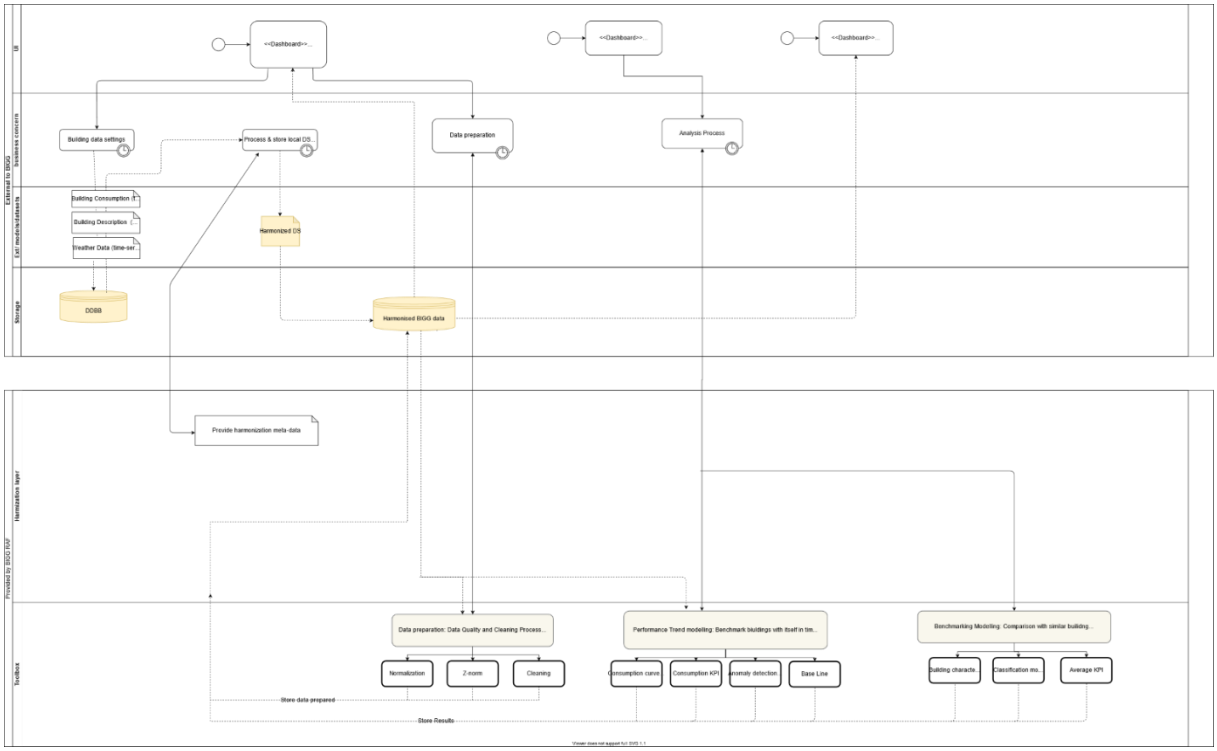
The table above summerizes the main result carried out in the first task of WP2. It gives a first high level framework and together with the description done by the UCs owners (with the diagrams provided in annexes) this first result will shape the future work of the next tasks of this work package. The next step will be to deepen the analysis of the exchanged data among these services and to specify in parallel the technical architecture that will support and orchestrate all the components needed to support these 15 UCs. A strong interaction is envisage with the work packages focusing on the commucation layer (WP3), on harmonisation of data mechanisms (WP4) and on the development of a toolbox for the processing of data (WP5).

We can see similarities in the different workflows of a different use cases justifying the BIGG project existence and defining direction for further targeted explorations: this first analysis needs to be detailed in latter work in order to produce more technical specifications of the BIGG common features.

V. ANNEXES

V.1. UC1

V.1.1. BPMN diagram



V.1.2. Specification of processes

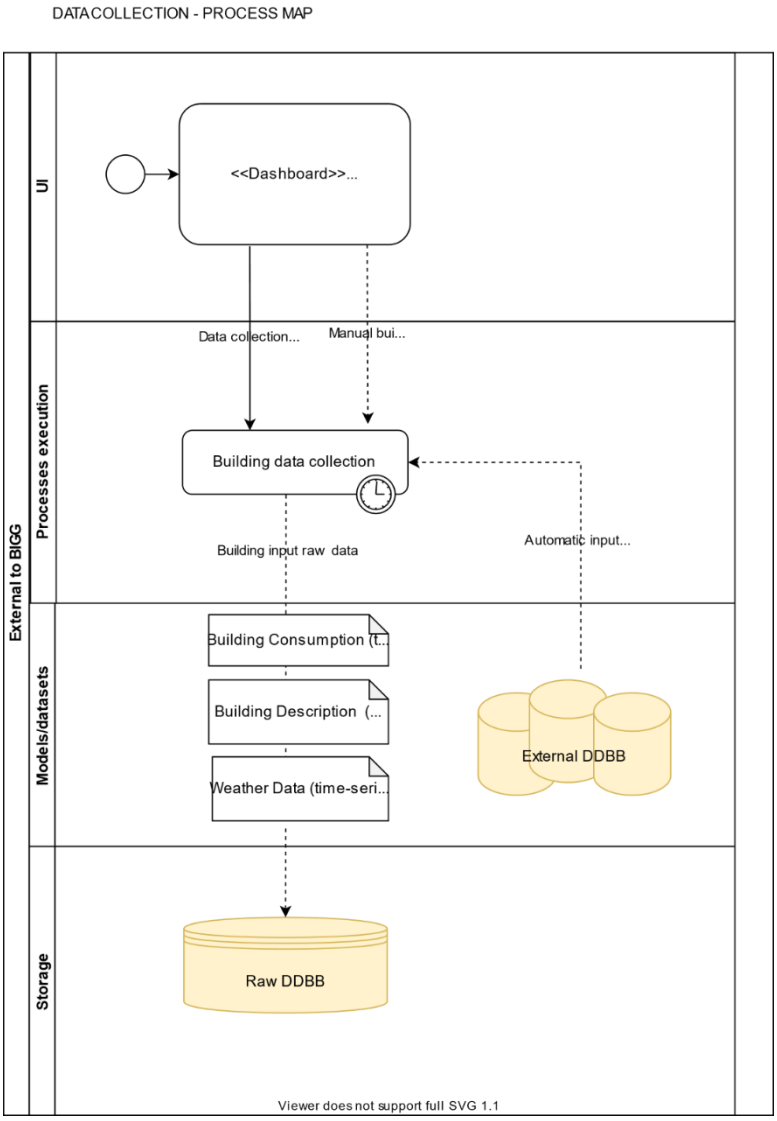
Process Map : <Data Collection>

The data required for the UC1 can come from different data sources. These data sources can be loaded from existing databases or uploaded manually by the user.

In the case of manual data upload, the end user will be able to load, modify or delete data from the UI, whether this is static (descriptive characteristics of the building) or time series (energy consumption).

In the case of data from existing databases, these can be collected directly through APIs or similar, the administrator user will have access from the UI to the import control of this data.

The building characteristics data, energy consumption or climate data will be systematically stored in raw in a Database. This Database will be located in the infrastructure of the existing ENMA application (owned by CIMNE).



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : <UI>

In the UI the end user, depending on certain permissions, can upload, modify or delete the building data manually (through web forms or by uploading CSV or Excel files, with predefined format), or configure the collection of data from databases. existing data, external to this Pool

<Dashboard: Management, loading, updating and configuration of input data.> **[ID:<BC1-UC01-001>]**

Type	*Use task*	
Name	*Dashboard: Management, loading, updating and configuration of input data*	
Documentation	*In the UI the end user, depending on certain permissions, can upload, modify or delete the building data manually (through web forms or by uploading CSV or*	

	Excel files, with predefined format), or configure the collection of data from databases. existing data, external to this Pool

LANE : < Processes execution>

<Building Data Collection> **[ID:<BC1-UC01-002>]**

Type	*Task type*	
Name	*Building data collection*	
Documentation	*The building data collection process is responsible for managing the data entry into the system.* *For manual data, the new data will be loaded and the modification or deletion of the data will be managed, according to the user's request.* *Depending on the configuration assigned by the user, timed calls will be launched to the different APIs to automatically retrieve data from external databases.*	

Exchange Requirement Data Objects

<Manual Input Raw Building data>

Type	Data Object
Name	*Manual raw building data*
Documentation	*Data uploaded, modified or deleted manually by the user in the user interface will be passed to the data collection process.*

<Automatic input Raw Building data>

Type	Data Object
Name	*Automatic raw building data*
Documentation	*The existing data in the different external databases will be processed by the data collection process. This data can be just new data or all data depending on the configuration that the user wants.*

LANE : < Models/Data Sets>

Library Data Objects

<Building Description >

Type	Data Object
Name	*Building Description*
Documentation	*The building description data object is semi-static data (unique values updated at each change of state (medium-long term)). This data can come from different sources which can be different external databases, documents or user interface forms. Depending on the source of the data, it will have different data models.* *These data describe the physical or typological characteristics of each of the buildings.* *Basic data summary:* *User identifier, building name, address, state, province, municipality, postal code, type of use, construction characteristics, construction area, number of floors, cadastral reference, etc.* *Installation data: cooling source, heating source, lighting types, etc.* *Cadastral data: all the cadastral information available.*

<Building Consumption >

Type	Data Object
Name	*Building consumption*
Documentation	*Building energy, electricity and gas consumption data can come from different sources (utility companies, energy information and management services, control systems, billing or manuals). In addition, these can have a heterogeneous aggregation, and can be hourly, monthly or both.* *Depending on the source of the data, it will have different data models.*

<Weather data >

Type	Data Object
Name	*Weather data*
Documentation	*The system will continuously retrieve data from meteorological stations from different stations located throughout Europe, with mostly hourly aggregation.* *The data collected can be:* *- Temperature, relative humidity, direct solar radiation in a horizontal plane, diffuse solar radiation, wind speed, wind direction, precipitation, etc.*

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Raw Data Base>

Type	Data Object
Name	*Raw Data Base*
Documentation	*Raw data is systematically stored in ENMA databases. The model of each of the data will be defined according to each of the source providers. The database is Apache HBASE (HBase is an open source non-relational distributed database)*

Exchange Requirement Data Objects

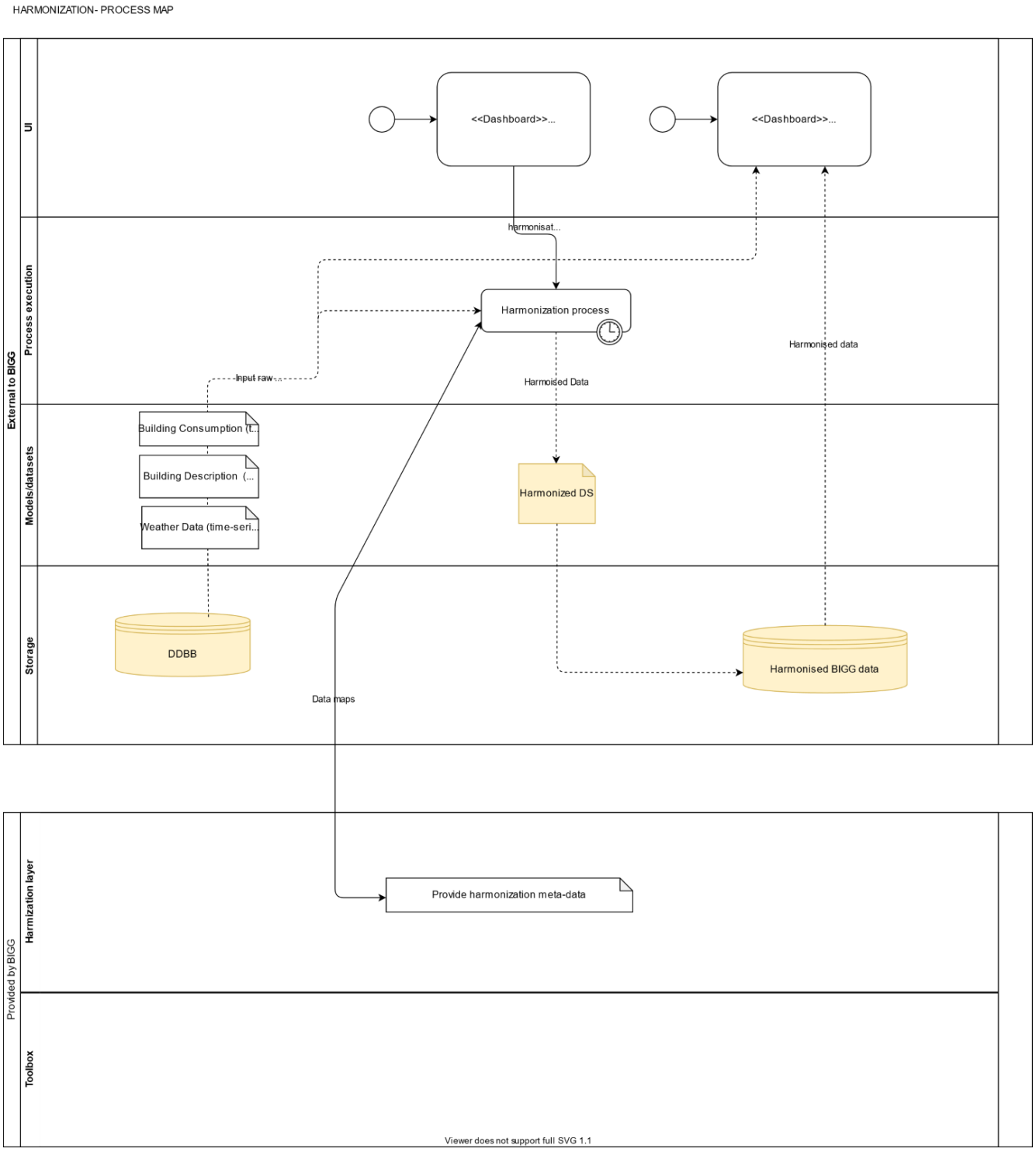
< Building input raw data>

Type	Data Object
Name	*Manual raw building data*
Documentation	*All data from the data collection process is stored in the Raw Database.*

A process map will also have a graphical version in the BPMN notation. This should be inserted before the pool and lane descriptions

Process Map : <Harmonization>

The harmonization process map is the process in charge of standardizing the raw data with the format of the BIGG data model. This standardization allows, on the one hand, to use BIGG's analytical developments and at the same time facilitates the exchange of data with other external systems or applications.



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Base>

Type	Data Object
Name	*Harmonized Data Base*
Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format. The database is Apache HBASE (HBase is an open source non-relational distributed database)*

Exchange Requirement Data Objects

<Raw Building data>

Type	Data Object
Name	*Raw building data*
Documentation	*Raw data is extracted from raw databases and sent to the data harmonization process when required.*

<Harmonized Building data>

Type	Data Object
Name	*Harmonized building data*
Documentation	*The harmonized data is extracted from the data harmonization process and is systematically stored in the harmonized database.*

LANE : < Processes execution>

<Harmonization process > [ID:< BC1-UC01-003>]

Type	*Task type*	
Name	*Harmonization process*	
Documentation	*The harmonization process is responsible for transforming the raw data into harmonized data in BIGG format. This process begins on a recurring basis and is controlled and managed by the administrator user from the user interface.* *The execution processes are the followings:* *- Retrieves raw input data that has not been harmonized so far, from the raw database.* *- Retrieves the data maps made in BIGG.* *- Assign the raw data to the BIGG data model.* *- Save the harmonized input data in the harmonized database*	

LANE : < Models/Data Sets>

Library Data Objects

<Building Description >

Type	Data Object
Name	*Building Description*
Documentation	*Same data and description as in the data collection process.*

<Building Consumption >

Type	Data Object
Name	*Building consumption*
Documentation	*Same data and description as in the data collection process.*

<Weather data >

Type	Data Object
Name	*Weather data*
Documentation	*Same data and description as in the data collection process.*

<Harmonized Datasets >

Type	Data Object
Name	*Harmonized Datasets*
Documentation	*The raw data (Building description, Building consumption and Weather data) harmonized and standardized on BIGG data model*

LANE : <UI>

In the user interface, the end user, depending on certain permissions, can manage and configure the execution of the harmonization process, and can explore both raw and harmonized data.

<Dashboard: Harmonization process management> **[ID:< BC1-UC01-004>]**

Type	*Use task*	
Name	*Dashboard: Harmonization process management*	
Documentation	*From the user interface, the administrator user can manage and control the harmonization process. You can configure and modify certain process parameters (execution times, mapping updates, etc.)*	

<Dashboard: Exploring raw and harmonized data> **[ID:< BC1-UC01-005>]**

Type	*Use task*	
Name	*Dashboard: Exploring raw and harmonized data*	
Documentation	*From the user interface, the user can view and explore the raw data that the system has ingested, as well as the result of the harmonization process. By controlling how much data has actually been harmonized, harmonization issues, even once the data has been normalized, the user can control and manage the data allocation in each building.*	

POOL : <Provided by BIGG >

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Harmonization layer>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies.

Library Data Objects

< Provide harmonization meta-data>

Type	Data object	
Name	*Provide harmonization meta-data*	
Documentation	*The BIGG harmonization layer has data mappings to transform the raw data into BIGG harmonized data, developed in the project. These data mappings will go through the harmonization process of the ENMA system so that you can execute them and harmonize the data.*	

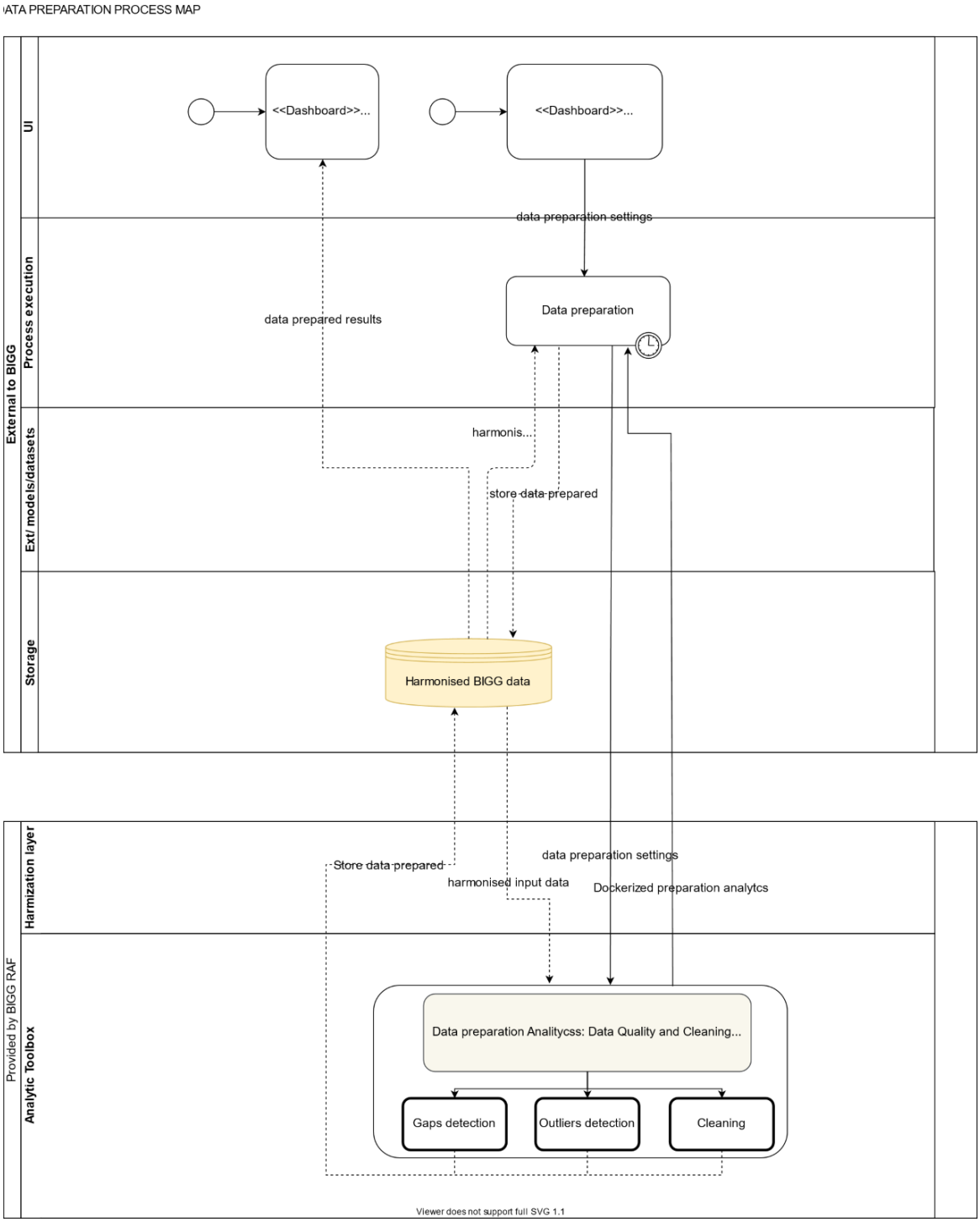
Exchange Requirement Data Objects

< Data mapping >

Type	Data Object
Name	*Data mapping*
Documentation	*Data mapping is the process of matching fields from one database to another. The data mappings (the matching fields from each raw data set to Bigg data model) contained in the BIGG system will be subjected to the ENMA harmonization process so that they are carried out on the data to be harmonized. This exchange is only necessary the first time or if there is any update in the data mapping.*

Process Map : <Data preparation>

The data preparation process is responsible for verifying the quality, purifying and transforming the input data so that it can be used without problems by the analytical processes.



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < UI>

In the user interface, the end user, depending on certain permissions, can manage, configure the execution of the data preparation process and can also explore results obtained from this process.

<Dashboard: Data preparation management> **[ID:< BC1-UC01-006>]**

Type	*Use task*	
Name	*Dashboard: Data preparation management*	
Documentation	*The end user, with administrator permissions, can from the UI manage and control the execution of preparation process. the user will be able to time the execution of the process, configure some parameters and control the execution of these processes (running, finalised, crashed, etc.).*	

<Dashboard: Data preparation visualization> **[ID:< BC1-UC01-007>]**

Type	*Use task*	
Name	*Dashboard: Data preparation visualization*	
Documentation	*In the user interface, the user can explore the results of the data preparation process, get the prepared data, see the differences between the raw data, and assess the quality of the data in the system.*	

LANE : < Processes execution>

<Data preparation > **[ID:< BC1-UC01-008>]**

Type	*Task type*	
Name	*Data preparation*	
Documentation	*This process is in charge of verifying the quality, purifying and transforming the previously harmonized data. It is a timed process, whose execution is controlled by the user managed from the UI. Once this process is applied, the data is ready to be processed by analytical processes. Depending on the need in each case, this process can be carried out in BIGG's own infrastructure or in ENMA's infrastructure (avoiding the movement of large amounts of data). If the process runs in ENMA, the analytics developed in the BIGG (dockerized) system will also be used.*	

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Bases>

Type	Data Object
Name	*Harmonized Data Bases*
Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format. The Harmonised databases are a combination of data marts (based on MongoDB: a document database, which means it stores data in JSON-like documents) and long-term DB (based on Apache HBASE: HBase is an open source non-relational distributed database)*

POOL : <Provided by BIGG >

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Analytics Toolbox>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies.

<Data preparation analytics > **[ID:< BC1-UC01-009>]**

Type	*Task type*	
Name	*Data preparation analytics*	
Documentation	*The data preparation process that is contained in the BIGG analysis toolbox can be executed in the BIGG infrastructure itself or in the ENMA infrastructure, depending on the needs of each case.* *This process retrieves the harmonized input data from the buildings and prepares it for further analysis.* *These processes are segmented by data type, for example, static data, time series of consumption or time series of meteorological data, since the processes can be different in each case.* *The processes could be:* *- data format* *- data transformation* *- detection of data gaps* *- outlier detection* *- data filling* *- data normalization*	

	- data modification *The results can vary from the data prepared itself to global performance indicators such as:* *- Buildings that have or have not passed the quality process* *- Buildings whose data has been corrected* *- Time series of consumption with gaps.* *- Meteorological data with gaps.* *- Static parameters with outliers* *- etc.*

Exchange Requirement Data Objects

< data prepared results >

Type	Data Object
Name	*Data prepared results*
Documentation	*The results of the analytical data preparation process are collected from the Harmonized DDBB and presented to users in the user interface.*

< Harmonised input data >

Type	Data Object
Name	*Harmonized input data*
Documentation	*The harmonized input data is extracted from the Harmonized DDBB and served to the data preparation execution process. Either in the BIGG or ENMA infrastructure.*

< Stored data prepared >

Type	Data Object
Name	*Stored data prepared*
Documentation	*The results of the data preparation process are stored back to the Harmonised DDBB.*

< Data preparation settings >

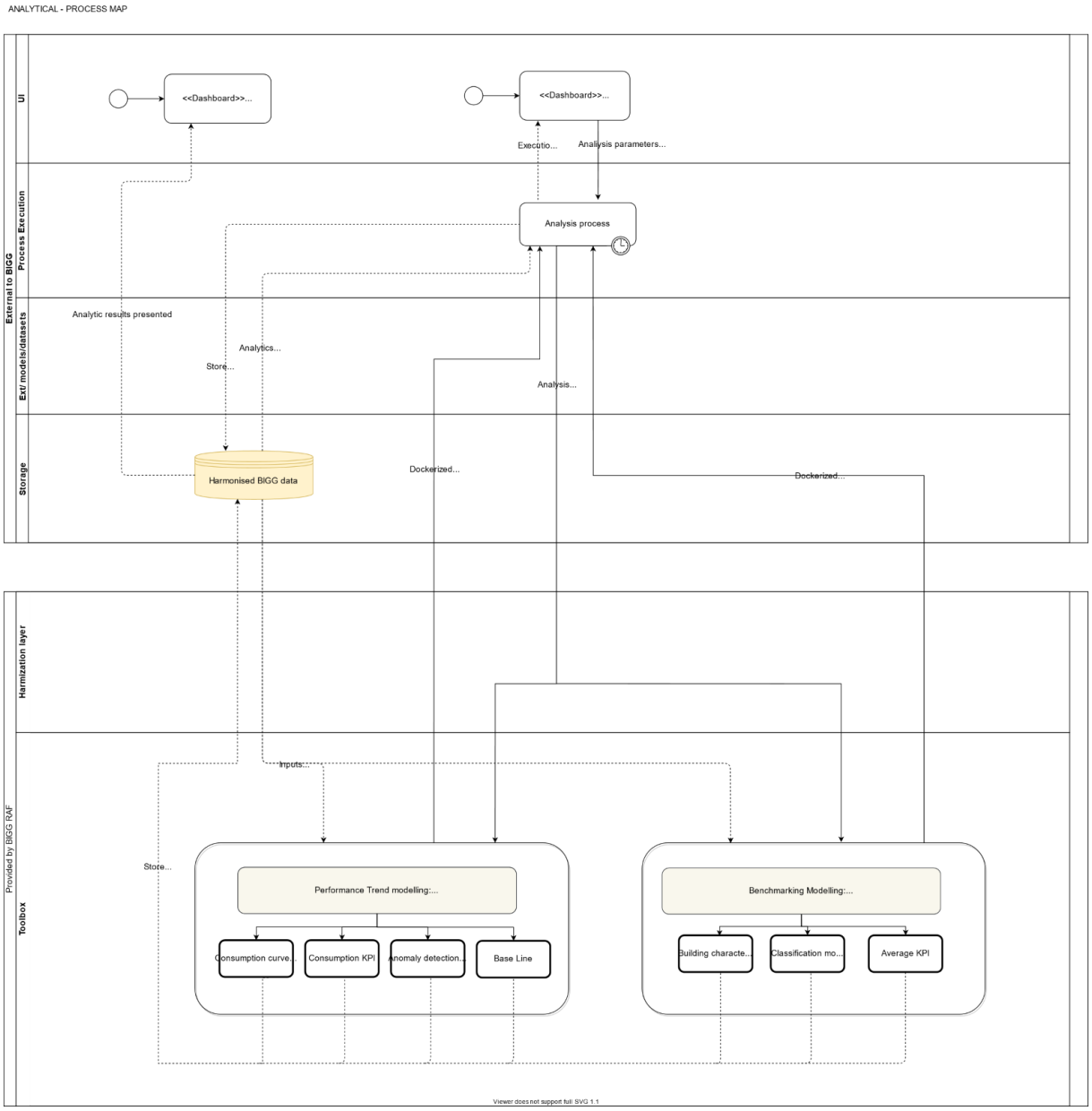
Type	Data Object
Name	*Data preparation settings*
Documentation	*The configuration parameters are passed to the process of running the data preparation process from the user interface.*

< Dockerized preparation data >

Type	Data Object
Name	*Dockerized preparation data*
Documentation	*In the event that the execution is carried out outside the BIGG infrastructure, the analysis tools, previously dockerized, will be imported to be deployed and executed in the external infrastructure.*

Process Map : <Analytical Process>

The analytical process is responsible for extracting the data, previously prepared, extracting the harmonized DDBB, obtaining the expected analytical results for the use case, saving them in the harmonized DDBB and presenting them to the end user.


ANALYTICAL - PROCESS MAP
External to BIGG
UI
Process Execution
Ext/ models/datasets
Storage
<<Dashboard>>...
<<Dashboard>>...
Executio...
Analysis parameters...
Analysis process
Analytic results presented
Analytics...
Store...
Analysis...
Harmonised BIGG data
Dockerized...
Dockerized...
Provided by BIGG RAF
Harmization layer
Toolbox
Inputs...
Store...
Performance Trend modelling:...
Consumption curve...
Consumption KPI
Anomaly detection...
Base Line
Benchmarking Modelling:...
Building characte...
Classification mo...
Average KPI
Viewer does not support full SVG 1.1

POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < UI>

In the user interface, the end user, depending on certain permissions, can manage, configure the execution of the data Analysis process and can also explore results obtained from this process.

<Dashboard: View/presentation of analytic results> **[ID:< BC1-UC01-010>]**

Type	*Use task*	
Name	*Dashboard: View/presentation of analytic results*	
Documentation	*The end user can explore the analytical results from the user interface. The presentation of these results will be simple and easily understandable. The user can easily filter the results to obtain the visualizations that he needs at all times.*	

<Dashboard: Analysis management> **[ID:< BC1-UC01-011>]**

Type	*Use task*	
Name	*Dashboard: Analysis management*	
Documentation	*The end user, with administrator permissions, can from the UI manage and control the execution of Analytics process. The user will be able to time the execution of the process, configure some parameters and control the execution of these processes (running, finalised, crashed, etc.).*	

LANE : < Processes execution>

<Analysis process > **[ID:< BC1-UC01-012>]**

Type	*Task type*	
Name	*Analysis process*	
Documentation	*The analysis process is responsible for managing and executing the analysis processes on the prepared data, based on the parameterization established by the user in the UI.* *For these use cases, the Performance Trend and Benchmarking processes are launched. The processes can be run on the BIGG infrastructure or on the local platform. If they are launched in the local infrastructure, the analytics developed in BIGG are used, which are imported dockerized to be deployed and executed.*	

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Bases>

Type	Data Object
Name	*Harmonized Data Bases*
Documentation	*In the harmonized databases are the data ready to be used by the analytical processes. The results of the analytical processes will also be stored in these databases. In the case of the Benchmarking analytical process, some of the results of the Performance Trend analytical process are used as inputs, which are also available from the DDBB.* *The Harmonised databases are a combination of data marts (based on MongoDB: a document database, which means it stores data in JSON-like documents) and long-term DB (based on Apache HBASE: HBase is an open source non-relational distributed database)*

POOL : <Provided by BIGG >

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Analytics Toolbox>

Analytics Toolbox contains the analysis developed in the BIGG project for this use case, the analysis are:

- Performance Trend modelling (history of each building)

- Benchmarking modelling (comparison with similar buildings)

Depending on the needs and volume of processing, this lane can be used for run analysis or simply to store the coupled analysis so they can be deployed and run on other infrastructures.

<Performance Trend modelling > [ID:< BC1-UC01-013>]

Type	*Task type*	
Name	*Performance Trend modelling*	
Documentation	*In Performance Trend modelling, the behaviour of each of the buildings is analysed according to its consumption history and its evolution.* *Running these analyses can be done on BIGG's own infrastructure or on the existing ENMA platform.* *The input data for this model is the prepared data that is stored in the harmonized databases.* *This modelling process can be divided into the following threads:* *- Clustering of historical consumption curves* *- Obtaining the KPIs of consumption. (consumption from Monday to Friday, base load consumption, dependence on heating and cooling consumption, etc.)* *- Creation of best baseline models.*	

	- Simulation of baseline *-Trend evaluation* *-Anomalies detection in the consumption performance profile.*

<Benchmarking modelling > [ID:< BC1-UC01-014>]

Type	*Task type*	
Name	*Benchmarking modelling*	
Documentation	*In Benchmarking modelling, the behaviour of buildings is analysed in comparison to similar buildings.* *Running these analyses can be done on BIGG's own infrastructure or on the existing ENMA platform.* *The input data for this model are the data from the data preparation process and the results of the Benchmarking modelling, both stored in the harmonized databases.* *This modelling process can be divided into the following threads:* *- Grouping of building characteristics (by consumption KPI, building characteristics)* *- Classification modelling* *- Building qualification* *- Obtaining average KPI's.*	

Exchange Requirement Data Objects

< Analytic results presented>

Type	Data Object
Name	*Analytic results presented*
Documentation	*The results of the analysis, stored in the harmonized DDBBs, will be extracted and presented to the end users.*

< Execution summary >

Type	Data Object
Name	*Execution summary*
Documentation	*After the analytic has launched, the end user receives the processing summary in the user interface. (Successfully executed processes, unfinished processes, crashed processes, etc.).*

< Analytics Inputs >

Type	Data Object
Name	*Analytics Inputs*
Documentation	*The necessary input data is sent to the analysis process from the harmonized databases.* *For Performance Trend modelling, these data are:* *- Historical energy consumption (time series)* *- Building characteristics* *- Weather data* *For Benchmarking modelling, these data are:* *- Historical energy consumption (time series)* *- Building characteristics* *- Weather data* *- KPI of consumption*

< Stored Analysis results >

Type	Data Object
Name	*Stored Analysis results*
Documentation	*The results of the analytical processes are stored in the harmonized databases.* *For Performance Trend modelling, the main results are:* *- Segmentation of consumption* *- KPI's of consumption* *- Periods with abnormal consumption* *- Best baseline model* *For Benchmarking modelling, the main results are:* *- Qualification of buildings.* *- Assignment of similarity group.* *- Average KPIs*

< Analysis parameters settings >

Type	Data Object
Name	*Analysis parameters settings*
Documentation	*From the UI the end user can modify some parameters that are used in the analysis process.*

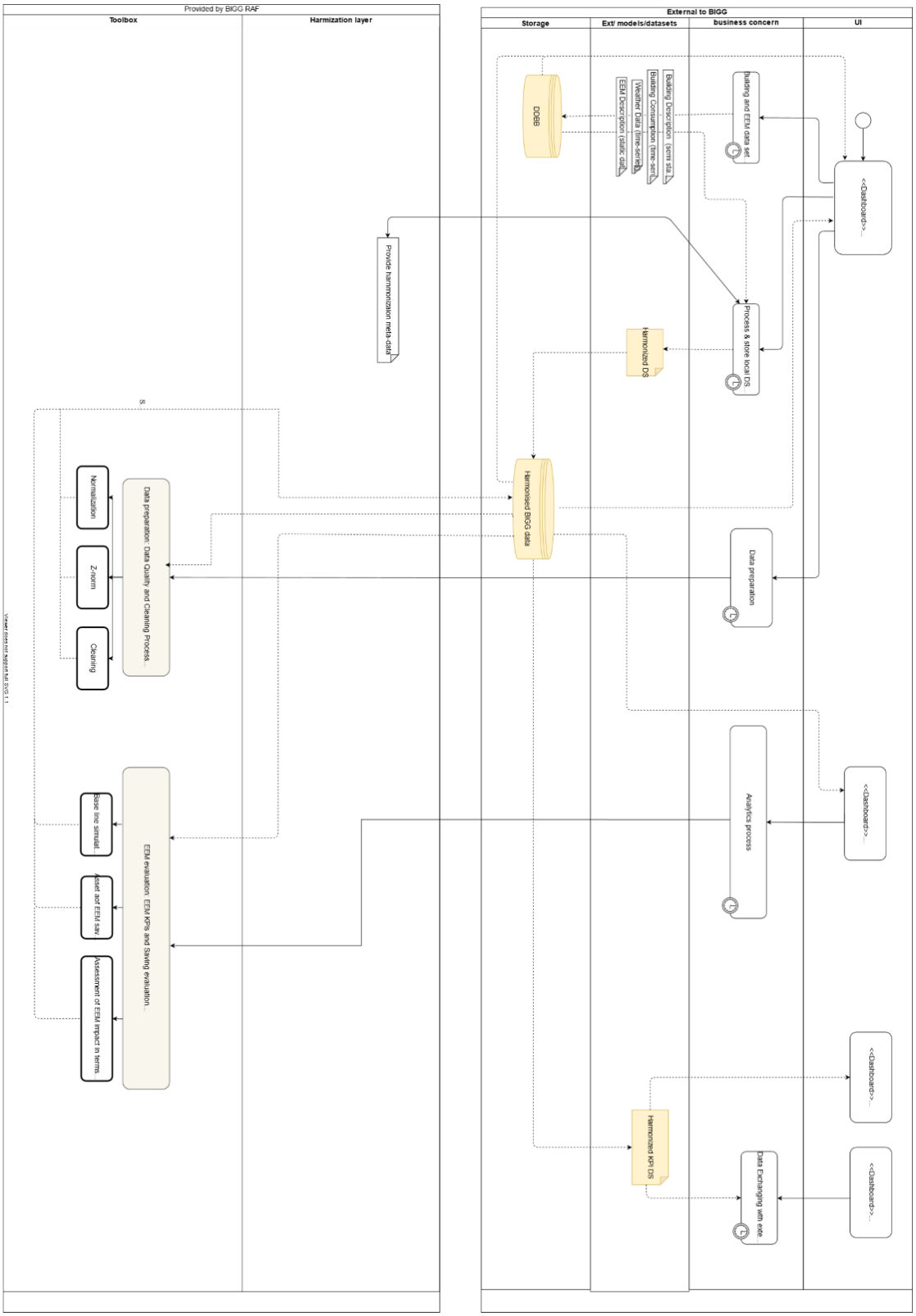
	For Performance Trend modelling, these can be (among others): *- Selection of the transition period and verification.* *- Selection of excluded periods.* *For Benchmarking modelling they can be (among others):* *- Pre-selection of buildings to process* *- Set some parameters (type of use, year of construction, etc.)*

< Dockerized Analytics >

Type	Data Object
Name	*Dockerized Analytics*
Documentation	*In the event that the execution is carried out outside the BIGG infrastructure, the analysis tools, previously dockerized, will be imported to be deployed and executed in the external infrastructure.*

V.2. UC2

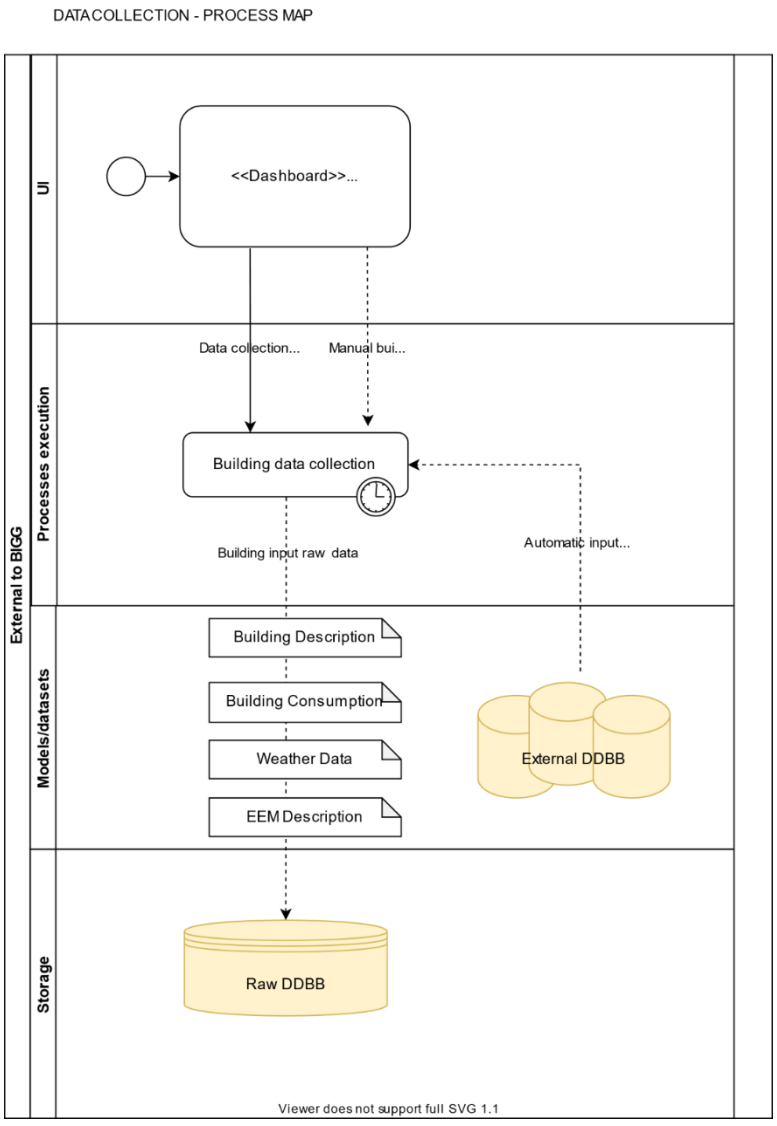
V.2.1. BPMN diagram



V.2.2. Specification of processes

Process Map : <Data Collection>

The data required for the UC1 can come from different data sources. These data sources can be loaded from existing databases or uploaded manually by the user.
In the case of manual data upload, the end user will be able to load, modify or delete data from the UI, whether this is static (descriptive characteristics of the building) or time series (energy consumption).
In the case of data from existing databases, these can be collected directly through APIs or similar, the administrator user will have access from the UI to the import control of this data.
The building characteristics data, energy consumption or climate data will be systematically stored in raw in a Database. This Database will be located in the infrastructure of the existing ENMA application (owned by CIMNE).



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : <UI>

In the UI the end user, depending on certain permissions, can upload, modify or delete the building data manually (through web forms or by uploading CSV or Excel files, with predefined format), or configure the collection of data from databases. existing data, external to this Pool

<Dashboard: Management, loading, updating and configuration of input data.> **[ID: <BC1-UC2-001>]**

Type	*User task*	
Name	*Dashboard: Management, loading, updating and configuration of input data*	
Documentation	*In the UI the end user, depending on certain permissions, can upload, modify or delete the building data manually (through web forms or by uploading CSV or Excel files, with predefined format), or configure the collection of data from databases. existing data, external to this Pool*	

LANE : < Processes execution>

<Building Data Collection> **[ID:<BC1-UC2-002>]**

Type	*Task type*	
Name	*Building data collection*	
Documentation	*The building data collection process is responsible for managing the data entry into the system.* *For manual data, the new data will be loaded and the modification or deletion of the data will be managed, according to the user's request.* *Depending on the configuration assigned by the user, timed calls will be launched to the different APIs to automatically retrieve data from external databases.*	

Exchange Requirement Data Objects

<Manual Input Raw Building data>

Type	Data Object
Name	*Manual raw building data*
Documentation	*Data uploaded, modified or deleted manually by the user in the user interface will be passed to the data collection process.*

<Automatic input Raw Building data>

Type	Data Object
Name	*Automatic raw building data*
Documentation	*The existing data in the different external databases will be processed by the data collection process. This data can be just new data or all data depending on the configuration that the user wants.*

LANE : < Models/Data Sets>

Library Data Objects

<Building Description >

Type	Data Object
Name	*Building Description*
Documentation	*The building description data object is semi-static data (unique values updated at each change of state (medium-long term)). This data can come from different sources which can be different external databases, documents or user interface forms. Depending on the source of the data, it will have different data models. These data describe the physical or typological characteristics of each of the buildings.* *Basic data summary:* *User identifier, building name, address, state, province, municipality, postal code, type of use, construction characteristics, construction area, number of floors, cadastral reference, etc.* *Installation data: cooling source, heating source, lighting types, etc.* *Cadastral data: all the cadastral information available.*

<Building Consumption >

Type	Data Object
Name	*Building consumption*
Documentation	*Building energy, electricity and gas consumption data can come from different sources (utility companies, energy information and management services, control systems, billing or manuals). In addition, these can have a heterogeneous aggregation, and can be hourly, monthly or both. Depending on the source of the data, it will have different data models.*

<Weather data >

Type	Data Object
Name	*Weather data*
Documentation	*The system will continuously retrieve data from meteorological stations from different stations located throughout Europe, with mostly hourly aggregation. The data collected can be: - Temperature, relative humidity, direct solar radiation in a horizontal plane, diffuse solar radiation, wind speed, wind direction, precipitation, etc.*

<EEM Description >

Type	Data Object
Name	*EEM Description*
Documentation	*The system will continuously retrieve from the user interface with a web form or a model document (csv or excel) the energy efficiency measures implemented in the buildings. The information for each EEM could be: - Name of the EEM, description, application date, economic investment, etc.*

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Raw Data Base>

Type	Data Object
Name	*Raw Data Base*
Documentation	*Raw data is systematically stored in ENMA databases. The model of each of the data will be defined according to each of the source providers. The database is Apache HBASE (HBase is an open source non-relational distributed database)*

Exchange Requirement Data Objects

< Building input raw data>

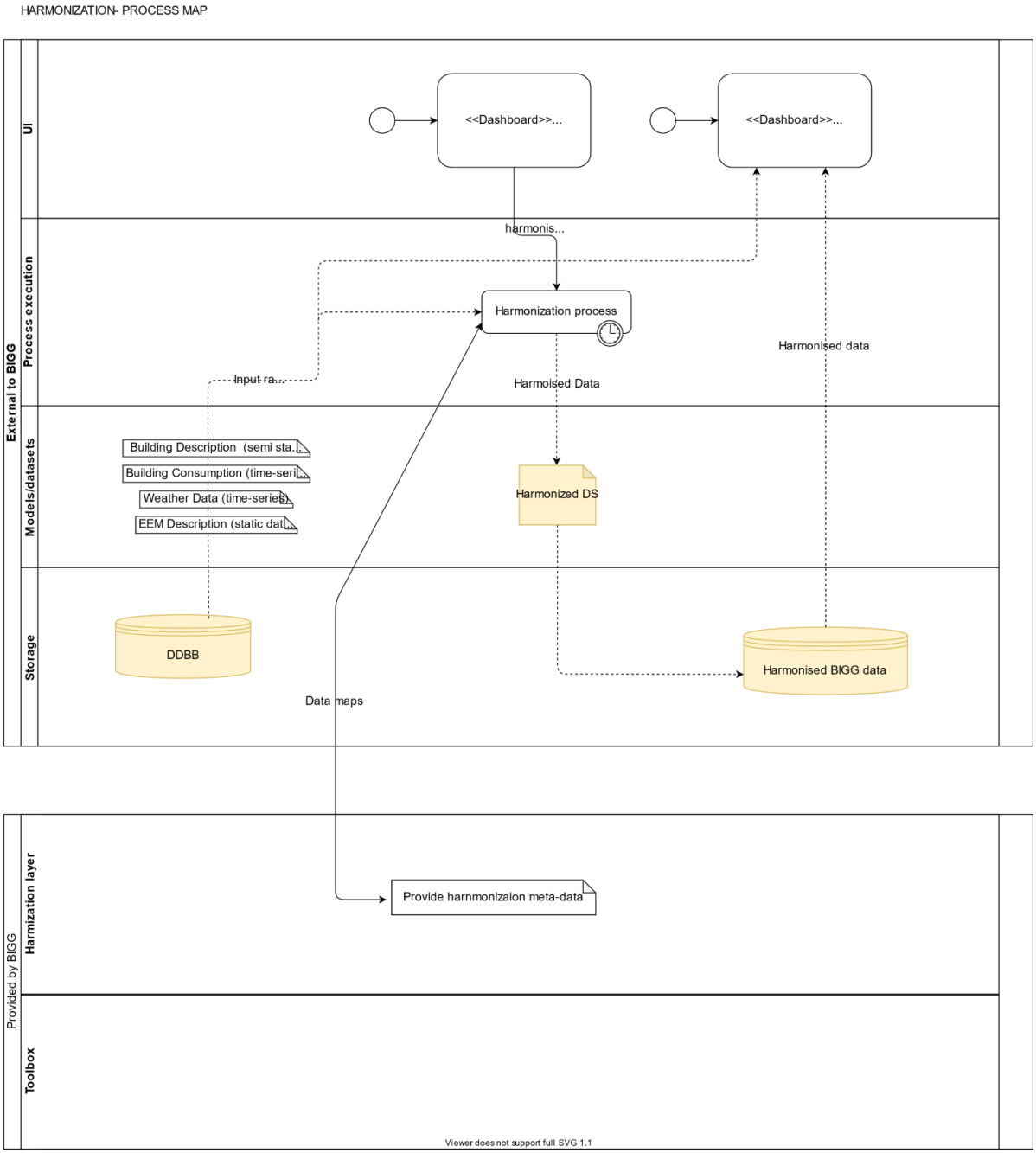
Type	Data Object
Name	*Manual raw building data*
Documentation	*All data from the data collection process is stored in the Raw Database.*

A process map will also have a graphical version in the BPMN notation. This should be inserted before the pool and lane descriptions

Page Break

Process Map : <Harmonization>

The harmonization process map is the process in charge of standardizing the raw data with the format of the BIGG data model. This standardization allows, on the one hand, to use BIGG's analytical developments and at the same time facilitates the exchange of data with other external systems or applications.



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Base>

Type	Data Object
Name	*Harmonized Data Base*
Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model*

	format. The database is Apache HBASE (HBase is an open source non-relational distributed database)

Exchange Requirement Data Objects
<Raw Building data>

Type	Data Object
Name	*Raw building data*
Documentation	*Raw data is extracted from raw databases and sent to the data harmonization process when required.*

<Harmonized Building data>

Type	Data Object
Name	*Harmonized building data*
Documentation	*The harmonized data is extracted from the data harmonization process and is systematically stored in the harmonized database.*

LANE : < Processes execution>

<Harmonization process > [ID:<BC1-UC2-003>]

Type	*Task type*	
Name	*Harmonization process*	
Documentation	*The harmonization process is responsible for transforming the raw data into harmonized data in BIGG format. This process begins on a recurring basis and is controlled and managed by the administrator user from the user interface. This process that runs is as follows:* *- Retrieves raw input data that has not been harmonized so far, from the raw database.* *- Retrieves the raw data maps made in BIGG.* *- Assign the raw data to the BIGG data model.* *- Save the harmonized input data in the harmonized database.*	

LANE : < Models/Data Sets>

Library Data Objects
<Building Description >

Type	Data Object
Name	*Building Description*
Documentation	*Same data and description as in the data collection process.*

<Building Consumption >

Type	Data Object
Name	*Building consumption*
Documentation	*Same data and description as in the data collection process.*

<Weather data >

Type	Data Object
Name	*Weather data*
Documentation	*Same data and description as in the data collection process.*

<EEM Description >

Type	Data Object

Name	*EEM Description*
Documentation	*Same data and description as in the data collection process.*

<Harmonized Datasets >

Type	Data Object
Name	*Harmonized Datasets*
Documentation	*The raw data (Building description, building consumption and Weather data and EEM description) harmonized and standardized on BIGG data model*

LANE : <UI>

In the user interface, the end user, depending on certain permissions, can manage and configure the execution of the harmonization process, and can explore both raw and harmonized data.

<Dashboard: Harmonization process management> **[ID:<BC1-UC2-004>]**

Type	*User task*	
Name	*Dashboard: Harmonization process management*	
Documentation	*From the user interface, the administrator user can manage and control the harmonization process. You can configure and modify certain process parameters (execution times, mapping updates, etc.)*	

<Dashboard: Exploring raw and harmonized data> **[ID:<BC1-UC2-005>]**

Type	*User task*	
Name	*Dashboard: Exploring raw and harmonized data*	
Documentation	*From the user interface, the user can view and explore the raw data that the system has ingested, as well as the result of the harmonization process. By controlling how much data has actually been harmonized, harmonization issues, even once the data has been normalized, the user can control and manage the data allocation in each building.*	

POOL : <Provided by BIGG >

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Harmonization layer>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies.

Library Data Objects
<Provide harmonization meta-data >

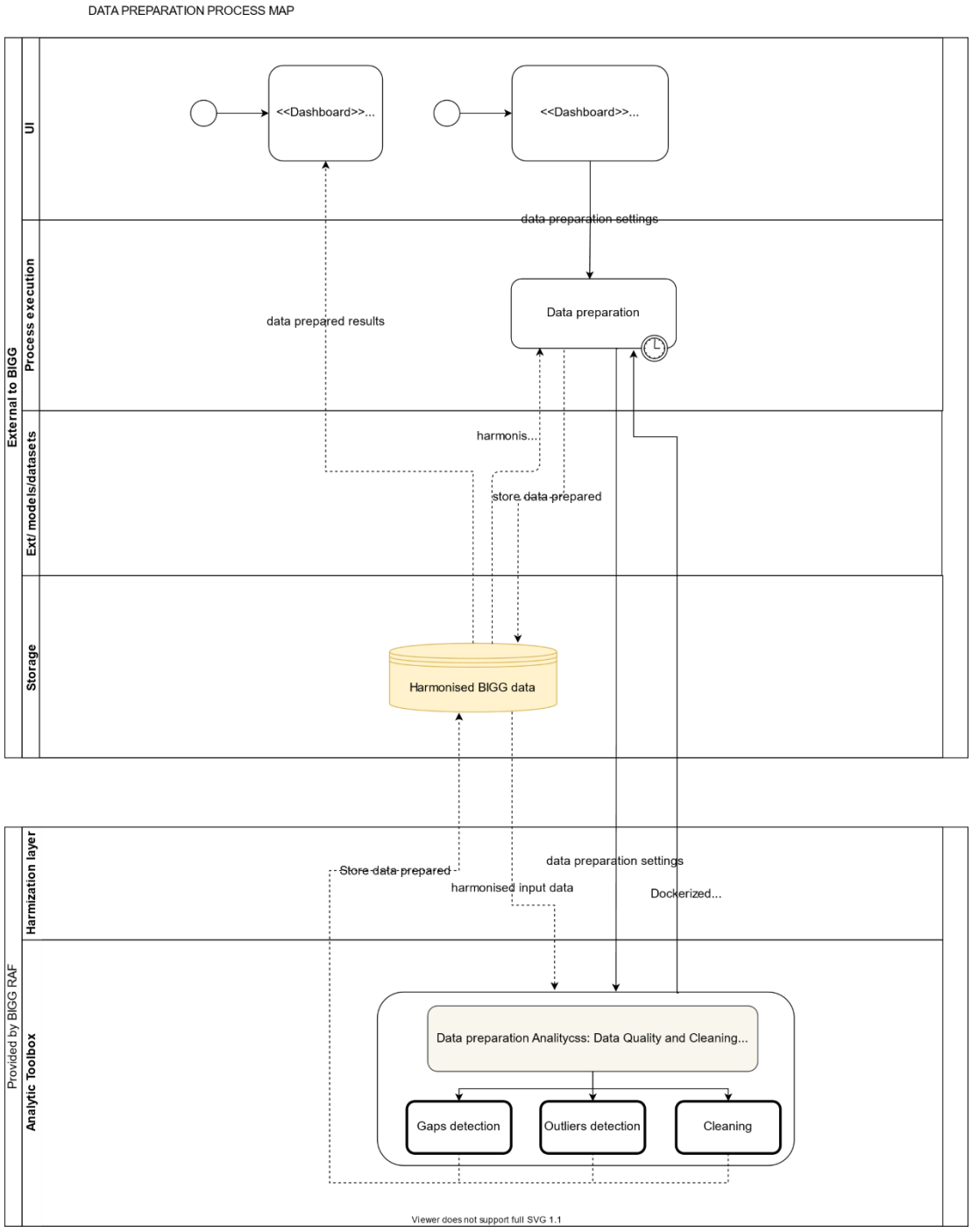
Type	Data object	
Name	*Provide harmonization meta-data*	
Documentation	*BIGG's harmonization layer has the raw data mappings developed in the project. These data mappings will go through the harmonization process of the ENMA system so that you can execute them and harmonize the data. For this use case, special attention will be paid to alignment with the schema and taxonomy used in EFFIG-DEEP to harmonize EEMs.*	

Exchange Requirement Data Objects
< Data mapping >

Type	Data Object
Name	*Data mapping*
Documentation	*The data mappings contained in the BIGG system will be subjected to the ENMA harmonization process so that they are carried out on the data to be harmonized.*

Process Map : <Data preparation>

The data preparation process is responsible for verifying the quality, purifying and transforming the input data so that it can be used without problems by the analytical processes.



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < UI>

In the user interface, the end user, depending on certain permissions, can manage, configure the execution of the data preparation process and can also explore results obtained from this process.

<Dashboard: Data preparation management> **[ID:< BC1-UC2-006>]**

Type	*User task*	
Name	*Dashboard: Data preparation management*	
Documentation	*The end user, with administrator permissions, can from the UI manage and control the execution of preparation process. the user will be able to time the execution of the process, configure some parameters and control the execution of these processes (running, finalised, crashed, etc.).*	

<Dashboard: Data preparation visualization> **[ID:< BC1-UC2-007>]**

Type	*User task*	
Name	*Dashboard: Data preparation visualization*	
Documentation	*In the user interface, the user can explore the results of the data preparation process, get the prepared data, see the differences between the raw data, and assess the quality of the data in the system.*	

LANE : < Processes execution>

<Data preparation > **[ID:< BC1-UC2-008>]**

Type	*Task type*	
Name	*Data preparation*	
Documentation	*This process is in charge of verifying the quality, purifying and transforming the previously harmonized data. It is a timed process, whose execution is controlled by the user managed from the UI. Once this process is applied, the data is ready to be processed by analytical processes. Depending on the need in each case, this process can be carried out in BIGG's own infrastructure or in ENMA's infrastructure (avoiding the movement of large amounts of data). If the process runs in ENMA, the analytics developed in the BIGG (dockerized) system will also be used.*	

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Bases>

Type	Data Object
Name	*Harmonized Data Bases*
Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format. The Harmonised databases are a combination of data marts (based on MongoDB: a document database, which means it stores data in JSON-like documents) and long-term DB (based on Apache HBASE: HBase is an open source non-relational distributed database)*

POOL : <Provided by BIGG >

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Analytics Toolbox>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies.

<Data preparation analytics > [ID:< BC1-UC2-009>]

Type	*Task type*	
Name	*Data preparation analytics*	
Documentation	*The data preparation process that is contained in the BIGG analysis toolbox can be executed in the BIGG infrastructure itself or in the ENMA infrastructure, depending on the needs of each case.* *This process retrieves the harmonized input data from the buildings and prepares it for further analysis.* *These processes are segmented by data type, for example, static data, time series of consumption or time series of meteorological data, since the processes can be different in each case.* *The processes could be:* *- data format* *- data transformation* *- detection of data gaps* *- outlier detection* *- data filling* *- data normalization* *- data modification* *The results can vary from the data prepared itself to global performance indicators such as:* *- Buildings that have or have not passed the quality process* *- Buildings whose data has been corrected* *- Time series of consumption with gaps.* *- Meteorological data with gaps.* *- Building and EEM static parameters with outliers* *- etc...*	

Exchange Requirement Data Objects

< data prepared results >

Type	Data Object
Name	*Data prepared results*
Documentation	*The results of the analytical data preparation process are collected from the Harmonized DDBB and presented to users in the user interface.*

< Harmonised input data >

Type	Data Object
Name	*Harmonized input data*
Documentation	*The harmonized input data is extracted from the Harmonized DDBB and served to the data preparation execution process. Either in the BIGG or ENMA infrastructure.*

< Stored data prepared >

Type	Data Object
Name	*Stored data prepared*
Documentation	*The results of the data preparation process are stored back to the Harmonised DDBB.*

< Data preparation settings >

Type	Data Object
Name	*Data preparation settings*

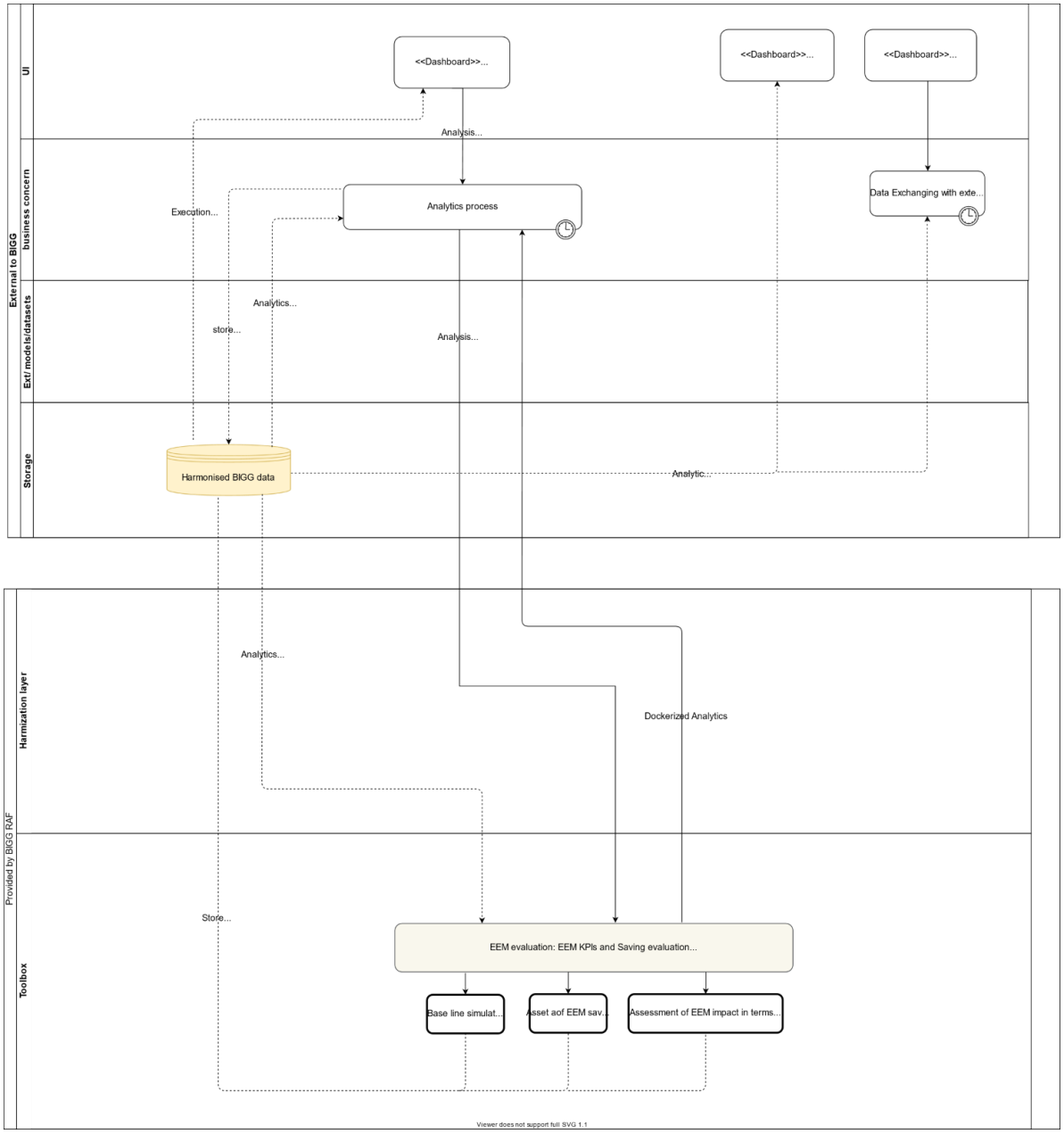
Documentation	*The configuration parameters are passed to the process of running the data preparation process from the user interface.*

< Dockerized preparation analytics >

Type	Data Object
Name	*Dockerized preparation analytics*
Documentation	*In the event that the execution is carried out outside the BIGG infrastructure, the analysis tools, previously dockerized, will be imported to be deployed and executed in the external infrastructure.*

Process Map : <Analytical Process>

The analytical process is responsible for extracting the data, previously prepared, extracting the harmonized DDBB, obtaining the expected analytical results for the use case, saving them in the harmonized DDBB and presenting them to the end user.



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < UI>

In the user interface, the end user, depending on certain permissions, can manage, configure the execution of the data Analysis process and can also explore results obtained from this process.

<Dashboard: View/presentation of analytic results> **[ID:< BC1-UC2-010>]**

Type	*User task*	
Name	*Dashboard: View/presentation of analytic results*	
Documentation	*The end user can explore the analytical results from the user interface. The presentation of these results will be simple and easily understandable. The user can easily filter the results to obtain the visualizations that he needs at all times.*	

<Dashboard: Analysis management> **[ID:< BC1-UC2-011>]**

Type	*User task*	
Name	*Dashboard: Analysis management*	
Documentation	*The end user, with administrator permissions, can from the UI manage and control the execution of Analytics process. The user will be able to time the execution of the process, configure some parameters and control the execution of these processes (running, finalised, crashed, etc.).*	

LANE : < Processes execution>

<Analysis process > **[ID:< BC1-UC2-012>]**

Type	*Task type*	
Name	*Analysis process*	
Documentation	*The analysis process is responsible for managing and executing the analysis processes on the prepared data, based on the parameterization established by the user in the UI.* *The processes can be run on the BIGG infrastructure or on the local platform. If they are launched in the local infrastructure, the analytics developed in BIGG are used, which are imported dockerized to be deployed and executed.*	

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Bases>

Type	Data Object
Name	*Harmonized Data Bases*
Documentation	*In the harmonized databases are the data ready to be used by the analytical processes. The results of the analytical processes will also be stored in these databases.* *The Harmonised databases are a combination of data marts (based on MongoDB: a document database, which means it stores data in JSON-like documents) and long-term DB (based on Apache HBASE: HBase is an open source non-relational distributed database)*

POOL : <Provided by BIGG >

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Analytics Toolbox>

Analytics Toolbox contains the analysis developed in the BIGG project for this use case, the analysis are:
- EEM evaluation (EEM KPIs and Savings evaluation)
Depending on the needs and volume of processing, this lane can be used for run analysis or simply to store the coupled analysis so they can be deployed and run on other infrastructures.

<EEM evaluation > [ID:< BC1-UC2-013>]

Type	*Task type*	
Name	*EEM evaluation*	
Documentation	*In EEM evaluation, the behaviour of each of the buildings is analysed according to its consumption history and its evolution.* *The execution of these analyses can be done on BIGG's own infrastructure or on the existing ENMA platform.* *The input data for this model is the prepared data that is stored in the harmonized databases.* *This modelling process can be divided into the following threads:* • *Steady state features transformation* • *Dynamic features transformation* • *Cross-validation data partitioning* • *Parameters optimization* *The expected results are:* • *Best Base line model for each period* • *Asses the energy savings of each EEM* • *Assessment of EEMs impact in terms of variation in optimal parameters and energy savings*	

Exchange Requirement Data Objects

< Analytic results presented>

Type	Data Object
Name	*Analytic results presented*
Documentation	*The results of the analysis, stored in the harmonized DDBBs, will be extracted and presented to the end users.*

< Execution summary >

Type	Data Object
Name	*Execution summary*
Documentation	*After the analytic has launched, the end user receives the processing summary in the user interface. (Successfully executed processes, unfinished processes, crashed processes, etc.).*

< Analytics Inputs >

Type	Data Object
Name	*Analytics Inputs*
Documentation	*The necessary input data is sent to the analysis process from the harmonized databases.* *For EEM evaluation the necessary inputs are:* *- Historical energy consumption (time series)* *- Building characteristics* *- Weather data* *- EEM description parameters*

< Stored Analysis results >

Type	Data Object
Name	*Stored Analysis results*
Documentation	*The results of the analytical processes are stored in the harmonized databases.* *For EEM evaluation the main results are:* • *Best Base line models* • *Assess the energy savings of each EEM* • *Assessment of EEM's impact in terms of variation in optimal parameters and energy savings*

< Analysis parameters settings >

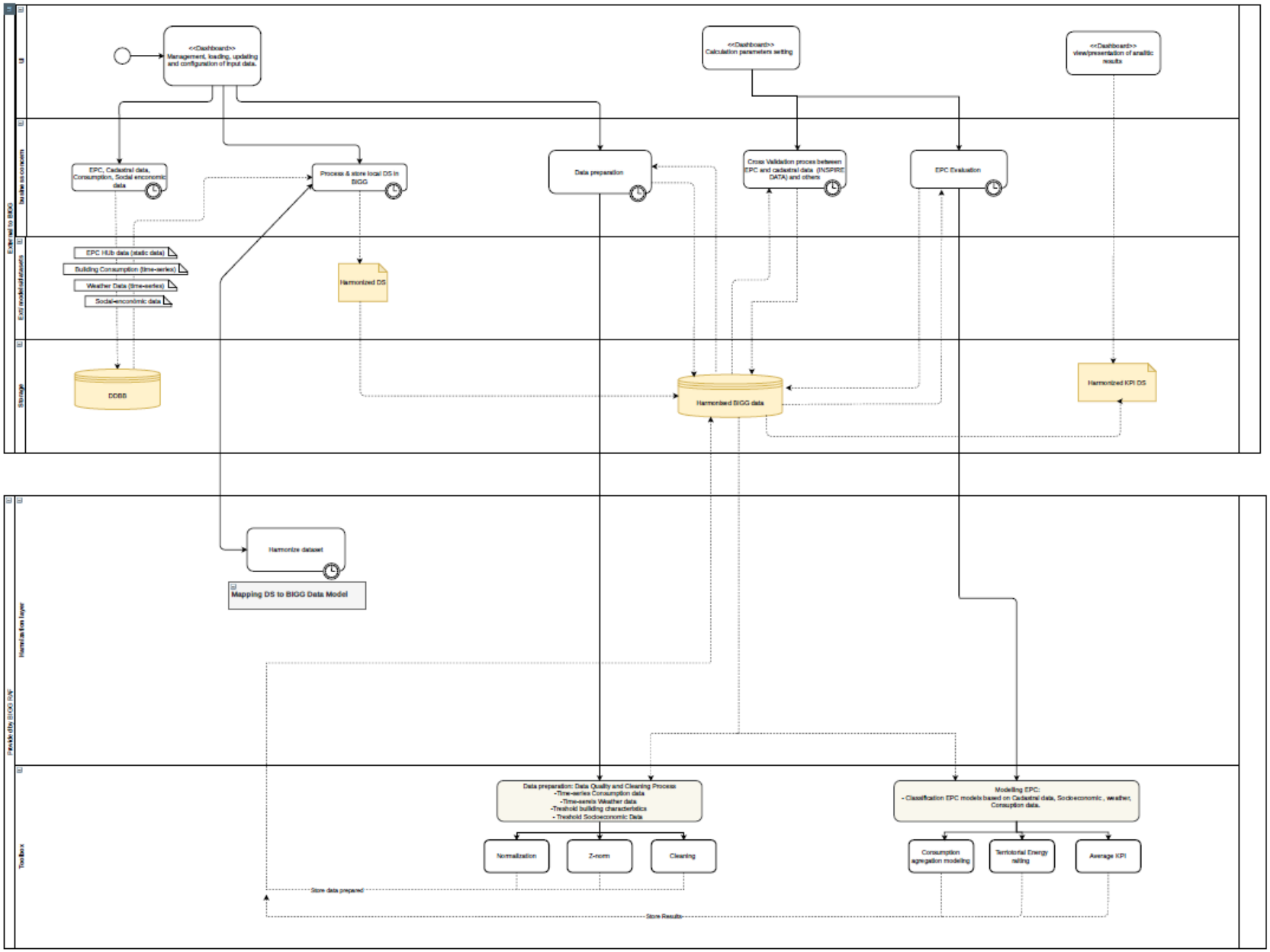
Type	Data Object
Name	*Analysis parameters settings*
Documentation	*From the UI the end user can modify some parameters that are used in the analysis process.* *For EEM evaluation, these can be (among others):* *- Selection of the transition period and verification.* *- Selection of excluded periods.*

< Dockerized Analytics >

Type	Data Object
Name	*Dockerized Analytics*
Documentation	*In the event that the execution is carried out outside the BIGG infrastructure, the analysis tools, previously dockerized, will be imported to be deployed and executed in the external infrastructure.*

V.3. UC3

V.3.1. BPMN diagram


External to BIGG
UI
Backend environment
Exist models/datasets
Storage
<<Dashboard>>
Management, loading, updating and configuration of input data.
<<Dashboard>>
Calculation parameters setting
<<Dashboard>>
view/presentation of analitic results
EPC, Cadastral data, Consumption, Social enconomic data
Process & store local DS in BIGG
Data preparation
Cross Validation proces between EPC and cadastral data (INSPIRE DATA) and others
EPC Evaluation
EPC HUb data (static data)
Building Consumption (time-series)
Weather Data (time-series)
Social-enconomic data
Harmonized DS
DDBB
Harmonized BIGG data
Harmonized KPI DS
Flexible city BIGG App
Harmonization layer
Toolbox
Harmonize dataset
Mapping DS to BIGG Data Model
Data preparation: Data Quality and Cleaning Process
-Time-series Consumption data
-Time-series Weather data
-Treshold building characteristics
- Treshold Socioeconomic Data
Normalization
Z-norm
Cleaning
Modelling EPC:
- Classification EPC models based on Cadastral data, Socioeconomic , weather, Consuption data.
Consumption agregation modeling
Territorial Energy rating
Average KPI
Store data prepared
Store Results

V.3.2. Specification of Processes

*To develop the "*Integration of INSPIRE spatial data with Energy Performance Certification (EPC)*" process, it has been divided into the following process maps in order to deepen its description.*

The breakdown of the process maps presented in this document are:

- Data collection

- Data harmonization

- Data preparation

- Analytical process

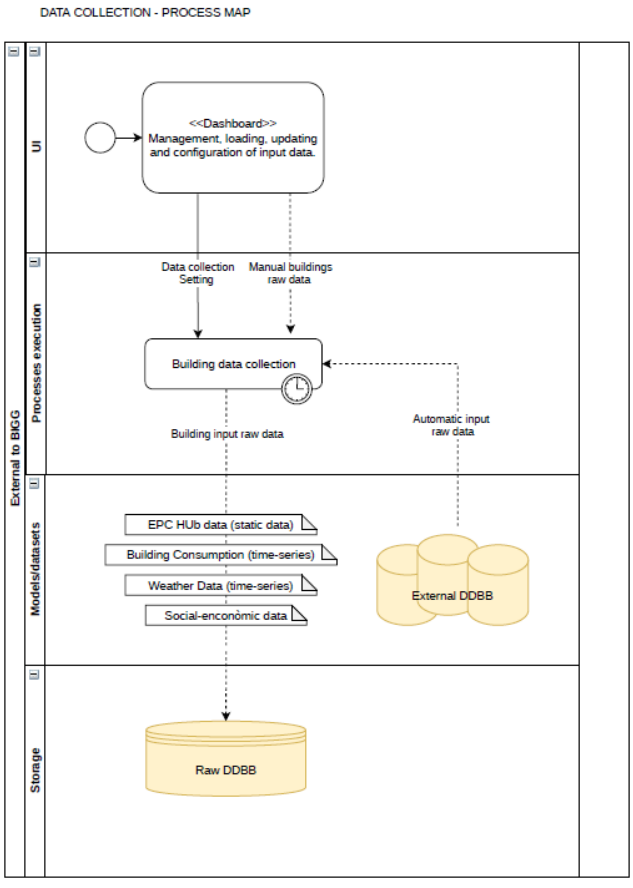
Process Map : <Data Collection>

The data required for the UC3 can come from different data sources. These data sources can be loaded from existing databases or uploaded manually by the user.

In the case of manual data upload, the end user will be able to load, modify or delete data from the UI.

In the case of data from existing databases, these can be collected directly through APIs or similar, the administrator user will have access from the UI to the import control of this data.

The EPC data, Weather, cadastral data will be systematically stored in raw in a Database. This Database will be located in the infrastructure of the existing ENMA application (owned by CIMNE).



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : <UI>

In the UI the end user, depending on certain permissions, can upload, modify or delete the building data manually (through web forms or by uploading CSV or Excel files, with predefined format), or configure the collection of data from databases. existing data, external to this Pool

<Dashboard: Management, loading, updating and configuration of input data.> [ID:<xx>]

Type	<type> *Use task*	
Name	*Dashboard: Management, loading, updating and configuration of input data*	
Documentation	*In the UI the end user, depending on certain permissions, can upload, modify or delete the building data manually (through web forms or by uploading CSV or Excel files, with predefined format), or configure the collection of data from databases. existing data, external to this Pool*	

LANE : < Processes execution>

<Building Data Collection> [ID:<BC2-UC03-001>]

Type	*Task type*	
Name	*Building data collection*	
Documentation	*The building data collection process is responsible for managing the data entry into the system.* *For manual data, the new data will be loaded and the modification or deletion of the data will be managed, according to the user's request.* *Depending on the configuration assigned by the user, timed calls will be launched to the different APIs to automatically retrieve data from external databases.*	

Exchange Requirement Data Objects

<Manual Input Raw Building data>

Type	Data Object
Name	*Manual raw building data*
Documentation	*Data uploaded, modified or deleted manually by the user in the user interface will be passed to the data collection process.*

<Automatic input Raw Building data>

Type	Data Object
Name	*Automatic raw building data*
Documentation	*The existing data in the different external databases will be processed by the data collection process. This data can be just new data or all data depending on the configuration that the user wants.*

LANE : < Models/Data Sets>

Library Data Objects

<EPC Hub data>

Type	Data Object
Name	<data object name> *EPC Hub data*
Documentation	<insert description of the data object> *All building Energy Performance Certificates are available on Open Data services of Catalan government.* *EPC hub data summary:* *Building ubication certification.* *Cadastral reference* Building description values (facade transmittance, windows transmittance, Electric Vehicle, thermal solar system, photovoltaic solar system, etc…) Primary energy (general, by sources) Qualification of primary energy consumption (general, by sources) Heating and cooling energy demand Etc…

<Building Consumption >

Type	Data Object
Name	<data object name> *Building consumption*
Documentation	<insert description of the data object> *Building, electricity energy consumption data can come from utility companies (Datadis) aggregated by Postal code. This data can be*

	obtained, as average, for use type (residential, industrial, access tariff) and these time series can be in hourly base.

<Weather data >

Type	Data Object
Name	<data object name> *Weather data*
Documentation	<insert description of the data object> *The system will continuously retrieve data from meteorological stations from different stations located throughout Europe, with mostly hourly aggregation.* *The data collected can be:* *- Temperature, relative humidity, direct solar radiation in a horizontal plane, diffuse solar radiation, wind speed, wind direction, precipitation, etc.*

<Socio-economic data >

Type	Data Object
Name	<data object name> *Socio-economic data*
Documentation	<insert description of the data object> *Economic rent for person average/aggregated by postal code will be collected from INE (Instituto Nacional de Estadistica).*

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Raw Data Base>

Type	Data Object
Name	<data object name> *Raw Data Base*
Documentation	<insert description of the data object> *Raw data is systematically stored in ENMA databases. The model of each of the data will be defined according to each of the source providers. The database is Apache HBASE (HBase is an open source non-relational distributed database)*

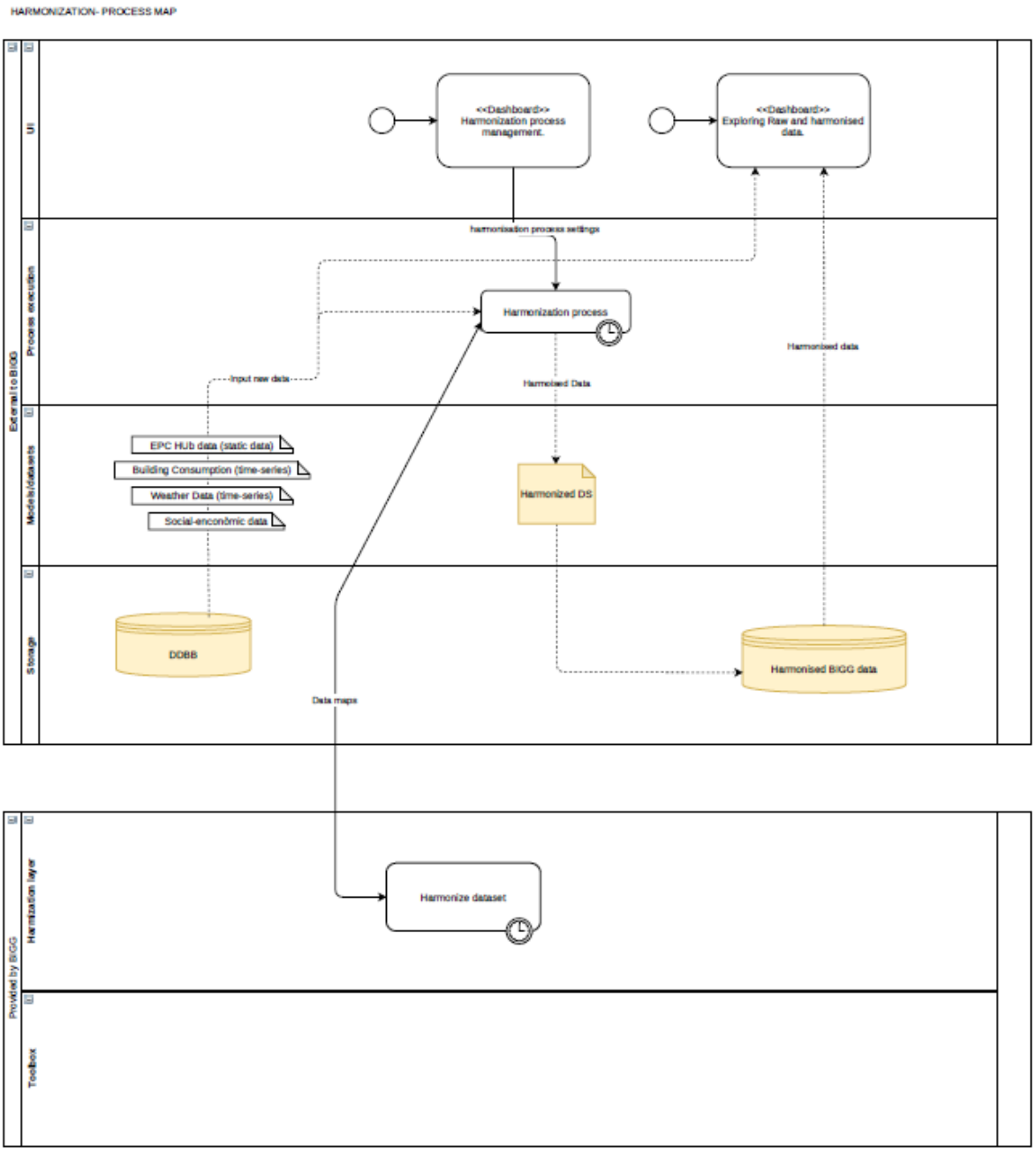
Exchange Requirement Data Objects

< Building input raw data>

Type	Data Object
Name	*Manual raw building data*
Documentation	*All data from the data collection process is stored in the Raw Database.*

Process Map : <Harmonization>

The harmonization process map is the process in charge of standardizing the raw data with the format of the BIGG data model. This standardization allows, on the one hand, to use BIGG's analytical developments and at the same time facilitates the exchange of data with other external systems or applications.



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Base>

Type	Data Object
Name	<data object name> *Harmonized Data Base*
Documentation	<insert description of the data object> *Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format. The database is Apache HBASE (HBase is an open source non-relational distributed database)*

Exchange Requirement Data Objects

<Raw Building data>

Type	Data Object
Name	*Raw building data*
Documentation	*Raw data is extracted from raw databases and sent to the data harmonization process when required.*

<Harmonized Building data>

Type	Data Object
Name	*Harmonized building data*
Documentation	*The harmonized data is extracted from the data harmonization process and is systematically stored in the harmonized database.*

LANE : < Processes execution>

<Harmonization process > [ID:<BC2-UC03-002>]

Type	*Task type*	
Name	*Harmonization process*	
Documentation	*The harmonization process is responsible for transforming the raw data into harmonized data in BIGG format. This process begins on a recurring basis and is controlled and managed by the administrator user from the user interface.*	

	This process that runs is as follows: *- Retrieves raw input data that has not been harmonized so far, from the raw database.* *- Retrieves the raw data maps made in BIGG.* *- Assign the raw data to the BIGG data model.* *- Save the harmonized input data in the harmonized database.*

LANE : < Models/Data Sets>

Library Data Objects

<Building Description >

Type	Data Object
Name	*Building Description*
Documentation	*Same data and description as in the data collection process.*

<Building Consumption >

Type	Data Object
Name	*Building consumption*
Documentation	*Same data and description as in the data collection process.*

<Weather data >

Type	Data Object
Name	*Weather data*
Documentation	*Same data and description as in the data collection process.*

<Harmonized Datasets >

Type	Data Object
Name	*Harmonized Datasets*
Documentation	*The raw data (Building description, building consumption and Weather data) harmonized and standardized on BIGG data model*

LANE : <UI>

In the user interface, the end user, depending on certain permissions, can manage and configure the execution of the harmonization process, and can explore both raw and harmonized data.

<Dashboard: Harmonization process management> [ID:<BC2-UC3-003>]

Type	*Use task*	
Name	*Dashboard: Harmonization process management*	
Documentation	*From the user interface, the administrator user can manage and control the harmonization process. You can configure and modify certain process parameters (execution times, mapping updates, etc.)*	

<Dashboard: Exploring raw and harmonized data> [ID:< BC2-UC03-004>]

Type	*Use task*	
Name	*Dashboard: Exploring raw and harmonized data*	
Documentation	*From the user interface, the user can view and explore the raw data that the system has ingested, as well as the result of the harmonization process. By controlling how much data has actually been harmonized, harmonization issues, even once the data has been normalized, the user can control and manage the data allocation in each building.*	

POOL : <Provided by BIGG >

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Harmonization layer>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies.

<Harmonized Dataset > [ID:<BC2-UC03-005>]

Type	*Task type*	
Name	*Harmonized dataset*	
Documentation	*BIGG's harmonization layer has the raw data mappings developed in the project. These data mappings will go through*	

	the harmonization process of the ENMA system so that you can execute them and harmonize the data.

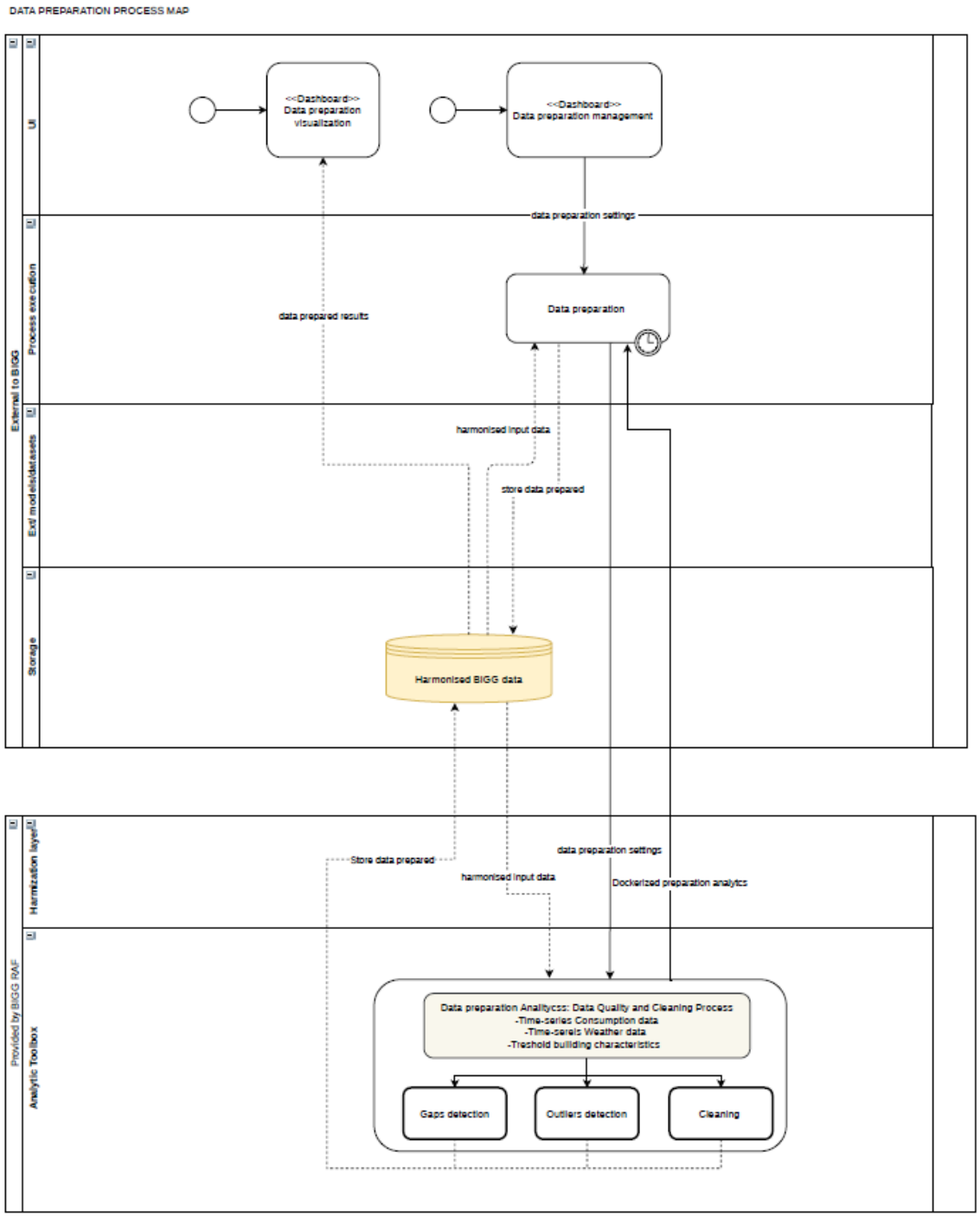
Exchange Requirement Data Objects

< Data mapping >

Type	Data Object
Name	*Data mapping*
Documentation	*The data mappings contained in the BIGG system will be subjected to the ENMA harmonization process so that they are carried out on the data to be harmonized.*

Process Map : <Data preparation>

The data preparation process is responsible for verifying the quality, purifying and transforming the input data so that it can be used without problems by the analytical processes.



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < UI>

In the user interface, the end user, depending on certain permissions, can manage, configure the execution of the data preparation process and can also explore results obtained from this process.

<Dashboard: Data preparation management> [ID:< BC2-UC03-006>]

Type	*Use task*	
Name	*Dashboard: Data preparation management*	
Documentation	*The end user, with administrator permissions, can from the UI manage and control the execution of preparation process. the user will be able to time the execution of the process, configure some parameters and control the execution of these processes (running, finalised, crashed, etc.).*	

<Dashboard: Data preparation visualization> [ID:<xx>]

Type	*Use task*	
Name	*Dashboard: Data preparation visualization*	
Documentation	*In the user interface, the user can explore the results of the data preparation process, get the prepared data, see the differences between the raw data, and assess the quality of the data in the system.*	

LANE : < Processes execution>

<Data preparation > [ID:< BC2-UC03-007>]

Type	*Task type*	
Name	*Data preparation*	
Documentation	*This process is in charge of verifying the quality, purifying and transforming the previously harmonized data. It is a timed process, whose execution is controlled by the user managed from the UI. Once this process is applied, the data is ready to be processed by analytical processes. Depending on the need in each case, this process can be carried out in BIGG's own infrastructure or in ENMA's infrastructure (avoiding the movement of large amounts of data). If the process runs in ENMA, the analytics developed in the BIGG (dockerized) system will also be used.*	

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Bases>

Type	Data Object
Name	*Harmonized Data Bases*

Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format. The Harmonised databases are a combination of data marts (based on MongoDB: a document database, which means it stores data in JSON-like documents) and long-term DB (based on Apache HBASE: HBase is an open source non-relational distributed database)*

POOL : <Provided by BIGG >

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Analytics Toolbox>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies.

<Data preparation analytics > [ID:< BC2-UC03-008>]

Type	*Task type*	
Name	*Data preparation analytics*	
Documentation	*The data preparation process that is contained in the BIGG analysis toolbox can be executed in the BIGG infrastructure itself or in the ENMA infrastructure, depending on the needs of each case.* *This process retrieves the harmonized input data from the buildings and prepares it for further analysis.* *These processes are segmented by data type, for example, static data, time series of consumption or time series of meteorological data, since the processes can be different in each case.* *The processes could be:* *- data format* *- data transformation* *- detection of data gaps* *- outlier detection* *- data filling* *- data normalization* *- data modification* *The results can vary from the data prepared itself to global performance indicators such as:* *- Buildings that have or have not passed the quality process*	

	- Buildings whose data has been corrected *- Time series of consumption with gaps.* *- Meteorological data with gaps.* *- EPC static parameters with outliers* *- etc...*

Exchange Requirement Data Objects

< data prepared results >

Type	Data Object
Name	*Data prepared results*
Documentation	*The results of the analytical data preparation process are collected from the Harmonized DDBB and presented to users in the user interface.*

< Harmonised input data >

Type	Data Object
Name	*Harmonized input data*
Documentation	*The harmonized input data is extracted from the Harmonized DDBB and served to the data preparation execution process. Either in the BIGG or ENMA infrastructure.*

< Stored data prepared >

Type	Data Object
Name	*Stored data prepared*
Documentation	*The results of the data preparation process are stored back to the Harmonised DDBB.*

< Data preparation settings >

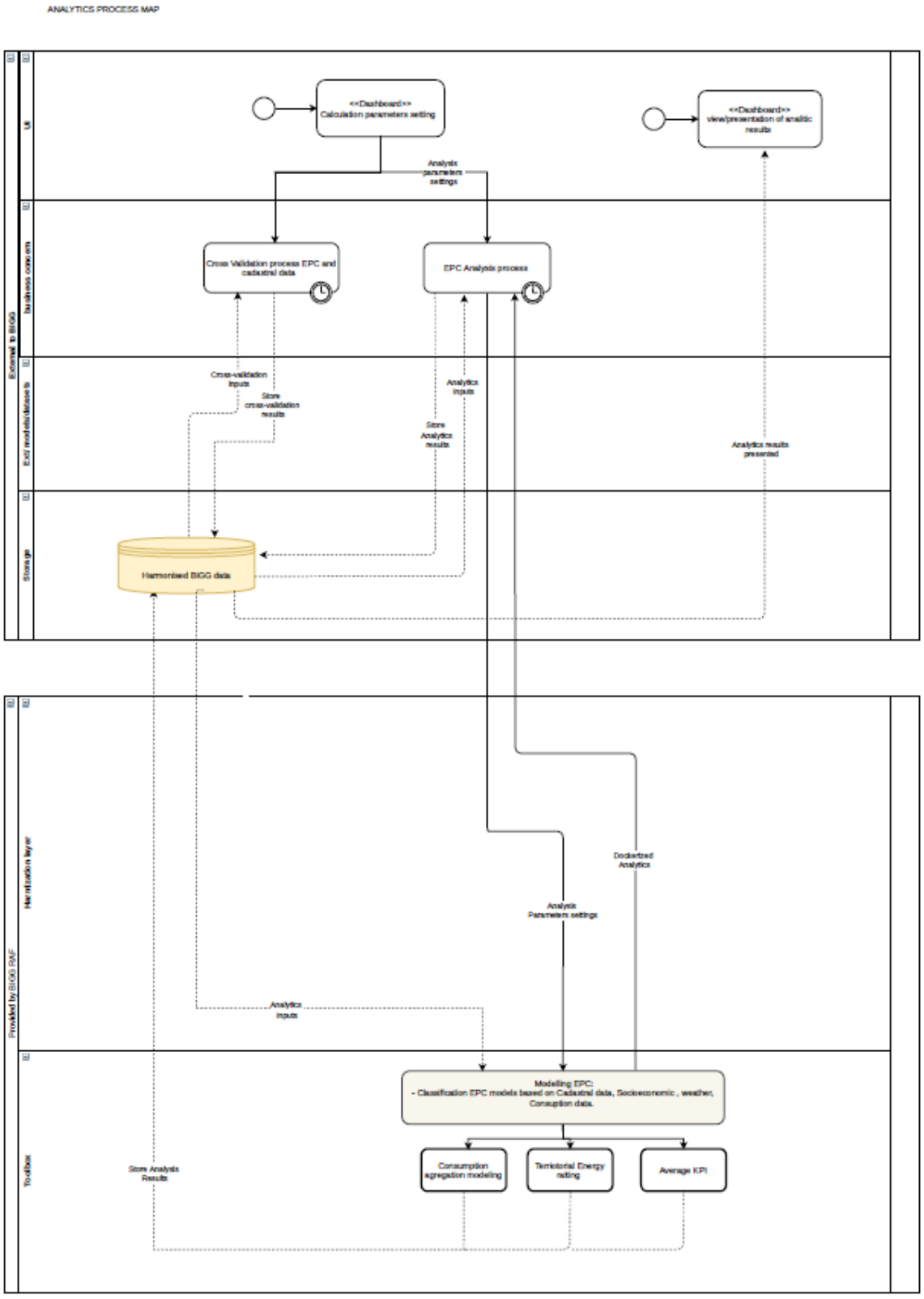
Type	Data Object
Name	*Data preparation settings*
Documentation	*The configuration parameters are passed to the process of running the data preparation process from the user interface.*

< Dockerized preparation analytics >

Type	Data Object
Name	*Dockerized preparation analytics*
Documentation	*In the event that the execution is carried out outside the BIGG infrastructure, the analysis tools, previously dockerized, will be imported to be deployed and executed in the external infrastructure.*

Process Map : <Analytical Process>

The analytical process is responsible for extracting the data, previously prepared, extracting the harmonized DDBB, obtaining the expected analytical results for the use case, saving them in the harmonized DDBB and presenting them to the end user.



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < UI>

In the user interface, the end user, depending on certain permissions, can manage, configure the execution of the data Analysis process and can also explore results obtained from this process.

<Dashboard: View/presentation of analytic results> [ID:< BC2-UC03-009>]

Type	*Use task*	
Name	*Dashboard: View/presentation of analytic results*	
Documentation	*The end user can explore the analytical results from the user interface. The presentation of these results will be simple and easily understandable. The user can easily filter the results to obtain the visualizations that he needs at all times.*	

<Dashboard: Analysis management> [ID:< BC2-UC03-010>]

Type	*Use task*	
Name	*Dashboard: Analysis management*	
Documentation	*The end user, with administrator permissions, can from the UI manage and control the execution of Analytics process. The user will be able to time the execution of the process, configure some parameters and control the execution of these processes (running, finalised, crashed, etc.).*	

LANE : < Processes execution>

<Cross-Validation process EPC/Cadastral data > [ID:< BC2-UC03-011>]

Type	*Task type*	
Name	*Cross-validation process EPC/cadastral data*	
Documentation	*The cross-validation process is the process in charge of verifying the coherence between the EPC data and the cadastral information. These processes will run on the local platform.*	

<Analysis process > [ID:< BC2-UC03-012>]

Type	*Task type*	
Name	*Analysis process*	
Documentation	*The analysis process is responsible for managing and executing the analysis processes on the prepared data, based on the parameterization established by the user in the UI.* *For these use cases, the EPC modelling. The processes can be run on the BIGG infrastructure or on the local platform. If they are launched in the local*	

	infrastructure, the analytics developed in BIGG are used, which are imported dockerized to be deployed and executed.

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Bases>

Type	Data Object
Name	*Harmonized Data Bases*
Documentation	*In the harmonized databases are the data ready to be used by the analytical processes. The results of the analytical processes will also be stored in these databases.* *The Harmonised databases are a combination of data marts (based on MongoDB: a document database, which means it stores data in JSON-like documents) and long-term DB (based on Apache HBASE: HBase is an open source non-relational distributed database)*

POOL : <Provided by BIGG >

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Analytics Toolbox>

Analytics Toolbox contains the analysis developed in the BIGG project for this use case, the analysis is:

- EPC modelling

- EPC Indicators predictor Based on weather data and cadastre data

- Characterise the theoretical energy consumption/demand of a certain subset of buildings

- Characterise the correlation between geographical areas and EPC indicators.

-Comparing actual geographically aggregated consumption data and socio-economic data

Depending on the needs and volume of processing, this lane can be used for run analysis or simply to store the coupled analysis so they can be deployed and run on other infrastructures.

<Modelling EPC> [ID:< BC2-UC03-013>]

Type	*Task type*	
Name	*Modelling EPC*	
Documentation	*In EPC Modelling, the large volume of data stored from the energy efficiency certificates of the Catalan HUB will be treated together with other data such as cadastral, meteorological or socioeconomic.*	

	The execution of these analyses can be done on BIGG's own infrastructure or on the existing ENMA platform (with preference of the last one) . *The input data for this model is the prepared data that is stored in the harmonized databases.* *This modelling process can be divided into the following threads:* *Steady state features transformation* *Dynamic features transformation* *Best ML model selection* *The expected results are:* *EPC Indicators predictor Based on weather data and cadastre data* *Characterise the theoretical energy consumption/demand of a certain subset of buildings* *Characterise the correlation between geographical areas and EPC indicators* *Comparing actual geographically aggregated consumption data and socio-economic data*

Exchange Requirement Data Objects

< Analytic results presented>

Type	Data Object
Name	*Analytic results presented*
Documentation	*The results of the analysis, stored in the harmonized DDBBs, will be extracted and presented to the end users.*

< Execution summary >

Type	Data Object
Name	*Execution summary*
Documentation	*After the analytic has launched, the end user receives the processing summary in the user interface. (Successfully executed processes, unfinished processes, crashed processes, etc.).*

< Analytics Inputs >

Type	Data Object
Name	*Analytics Inputs*

Documentation	*The necessary input data is sent to the analysis process from the harmonized databases.* *For Modelling EPC, the necessary inputs are:* *- Aggregated historical energy consumption (time series)* *- Socio-economic data* *- Weather data* *- EPC description parameters*

< Stored Analysis results >

Type	Data Object
Name	*Stored Analysis results*
Documentation	*The results of the analytical processes are stored in the harmonized databases.* *For EPC modelling the main results are:* *EPC Indicators predictor Based on weather data and cadastre data* *Characterise the theoretical energy consumption/demand of a certain subset of buildings* *Characterise the correlation between geographical areas and EPC indicators* *Comparing actual geographically aggregated consumption data and socio-economic data*

< Analysis parameters settings >

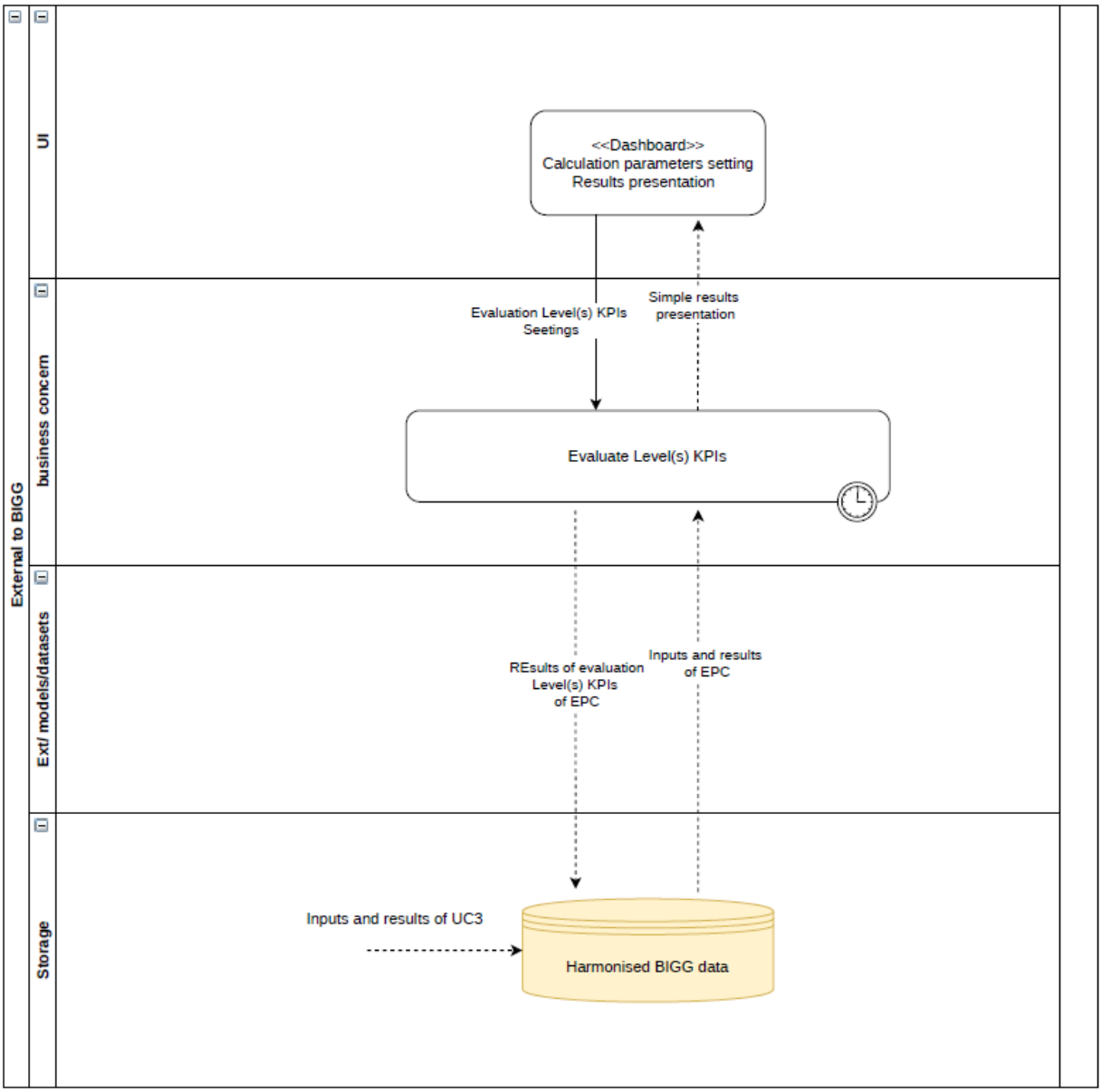
Type	Data Object
Name	*Analysis parameters settings*
Documentation	*From the UI the end user can modify some parameters that are used in the analysis process.* *For EEM evaluation, these can be (among others):* *- Selection of region to be analysed.* *- Selection of excluded time periods.* *- etc…*

< Dockerized Analytics >

Type	Data Object
Name	*Dockerized Analytics*
Documentation	*In the event that the execution is carried out outside the BIGG infrastructure, the analysis tools, previously dockerized, will be imported to be deployed and executed in the external infrastructure.*

V.4. UC4

V.4.1. BPMN diagram



V.4.2. Specification of processes

certain permission manage and control the results of certain analytics applied over data.

<Dashboard: Calculation parameter setting> [ID:< BC2-UC04-001>]

Type	*Use task*	
Name	*Dashboard: Calculation parameters setting*	

Documentation	*From the user interface, the end user with certain permissions could configure the parameters and manage the Level(s) KPI's.*

LANE : < Processes execution>

<Evaluate Level(s) KPI's> [ID:<BC2-UC04-002>]

Type	*Task type*	
Name	*Evaluation Level(s) KPI's*	
Documentation	The results of EPC, stored by UC3, are used to evaluate the possibility of Level(S) KPI's calculations.	

Exchange Requirement Data Objects

<Input and results of EPC>

Type	Data Object
Name	*Input and results of EPC*
Documentation	*The data stored in DDBB is passed to evaluated Level(s) KPI's process*

<Results of Evaluation Level(s) KPIs of EPC>

Type	Data Object
Name	*Results of evaluation Level(s) KPIs of EPC*
Documentation	*The results of evaluation of Level(s) KPIs are stored in the DDBB*

<Simple results presentation >

Type	Data Object
Name	*Simple results presentation*
Documentation	*The results are presented in the UI the analytics of results.*

<Evaluation Level(s) KPIs settings >

Type	Data Object

Name	*Evaluation Level(s) KPIs settings*
Documentation	*Some parameters of configuration the evaluation Level(s) KPIs processes are provided from UI.* *Selection of buildings* *Selection of KPIs to be calculated* *etc*

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Bases>

Type	Data Object
Name	*Harmonized Data Bases*
Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format. The Harmonised databases are a combination of data marts (based on MongoDB: a document database, which means it stores data in JSON-like documents) and long-term DB (based on Apache HBASE: HBase is an open source non-relational distributed database)*

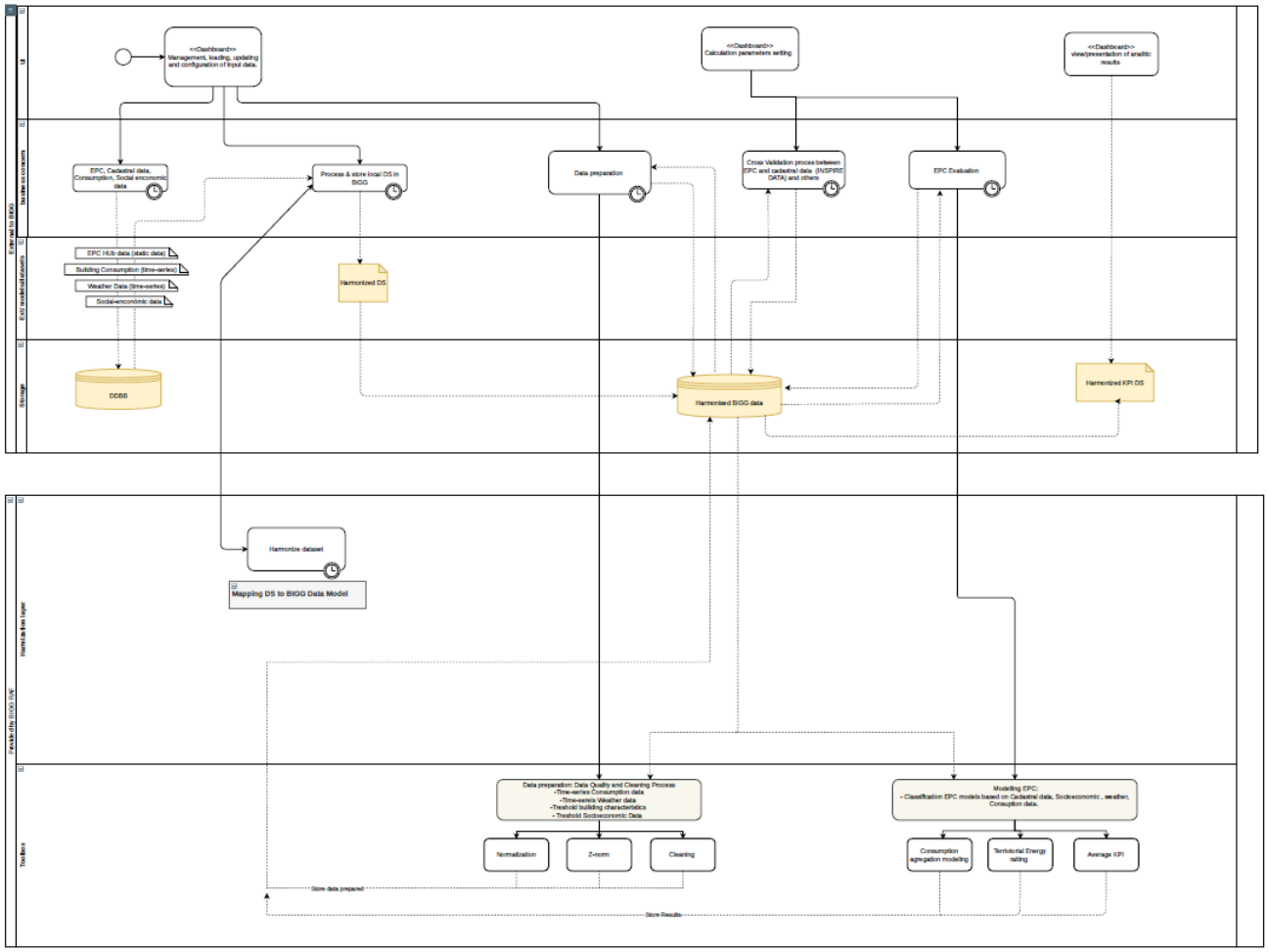
Exchange Requirement Data Objects

<Inputs and results of UC2>

Type	Data Object
Name	*Inputs and results of UC3*
Documentation	*The inputs and results of UC3, that must be used in this UC4, are stored in a data base in harmonized and prepared way.*

V.5. UC5

V.5.1. BPMN diagram



V.5.2. Specification of processes

*To develop the "*Interoperability between BIM, BMS, CMMS and simulation engines *" process, it has been divided into the following process maps in order to deepen its description.*

The breakdown of the process maps presented in this document are:

- Data collection

- Data harmonization

- Data preparation

- Analytical process

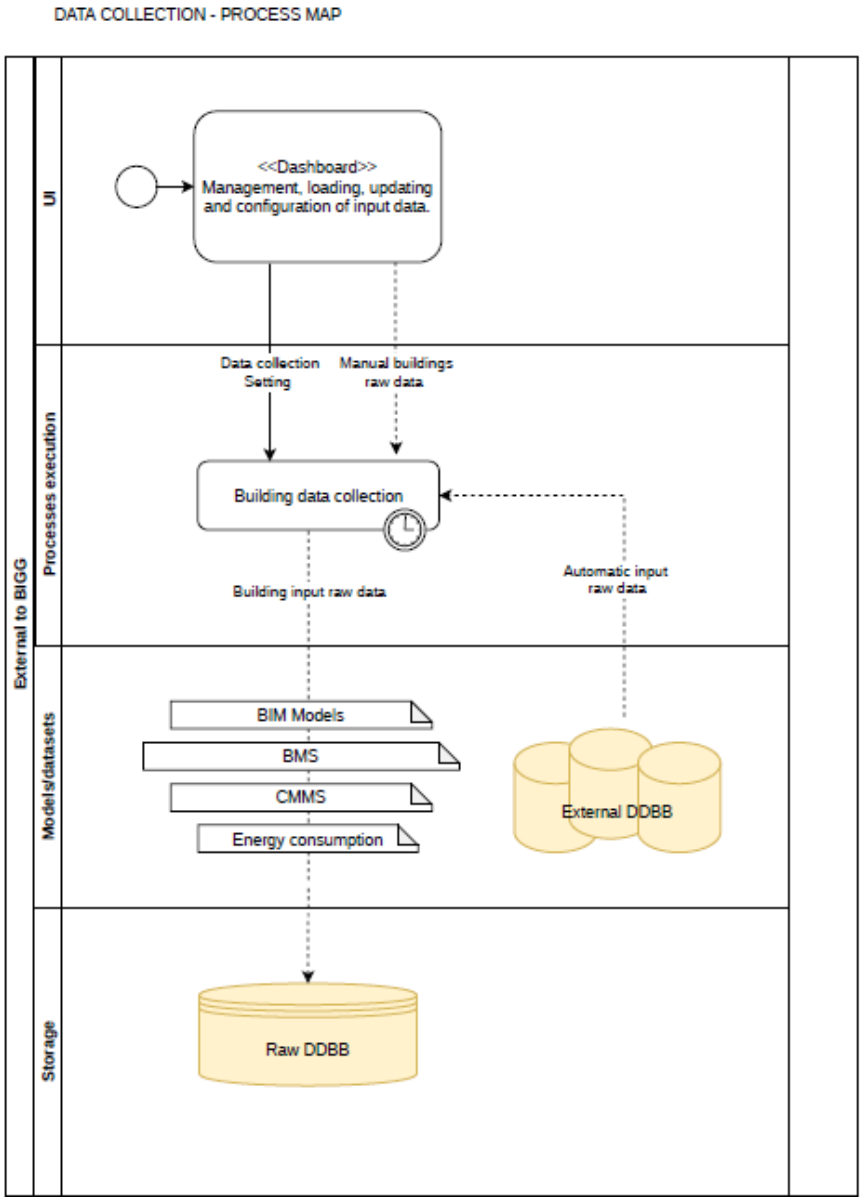
Process Map : <Data Collection>

The data required for the UC3 can come from different data sources. These data sources can be loaded from existing databases or uploaded manually by the user.

In the case of manual data upload, the end user will be able to load, modify or delete data from the UI.

In the case of data from existing databases, these can be collected directly through APIs or similar, the administrator user will have access from the UI to the import control of this data.

The EPC data, Weather, cadastral data will be systematically stored in raw in a Database. This Database will be located in the infrastructure of the existing ENMA application (owned by CIMNE).



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : <UI>

In the UI the end user, depending on certain permissions, can upload, modify or delete the building data manually (through web forms or by uploading CSV or Excel files, with predefined format), or configure the collection of data from databases. existing data, external to this Pool

<Dashboard: Management, loading, updating and configuration of input data.> [ID:<xx>]

Type	*Use task*	
Name	*Dashboard: Management, loading, updating and configuration of input data*	
Documentation	*In the UI the end user, depending on certain permissions, can upload, modify or delete the building data manually (through web forms or by uploading CSV or Excel files, with predefined format), or configure the collection of data from databases. existing data, external to this Pool*	

LANE : < Processes execution>

<Building Data Collection> [ID:<BC3-UC05-001>]

Type	*Task type*	
Name	*Building data collection*	
Documentation	*The building data collection process is responsible for managing the data entry into the system.* *For manual data, the new data will be loaded and the modification or deletion of the data will be managed, according to the user's request.* *Depending on the configuration assigned by the user, timed calls will be launched to the different APIs to automatically retrieve data from external databases.*	

Exchange Requirement Data Objects

<Manual Input Raw Building data>

Type	Data Object
Name	*Manual raw building data*
Documentation	*Data uploaded, modified or deleted manually by the user in the user interface will be passed to the data collection process.*

<Automatic input Raw Building data>

Type	Data Object

Name	*Automatic raw building data*
Documentation	*The existing data in the different external databases will be processed by the data collection process. This data can be just new data or all data depending on the configuration that the user wants.*

LANE : < Models/Data Sets>

Library Data Objects

<BIM Models>

Type	Data Object
Name	*BIM Models*
Documentation	*In the design phase, the creation of BIM (Building Information Modelling) models is mandatory for all newly constructed buildings owned by the Generalitat de Catalunya.* *Some of them will be uploaded to the BIGG project to harmonize and merge with all other use case sources.*

<BMS >

Type	Data Object
Name	*BMS*
Documentation	*Most of the buildings managed by infraestructrues.cat have building management systems. These are proprietary systems from different vendors (Schneider, Controlli, siemens, Danfoss, Sauter, etc ...) but they are all connected to a common gateway (IXON) that collects all data in the BACnet protocol. These will be collected in the BIGG project and stored in raw.*

<CMMS>

Type	Data Object
Name	*CMMS*
Documentation	*Most of the buildings managed by infraestructrues.cat are integrated into a computerized maintenance management system (CMMS) (MANTEST). In this system, the building inventory and all maintenance actions (scheduled and executed) are stored for all buildings.* *All of these will be integrated into the raw DDBB that will be used in the project.*

<Energy Consumption >

Type	Data Object
Name	*Energy consumption*
Documentation	*The electric and gas energy consumption is collected for all buildings and stored.*

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Raw Data Base>

Type	Data Object
Name	*Raw Data Base*
Documentation	*Raw data is systematically stored in ENMA databases. The model of each of the data will be defined according to each of the source providers. The database is Apache HBASE (HBase is an open source non-relational distributed database)*

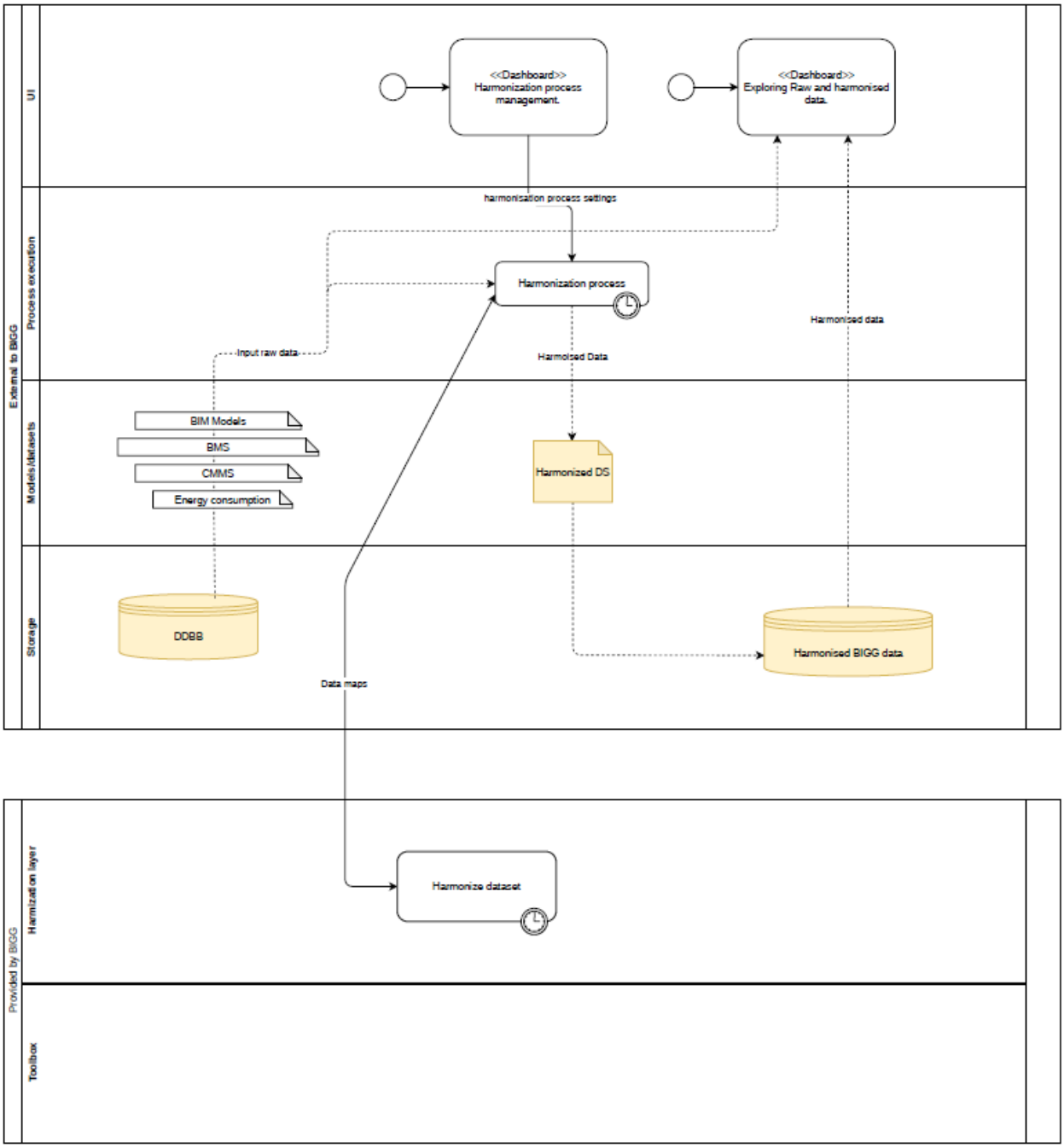
Exchange Requirement Data Objects

< Building input raw data>

Type	Data Object
Name	*Manual raw building data*
Documentation	*All data from the data collection process is stored in the Raw Database.*

Process Map : <Harmonization>

The harmonization process map is the process in charge of standardizing the raw data with the format of the BIGG data model. This standardization allows, on the one hand, to use BIGG's analytical developments and at the same time facilitates the exchange of data with other external systems or applications.



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Base>

Type	Data Object
Name	*Harmonized Data Base*
Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format. The database is Apache HBASE (HBase is an open source non-relational distributed database)*

Exchange Requirement Data Objects

<Raw Building data>

Type	Data Object
Name	*Raw building data*
Documentation	*Raw data is extracted from raw databases and sent to the data harmonization process when required.*

<Harmonized Building data>

Type	Data Object
Name	*Harmonized building data*
Documentation	*The harmonized data is extracted from the data harmonization process and is systematically stored in the harmonized database.*

LANE : < Processes execution>

<Harmonization process > [ID:<BC3-UC05-002>]

Type	*Task type*	
Name	*Harmonization process*	
Documentation	*The harmonization process is responsible for transforming the raw data into harmonized data in BIGG format. This process begins on a recurring basis and is controlled and managed by the administrator user from the user interface.* *This process that runs is as follows:* *- Retrieves raw input data that has not been harmonized so far, from the raw database.* *- Retrieves the raw data maps made in BIGG.* *- Assign the raw data to the BIGG data model.* *- Save the harmonized input data in the harmonized database.*	

LANE : < Models/Data Sets>

Library Data Objects

<BIM Models >

Type	Data Object
Name	*BIM Models*
Documentation	<insert description of the data object> *Same data and description as in the data collection process.*

<BMS >

Type	Data Object
Name	*BMS*
Documentation	<insert description of the data object> *Same data and description as in the data collection process.*

<CMMS >

Type	Data Object
Name	*CMMS*
Documentation	*Same data and description as in the data collection process.*

<Energy Consumption >

Type	Data Object
Name	*Energy consumptions*
Documentation	*Same data and description as in the data collection process*

LANE : <UI>

In the user interface, the end user, depending on certain permissions, can manage and configure the execution of the harmonization process, and can explore both raw and harmonized data.

<Dashboard: Harmonization process management> [ID:<BC3-UC5-003>]

Type	*Use task*	
Name	*Dashboard: Harmonization process management*	

Documentation	*From the user interface, the administrator user can manage and control the harmonization process. You can configure and modify certain process parameters (execution times, mapping updates, etc.)*

<Dashboard: Exploring raw and harmonized data> [ID:< BC3-UC05-004>]

Type	*Use task*	
Name	*Dashboard: Exploring raw and harmonized data*	
Documentation	*From the user interface, the user can view and explore the raw data that the system has ingested, as well as the result of the harmonization process. By controlling how much data has actually been harmonized, harmonization issues, even once the data has been normalized, the user can control and manage the data allocation in each building.*	

POOL : <Provided by BIGG >

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Harmonization layer>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies.

<Harmonized Dataset > [ID:<BC3-UC05-005>]

Type	*Task type*	
Name	*Harmonized dataset*	
Documentation	*BIGG's harmonization layer has the raw data mappings developed in the project. These data mappings will go through the harmonization process of the ENMA system so that you can execute them and harmonize the data.*	

Exchange Requirement Data Objects

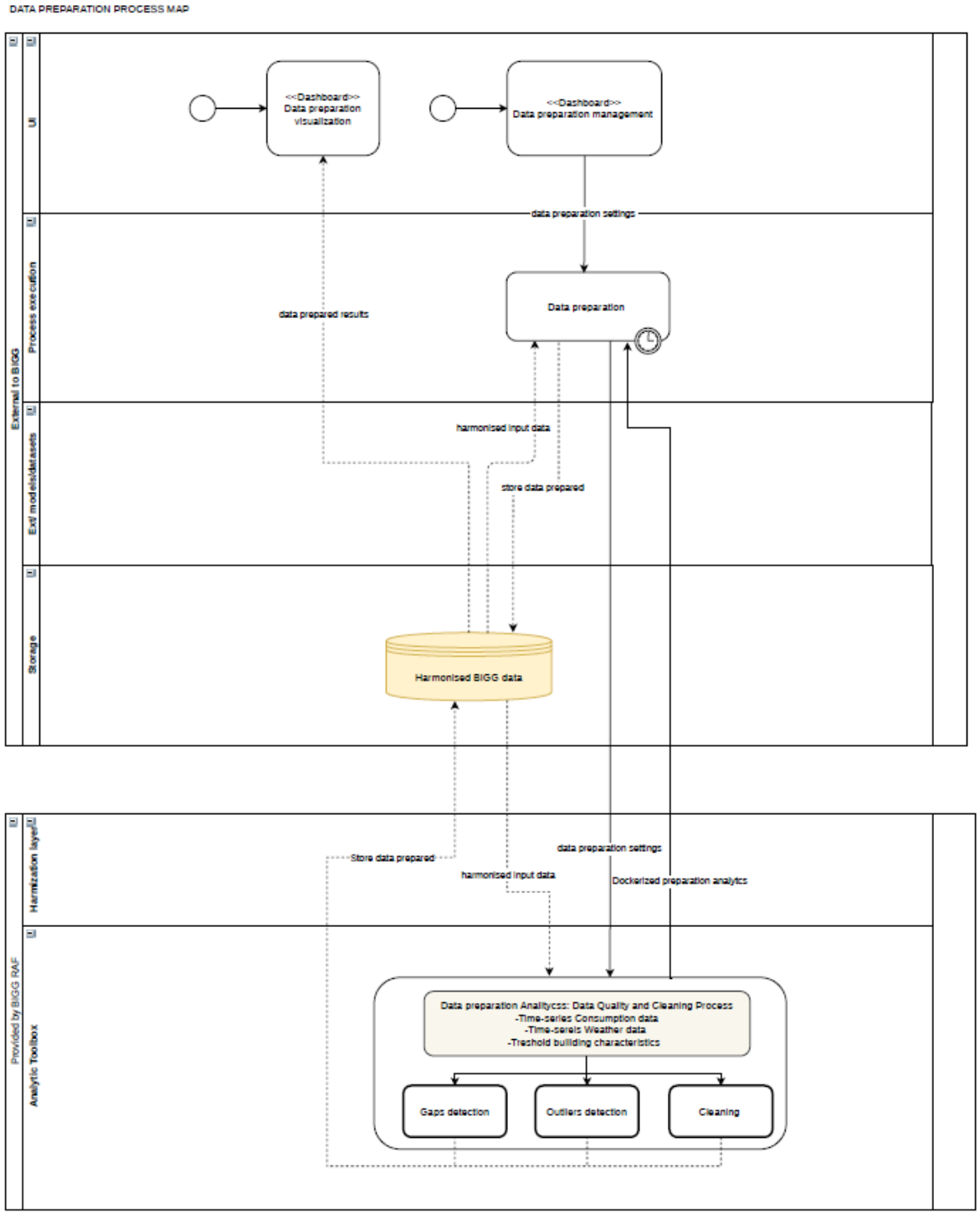
< Data mapping >

Type	Data Object
Name	*Data mapping*

Documentation	*The data mappings contained in the BIGG system will be subjected to the ENMA harmonization process so that they are carried out on the data to be harmonized.*

Process Map : <Data preparation>

The data preparation process is responsible for verifying the quality, purifying and transforming the input data so that it can be used without problems by the analytical processes.



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < UI>

In the user interface, the end user, depending on certain permissions, can manage, configure the execution of the data preparation process and can also explore results obtained from this process.

<Dashboard: Data preparation management> [ID:< BC3-UC05-006>]

Type	*Use task*	
Name	*Dashboard: Data preparation management*	
Documentation	*The end user, with administrator permissions, can from the UI manage and control the execution of preparation process. the user will be able to time the execution of the process, configure some parameters and control the execution of these processes (running, finalised, crashed, etc.).*	

<Dashboard: Data preparation visualization> [ID:<BC3-UC05-007>]

Type	*Use task*	
Name	*Dashboard: Data preparation visualization*	
Documentation	*In the user interface, the user can explore the results of the data preparation process, get the prepared data, see the differences between the raw data, and assess the quality of the data in the system.*	

LANE : < Processes execution>

<Data preparation > [ID:< BC3-UC05-008>]

Type	*Task type*	
Name	*Data preparation*	
Documentation	*This process is in charge of verifying the quality, purifying and transforming the previously harmonized data. It is a timed process, whose execution is controlled by the user managed from the UI. Once this process is applied, the data is ready to be processed by analytical processes. Depending on the need in each case, this process can be carried out in BIGG's own infrastructure or in ENMA's infrastructure (avoiding the movement of large amounts of data). If the process runs in ENMA, the analytics developed in the BIGG (dockerized) system will also be used.*	

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Bases>

Type	Data Object
Name	*Harmonized Data Bases*
Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format. The Harmonised databases are a combination of data marts (based on MongoDB: a document database, which means it stores data in JSON-like documents) and long-term DB (based on Apache HBASE: HBase is an open source non-relational distributed database)*

POOL : <Provided by BIGG >

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Analytics Toolbox>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies.

<Data preparation analytics > [ID:< BC3-UC05-009>]

Type	*Task type*	
Name	*Data preparation analytics*	
Documentation	*The data preparation process that is contained in the BIGG analysis toolbox can be executed in the BIGG infrastructure itself or in the ENMA infrastructure, depending on the needs of each case.* *This process retrieves the harmonized input data from the buildings and prepares it for further analysis.* *These processes are segmented by data type, for example, static data, time series of consumption or time series of meteorological data, since the processes can be different in each case.* *The processes could be:* *- data format* *- data transformation*	

	- detection of data gaps *- outlier detection* *- data filling* *- data normalization* *- data modification* *The results can vary from the data prepared itself to global performance indicators such as:* *- Buildings that have or have not passed the quality process* *- Buildings whose data has been corrected* *- Time series of sensors with gaps.* *- Inventory static parameters with outliers* *- etc...*

Exchange Requirement Data Objects

< data prepared results >

Type	Data Object
Name	*Data prepared results*
Documentation	*The results of the analytical data preparation process are collected from the Harmonized DDBB and presented to users in the user interface.*

< Harmonised input data >

Type	Data Object
Name	*Harmonized input data*
Documentation	*The harmonized input data is extracted from the Harmonized DDBB and served to the data preparation execution process. Either in the BIGG or ENMA infrastructure.*

< Stored data prepared >

Type	Data Object
Name	*Stored data prepared*
Documentation	*The results of the data preparation process are stored back to the Harmonised DDBB.*

< Data preparation settings >

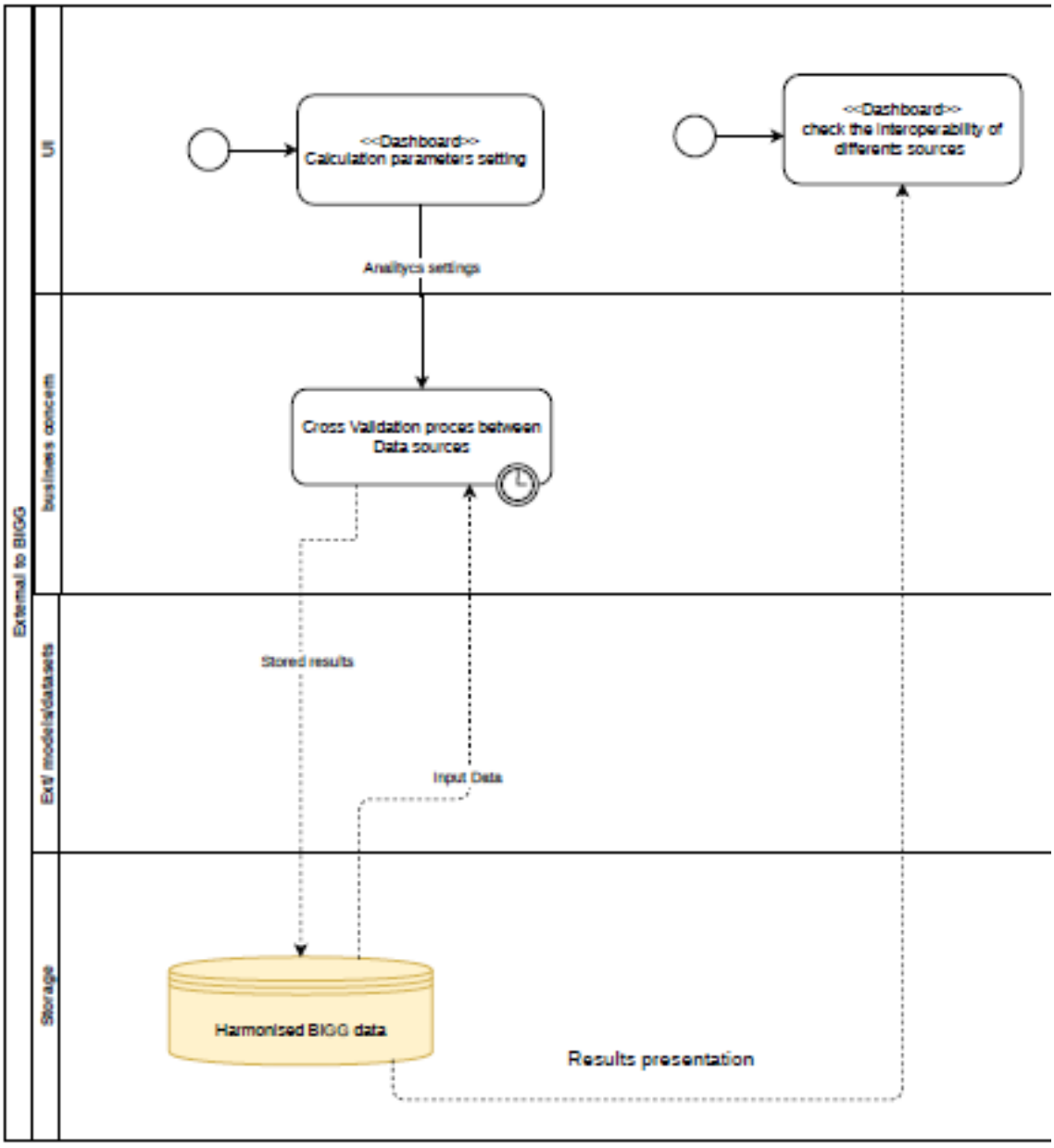
Type	Data Object
Name	*Data preparation settings*
Documentation	*The configuration parameters are passed to the process of running the data preparation process from the user interface.*

< Dockerized preparation analytics >

Type	Data Object
Name	*Dockerized preparation analytics*
Documentation	*In the event that the execution is carried out outside the BIGG infrastructure, the analysis tools, previously dockerized, will be imported to be deployed and executed in the external infrastructure.*

Process Map : <Analytical Process>

The analytical process is responsible for extracting the data, previously prepared, extracting the harmonized DDBB, obtaining the expected analytical results for the use case, saving them in the harmonized DDBB and presenting them to the end user.



POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : < UI>

In the user interface, the end user, depending on certain permissions, can manage, configure the execution of the data Analysis process and can also explore results obtained from this process.

<Dashboard: View/presentation of analytic results> [ID:< BC3-UC05-010>]

Type	*Use task*	
Name	*Dashboard: View/presentation of analytic results*	
Documentation	*The end user can explore the analytical results from the user interface. The presentation of these results will be simple and easily understandable. The user can easily filter the results to obtain the visualizations that he needs at all times.*	

<Dashboard: Analysis management> [ID:< BC3-UC05-011>]

Type	*Use task*	
Name	*Dashboard: Analysis management*	
Documentation	*The end user, with administrator permissions, can from the UI manage and control the execution of Analytics process. The user will be able to time the execution of the process, configure some parameters and control the execution of these processes (running, finalised, crashed, etc.).*	

LANE : < Processes execution>

<Cross-Validation process > [ID:< BC3-UC05-012>]

Type	*Task type*	
Name	*Cross-validation process EPC/cadastral data*	
Documentation	*The cross-validation process is the process in charge of verifying the coherence between the different data sources that are collected for each building.. These processes will run on the local platform.*	

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Bases>

Type	Data Object
Name	*Harmonized Data Bases*

Documentation	*In the harmonized databases are the data ready to be used by the analytical processes. The results of the analytical processes will also be stored in these databases.* *The Harmonised databases are a combination of data marts (based on MongoDB: a document database, which means it stores data in JSON-like documents) and long-term DB (based on Apache HBASE: HBase is an open source non-relational distributed database)*

Exchange Requirement Data Objects

< Analytic results presented>

Type	Data Object
Name	*Analytic results presented*
Documentation	*The results of the analysis, stored in the harmonized DDBBs, will be extracted and presented to the end users.*

< Execution summary >

Type	Data Object
Name	*Execution summary*
Documentation	*After the analytic has launched, the end user receives the processing summary in the user interface. (Successfully executed processes, unfinished processes, crashed processes, etc.).*

< Analytics Inputs >

Type	Data Object
Name	*Analytics Inputs*
Documentation	*The necessary input data is sent to the analysis process from the harmonized databases.* *For Cross-validation, the necessary inputs are:* *- BMS data* *- BIM data* *- Energy consumption data* *- CMMs data*

< Stored Analysis results >

Type	Data Object
Name	*Stored Analysis results*

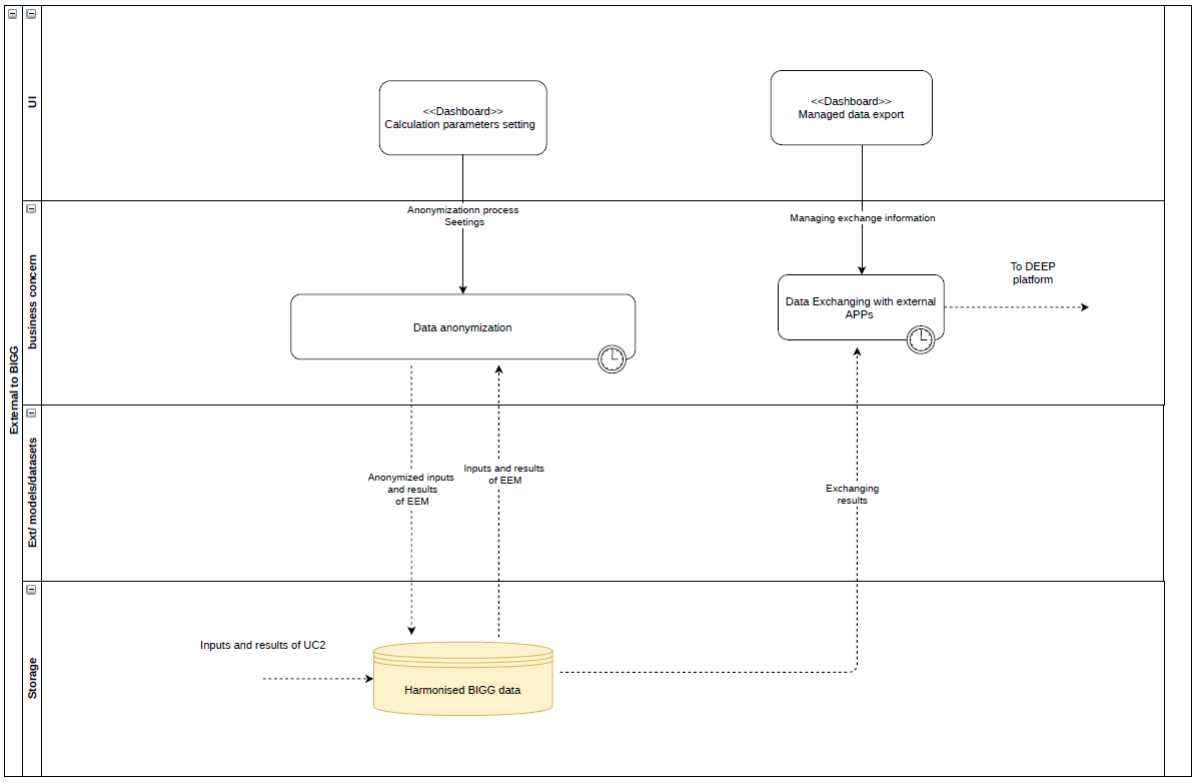
Documentation	*The results of the analytical processes are stored in the harmonized databases.* *For Cross-Validation the main results are:* *Conflicted values* *Not correlated values.* *Etc..*

< Analysis parameters settings >

Type	Data Object
Name	*Analysis parameters settings*
Documentation	*From the UI the end user can modify some parameters that are used in the analysis process.* *For Corss-Validation, these can be (among others):* *- Selection of buildings analysed.* *- Selection of period of execution.* *- etc…*

V.6. UC6

V.6.1. BPMN diagram



V.6.2. Specification of processes

POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : <UI>

In the user interface, the end user, depending on certain permissions, can manage, configure the execution of the data Analysis process and can also explore results obtained from this process.

<Dashboard: Calculation parameter setting> [ID:< BC3-UC06-001>]

Type	*Use task*	
Name	*Dashboard: Calculation parameters setting*	
Documentation	*From the user interface, the end user with certain permissions could configure the parameters and manage the anonymization process of the data to be exported.*	

<Dashboard: Manage data export> [ID:< BC3-UC06-002>]

Type	*Use task*	
Name	*Dashboard: Manage data export*	
Documentation	*From the user interface, the end user with certain permissions could configure the parameters and manage the exported process.*	

LANE : < Processes execution>

<Data anonymization> [ID:<BC3-UC06-003>]

Type	*Task type*	
Name	*Data anonymization*	
Documentation	The input data and the results of the EEM evaluation process that are stored in the DDBB will be anonymized at the user's request and prepared to be exported to the DEEP platform.	

<Data exchanging with external app> [ID:<BC3-UC06-004>]

Type	*Task type*	
Name	*Date exchanging with external app*	
Documentation		

	The anonymized inputs and results are prepared and exchanged with an external application (EFFIG DEEP), in case the exchange process is carried out through an API, this will be launched, in case the exchange process is done with a document , this will be created and available to the end user

Exchange Requirement Data Objects

<Input and results of EEM>

Type	Data Object
Name	*Input and results of EEM*
Documentation	*The data stored in DDBB is passed to the anonymization process*

<Anonymized input and results>

Type	Data Object
Name	*Anonymized input and results*
Documentation	*The results of anonymization process are stored in the DDBB*

<Exchanging results >

Type	Data Object
Name	*Exchanging results*
Documentation	*The anonymized results are extorted from de DDBB and passed to Data exchanging process*

<Anonymized process settings >

Type	Data Object
Name	*Anonymized process settings*
Documentation	*Some parameters of configuration the anonymization processes are provided from UI.* *Selection of buildings* *Selection of anonymization typology* *Selection of EEM typology* *etc*

<Managing exchanging information>

Type	Data Object
Name	*Anonymized process settings*
Documentation	*Some parameters of configuration of anonymization process are provided from UI*

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Bases>

Type	Data Object
Name	*Harmonized Data Bases*
Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format. The Harmonised databases are a combination of data marts (based on MongoDB: a document database, which means it stores data in JSON-like documents) and long-term DB (based on Apache HBASE: HBase is an open source non-relational distributed database)*

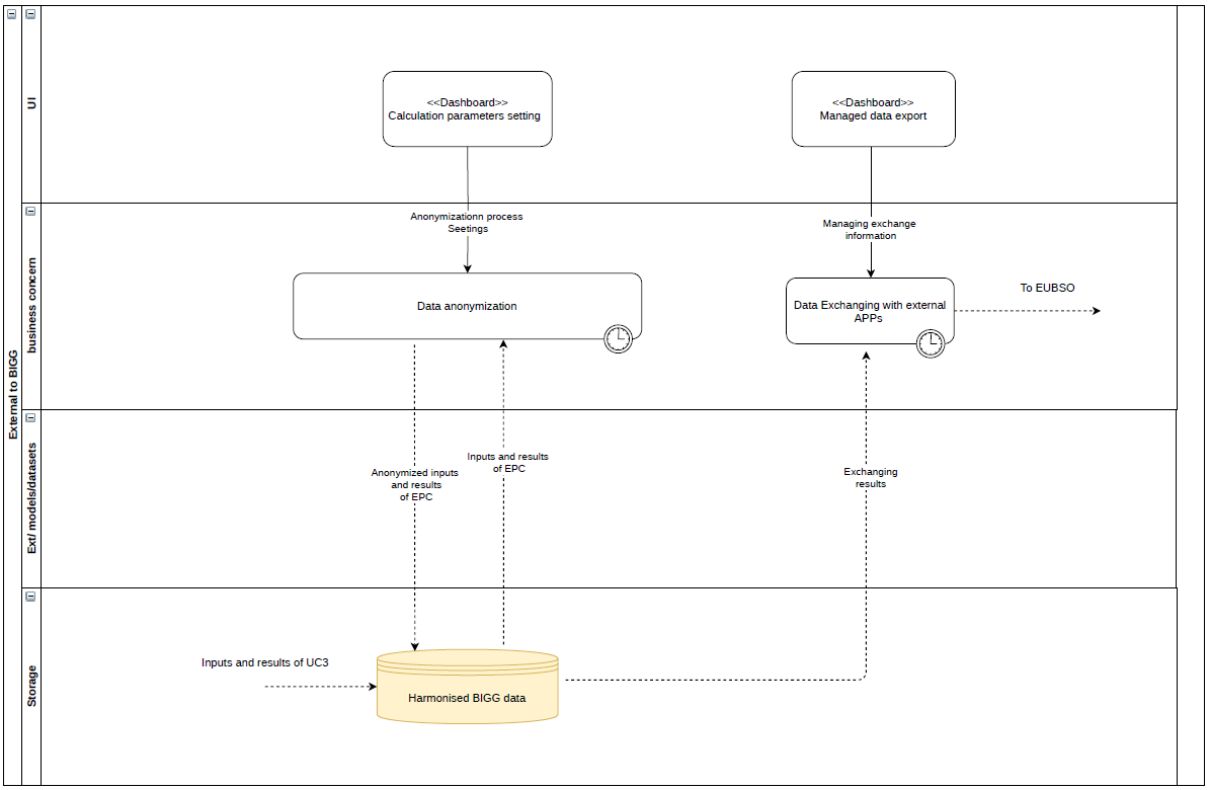
Exchange Requirement Data Objects

<Inputs and results of UC2>

Type	Data Object
Name	*Inputs and results of UC2*
Documentation	*The inputs and results of UC2, that must be used in this UC6, are stored in a data base in harmonized and prepared way.*

V.7. UC7

V.7.1. BPMN diagram



V.7.2. Specification of processes

POOL : <External to BIGG >

This process will be carried out entirely in the group <External to BIGG>. <External to BIGG> is understood as the processes that are developed in the existing ENMA application, owned by CIMNE.

LANE : <UI>

In the user interface, the end user, depending on certain permissions, can manage, configure the execution of the data Analysis process and can also explore results obtained from this process.

<Dashboard: Calculation parameter setting> [ID:< BC3-UC07-001>]

Type	*Use task*	
Name	*Dashboard: Calculation parameters setting*	

Documentation	*From the user interface, the end user with certain permissions could configure the parameters and manage the anonymization process of the data to be exported.*

<Dashboard: Manage data export> [ID:< BC3-UC07-002>]

Type	*Use task*	
Name	*Dashboard: Manage data export*	
Documentation	*From the user interface, the end user with certain permissions could configure the parameters and manage the exported process.*	

LANE : < Processes execution>

<Data anonymization> [ID:<BC3-UC07-003>]

Type	*Task type*	
Name	*Data anonymization*	
Documentation	The input data and the results of the EEM evaluation process that are stored in the DDBB will be anonymized at the user's request and prepared to be exported to the EUBSO platform.	

<Data exchanging with external app> [ID:<BC3-UC07-004>]

Type	*Task type*	
Name	*Date exchanging with external app*	
Documentation	The anonymized inputs and results are prepared and exchanged with an external application (EUBSO), in case the exchange process is carried out through an API, this will be launched, in case the exchange process is done with a document, this will be created and available to the end user	

Exchange Requirement Data Objects

<Input and results of EEM>

Type	Data Object
Name	*Input and results of EEM*
Documentation	*The data stored in DDBB is passed to the anonymization process*

<Anonymized input and results>

Type	Data Object
Name	*Anonymized input and results*
Documentation	*The results of anonymization process are stored in the DDBB*

<Exchanging results >

Type	Data Object
Name	*Exchanging results*
Documentation	*The anonymized results are extorted from de DDBB and passed to Data exchanging process*

<Anonymized process settings >

Type	Data Object
Name	*Anonymized process settings*
Documentation	*Some parameters of configuration the anonymization processes are provided from UI.* *Selection of buildings* *Selection of anonymization typology* *Selection of EPC typology* *etc*

<Managing exchanging information>

Type	Data Object
Name	*Anonymized process settings*
Documentation	*Some parameters of configuration of anonymization process are provided from UI*

LANE : < Storage>

The Storage LANE refers to the different databases that the ENMA system has.

Library Data Objects

<Harmonized Data Bases>

Type	Data Object
Name	*Harmonized Data Bases*
Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format. The Harmonised databases are a combination of data marts (based on MongoDB: a document database, which means it stores data in JSON-like documents) and long-term DB (based on Apache HBASE: HBase is an open source non-relational distributed database)*

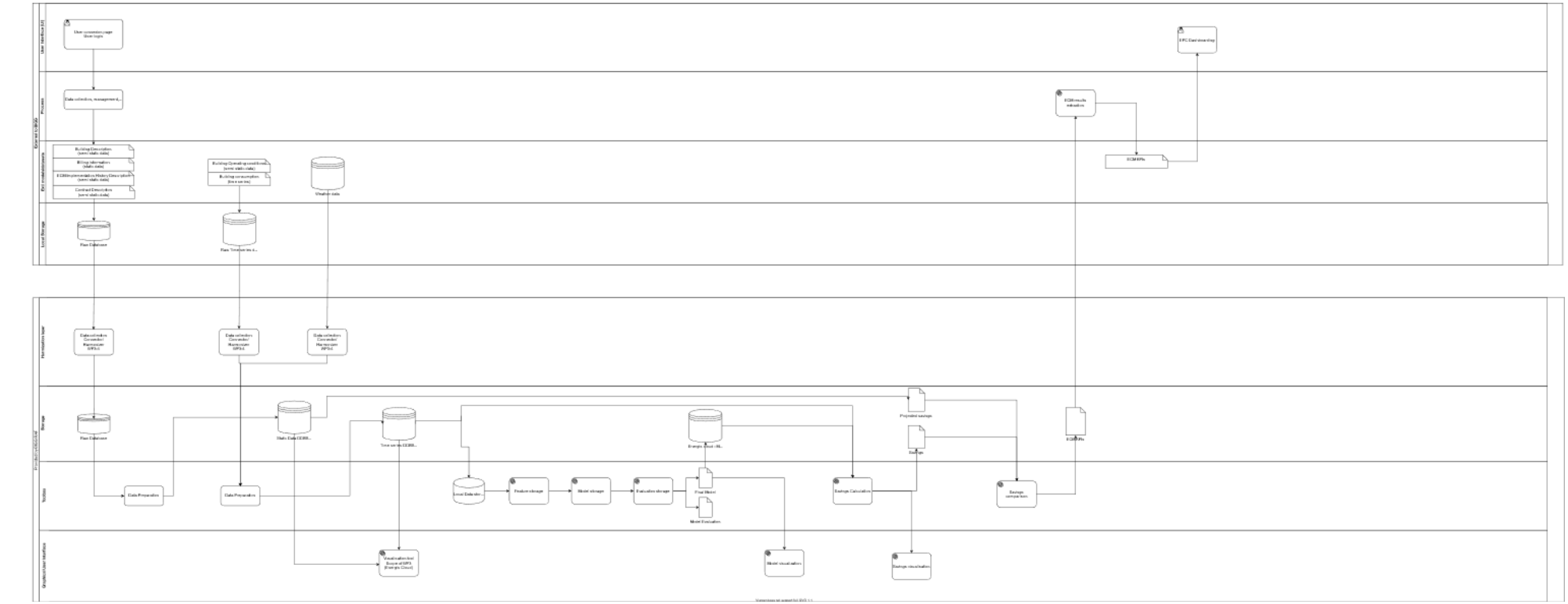
Exchange Requirement Data Objects

<Inputs and results of UC3>

Type	Data Object
Name	*Inputs and results of UC3*
Documentation	*The inputs and results of UC3, that must be used in this UC7, are stored in a data base in harmonized and prepared way.*

V.8. UC8

V.8.1. BPMN diagram


External to BIPV
User Interface (UI)
Process
Key considerations
Local Storage
User connection page
Show login
Data collection, management a...
Building Consumption
[semi-static data]
Billing information
[Static data]
ECEM implementation History Consumption
[semi-static data]
Contract Consumption
[semi-static data]
Building Consumption
[live session]
Building Consumption
[live session]
Weather data
Raw Database
Raw Timeseries a...
ECEM results
extraction
ECEM KPIs
EPC Dashboarding
Formulation layer
Storage
Toolbox
Visualization layer
Data collection
Convention
Harmonizer
(KPIs)
Data collection
Convention
Harmonizer
(KPIs)
Data collection
Convention
Harmonizer
(KPIs)
Raw Database
Data Preparation
Data Preparation
Static Data (EEBR...
Time series (EEBR...
Local Data stor...
Feature storage
Model storage
Evaluation storage
Final Model
Model Evaluation
Energy Cloud (SL...
Projected savings
Savings
Savings Calculation
Savings
comparison
ECEM KPIs
Visualization tool
Source of KPI
(Energis Cloud)
Model visualization
Savings visualization

V.8.2. Specification of processes

To develop the "Collection, Storage, viewing and Management of the EPCo related data " process, the following steps will be needed:

- Data collection

- Data harmonization

- Data preparation

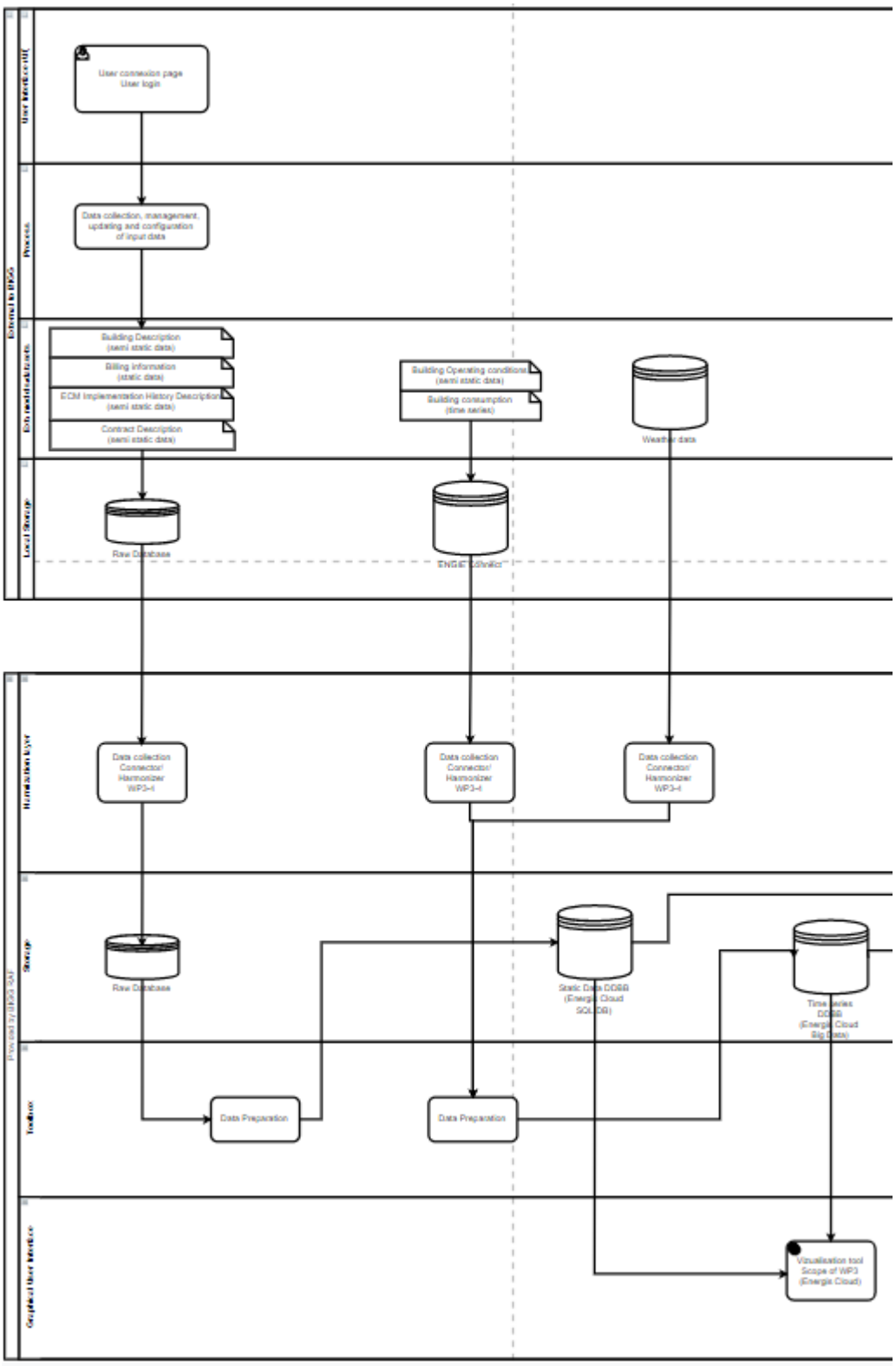
Process Map : <Data collection>

The data required for the UC8 can come from different data sources. These data sources can be loaded from existing databases or uploaded manually by the user.

In the case of manual data upload, the end user will be able to load, modify or delete data from the UI, whether this is static (descriptive characteristics of the building) or semi static (electricity tariffs).

In the case of data from existing databases (Time series (energy consumption, outdoor temperature), these can be collected directly through APIs or similar, the administrator user will have access from the UI to the import control of this data.

The building characteristics data, energy consumption or climate data will be systematically stored in raw in a Database. This Database will be located in the infrastructure of the existing Energis environment (owned by Energis).


User connexion page
User login
Data collection, management, updating and configuration of input data
Building Description (semi static data)
Billing information (static data)
ECM Implementation History Description (semi static data)
Contract Description (semi static data)
Building Operating conditions (semi static data)
Building consumption (time series)
Weather data
Raw Database
ENGIE Cofely
Data collection Connector/ Harmonizer WP3-I
Data collection Connector/ Harmonizer WP3-I
Data collection Connector/ Harmonizer WP3-I
Raw Database
Static Data DDBB (Energis Cloud SQL DB)
Time series DDBB (Energis Cloud Big Data)
Data Preparation
Data Preparation
Visualisation tool Scope of WP3 (Energis Cloud)

POOL : <External to BIGG>

Processes that will need to be undertaken by the user to enable the BIGG platform capabilities.

LANE : <User Interface (UI)>

In the UI the end user, depending on certain permissions, can upload, modify or delete the building data manually (through web forms or by uploading CSV or Excel files, with predefined format), or configure the collection of data from databases. existing data, external to this Pool

LANE : < Processes >

LANE : < Models/Data Sets >

This LANE describes the models and data items that must be managed in the process.

LANE : < Local Storage >

The Storage LANE refers to the different databases that the Energis Cloud service provides.

POOL : < Provided by BIGG platform>

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Harmonization layer>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies. This essentially refers to the scope of work covered in the Work Package 4 of the BIGG project organization.

LANE : < Storage >

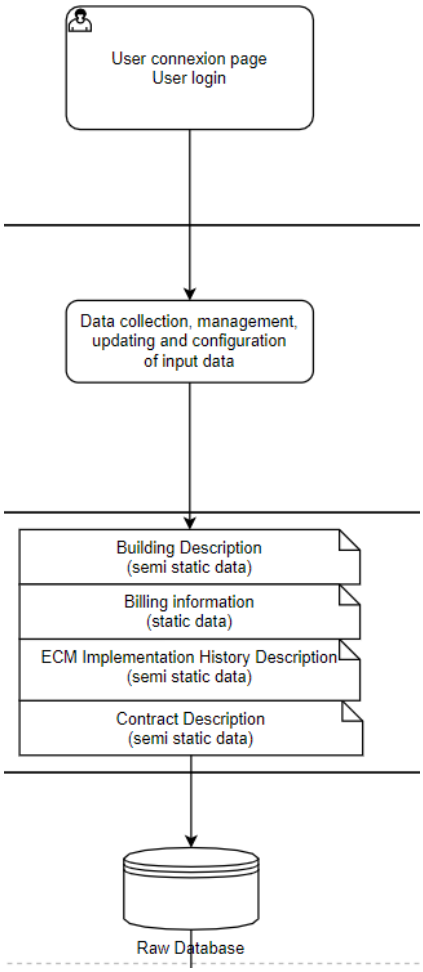
In the storage LANE, all the database and data object that will be stored on the BIGG storage service.

LANE : < Toolbox >

This LANE describes the Artificial intelligence modules and features performed by the BIGG platform to transform, evaluate, analyse, model, assess and process the data on the platform. This essentially refers to the scope of work covered in the Work Package 5 of the BIGG project organization.

LANE : < Graphical User Interface >

In the graphical user interface lane, the document described the services associated with the visualisation capabilities provided by the BIGG service. This essentially refers to the scope of work covered in the Work Package 3 of the BIGG project organization.



< Process > [ID:<BC4-UC08-001>]

Type	User Interface	
Name	Login and user connexion	
Documentation	*The end user connects to the service via a web portal and a user profile (login/password). From this online platform, the user will access the functions to collect data from different sources.*	

<Process > [ID:<BC4-UC08-002>]

Type	Data Collection	
Name	Data collection, management, updating and configuration of input data	
Documentation	*The end user will have the capabilities to link different data sources to the platform. To review the available data sources already uploaded, manage existing linked data sources and modify/update them if necessary*	

<Data Object Name>

Type	Data Object
Name	*Building description*
Documentation	*Information relative to the building and necessary to describe the building and identify it. Building ID, address, name, surface, coordinates, owner, manager, occupancy, Systems…*

<Data Object Name>

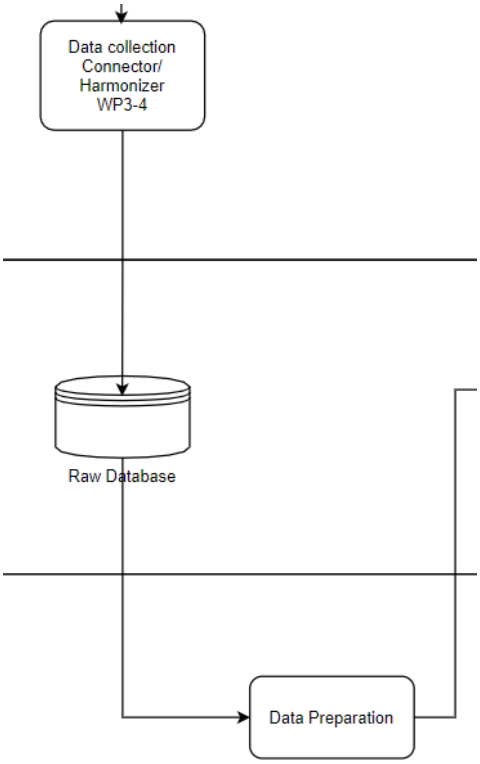
Type	Data Object
Name	*Billing information*
Documentation	*Static data* *Billing history information of the building. Repository of energy consumption (electricity, gas, water, biomasse, fuel, district heat, district cooling, steam …) , tariffs, contractual conditions (subscribed capacity)*

<Data Object Name>

Type	Data Object
Name	*ECM information*
Documentation	*Semi Static data* *Information relative to the Energy Conservation Measures already implemented at the site. Nature of the measure, date of implementation, scope of the modification…*

<Data Object Name>

Type	Data Object
Name	*Contract description*
Documentation	*Semi Static data* *Details of existing contracts having an impact on existing building operation such as maintenance contracts, energy supply contracts contracted by the building owner or the asset manager (terms of the contract, begin and end date, savings target, baseline definition…)*



<Process > [ID:<BC4-UC08-003>]

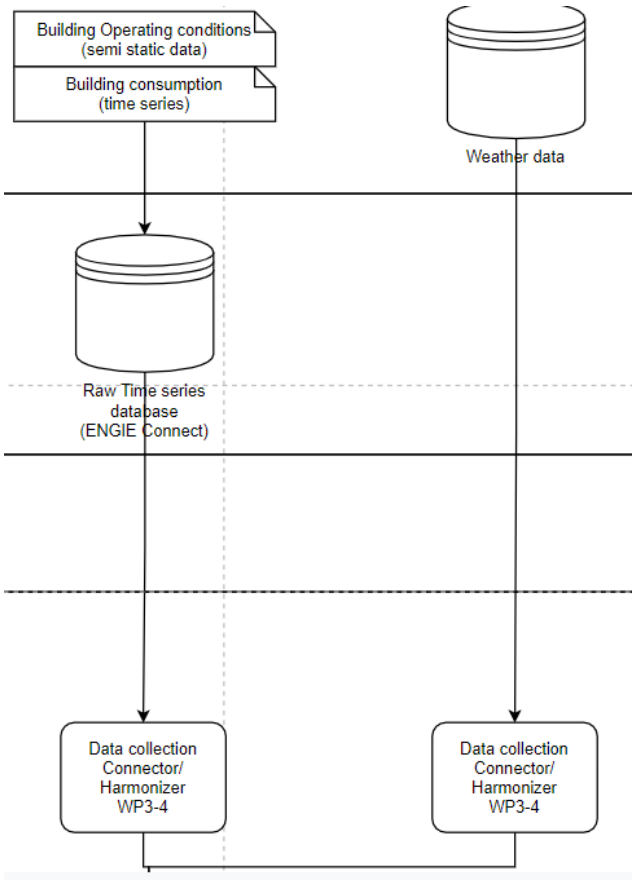
Type	Connector and harmonization	
Name	Data collection, connector and harmonization	
Documentation	*Data collection from the local service raw database to the BIGG platform storage. The data is collected either directly from a manual upload or through the BIGG platform API service. The data is collected and harmonized to match the BIGG data Standard 4 Building.*	

<Data Object Name>

Type	Data Object
Name	*Raw Database*
Documentation	*The static and semi static data describing the buildings, the contracts, the ECM, the billing information and historical consumption data are stored after they have been processed by the BIGG connector.*

<Process > [ID:<BC4-UC08-004>]

Type	Data Preparation	
Name	Data Preparation	
Documentation	*The raw database is then processed to make sure all the data stored is usable and processable through the BIGG service. The process aims at removing the outliers, missing items and inconsistent data.*	



<Data Object Name>

Type	Data Object
Name	*Building Operating conditions* *(semi static data)*
Documentation	All time series data streams collected on site through sensors and devices connected to the local service monitoring system or BMS.

<Data Object Name>

Type	Data Object
Name	*Building consumption* *(time series)*
Documentation	All time series data streams collected on site through sensors and devices connected to the local service monitoring system or BMS.

<Data Object Name>

Type	Data Object
Name	*Building consumption* *(time series)*
Documentation	All consumption time series data streams collected on site through meters and sub meters connected to the local service monitoring system or BMS. This data also includes consumption data streams coming from utility services or APIs.

<Process > **[ID:<BC4-UC08-005>]**

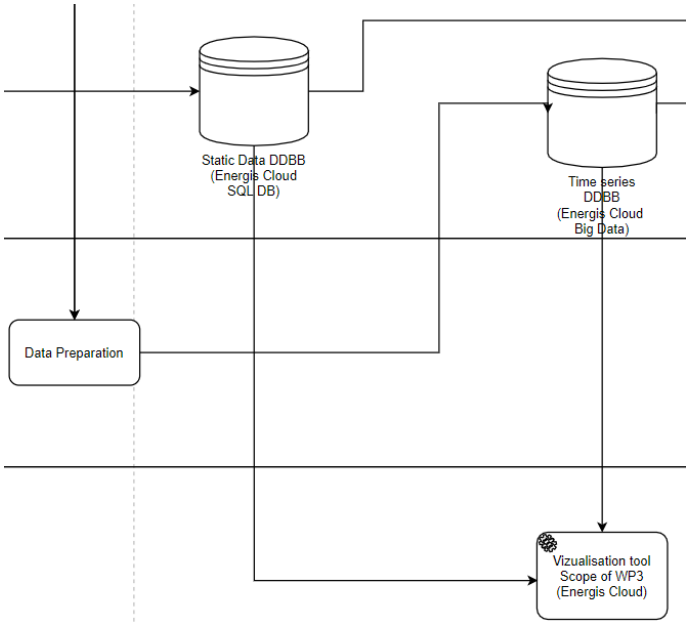
Type	Connector and harmonization	
Name	Data collection, connector and harmonization	
Documentation	*Data collection from the local service raw database to the BIGG platform storage. The data is collected either directly from a manual upload or through the BIGG platform API service. The data is collected and harmonized to match the BIGG data Standard 4 Building.*	

<Data Object Name>

Type	Data Object
Name	*Weather data*
Documentation	*The historical data regarding weather data (temperature, relative humidity, heating degree days, cooling degree days…)*

<Process > **[ID:<BC4-UC08-006>]**

Type	Connector and harmonization	
Name	Data collection, connector and harmonization	
Documentation	*Data collection from the weather data service database to the BIGG platform storage. The data is collected through the BIGG platform API. The data is collected and harmonized to match the BIGG data Standard 4 Building.*	



Library Data Objects

<Process > **[ID:<BC4-UC08-007>]**

Type	Data Preparation	
Name	Data Preparation	
Documentation	*The time series database is then processed to make sure all the data stored is usable and processable through the BIGG service. The process aims at removing the outliers, missing items and inconsistent data.*	

<Data Object Name>

Type	Data Object
Name	*Static Data DDBB* *(Energis Cloud SQL DB)*
Documentation	After the static raw database has been cleansed and prepared to matche the requirements of the AI data analysis service, the data is stored on the BIGG storage.

<Data Object Name>

Type	Data Object
Name	*Time series DDBB* *(Energis Cloud* *Big Data)*

Documentation	After the time series raw database has been cleansed and prepared to matche the requirements of the AI data analysis service, the data is stored on the BIGG storage.

<Process > **[ID:<BC4-UC08-008>]**

Type	User Interface	
Name	Vizualisation tool Scope of WP3 (Energis Cloud)	
Documentation	*The time series database is then processed to make sure all the data stored is usable and processable through the BIGG service. The process aims at removing the outliers, missing items and inconsistent data.*	

V.9. UC9

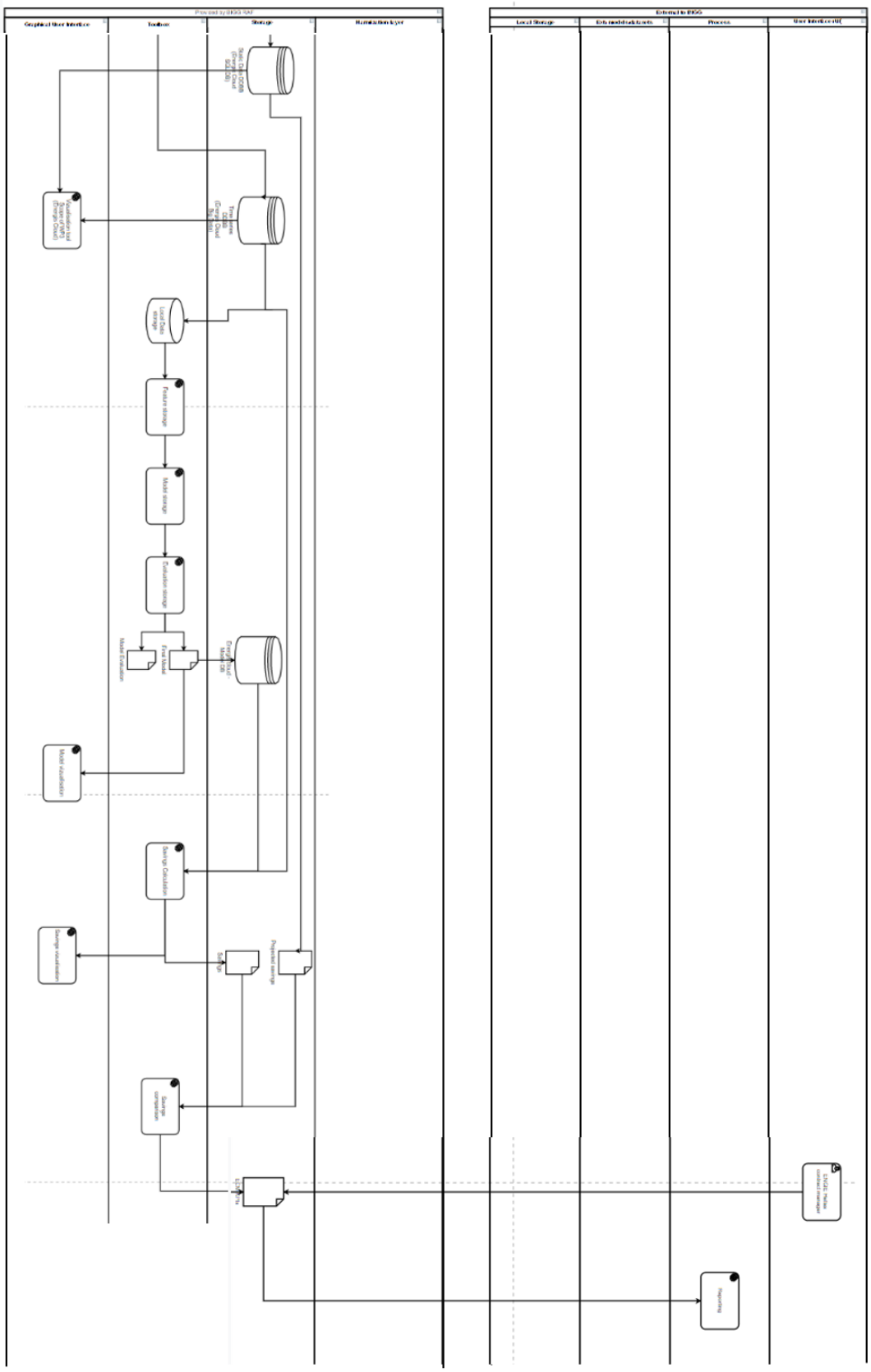
V.9.1. BPMN diagram

See V.8.1.

V.9.2. Specification of processes

To develop the "Collection, Storage, viewing and Management of the EPCo related data " process, the following steps will be needed:

- Prepared Data collection (data prepared through Use Case 8 workflow)
- Data storage
- Feature storage
- Model Storage
- Evaluation Storage
- Model Storage
- Savings calculation



POOL : <External to BIGG>

Processes that will need to be undertaken by the user to enable the BIGG platform capabilities.

LANE : <User Interface (UI)>

In the UI the end user, depending on certain permissions, can upload, modify or delete the building data manually (through web forms or by uploading CSV or Excel files, with predefined format), or configure the collection of data from databases. existing data, external to this Pool

LANE : < Processes >

LANE : < Models/Data Sets >

This LANE describes the models and data items that must be managed in the process.

LANE : < Local Storage >

The Storage LANE refers to the different databases that the Energis Cloud service provides.

POOL : < Provided by BIGG platform>

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Harmonization layer>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies. This essentially refers to the scope of work covered in the Work Package 4 of the BIGG project organization.

LANE : < Storage >

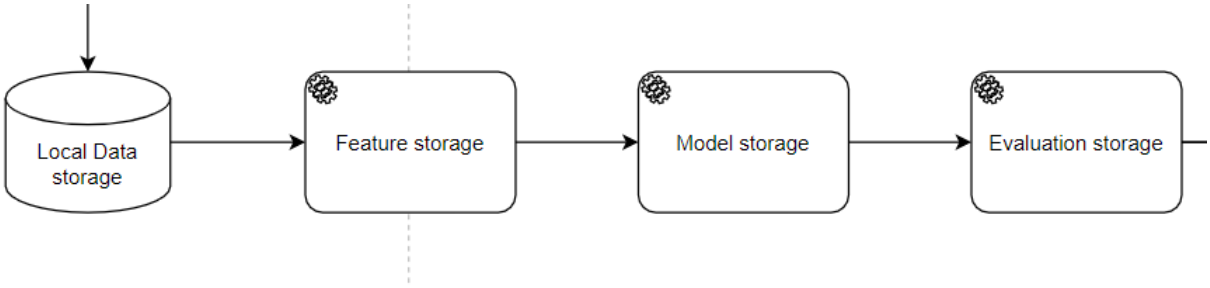
In the storage LANE, all the database and data object that will be stored on the BIGG storage service.

LANE : < Toolbox >

This LANE describes the Artificial intelligence modules and features performed by the BIGG platform to transform, evaluate, analyse, model, assess and process the data on the platform. This essentially refers to the scope of work covered in the Work Package 5 of the BIGG project organization.

LANE : < Graphical User Interface >

In the graphical user interface lane, the document described the services associated with the visualisation capabilities provided by the BIGG service. This essentially refers to the scope of work covered in the Work Package 3 of the BIGG project organization.



<Data Object Name>

Type	Data Object
Name	*Local data storage*
Documentation	*The time series data is stored locally (where the AI toolbox process is taking place) so that the AI toolbox computation can be performed quickly and efficiently.*

< Process > **[ID:<BC4-UC09-001>]**

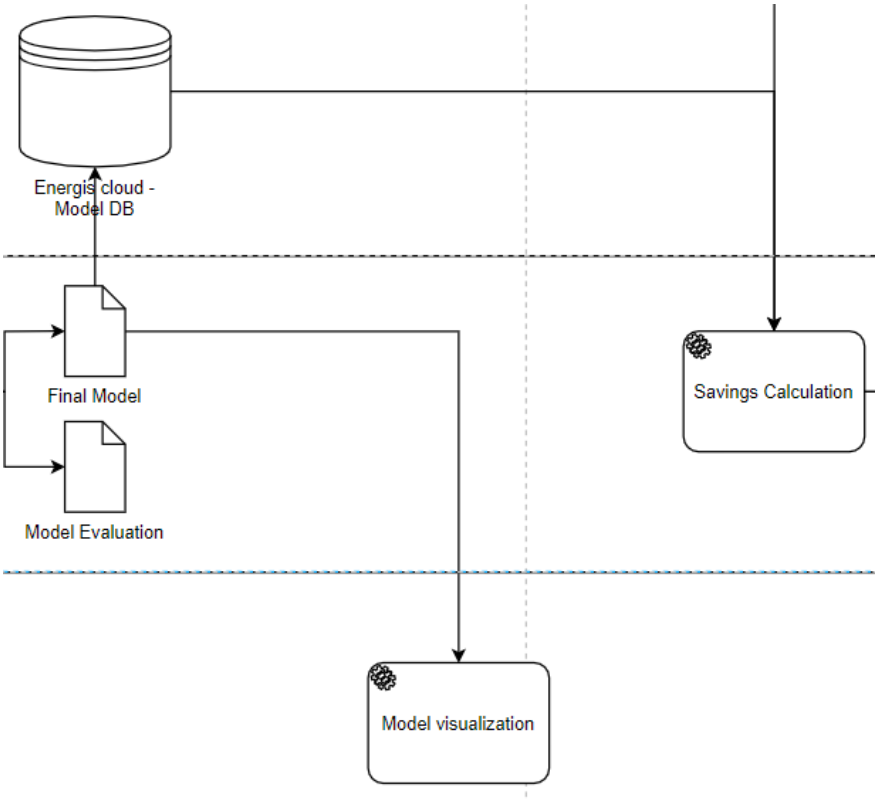
Type	AI Process	
Name	Feature Storage	
Documentation	*The data is analysed and certain patterns and characteristics are extracted from the data and stored for future usage in AI algorithms (e.g. weekly consumption pattern).*	

<Process > **[ID:<BC4-UC09-002>]**

Type	AI Process	
Name	Model storage	
Documentation	*A regression model is identified based on the input data, the previously extracted features and an AI model library. This model library consists of several regression models that must thereafter be compared with each other.*	

<Process > **[ID:<BC4-UC09-003>]**

Type	AI Process
Name	*Evaluation storage*
Documentation	*The quality of each model is evaluated based on certain statistical criteria (e.g. root mean square error, bias, etc.) Evaluation is used to select the best model and is also a quality certificate for the best model that is provided to the user.*



<Data Object Name>

Type	Data Object
Name	*Energis cloud model database*
Documentation	*Storage of the evaluated models and their evaluations for future use in the calculation of the savings*

<Data Object Name>

Type	Data Object
Name	*Final Model*
Documentation	*The final model is the mathematical representation of the building consumption over time with its dependencies to the different influence factors. The model is being challenged over time but the final version of the model is stored as a semi static data on the BIGG storage system.*

<Data Object Name>

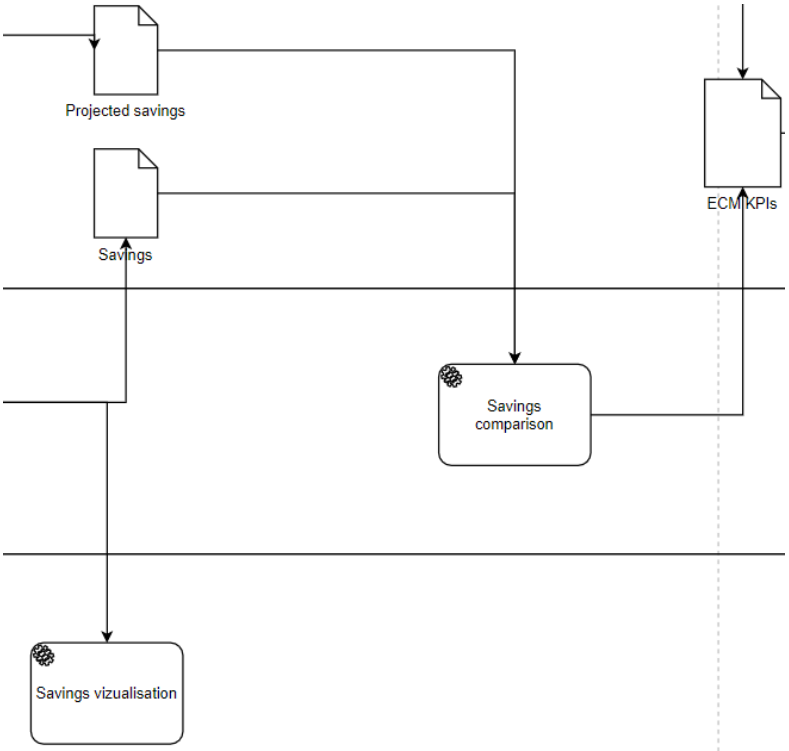
Type	Data Object
Name	*Model Evaluation*
Documentation	*Beside the final model, the model evaluation also is stored so that the following version of the model can be challenged and compared .*

<Process > [ID:<BC4-UC09-004>]

Type	Data visualization	
Name	Model Visualization	
Documentation	*Data collection from the local service raw database to the BIGG platform storage. The data is collected either directly from a manual upload or through the BIGG platform API service. The data is collected and harmonized to match the BIGG data Standard 4 Building.*	

<Process > **[ID:<BC4-UC09-005>]**

Type	AI Process	
Name	Savings calculation	
Documentation	*Once the building consumption can be modelled accurately, it is used to assess the savings of a defined ECM. The building modelled consumption is calculated from the date of implementation of the ECM and then compared with the real measured consumption of the asset.*	



<Data Object Name>

Type	Data Object
Name	Projected savings
Documentation	Projected savings are identified from the ECM nature, the implemented date and assessed impact on the site consumption. This is a predefined number that is defined contractually and part of the EPCo. The projected savings are recovered from the building raw database as defined in the UC8 process flow.

<Data Object Name>

Type	Data Object
Name	*Savings*
Documentation	Savings is the data object that is produced from the comparison between the asset energy consumption model and the measured asset consumption. The results of this comparison is being stored as a data item in the BIGG data storage.

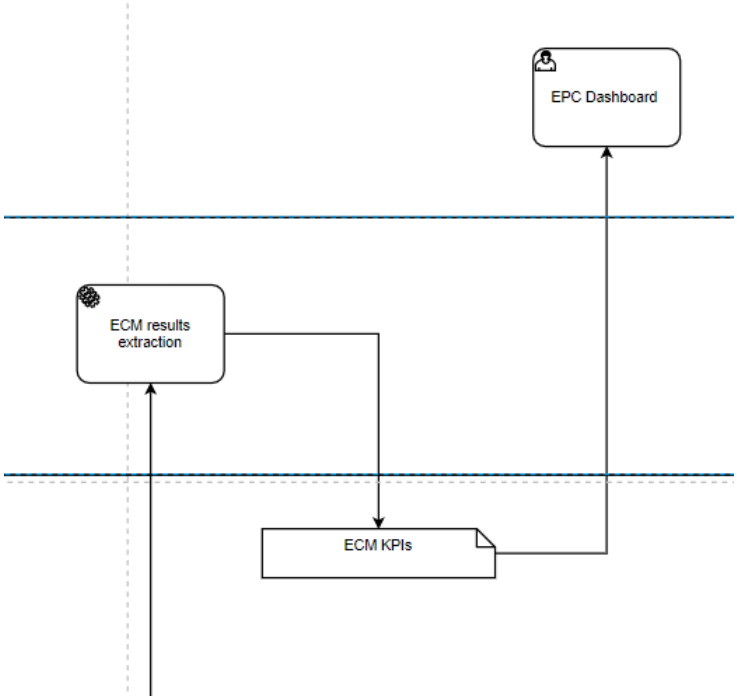
<Process > **[ID:<BC4-UC09-006>]**

Type	AI Process	
Name	Savings comparison	
Documentation	*Savings comparison is the process that compares the calculated savings with the projected ones. The results of this comparison are*	

	an indicator of the on-going situation of the EPCo. It allows to assess how the EPCo management team is performing.

<Data Object Name>

Type	Data Object
Name	*ECM KPI*
Documentation	From the Savings comparison results, the BIGG service outputs a report of the on-going conditions of the EPC contractor performances. The KPI will include for a given ECM, the ECM profitability (NPV, updated ROI projection), alignment with the contractual terms, projected compared with the real savings estimation.



Library Data Objects

<Process > [ID:<BC4-UC09-007>]

Type	Data extraction	
Name	ECM Results extraction	
Documentation	*The data from the BIGG platform analysis of the ECM performances is being retrieved by the local service. The extraction module allows for a choice of the extraction method and the format of extraction.*	

<Data Object Name>

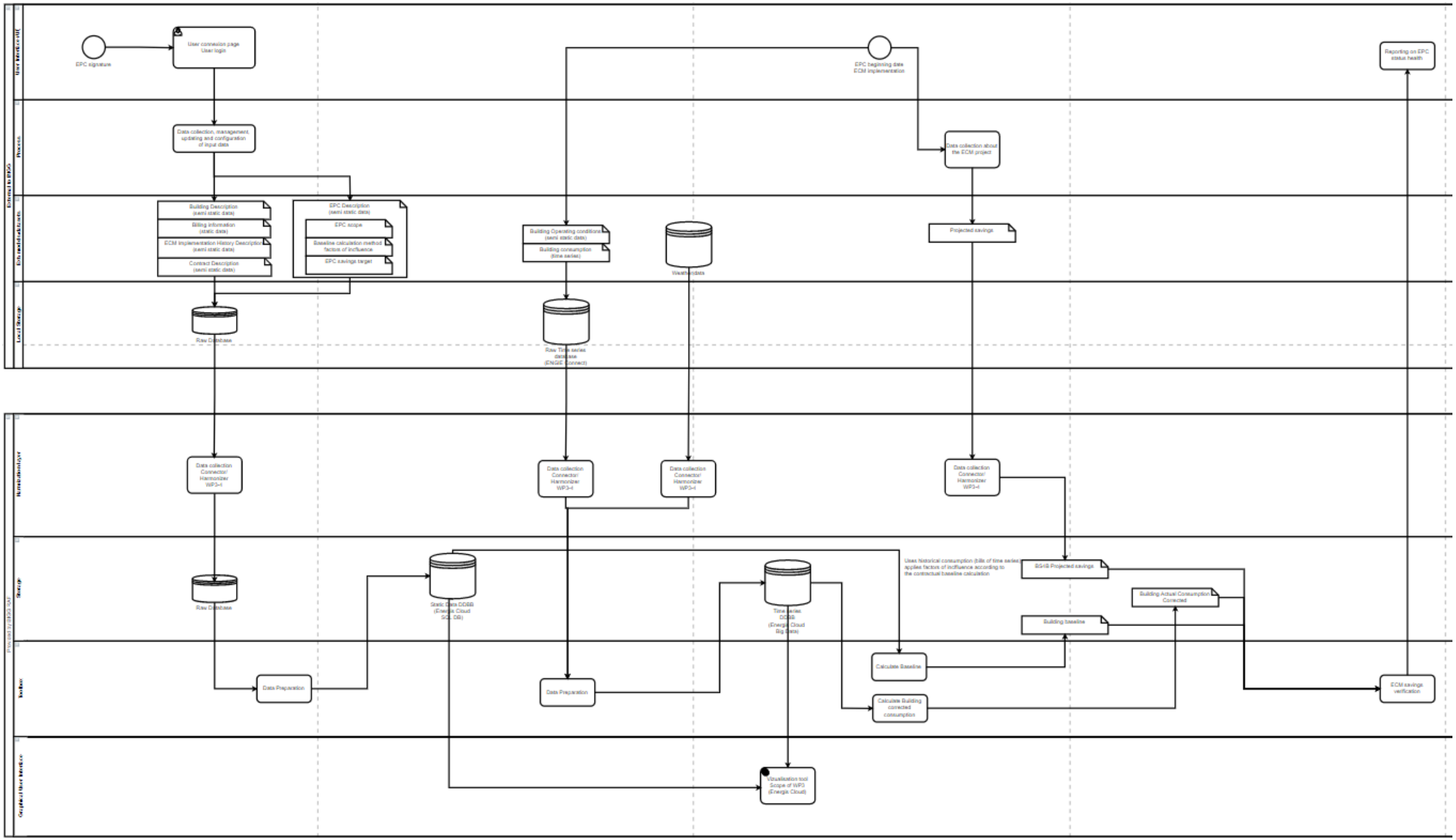
Type	Data Object
Name	*ECM KPIs*
Documentation	The indicators reflecting the ECM performances are exported and stored on a given format (chosen by the user on the extraction module). The resulted extraction is stored on the local service storage.

<Process > [ID:<BC4-UC09-008>]

Type	User Interface	
Name	EPC Dashboarding	
Documentation	*The local service allows the creation of a dashboard presenting the results produced by the BIGG platform analysis.*	

V.10. UC10

V.10.1. BPMN diagram


EPC signature
User connexion page
User login
Data collection, management, updating and configuration of input data
Building Description (semi static data)
Billing information (static data)
ECM implementation History Description (semi static data)
Contract Description (semi static data)
EPC Description (semi static data)
EPC scope
Baseline calculation method factors of influence
EPC savings target
Raw Database
Building Operating conditions (semi static data)
Building consumption (time series)
Weatherdata
Raw Time series database (ENGIE Connect)
EPC beginning date ECM implementation
Data collection about the ECM project
Projected savings
Reporting on EPC status health
Data collection Connector/ Harmonizer WP3-4
Data collection Connector/ Harmonizer WP3-4
Data collection Connector/ Harmonizer WP3-4
Data collection Connector/ Harmonizer WP3-4
Raw Database
Static Data CDBB (Energik Cloud SQL DB)
Time series DDBB (Energik Cloud Big Data)
Uses historical consumption (bills of time series) applies factors of influence according to the contractual baseline calculation
EG+S Projected savings
Building Actual Consumption Corrected
Building baseline
Data Preparation
Data Preparation
Calculate Baseline
Calculate Building corrected consumption
ECM savings verification
Visualization tool Scope of WP3 (Energik Cloud)

V.10.2. Specification of processes

<External to BIGG>

Processes that will need to be undertaken by the user to enable the BIGG platform capabilities.

LANE : <User Interface (UI)>

In the UI the end user, depending on certain permissions, can upload, modify or delete the building data manually (through web forms or by uploading CSV or Excel files, with predefined format), or configure the collection of data from databases. existing data, external to this Pool

LANE : < Processes >

LANE : < Models/Data Sets >

This LANE describes the models and data items that must be managed in the process.

LANE : < Local Storage >

The Storage LANE refers to the different databases that the Energis Cloud service provides.

POOL : < Provided by BIGG platform>

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE : <Harmonization layer>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies. This essentially refers to the scope of work covered in the Work Package 4 of the BIGG project organization.

LANE : < Storage >

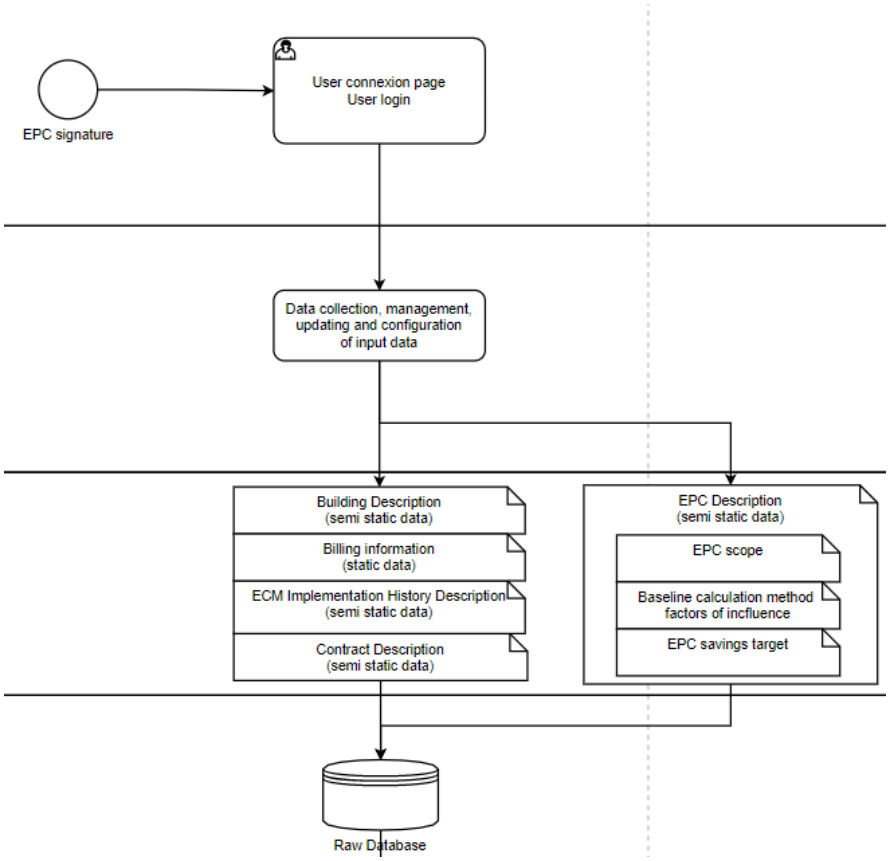
In the storage LANE, all the database and data object that will be stored on the BIGG storage service.

LANE : < Toolbox >

This LANE describes the Artificial intelligence modules and features performed by the BIGG platform to transform, evaluate, analyse, model, assess and process the data on the platform. This essentially refers to the scope of work covered in the Work Package 5 of the BIGG project organization.

LANE : < Graphical User Interface >

In the graphical user interface lane, the document described the services associated with the visualisation capabilities provided by the BIGG service. This essentially refers to the scope of work covered in the Work Package 3 of the BIGG project organization.



< Process > [ID:<BC4-UC010-001>]

Type	User Interface	
Name	Login and user connexion	
Documentation	*The end user connects to the service via a web portal and a user profile (login/password). From this online platform, the user will access the functions to collect data from different sources.*	

<Process > [ID:<BC4-UC010-002>]

Type	Data Collection	
Name	Data collection, management, updating and configuration of input data	
Documentation	*The end user will have the capabilities to link different data sources to the platform. To review the available data sources already uploaded, manage existing linked data sources and modify/update them if necessary*	

<Data Object Name>

Type	Data Object
Name	*Building description*

Documentation	*Information relative to the building and necessary to describe the building and identify it. Building ID, address, name, surface, coordinates, owner, manager, occupancy, Systems…*

<Data Object Name>

Type	Data Object
Name	*Billing information*
Documentation	*Static data* *Billing history information of the building. Repository of energy consumption (electricity, gas, water, biomasse, fuel, district heat, district cooling, steam …) , tariffs, contractual conditions (subscribed capacity)*

<Data Object Name>

Type	Data Object
Name	*ECM information*
Documentation	*Semi Static data* *Information relative to the Energy Conservation Measures already implemented at the site. Nature of the measure, date of implementation, scope of the modification…*

<Data Object Name>

Type	Data Object
Name	*Contract description*
Documentation	*Semi Static data* *Details of existing contracts having an impact on existing building operation such as maintenance contracts, energy supply contracts contracted by the building owner or the asset manager (terms of the contract, begin and end date, savings target, baseline definition…)*

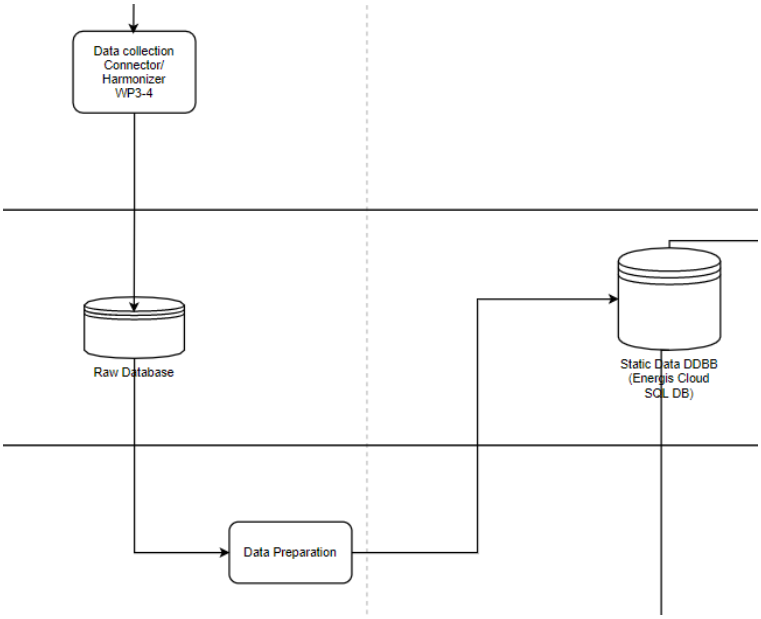
<Data Object Name>

Type	Data Object
Name	*EPCo description*
Documentation	*Semi Static data* *Details of existing Energy Performance contracts between the building owner and the energy service company. This data object includes all the data describing the terms of the on-going EPCo. It should include at least the following information:*

	The EPC technical scope describing the perimeter on which the EPCo applies (equipment, buildings, zones) Baseline calculation method. Details of the calculation process of the building baseline. This process usually leverages the IPMVP protocol to define the baseline conditions and the method to compare it to the actual building consumption EPCo savings target. Description of the commitment of the Energy services company in terms of savings through the EPCo. Definition of the projected savings and its impact on the Energy service company's annual fee.

<Data Object Name>

Type	Data Object
Name	*Rax Database*
Documentation	*Semi Static data* All the static and semi static data described above is centralized on a local database so it can be accessed by the BIGG service platform.



<Process > **[ID:<BC4-UC010-003>]**

Type	Connector and harmonization
Name	Data collection, connector and harmonization
Documentation	*Data collection from the local service raw database to the BIGG platform storage. The data is collected either directly from a manual upload or through the BIGG platform API service. The data is collected and harmonized to match the BIGG data Standard 4 Building.*

<Data Object Name>

Type	Data Object
Name	*Raw Database*
Documentation	*The static and semi static data describing the buildings, the contracts, the ECM, the billing information and historical consumption data are stored after they have been processed by the BIGG connector.*

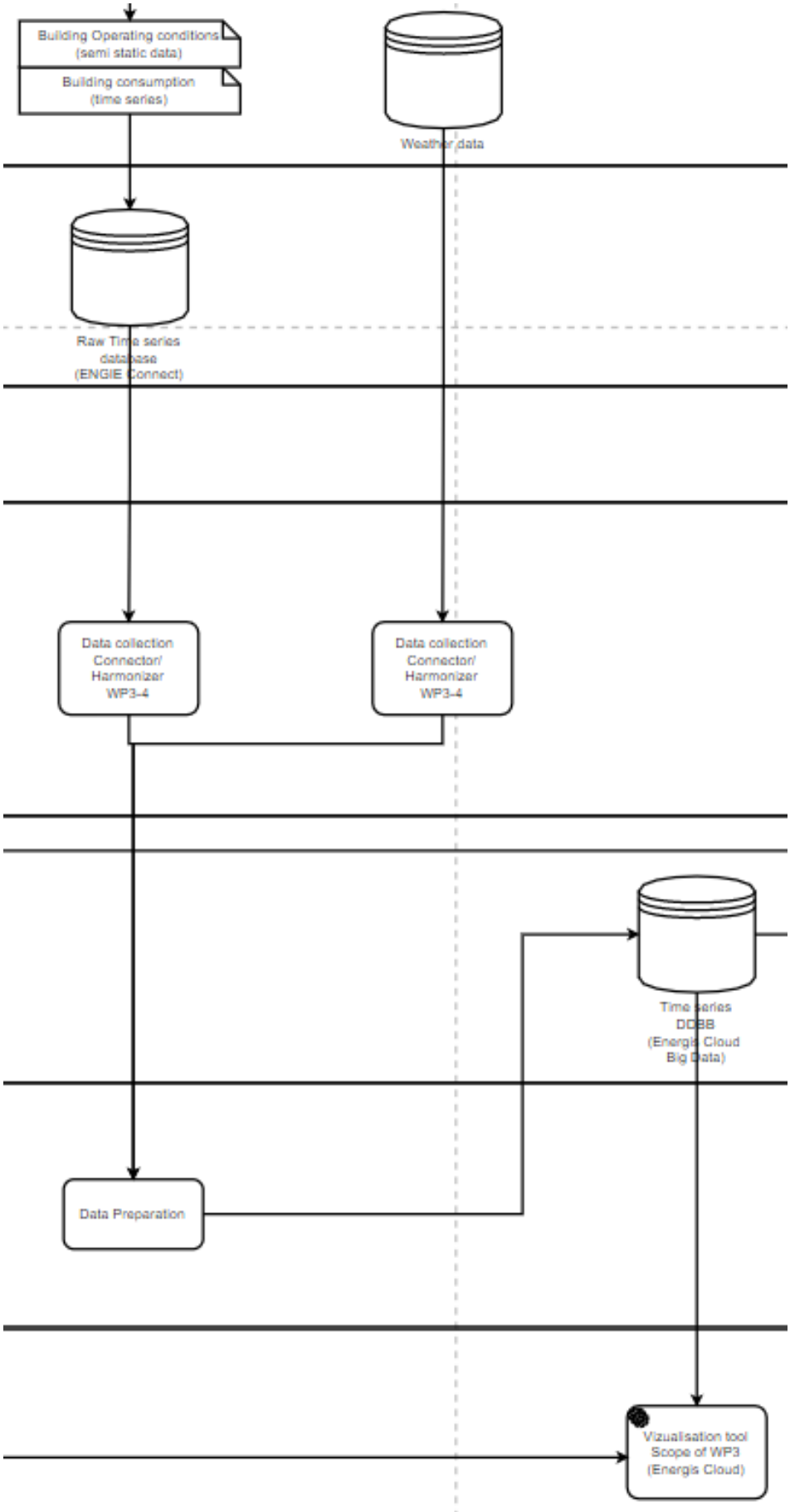
<Process > **[ID:<BC4-UC010-004>]**

Type	Data Preparation
Name	Data Preparation
Documentation	*The raw database is then processed to make sure all the data stored is usable and processable through the BIGG service. The process aims at removing the outliers, missing items and inconsistent data.*

<Data Object Name>

Type	Data Object
Name	*Static Database DDBB*

Documentation	*The static and semi static data describing the buildings, the contracts, the ECM, the billing information and historical consumption data are stored after they have been processed and harmonized by the BIGG platform connector.*


Building Operating conditions
(semi static data)
Building consumption
(time series)
Weather data
Raw Time series
database
(ENGIE Connect)
Data collection
Connector/
Harmonizer
WP3-4
Data collection
Connector/
Harmonizer
WP3-4
Time series
DDBB
(Energis Cloud
Big Data)
Data Preparation
Vizualisation tool
Scope of WP3
(Energis Cloud)

<Data Object Name>

Type	Data Object
Name	*Building Operating conditions* *(semi static data)*
Documentation	All time series data streams collected on site through sensors and devices connected to the local service monitoring system or BMS.

<Data Object Name>

Type	Data Object
Name	*Building consumption* *(time series)*
Documentation	All time series data streams collected on site through sensors and devices connected to the local service monitoring system or BMS.

<Data Object Name>

Type	Data Object
Name	*Raw time series data base*
Documentation	All consumption time series data streams collected on site through meters and sub meters connected to the local service monitoring system or BMS. This data also includes consumption data streams coming from utility services or APIs.

<Process > **[ID:<BC4-UC010-005>]**

Type	Connector and harmonization	
Name	Data collection, connector and harmonization	
Documentation	*Data collection from the local service raw database to the BIGG platform storage. The data is collected either directly from a manual upload or through the BIGG platform API service. The data is collected and harmonized to match the BIGG data Standard 4 Building.*	

<Data Object Name>

Type	Data Object
Name	*Weather data*
Documentation	*The historical data regarding weather data (temperature, relative humidity, heating degree days, cooling degree days…)*

<Process > **[ID:<BC4-UC10-006>]**

Type	Connector and harmonization	
Name	Data collection, connector and harmonization	
Documentation	*Data collection from the weather data service database to the BIGG platform storage. The data is collected through the BIGG platform API. The data is collected and harmonized to match the BIGG data Standard 4 Building.*	

Process > **[ID:< BC4-UC010-007>]**

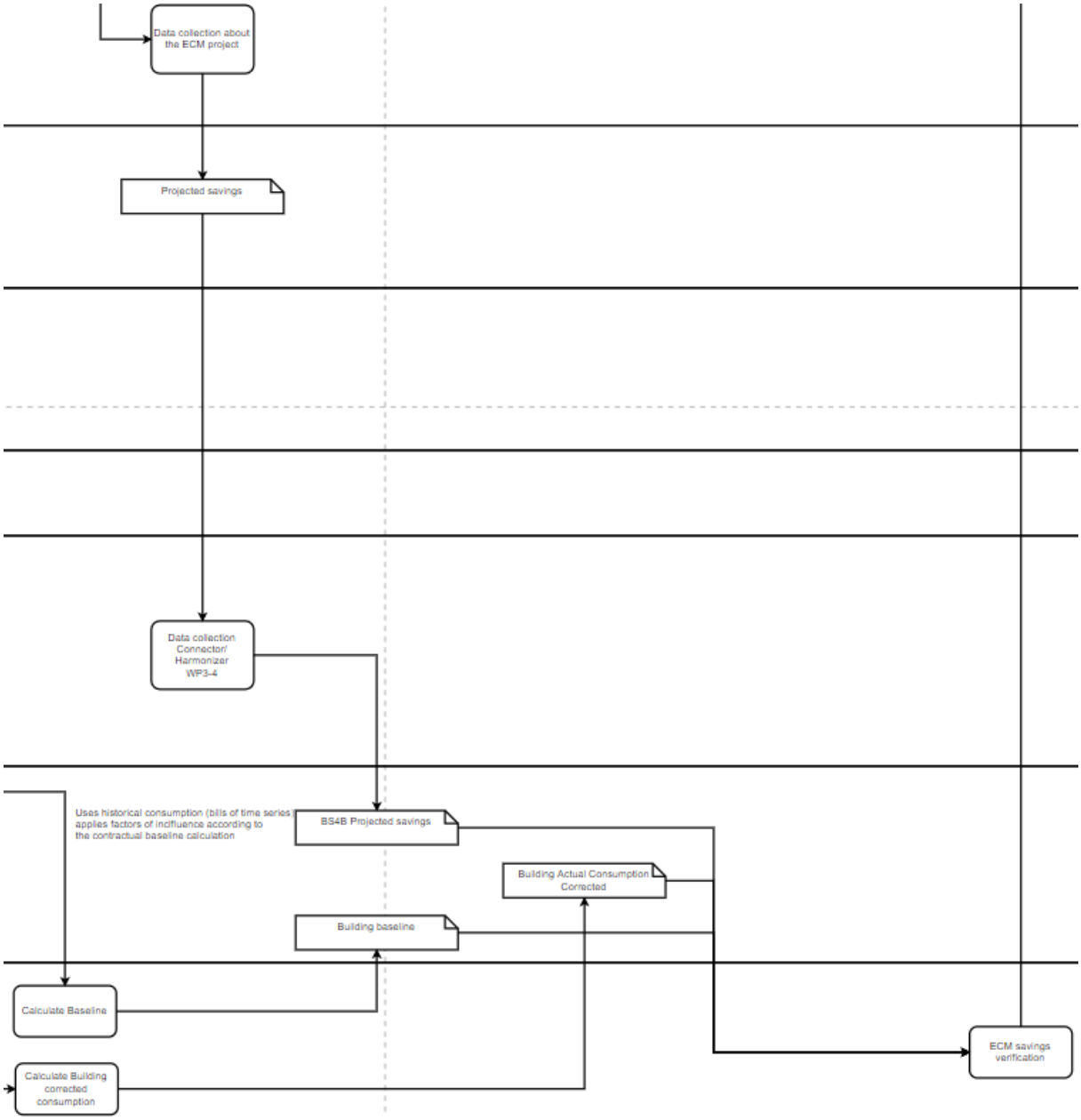
Type	Data Preparation	
Name	Data Preparation	
Documentation	*The time series database is then processed to make sure all the data stored is usable and processable through the BIGG service. The process aims at removing the outliers, missing items and inconsistent data.*	

<Data Object Name>

Type	Data Object
Name	*Time series DDBB*
Documentation	*The data stored ready for utilisation by the BIGG toolbox features*

Process > **[ID:< BC4-UC010-008>]**

Type	Data Visualization	
Name	Energis Cloud VisualizationTool	
Documentation	*Capability to visualize the different data streams directly on the platform.*	



Process > **[ID:< BC4-UC010-009>]**

Type	Data Collection	
Name	ECM Project data collection	
Documentation	*The energy service company collects all relevant information allowing to describe the ECM. This data should include:* Nature of the ECM Projected savings Date of ECM implementation ECM provider	

	Description of the equipment pre and post ECM *Description of the operation pre and post ECM*

<Data Object Name>

Type	Data Object
Name	*ECM description*
Documentation	*This data object being the repository of the information describing the ECM as mentioned in the ECM project data collection above*

Process > **[ID:< BC4-UC010-010>]**

Type	Data Harmonization	
Name	Data collection Connector/ Harmonizer WP3-4	
Documentation	*Data collection from the ECM data object to the BIGG platform storage. The data is collected through the BIGG platform API. The data is then harmonized to match the BIGG data Standard 4 Building...*	

<Data Object Name>

Type	Data Object
Name	*BS4B ECM description*
Documentation	*The ECM data object after harmonization. Now under the BIGG standard 4 Building format*

Process > **[ID:< BC4-UC010-011>]**

Type	AI Process	
Name	Baseline calculation	
Documentation	*Computation of the baseline calculation process as defined in the Energy performance contract definition to generate the building baseline data object*	

<Data Object Name>

Type	Data Object
Name	*BS4B Building baseline*
Documentation	*The data object describing the building baseline. Conditions describing the building energy consumption as projected under the contract defition*

Process > **[ID:< BC4-UC010-012>]**

Type	AI Process	
Name	Calculate Building corrected consumption	
Documentation	*Computation of the Building actual consumption after taking into account the impact of the factors of influence as described in the contract terms. The process of calculation will be following the IPMVP general methodology.to account for factors of influence such as occupancy and weather conditions.*	

<Data Object Name>

Type	Data Object
Name	*BS4B Building Actual Consumption* *Corrected*
Documentation	*The data object describing the building actual consumption and performances.*

Process > **[ID:< BC4-UC010-013>]**

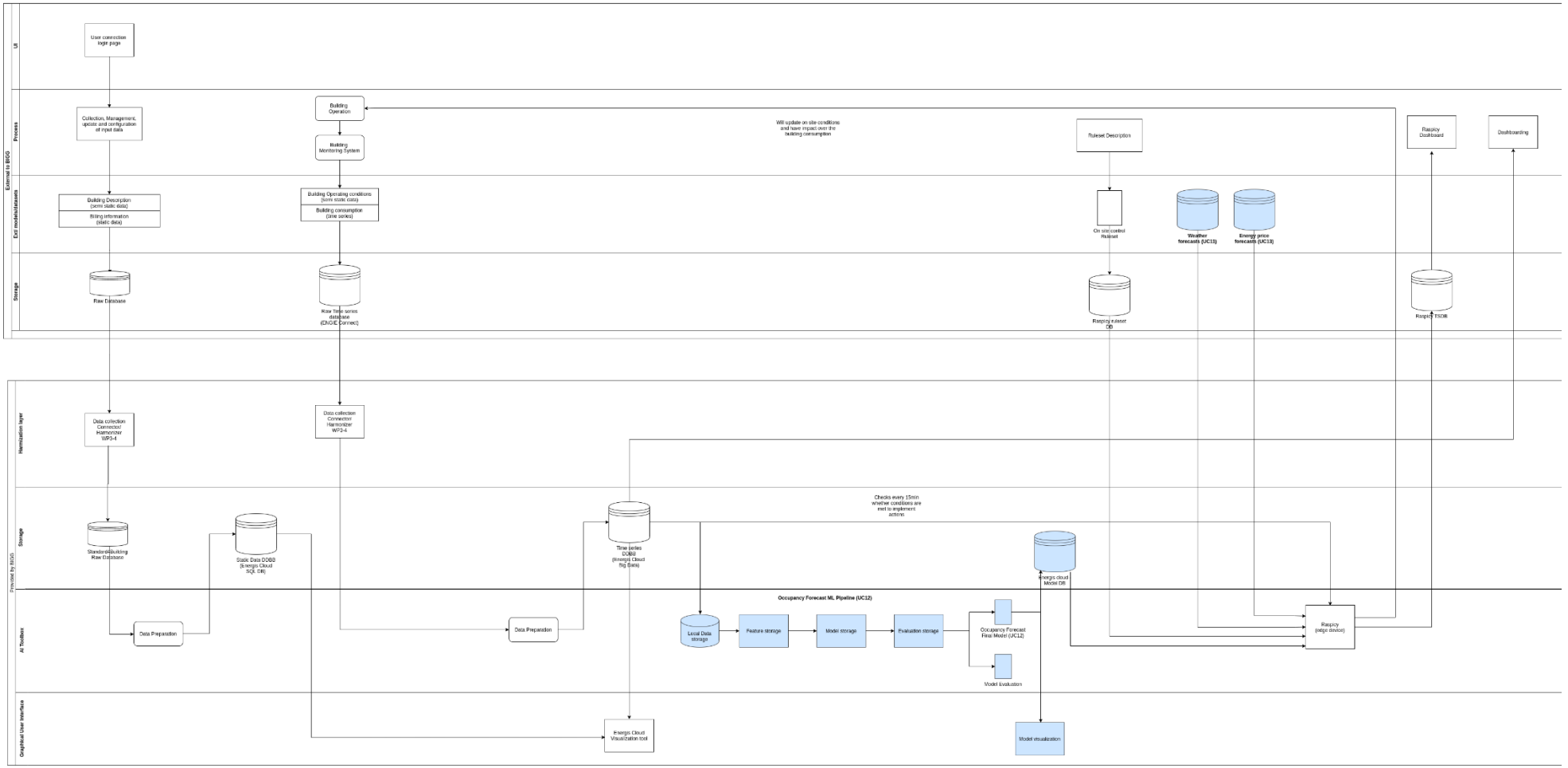
Type	AI Process	
Name	ECM savings verification	
Documentation	*Calculation process in which the BIGG platform compares the building baseline with the corrected building actual consumption to quantify the actual results of the ECM implementation in terms of energy savings. The savings are then compared to the projected savings (as agreed on the EPCo).The results enable the contract manager to track the overall status of the EPCo and make sure the current building performances are in line with the commitment taken as part of the EPCo.*	

Process > **[ID:< BC4-UC010-014>]**

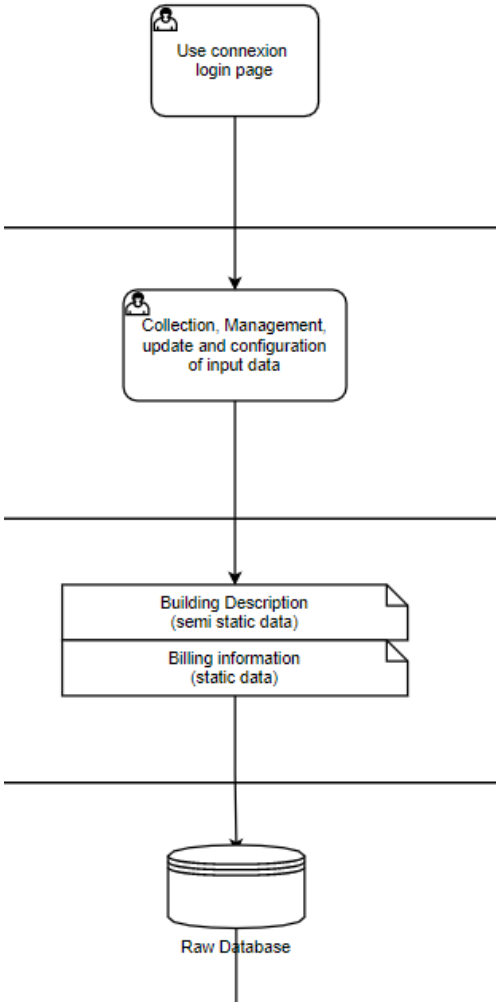
Type	Data visualization	
Name	Reporting on EPCo status health	
Documentation	*Reporting capabilities on the local application to display the results of the ECM actual savings calculation as well as the general status of the EPCo in terms of performances compared with the contractual commitments.*	

V.11. UC11-12-13

V.11.1. BPMN diagram


External to BIGG
UI
Process
Exist models/datasets
Storage
User connection login page
Collection, Management, update and configuration of input data
Building Description (semi static data)
Billing information (static data)
Raw Database
Building Operation
Building Monitoring System
Building Operating conditions (semi static data)
Building consumption (time series)
Raw Time series database (EN/CIE Connect)
Will update on site conditions and have impact over the building consumption
Ruleset Description
On site control Ruleset
Rasplcy ruleset DB
Weather forecasts (UC11)
Energy price forecasts (UC13)
Rasplcy Dashboard
Dashboarding
Rasplcy TSDB
Provided by BIGG
Harmization layer
Storage
AI Toolbox
Graphical User Interface
Data collection Connector/ Harmonizer WP3-4
Data collection Connector Harmonizer WP2-4
Standard Building Raw Database
Static Data DOB2 (Energis Cloud SQL DB)
Time series DOB2 (Energis Cloud Big Data)
Checks every 15min whether conditions are met to implement actions
Energis cloud Model DB
Occupancy Forecast ML Pipeline (UC12)
Data Preparation
Data Preparation
Local Data storage
Feature storage
Model storage
Evaluation storage
Occupancy Forecast Final Model (UC12)
Model Evaluation
Rasplcy (edge device)
Energis Cloud Visualization tool
Model visualization

V.11.2. Specification of processes



< Process > [ID:<BC5-UC13-001>]

Type	User Interface	
Name	Login and user connexion	
Documentation	*The end user connects to the service via a web portal and a user profile (login/password). From this online platform, the user will access the functions to collect data from different sources.*	

<Process > [ID:< BC5-UC13-002>]

Type	Data Collection	
Name	Data collection, management, updating and configuration of input data	

Documentation	*The end user will have the capabilities to link different data sources to the platform. To review the available data sources already uploaded, manage existing linked data sources and modify/update them if necessary*

<Data Object Name>

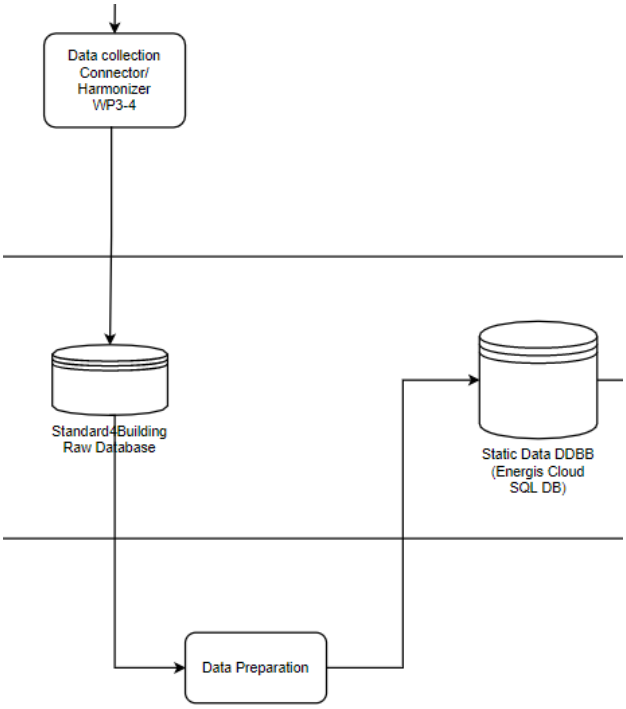
Type	Data Object
Name	*Building description*
Documentation	*Information relative to the building and necessary to describe the building and identify it. Building ID, address, name, surface, coordinates, owner, manager, occupancy, Systems…*

<Data Object Name>

Type	Data Object
Name	*Billing information*
Documentation	*Static data* *Billing history information of the building. Repository of energy consumption (electricity, gas, water, biomass, fuel, district heat, district cooling, steam …) , tariffs, contractual conditions (subscribed capacity)*

<Data Object Name>

Type	Data Object
Name	*Raw Database*
Documentation	*Static and Semi Static data* *A central repository of the data collected about the assets*



<Process > [ID:< BC5-UC13-003>]

Type	Connector and harmonization	
Name	Data collection, connector and harmonization	
Documentation	*Data collection from the local service raw database to the BIGG platform storage. The data is collected either directly from a manual upload or through the BIGG platform API service. The data is collected and harmonized to match the BIGG data Standard 4 Building.*	

<Data Object Name>

Type	Data Object
Name	*Raw Database*
Documentation	*The static and semi static data describing the buildings, the contracts, the ECM, the billing information and historical consumption data are stored after they have been processed and harmonized by the BIGG standard4buildingconnector.*

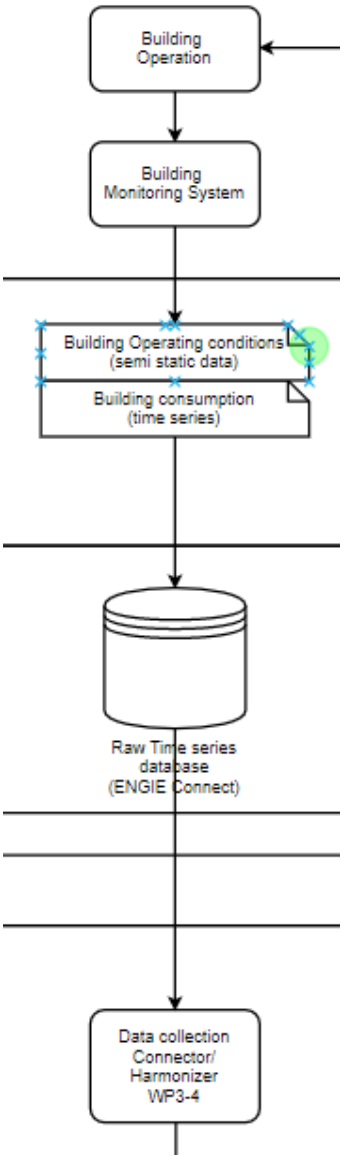
<Process > [ID:< BC5-UC13-004>]

Type	Data Preparation	
Name	Data Preparation	

Documentation	*The raw database is then processed to make sure all the data stored is usable and processable through the BIGG service. The process aims at removing the outliers, missing items and inconsistent data.*

<Data Object Name>

Type	Data Object
Name	*Static Data DDBB*
Documentation	*The data stored ready for utilisation by the BIGG toolbox features*



<Process > **[ID:< BC5-UC13-005>]**

Type	On-going process	
Name	Building operation	
Documentation	*Describes the on-site operation. The building equipment performing its designed actions to maintain comfort in the building. Building operation includes whole the on-going actions taking place in the building. It includes operation of the HVAC equipment, the lighting system, the elevators, the doors, windows but also the building users behaviour, the maintenance actions performed by the building maintenance team… All the on-going events that have an impact on the building total energy consumption.*	

<Process > **[ID:< BC5-UC13-006>]**

Type	On-going process	

Name	Building Monitoring System
Documentation	*Describes the devices, meters and sub-meters collecting data about the on-going building operation. All the devices that are collecting real time measurements of energy, water, temperature, occupancy, light level but also the status of the equipment on site, the valves open/close status, three way valve position, damper position, fan speed …*

<Data Object Name>

Type	Data Object
Name	*Building Operating conditions* *(semi static data)*
Documentation	All time series data streams collected on site through sensors and devices connected to the local service monitoring system or BMS.

<Data Object Name>

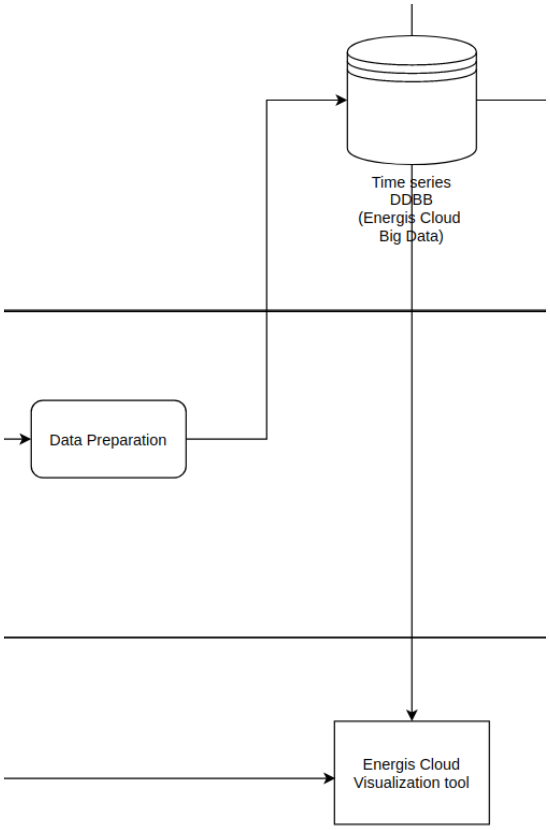
Type	Data Object
Name	*Building consumption* *(time series)*
Documentation	All time series data streams collected on site through sensors and devices connected to the local service monitoring system or BMS.

<Data Object Name>

Type	Data Object
Name	*Raw time series data base*
Documentation	All consumption time series data streams collected on site through meters and sub meters connected to the local service monitoring system or BMS. This data also includes consumption data streams coming from utility services or APIs.

<Process > **[ID:<BC5-UC13-007>]**

Type	Connector and harmonization	
Name	Data collection, connector and harmonization	
Documentation	*Data collection from the local service raw database to the BIGG platform storage. The data is collected either directly from a manual upload or through the BIGG platform API service. The data is collected and harmonized to match the BIGG data Standard 4 Building.*	



Process > **[ID:< BC5-UC13-008>]**

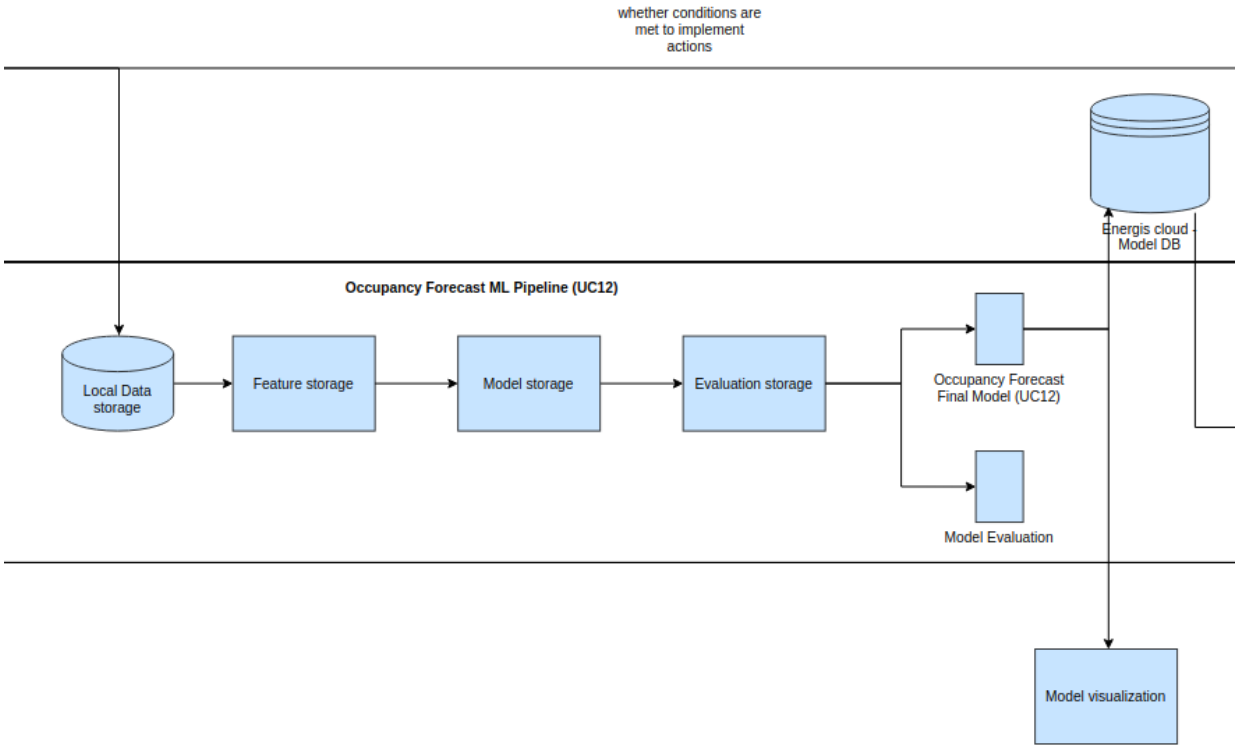
Type	Data Preparation	
Name	Data Preparation	
Documentation	*The time series database is then processed to make sure all the data stored is usable and processable through the BIGG service. The process aims at removing the outliers, missing items and inconsistent data.*	

<Data Object Name>

Type	Data Object
Name	*Time series DDBB*
Documentation	*The data stored ready for utilisation by the BIGG toolbox features*

Process > **[ID:< BC5-UC13-009>]**

Type	Data Visualization	
Name	Energis Cloud Visualization Tool	
Documentation	*Capability to visualize the different data streams directly on the platform.*	



<Data Object Name>

Type	Data Object
Name	*Local data storage*
Documentation	*The time series data is stored locally (where the AI toolbox process is taking place) so that the AI toolbox computation can be performed quickly and efficiently.*

< Process > [ID:<BC5-UC13-010>]

Type	AI Process	
Name	Feature Storage	
Documentation	*The data is analysed and certain patterns and characteristics are extracted from the data and stored for future usage in AI algorithms.*	

<Process > [ID:<BC5-UC13-011>]

Type	AI Process	
Name	Model storage	
Documentation	*A decision tree model is identified based on the input data, the previously extracted features and an AI model library. Different hyperparameters can be tested to select the best model to use in a specific case.*	

<Process > [ID:<BC5-UC13-012>]

Type	AI Process
Name	*Evaluation storage*
Documentation	*The quality of each model is evaluated based on certain statistical criteria (e.g. root mean square error, bias, etc.) Evaluation is used to select the best model and is also a quality certificate for the best model that is provided to the user.*

<Data Object Name>

Type	Data Object
Name	*Energis cloud model database*
Documentation	*Storage of the evaluated models and their evaluations for future use.*

<Data Object Name>

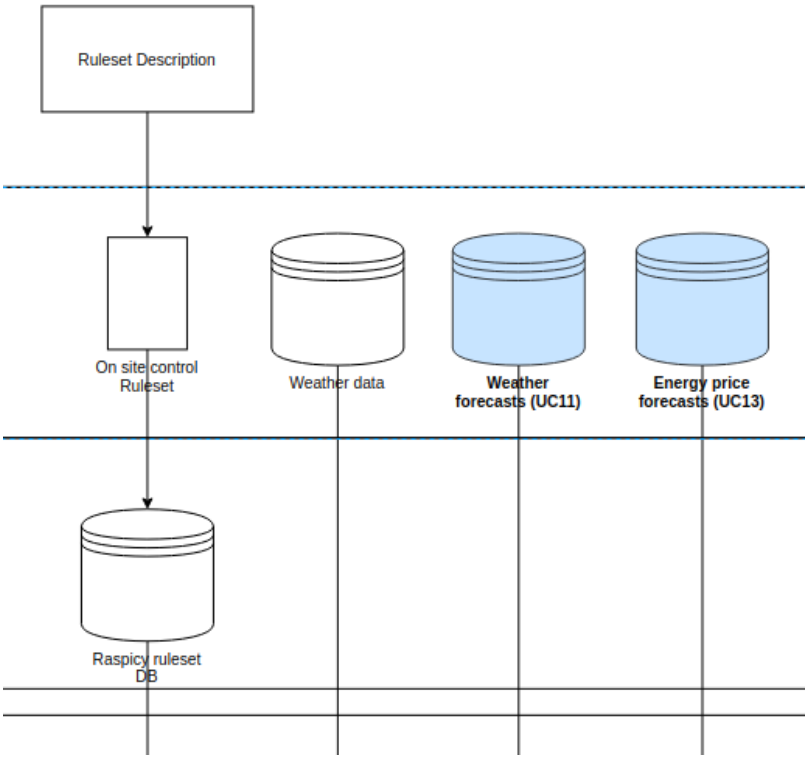
Type	Data Object
Name	*Occupancy forecast final model (UC12)*
Documentation	*The final model is a tabular representation of the occupancy forecast, that can be stored as a metric according to the specific Energis.cloud format and refreshed periodically.*

<Data Object Name>

Type	Data Object
Name	*Model Evaluation*
Documentation	*Beside the final model, the model evaluation also is stored so that the following version of the model can be challenged and compared.*

<Process > [ID:<BC5-UC13-013>]

Type	Data visualization	
Name	Model Visualization	
Documentation	*The current model can be visualized using different types of views.*	



Process > [ID:< BC5-UC13-014>]

Type	Data Collection	
Name	Ruleset description	
Documentation	*A description of the actions that can be done on the existing building control system. The ruleset library should define the identity of the device that can take action, the identity of the equipment upon which the action is undertaken, the nature of the action that can be undertaken, the amplitude of the modification that can be implemented and the minimum change value allowed.*	

<Data Object Name>

Type	Data Object
Name	*On site control Ruleset*
Documentation	*A data object centralizing the information about the ruleset as described in the process BC5-UC-11-009*

<Data Object Name>

Type	Data Object
Name	*Raspicy Ruleset DB*

Documentation	*A database that stores all the rulesets defined for a specific case or scenario. Only one of the stored rulesets will be active in a specific moment and used for the optimization.*

<Data Object Name>

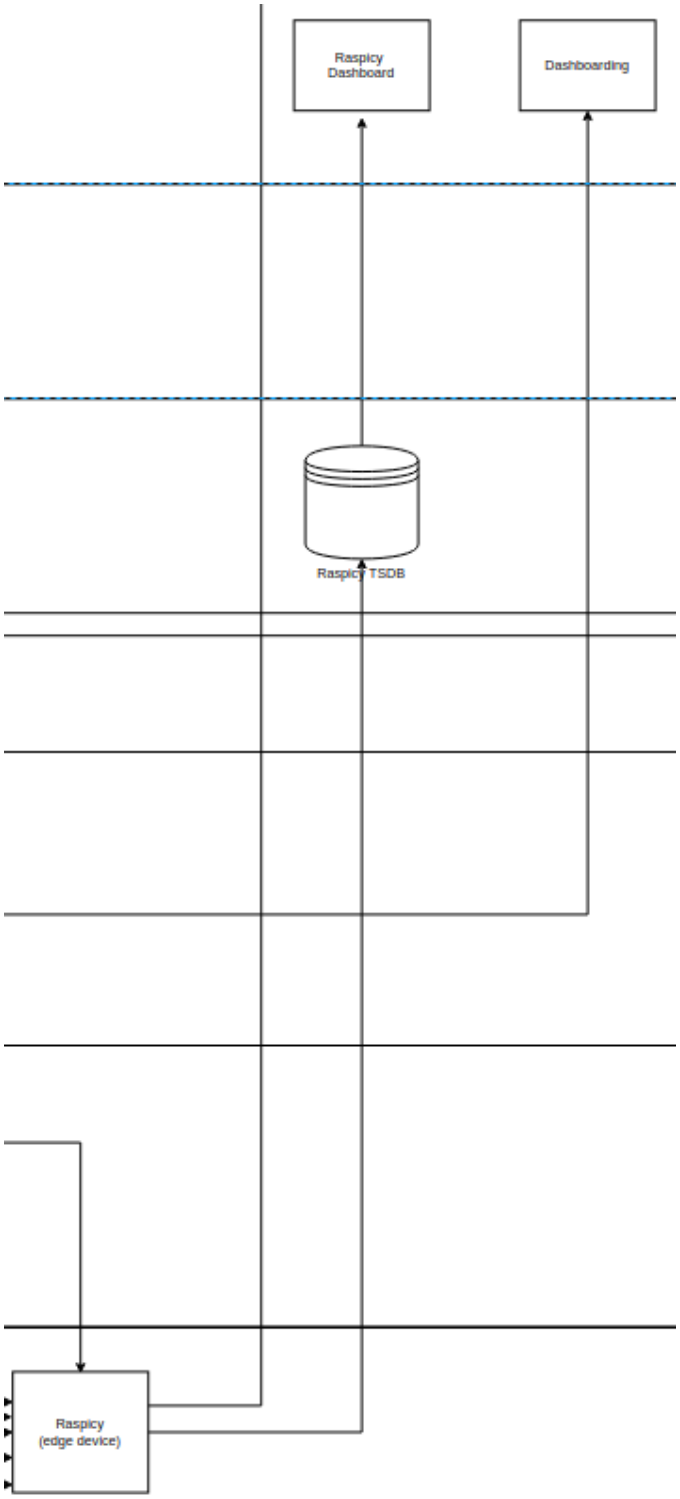
Type	Data Object
Name	*Weather data*
Documentation	*The data about the current weather at the location identified by a latitude and a longitude. These weather data are retrieved directly by the edge device if they are part of the ruleset.*

<Data Object Name>

Type	Data Object
Name	*Weather Forecast data (UC11)*
Documentation	*The Weather forecasts data giving details on the forecasted temperature, sunlight level and relative humidity for the next day. These weather data are retrieved directly by the edge device if they are part of the ruleset.*

<Data Object Name>

Type	Data Object
Name	*Energy price forecast data (UC13)*
Documentation	*The energy price forecasts data used in UC13 are retrieved from an external service through the defined APIs. These data are retrieved directly by the edge device if they are part of the ruleset.*



Process > **[ID:< BC5-UC13-015>]**

Type	Data Visualization	
Name	Raspicy Visualization Tool	

Documentation	*Capability to visualize the data (e.g., sensor data, weather data, actions, etc...) and rules used for the optimization on a local dashboard.*

<Data Object Name>

Type	Data Object
Name	*Raspicy TSDB*
Documentation	*Time series database that stores all the data related to the optimization performed by the Raspicy (e.g., sensor data, weather data, actions, etc...) as a time-series.*

Process > **[ID:< BC5-UC13-016>]**

Type	Data Visualization	
Name	Dashboarding	
Documentation	*Capability to define specific dashboards in a cloud platform to display the building operation and the impact of the implemented rules over the building control system*	

Process > **[ID:< BC5-UC13-0017>]**

Type	Control operation	
Name	Raspicy (edge device)	
Documentation	*The on-site cloud connected controller*	

V.12. UC14

V.12.1. BPMN diagram

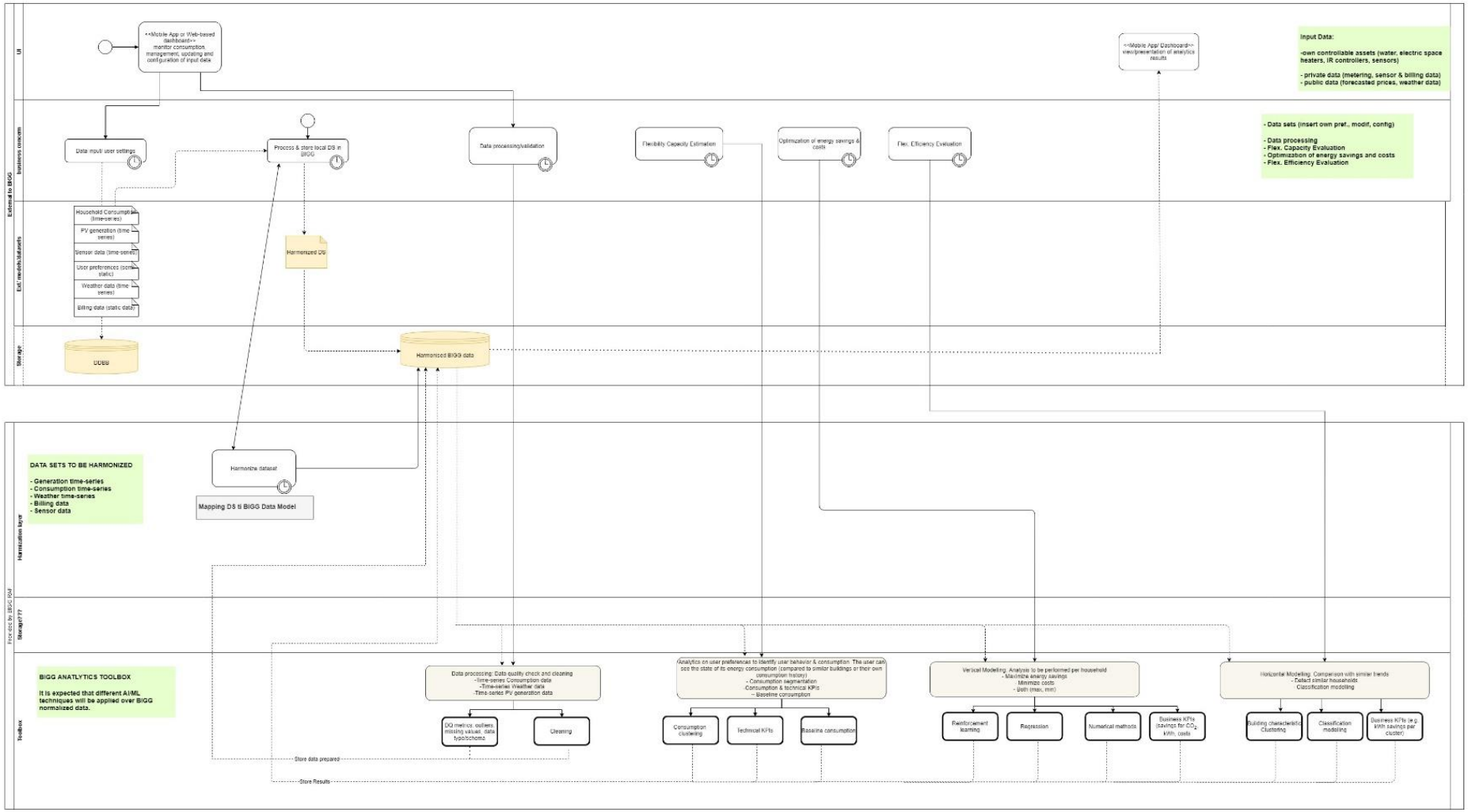
UC 14- DR for Electricity

UC-Description: The main objectives of this Use Case are as follows:

1. **Monitoring of real-time energy consumption**
2. **Estimation of flexibility potential/capacity**
3. **Maximization of energy savings and/or minimization of energy costs**
4. **Evaluation of flexibility efficiency**

Target buildings: Households

Target User: Residential end-users, Retailers



V.12.2. Specification of processes

To develop the "Electricity DR" process, the following process maps are needed:

- Data collection

- Data harmonization

- Data processing/validation

- Flexibility capacity estimation

- Optimization of energy savings and costs

- Flexibility efficiency estimation

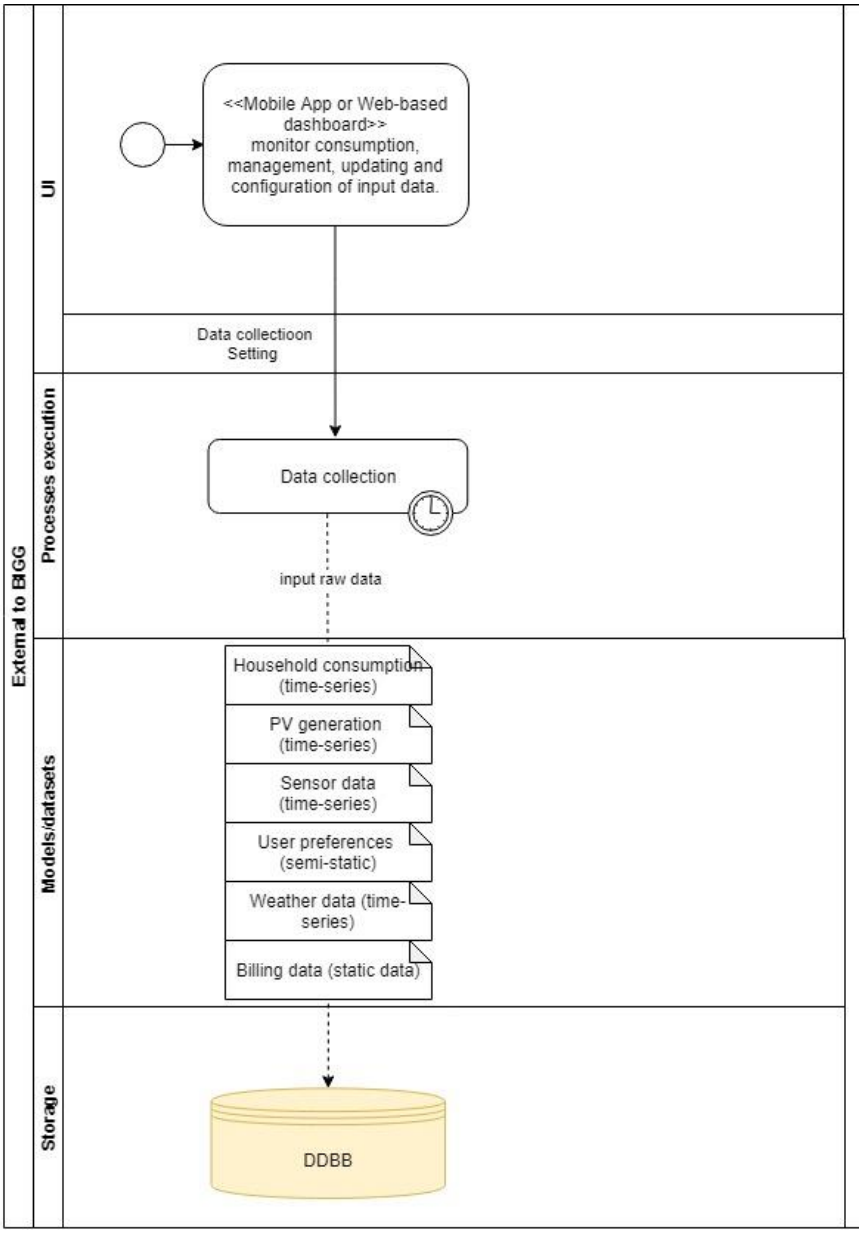
Process Map: <Data Collection>

The data required for the UC14 can come from different data sources. These data sources can be loaded from existing local databases/repositories or uploaded manually by the user (e.g. user preferences).

In the case of manual data upload, the end user will be able to load, modify or configure data from the UI, whether this is static (user preferences) or time series (energy consumption).

In the case of data from existing databases, these can be collected directly through HERON's REST API, the administrator user will have access from the UI to the import control of this data.

Power and energy measurements as well as sensor data are automatically stored HERON's time-series database (InfluxDB). The measurements can then be processed, queried, or visualized from anywhere and anytime using HERON's exposed REST API.



POOL : <External to BIGG >

This pool entails the processes which will be carried out entirely in the group <External to BIGG>.

LANE: <UI>

In the UI (mobile app and/or web-based platform) the end user will be able to visualise his real-time and historical consumption as well as insert his/her preferences. The user will utilise the UI to configure 1) his/her preferences with respect to which equipment can be considered for DR purposes; 2) whether opt-in/out for a DR event; 3) specify a time-schedule or rule upon which the activation of the controllable device is executed for a specified duration.

<Monitor, management, updating and configuration of input data> [ID:<BC6-UC14-01>]

Type	*Login & user settings*	
Name	*Monitor, management, updating and configuration of input data*	

Documentation	*The end user connects to his profile though the mobile app and will have access to the functions to retrieve and monitor data from different sources. The end user will also be capable of reviewing the available data sources already uploaded, manage existing data, insert and configure his/her own preferences as well as modify/update them if necessary.*

LANE: < Processes execution>

<Data Collection> [ID:< BC6-UC14-02>]

Type	*Data Collection*	
Name	*Data collection*	
Documentation	*The data collection process is responsible for managing the data entry into the system. Power and energy measurements as well as sensor data (near real-time) are pushed from the devices to HERON's backend infrastructure using the VerneMQ MQTT broker. An intermediate service of Telegraf is used to automatically store in real-time the MQTT messages to the respective format used in our time-series database (InfluxDB). The measurements can then be processed, queried, or visualized from anywhere and anytime using HERON's exposed REST API. The end users may login to their mobile app and/or dashboard front-end using their own credentials to monitor historical and real-time measurements for any time period. The front-end retrieves measurements data by executing the required API queries. Multiple meters/control relays/sensors can be grouped under a common project (tag) to manage data access based on specific user roles.*	

LANE: < Models/Data Sets>

This LANE describes the external and existing models and data that must be managed in the process.

Library Data Objects

<Household Consumption>

Type	Data Object
Name	*Household consumption*
Documentation	*Real-time electricity consumption data at total and relay/plug level from smart meters. Active and reactive power is provided with a 30-sec granularity whereas consumed energy is provided with a 1-min granularity. In addition, this data can have a heterogeneous aggregation, and can be provided with configured granularity as requested by the DR platform. A subset of the pilot participants will be also equipped with one relay to control (ON/OFF) their*

	water heaters/space heaters, etc. Actuations close to real-time will be also recorded.

<PV Generation>

Type	Data Object
Name	*PV generation*
Documentation	*Time-series generation from PVs installed on rooftops of some prosumers participating in the pilot cluster. Power and produced energy are provided with a 30-sec and a 1-min granularity, respectively. In addition, this data can have a heterogeneous aggregation, and can be provided with configured granularity as requested by the DR platform to perform production and energy gains calculations, such as on daily and/or monthly basis.*

<Sensor data>

Type	Data Object
Name	*Sensor data*
Documentation	*A subset of households will be also equipped with IoT sensors providing real-time readings (such as temperature) with configurable granularity.*

<User preferences>

Type	Data Object
Name	*User Preferences*
Documentation	*The user will be able to insert his preferences (semi-static data) through the mobile app, such as a time-schedule upon which the activation of a controllable device is executed for a specified duration or for shifting the operation of the controllable loads within the day.*

<Billing data>

Type	Data Object
Name	*Billing data*
Documentation	*Static data concerning the billing history information of the residential buildings, such as electricity consumption, charges, total electricity cost, contractual capacity etc.*

<Weather data >

Type	Data Object
Name	*Weather data*

Documentation	*The system will continuously retrieve publicly available data from meteorological stations located throughout Europe and/or national stations, with mostly hourly aggregation.* *The data collected could be:* *- Temperature, relative humidity, direct solar radiation in a horizontal plane, diffuse solar radiation, wind speed, wind direction, precipitation, etc.*

LANE: < Storage>

The Storage LANE refers to the different databases and data objects that will be stored in HERON's local system.

Library Data Objects

<Raw Database>

Type	Data Object
Name	*Raw Data Base*
Documentation	*Power and energy measurements as well as sensor data are pushed from the meters to the backend infrastructure using the VerneMQ MQTT broker. An intermediate service of Telegraf is used to automatically store in real-time the MQTT messages to the respective format used in HERON's time-series database, InfluxDB which is an open-source time-series database, deployed using container orchestration technologies. The measurements can then be processed, queried, or visualized from anywhere and anytime using HERON's exposed REST API.*

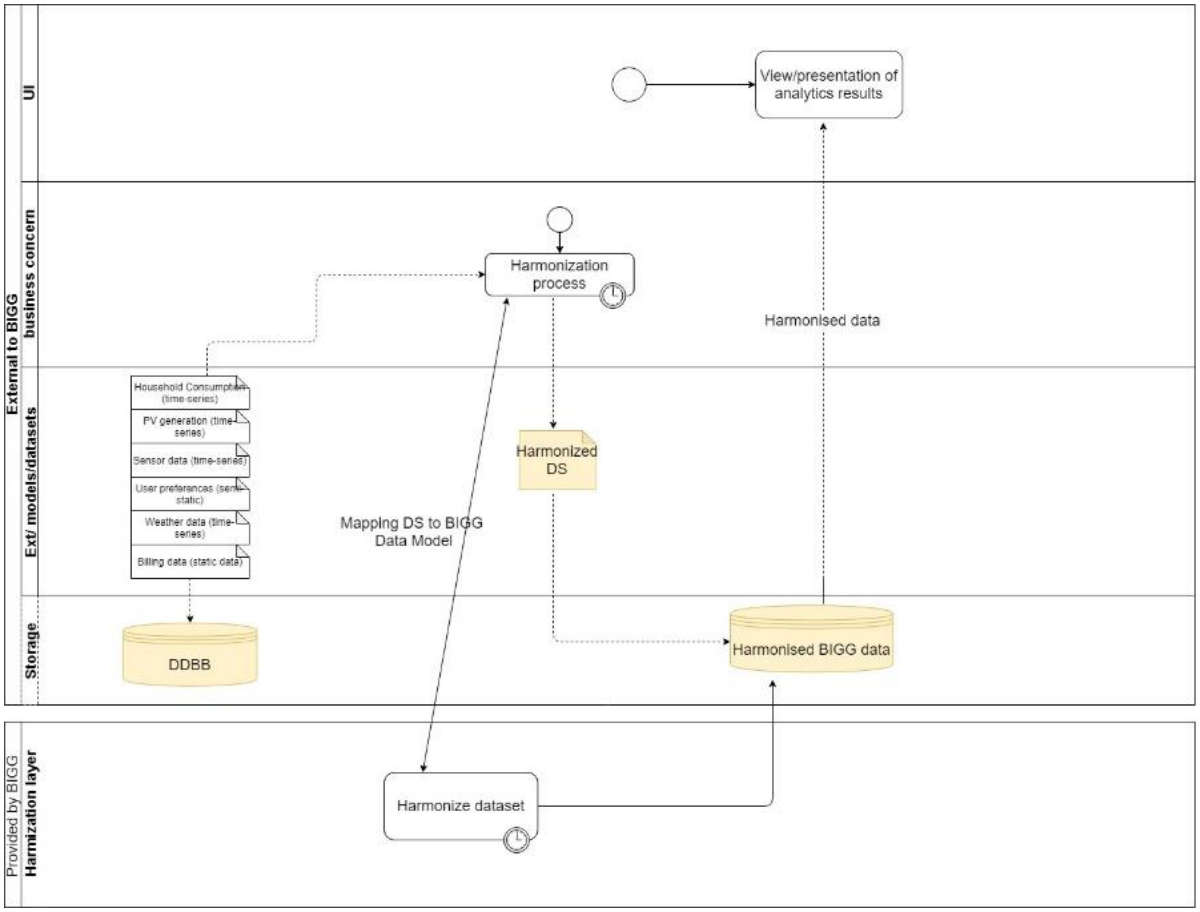
Exchange Requirement Data Objects

<Input raw data>

Type	Data Object
Name	*Time-series, static and semi-static information exchange (measurements, commands, etc.)*
Documentation	*All data from the data collection process is stored in the Raw Database.*

Process Map: <Harmonization>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies. This standardization allows, on the one hand, to use BIGG's analytical developments and at the same time facilitates the exchange of data with other external systems or applications.



POOL : <External to BIGG>

This pool entails the processes which will be carried out entirely in the group <External to BIGG>.

LANE: <Storage>

The Storage LANE refers to the different databases and data objects that will be stored in HERON's local system.

Library Data Objects

<Harmonized Data Base>

Type	Data Object
Name	*Harmonized Data Base*
Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format.*

Exchange Requirement Data Objects

<Input raw data>

Type	Data Object

Name	*Input raw data*
Documentation	*Raw data is extracted from raw databases and sent to the data harmonization process when required.*

<Harmonized raw data>

Type	Data Object
Name	*Harmonized raw data*
Documentation	*The harmonized data is extracted from the data harmonization process and is systematically stored in the harmonized database.*

LANE: < Processes execution>

<Harmonization process> [ID:< BC6-UC14-03>]

Type	*Harmonization*	
Name	*Harmonization process*	
Documentation	*The harmonization process is responsible for transforming the raw data into harmonized data in BIGG format.* *This process that runs is as follows:* *- Retrieves raw input data that has not been harmonized so far, from the raw database.* *- Retrieves the raw data maps made in BIGG.* *- Assign the raw data to the BIGG data model.* *- Save the harmonized input data in the harmonized database.*	

LANE: < Models/Data Sets>

This LANE describes the external and existing models and data that must be managed in the process.

Library Data Objects

Household Consumption>

Type	Data Object
Name	*Household consumption*
Documentation	*Same as in data collection process*

<PV Generation>

Type	Data Object

Name	*PV generation*
Documentation	*Same as in data collection process*

<Sensor data>

Type	Data Object
Name	*Sensor data*
Documentation	*Same as in data collection process*

<User preferences>

Type	Data Object
Name	*User Preferences*
Documentation	*Same as in data collection process*

<Billing data>

Type	Data Object
Name	*Billing data*
Documentation	*Same as in data collection process*

<Weather data>

Type	Data Object
Name	*Weather data*
Documentation	*Same as in data collection process*

<Harmonized Datasets>

Type	Data Object
Name	*Harmonized Datasets*
Documentation	*All input raw data harmonized and standardized in accordance with the BIGG data model*

LANE: <UI>

In the UI (mobile app and/or web-based platform) the end user will be able to explore both raw and harmonised data.

<Exploring raw and harmonized data> [ID:< BC6-UC14-04>]

Type	UI	

Name	*Exploring raw and harmonized data*
Documentation	*From the user interface, the user can view and explore the raw data that the system has ingested, as well as the result of the harmonization process.*

POOL : <Provided by BIGG>

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE: <Harmonization layer>

The harmonization layer provides the mappings of the raw data from the use case to the BIGG data model, as well as the mapping of the BIGG data model to other existing ontologies.

<Harmonise DS to BIGG model> [ID:< BC6-UC14-05>]

Type	*Harmonization*	
Name	*Harmonise DS to BIGG model*	
Documentation	*BIGG's harmonization layer has the raw data mappings developed in the project. These data mappings will go through the harmonization process of the ENMA system so that you can execute them and harmonize the data.*	

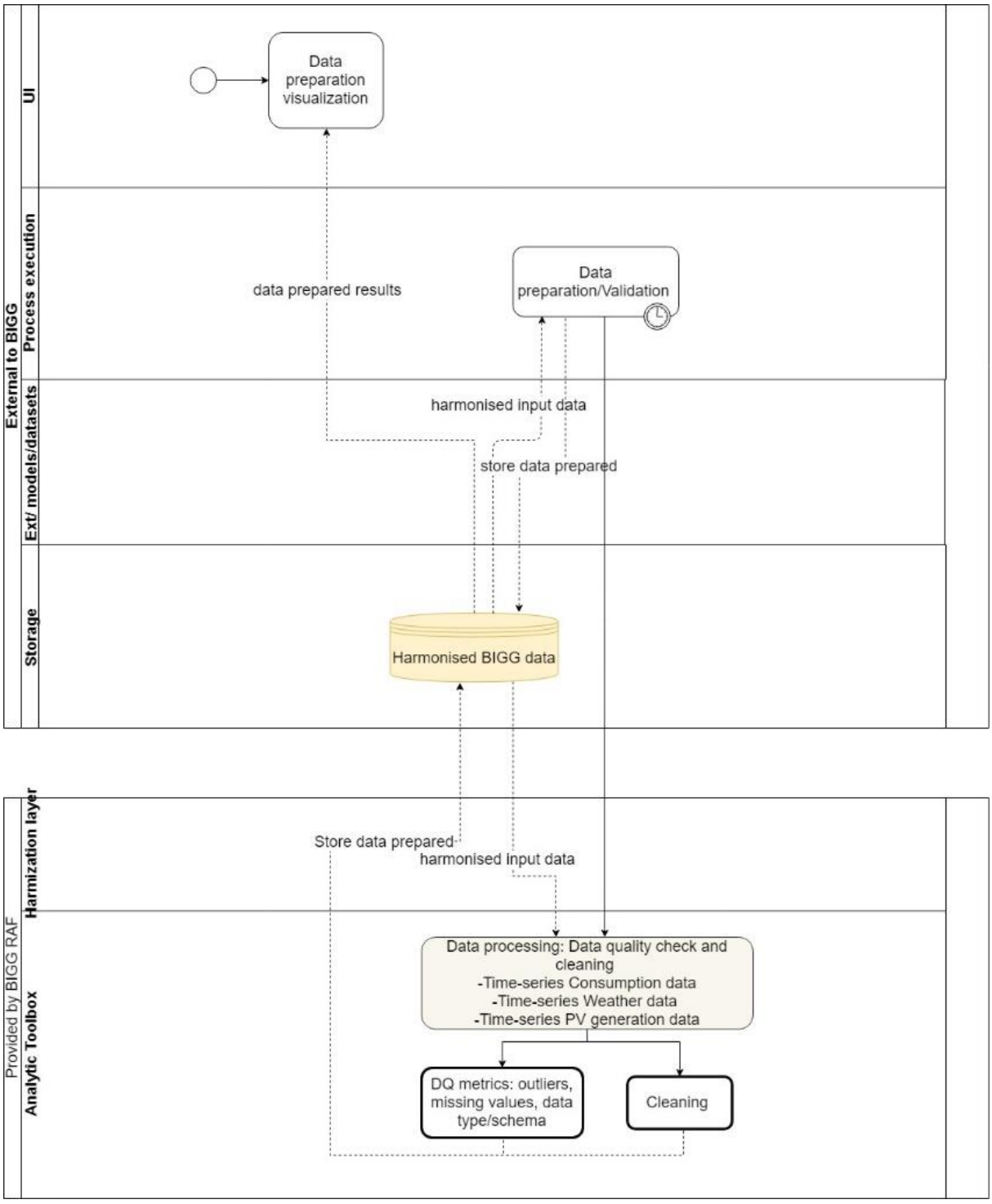
Exchange Requirement Data Objects

<Data mapping>

Type	Data Object
Name	*Data mapping*
Documentation	*The data mappings contained in the BIGG system will be subjected to the harmonization process so that they are carried out on the data to be harmonized.*

Process Map: <Data preparation>

The data preparation process is responsible for verifying the quality, purifying and transforming the input data so that it can be used without problems by the analytical processes.



POOL : <External to BIGG>

This pool entails the processes which will be carried out entirely in the group <External to BIGG>.

LANE: <UI>

In the UI (mobile app and/or web-based platform) the end user will be able to explore and visualise prepared data.

<Data preparation visualization> [ID:< BC6-UC14-06>]

Type	*Use task*	
Name	*Data preparation visualization*	
Documentation	*In the user interface, the user can explore the results of the data preparation process, get the prepared data, and potentially identify the differences between the raw data, and assess the quality of the data in the system.*	

LANE: <Processes execution>

<Data preparation> [ID:< BC6-UC14-07>]

Type	*Data preparation*	
Name	*Data preparation*	
Documentation	*This process is in charge of verifying that all previously harmonised data is usable and processable through the BIGG service. Once this process is applied, the data is ready to be processed by the BIGG AI toolbox. This process aims at removing the outliers, missing items and inconsistent data.*	

LANE: <Storage>

The Storage LANE refers to the different databases and data objects that will be stored in HERON's local system.

Library Data Objects

<Harmonized Database>

Type	Data Object
Name	*Harmonized Datasets*
Documentation	*Once the data is processed by the data harmonization process, it is stored in the harmonized database. In this database, all data is in the BIGG data model format.*

POOL : <Provided by BIGG>

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE: <Analytics Toolbox>

This LANE describes the AI modules and features performed by the BIGG platform to transform, evaluate, analyse, model, assess and process the data on the platform. This essentially refers to the scope of work covered in the Work Package 5 of the BIGG project.

<Data processing/validation> [ID:< BC6-UC14-08>]

Type	*Data processing/validation*	
Name	*Data processing/validation*	
Documentation	*The data preparation process that is performed by the BIGG analysis toolbox to make sure all the data stored is usable and processable through the BIGG service. The process aims at removing the outliers, missing items and inconsistent data This process retrieves the harmonized input data and prepares it for further analysis.* *These processes are segmented by data type, for example, static/semi-static data, time-series data since the processes can be different in each case.* *The data processing/validation could involve:* *- data format* *- data cleaning* *- detection of missing values* *- outlier detection* *- data filling* *- data normalization* *- data modification*	

Exchange Requirement Data Objects

< Data prepared results>

Type	Data Object
Name	*Data prepared results*
Documentation	*The results of the analytical data preparation process are collected by the harmonised dataset and presented to the UI..*

< Harmonised input data>

Type	Data Object
Name	*Harmonized input data*
Documentation	*The harmonized input data is published from the harmonised database to the data preparation execution process.*

< Stored data prepared>

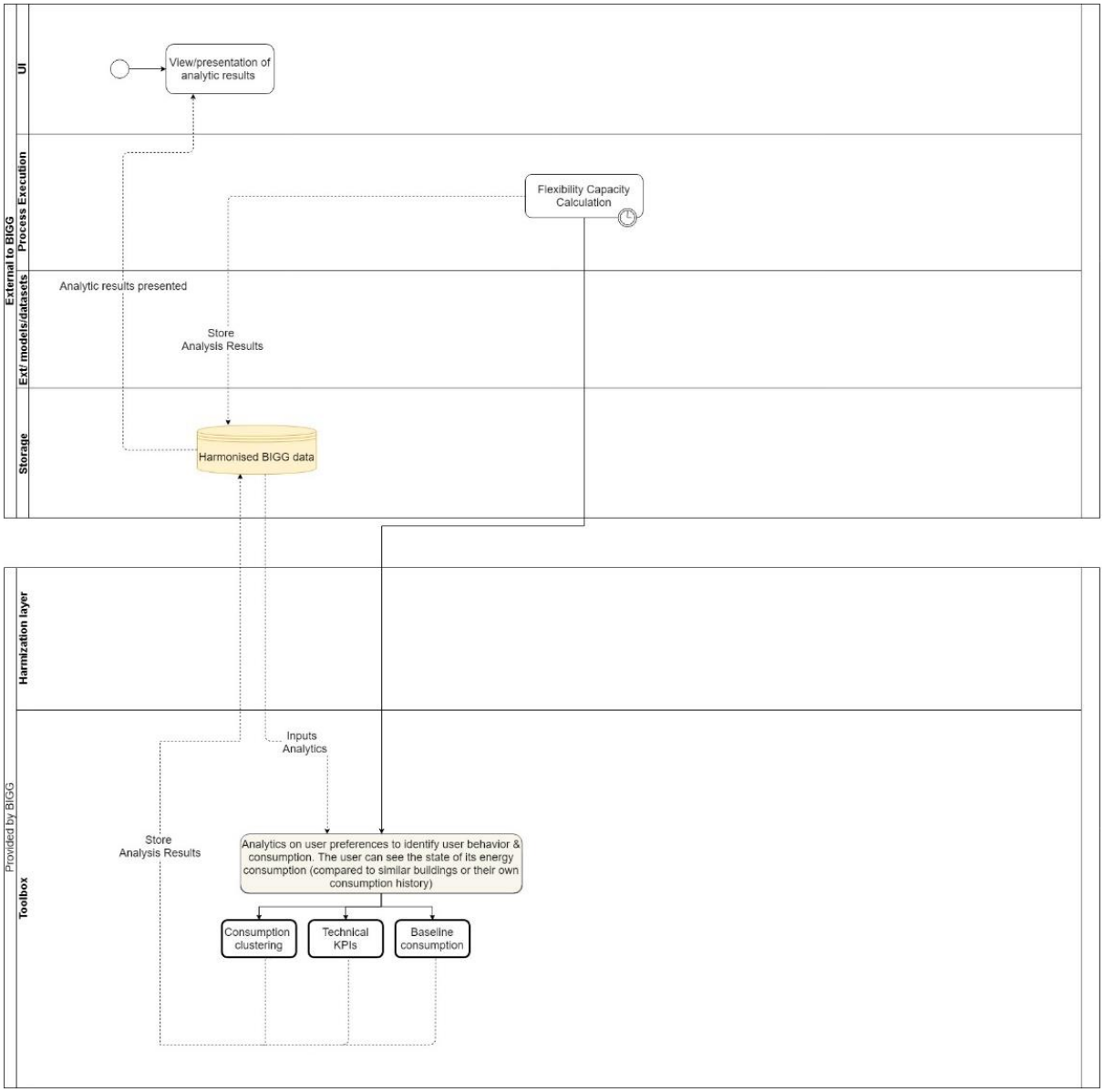
Type	Data Object
Name	*Stored data prepared*
Documentation	*The results of the data preparation process are stored back to the harmonised database.*

Process Map: <*Flexibility capacity estimation*>

This process involves the real-time monitoring of electricity consumption at household and/or relay level as well as of PV generation for prosumers. In addition, sensor data (such as temperature, humidity or movement) will be monitored to evaluate consumers' flexibility potential.

Analytics performed within the BIGG toolbox will incorporate user preferences to identify user behaviour & consumption patterns which will comprise input for benchmarking (consumption segmentation, baseline consumption). The user can also monitor the state of its energy consumption (compared to similar households or their own consumption history).

The analytical process is responsible for extracting the data, previously prepared, extracting the harmonized datasets, obtaining the expected analytical results for the use case, saving them in the harmonized database and presenting them to the end user.



POOL : <External to BIGG>

This pool entails the processes which will be carried out entirely in the group <External to BIGG>.

LANE: <UI>

In the UI (mobile app and/or web-based platform) the end user will be able to explore and visualise the results obtained from this process.

<View/presentation of analytic results> [ID:< BC6-UC14-09>]

Type	*View/presentation of analytic results*
Name	*View/presentation of analytic results*
Documentation	*The end user can explore the analytical results from the user interface. The presentation of these results will be simple and*

	easily understandable. The user can easily filter the results to obtain the visualizations that he/she needs.

LANE: <Processes execution>

<*Flexibility capacity estimation*> [ID:< BC6-UC14-10>]

Type	*Flexibility Capacity Estimation*
Name	*Flexibility Capacity Estimation*
Documentation	*Analysis will involve:* *a. Formulate clusters based on dynamic consumption data (e.g. heavy vs. light consumers, day vs night consumers etc.)* *b. Provide input to optimization (e.g. find the optimal combination of equipment/ smart appliances to fulfil the request).* *Analytics to be performed within the BIGG toolbox will incorporate user preferences to identify user behaviour & consumption patterns which will comprise input for benchmarking (consumption segmentation, baseline consumption). The user can also monitor the state of its energy consumption (compared to similar households or their own consumption history).* ***KPIs involved:*** *Metrics such as historical kWh consumed, CO_2 emissions, € spent, inventory of controllable devices based on their specs (Watt), # engaged users or # of participants that use the mobile app.*

LANE: <Storage>

The Storage LANE refers to the different databases and data objects that will be stored in HERON's local system.

Library Data Objects

<Harmonized Database>

Type	Data Object
Name	*Harmonized Database*
Documentation	*The harmonised data are used by the analytical processes. The results of the analytical processes will also be stored in the harmonised database.* *In this database, all data is in the BIGG data model format.*

POOL : <Provided by BIGG>

LANE: <Analytics Toolbox>

This LANE describes the AI modules and features performed by the BIGG platform to transform, evaluate, analyse, model, assess and process the data on the platform. This essentially refers to the scope of work covered in the Work Package 5 of the BIGG project.

<Analytics on user preferences> [ID:< BC6-UC14-11>]

Type	*Analytics on user preferences*	
Name	*Analytics on user preferences*	
Documentation	*Analytics on user preferences to identify user behaviour & consumption. The user can see the state of its energy consumption (compared to similar buildings or their own consumption history). The following threads could be considered:* *-Consumption segmentation* *-Clustering of historical consumption curves* *-Identifying relevant KPIs* *-Elaboration of baseline consumption.*	

Exchange Requirement Data Objects

< Analytic results presented>

Type	Data Object
Name	*Analytic results presented*
Documentation	*The results of the analysis, stored in the harmonized DDBBs, will be extracted and presented to the end users.*

<Analytics inputs>

Type	Data Object
Name	*Analytics Inputs*
Documentation	*The necessary input data is sent to the analysis process from the harmonized databases.* *This data could be:* *- Historical energy consumption (time-series)* *- Actuations ON/OFF* *- Weather data* *- Historical RES generation (time-series)*

<Stored analysis results>

Type	Data Object
Name	*Stored Analysis results*

Documentation	*The results of the analytical processes are stored in the harmonized database.* *The main results could be:* *Segmentation of consumption (heavy vs. light consumers, day vs night consumers), CO_2 emissions, € spent, inventory of controllable devices based on their specs (Watt), # engaged users or # of participants that use the mobile app, forecasts of load peaks etc.*

Process Map: <*Optimization of energy savings and costs*>

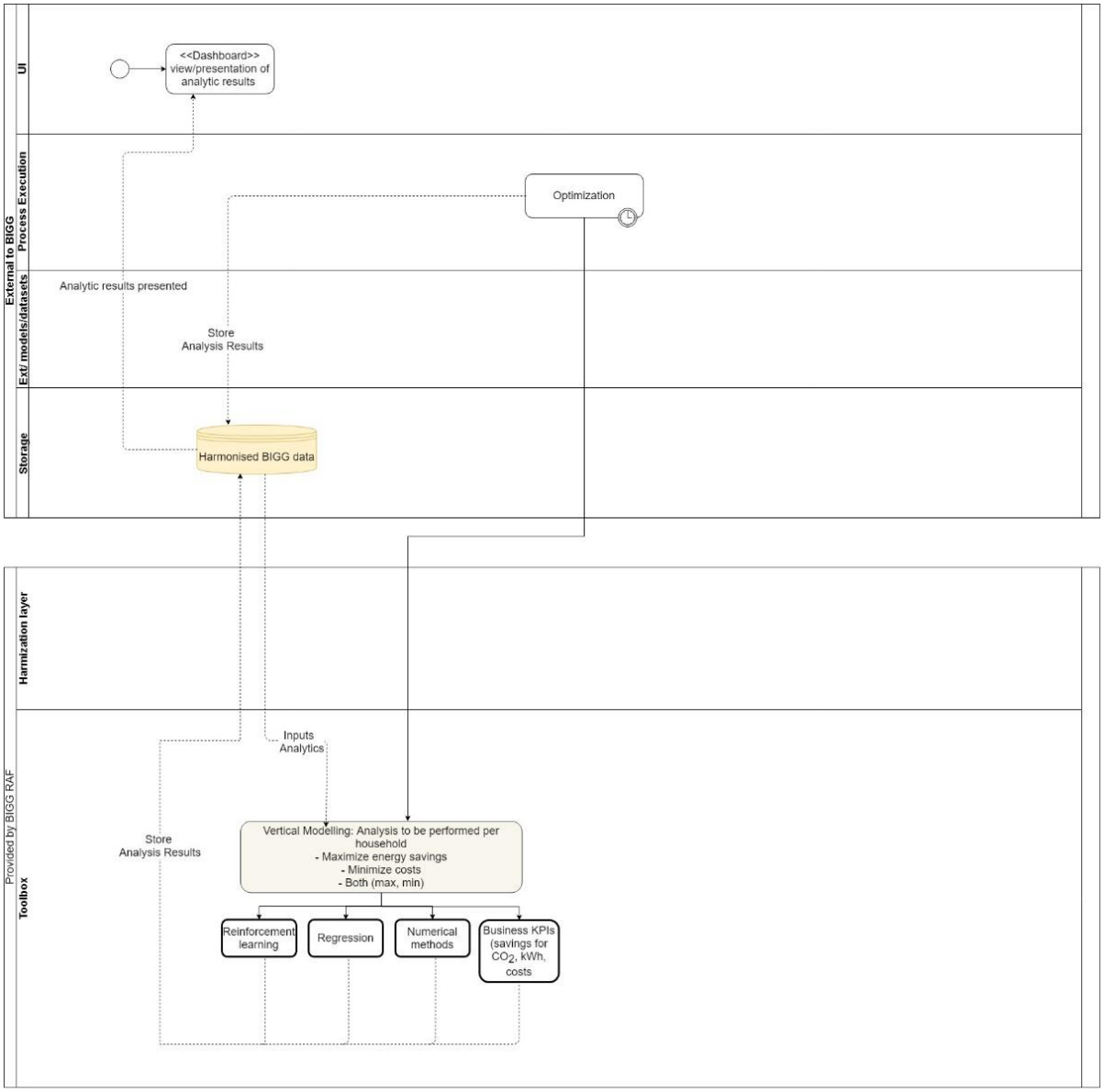
This process involves consumption management through the activation of DR events for minimizing the costs (based on market observed and forecasted data) ***and/or*** *maximizing savings. Vertical modelling for analysing energy consumption trends per household.*

Analytics performed within the BIGG toolbox may include reinforcement learning, regression and computational/numerical methods (optimization).

KPIs involved:

Evaluate the DR offer (before and after DR) using a set of metrics (CO_2, kWh, costs).

The analytical process is responsible for extracting the data, previously prepared, extracting the harmonized datasets, obtaining the expected analytical results for the use case, saving them in the harmonized database and presenting them to the end user.



POOL : <External to BIGG>

This pool entails the processes which will be carried out entirely in the group <External to BIGG>.

LANE: <UI>

In the UI (mobile app and/or web-based platform) the end user will be able to explore and visualise the results obtained from this process.

<View/presentation of analytic results> [ID:< BC6-UC14-12>]

Type	*View/presentation of analytic results*
Name	*View/presentation of analytic results*
Documentation	*The end user can explore the analytical results from the user interface. The presentation of these results will be simple and*

	easily understandable. The user can easily filter the results to obtain the visualizations that he/she needs.

LANE: <Processes execution>

<*Optimization*> [ID:< BC6-UC14-13>]

Type	*Optimization of energy savings and costs*
Name	*Optimization of energy savings and costs*
Documentation	*Analysis will involve:* *Consumption management through the activation of DR events for minimizing the costs (based on market observed and forecasted data)* ***and/or*** *maximizing savings. Vertical modelling for analysing energy consumption trends per household.* *Analytics performed within the BIGG toolbox may include reinforcement learning, regression and computational/numerical methods (optimization).* ***KPIs involved:*** *Evaluate the DR offer (before and after DR) using a set of metrics (CO_2, kWh, costs).*

LANE: <Storage>

The Storage LANE refers to the different databases and data objects that will be stored in HERON's local system.

Library Data Objects

<Harmonized Database>

Type	Data Object
Name	*Harmonized Database*
Documentation	*The harmonised data are used by the analytical processes. The results of the analytical processes will also be stored in the harmonised database.* *In this database, all data is in the BIGG data model format.*

POOL : <Provided by BIGG>

This process will be partially carried out in the ***POOL*** *<Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE: <Analytics Toolbox>

This LANE describes the AI modules and features performed by the BIGG platform to transform, evaluate, analyse, model, assess and process the data on the platform. This essentially refers to the scope of work covered in the Work Package 5 of the BIGG project.

<Vertical Modelling> [ID:< BC6-UC14-14>]

Type	*Vertical modelling*
Name	*Vertical modelling*
Documentation	*Analytics to identify user behaviour & consumption according to its consumption history and its evolution. The user can see the state of its energy consumption compared to its own consumption history. The input data for this model is the prepared data that is stored in the harmonized database.* *The following threads could be considered:* *-Clustering of historical consumption curves* *-Obtaining the KPIs of consumption (weekdays consumption vs weekend consumption, base load consumption etc.)* *-Detection of load peaks* *-Obtaining the KPIs re. energy and/or cost savings* *-Elaboration of baseline models.*

Exchange Requirement Data Objects

< Analytic results presented>

Type	Data Object
Name	*Analytic results presented*
Documentation	*The results of the analysis, stored in the harmonized DDBBs, will be extracted and presented to the end users.*

<Analytics inputs>

Type	Data Object
Name	*Analytics Inputs*
Documentation	*The necessary input data is sent to the analysis process from the harmonized databases.* *This data could be:* *- Historical energy consumption (time-series)* *- Actuations ON/OFF* *- Weather data* *- Historical RES generation (time-series)*

<Stored analysis results>

Type	Data Object
Name	*Stored Analysis results*
Documentation	*The results of the analytical processes are stored in the harmonized database.* *The main results could be:* *Evaluate the DR offer (before and after DR) using a set of metrics (CO_2, kWh, costs).*

Process Map: <*Flexibility Efficiency Evaluation*>

This process involves end-user decision related to opt-in/opt-out from the DR scheme.

Dynamic consumption clusters based on real-time consumption or sensor data.

Peak shavings vs. base load shifting (in pursuit of monetary gains or environmental signals).

Analytics performed within the BIGG toolbox may involve horizontal modelling for comparing households with similar trends in order to:

a. Detect similar households.

b. Classification modelling for rating the buildings' energy consumption in relation to similar buildings.

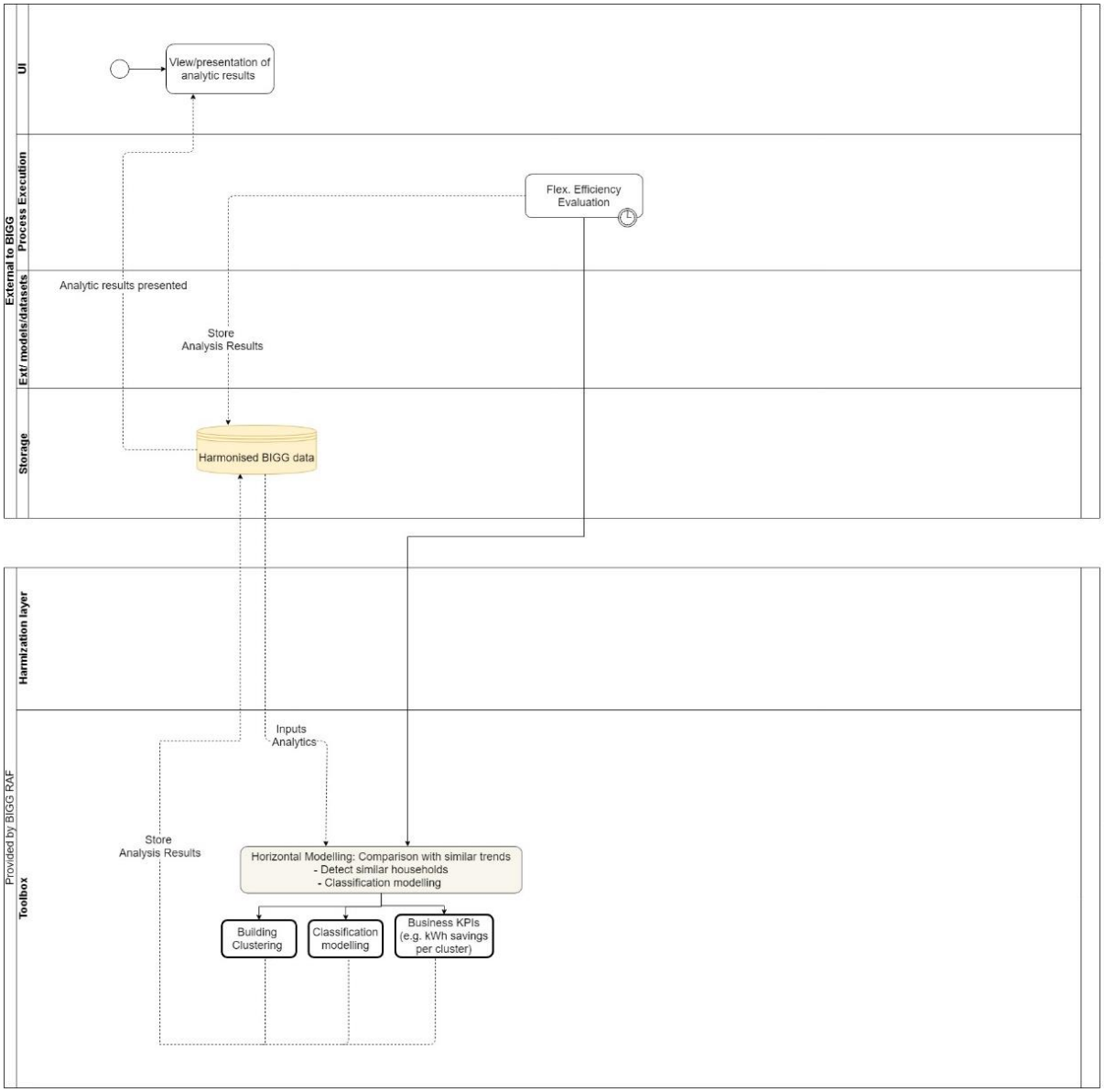
c. Identify clusters of end users based on their consumption behaviour.

KPIs involved:

Measure successful DR activations (cases where consumers do not override DR recommendation)

Analytics for the different user clusters e.g. number of peak shaving users vs. load shifters, volumes of energy savings per cluster as well as the number of active users.

The analytical process is responsible for extracting the data, previously prepared, extracting the harmonized datasets, obtaining the expected analytical results for the use case, saving them in the harmonized database and presenting them to the end user.



POOL : <External to BIGG>

This pool entails the processes which will be carried out entirely in the group <External to BIGG>.

LANE: <UI>

In the UI (mobile app and/or web-based platform) the end user will be able to explore and visualise the results obtained from this process.

<View/presentation of analytic results> [ID:< BC6-UC14-15>]

Type	*View/presentation of analytic results*	
Name	*View/presentation of analytic results*	
Documentation	*The end user can explore the analytical results from the user interface. The presentation of these results will be simple and easily understandable. The user can easily filter the results to obtain the visualizations that he/she needs.*	

LANE: <Processes execution>

<*Flexibility evaluation*> [ID:< BC6-UC14-16>]

Type	*Flexibility efficiency evaluation*	
Name	*Flexibility efficiency evaluation*	
Documentation	*Analysis will involve:* *This process involves end-user decision related to opt-in/opt-out from the DR scheme.* *Dynamic consumption clusters based on real-time consumption or sensor data.* *Peak shavings vs. base load shifting (in pursuit of monetary gains or environmental signals).* *Analytics performed within the BIGG toolbox may involve horizontal modelling for comparing households with similar trends in order to:* *a. Detect similar households.* *b. Classification modelling for rating the buildings' energy consumption in relation to similar buildings.* *c. Identify clusters of end users based on their consumption behaviour.* ***KPIs involved:*** *Measure successful DR activations (cases where consumers do not override DR recommendation)* *Analytics for the different user clusters e.g. number of peak shaving users vs. load shifters, volumes of energy savings per cluster as well as the number of active users.*	

LANE: <Storage>

The Storage LANE refers to the different databases and data objects that will be stored in HERON's local system.

Library Data Objects

<Harmonized Database>

Type	Data Object
Name	*Harmonized Database*
Documentation	*The harmonised data are used by the analytical processes. The results of the analytical processes will also be stored in the harmonised database.*

	In this database, all data is in the BIGG data model format.

POOL : <Provided by BIGG>

*This process will be partially carried out in the **POOL** <Provided by BIGG>. <Provided by BIGG> means the processes that take place in the BIGG system.*

LANE: <Analytics Toolbox>

This LANE describes the AI modules and features performed by the BIGG platform to transform, evaluate, analyse, model, assess and process the data on the platform. This essentially refers to the scope of work covered in the Work Package 5 of the BIGG project.

<Horizontal Modelling> [ID:< BC6-UC14-17>]

Type	*Horizontal modelling*	
Name	*Horizontal modelling*	
Documentation	*In horizontal modelling, the behaviour of end users is analysed in comparison to similar behaviours/trends.* *The input data for this model are the data from the data preparation process and the results of the vertical modelling, both stored in the harmonized databases.* *This modelling process can be divided into the following threads:* *- Grouping of households/end users (by consumption KPI)* *- Classification modelling* *- Obtaining average KPI's.*	

Exchange Requirement Data Objects

< Analytic results presented>

Type	Data Object
Name	*Analytic results presented*
Documentation	*The results of the analysis, stored in the harmonized DDBBs, will be extracted and presented to the end users.*

<Analytics inputs>

Type	Data Object
Name	*Analytics Inputs*
Documentation	*The necessary input data is sent to the analysis process from the harmonized databases.* *This data could be:* *- Historical energy consumption (time-series)*

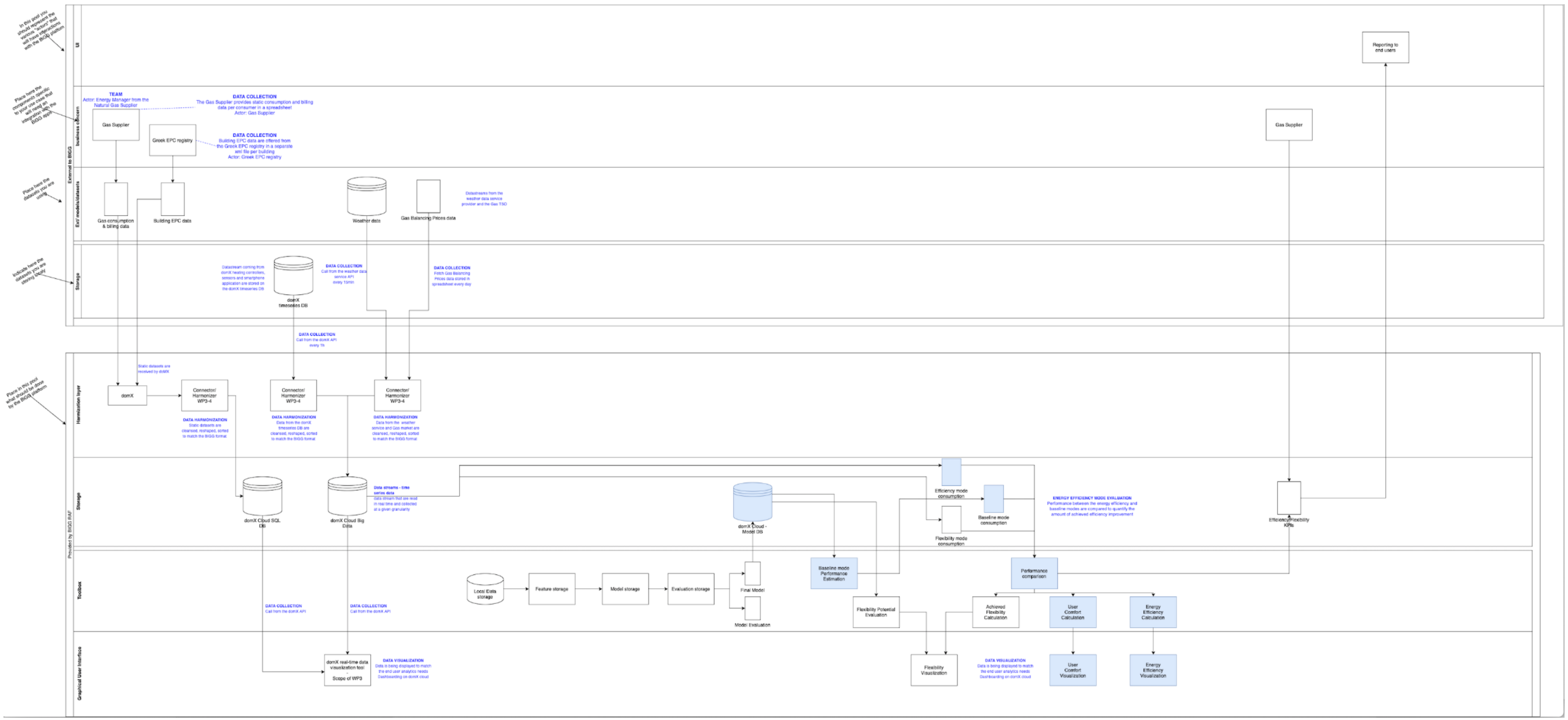
	- Actuations ON/OFF *- Weather data* *- Historical RES generation (time-series)*

<Stored analysis results>

Type	Data Object
Name	*Stored Analysis results*
Documentation	*The results of the analytical processes are stored in the harmonized database.* *The main results could be:* *Number of peak shaving users vs. load shifters, volumes of energy savings per cluster as well as the number of active users and number of successful activations within a pre-defined timeframe.*

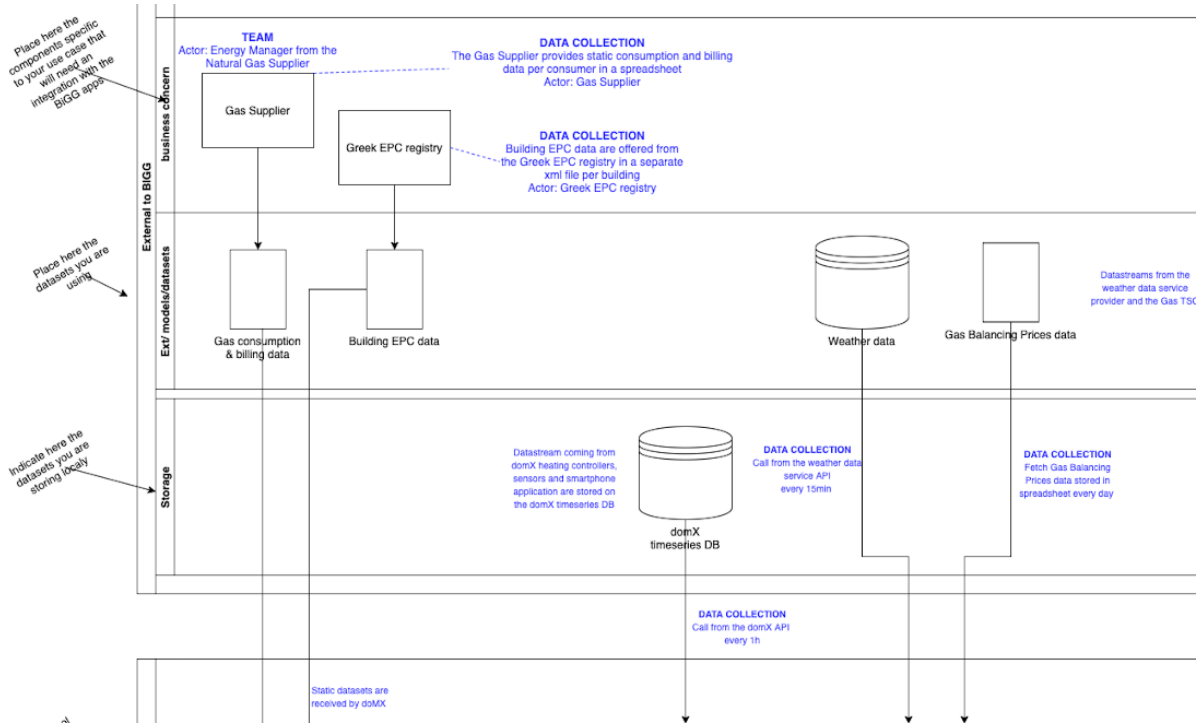
V.13. UC15

V.13.1. BPMN diagram


In this pool you should represent the various "actors" that will have interactions with the BIGG platform
Place here the components specific to your use case that will need an integration with the BIGG apps
Place here the datasets you are using
Indicate here the datasets you are storing locally
Place in this pool what should be done by the BIGG platform
External to BIGG
UI
business concern
Ext/ models/datasets
Storage
Reporting to end users
TEAM
Actor: Energy Manager from the Natural Gas Supplier
Gas Supplier
DATA COLLECTION
The Gas Supplier provides static consumption and billing data per consumer in a spreadsheet
Actor: Gas Supplier
Greek EPC registry
DATA COLLECTION
Building EPC data are offered from the Greek EPC registry in a separate xml file per building
Actor: Greek EPC registry
Gas Supplier
Gas consumption & billing data
Building EPC data
Weather data
Gas Balancing Prices data
Datastreams from the weather data service provider and the Gas TSO
Datastream coming from domX heating controllers, sensors and smartphone application are stored on the domX timeseries DB
domX timeseries DB
DATA COLLECTION
Call from the weather data service API every 15min
DATA COLLECTION
Fetch Gas Balancing Prices data stored in spreadsheet every day
DATA COLLECTION
Call from the domX API every 1h
Provided by BIGG RAF
Harmization layer
Storage
Toolbox
Graphical User Interface
Static datasets are received by doMX
domX
Connector/ Harmonizer WP3-4
DATA HARMONIZATION
Static datasets are cleansed, reshaped, sorted to match the BIGG format
Connector/ Harmonizer WP3-4
DATA HARMONIZATION
Data from the domX timeseries DB are cleansed, reshaped, sorted to match the BIGG format
Connector/ Harmonizer WP3-4
DATA HARMONIZATION
Data from the weather service and Gas market are cleansed, reshaped, sorted to match the BIGG format
domX Cloud SQL DB
domX Cloud Big Data
Data streams - time series data
data stream that are read in real time and collected at a given granularity
domX Cloud - Model DB
Efficiency mode consumption
Baseline mode consumption
Flexibility mode consumption
ENERGY EFFICIENCY MODE EVALUATION
Performance between the energy efficiency and baseline modes are compared to quantify the amount of achieved efficiency improvement
Efficiency/Flexibility KPIs
DATA COLLECTION
Call from the domX API
DATA COLLECTION
Call from the domX API
Local Data storage
Feature storage
Model storage
Evaluation storage
Final Model
Model Evaluation
Baseline mode Performance Estimation
Flexibility Potential Evaluation
Performance comparison
Achieved Flexibility Calculation
User Comfort Calculation
Energy Efficiency Calculation
domX real-time data visualization tool
-
Scope of WP3
DATA VISUALIZATION
Data is being displayed to match the end user analytics needs
Dashboarding on domX cloud
Flexibility Visualization
DATA VISUALIZATION
Data is being displayed to match the end user analytics needs
Dashboarding on domX cloud
User Comfort Visualization
Energy Efficiency Visualization

V.13.2. Specification of processes

Data Collection



<Data Object Name>

Type	Data Object
Name	*DomX timeseries DB*
Documentation	*Datastream coming from domX heating controllers, sensors and smartphone application are stored on the domX timeseries DB.*

<Data Object Name>

Type	Data Object
Name	*Weather data*
Documentation	*Datastream coming from the weather data service provider*

<Data Object Name>

Type	Data Object
Name	*Natural Gas Balancing Prices*
Documentation	*Natural Gas Balancing Prices data are collected in spreadsheet file*

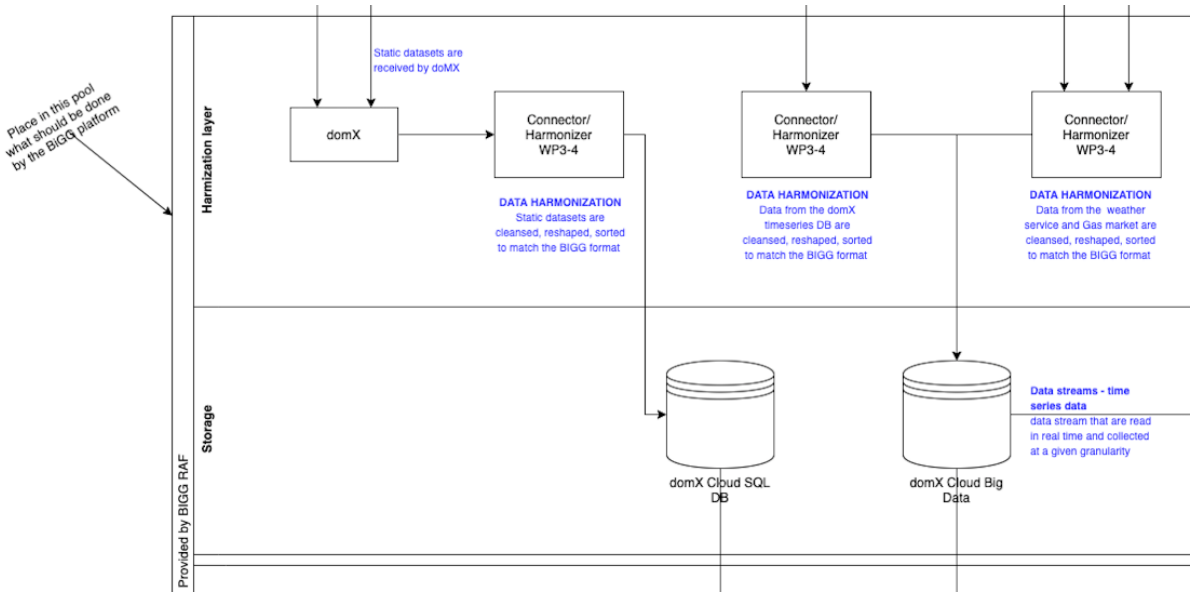
<Data Object Name>

Type	Data Object
Name	*Gas consumption & billing data*
Documentation	*The Gas Supplier provides static consumption and billing data per consumer in a spreadsheet file*

<Data Object Name>

Type	Data Object
Name	*Building EPCo data*
Documentation	*Static EPCo data are collected from the Greek EPCo registry in a xml file per building*

Data storage - harmonization



<Data Object Name>

Type	Data Object
Name	*DomX Cloud SQL DB*
Documentation	

<Data Object Name>

Type	Data Object
Name	*DomX Cloud Big Data*
Documentation	

<Process > [ID:<BC6-UC15-001>]

Type	Data collection Process
Name	*Static datasets collection*
Documentation	

<Process > [ID:<BC6-UC15-002>]

Type	Data collection Process
Name	*Real-time datastream collection*
Documentation	

<Process > [ID:<BC6-UC15-003>]

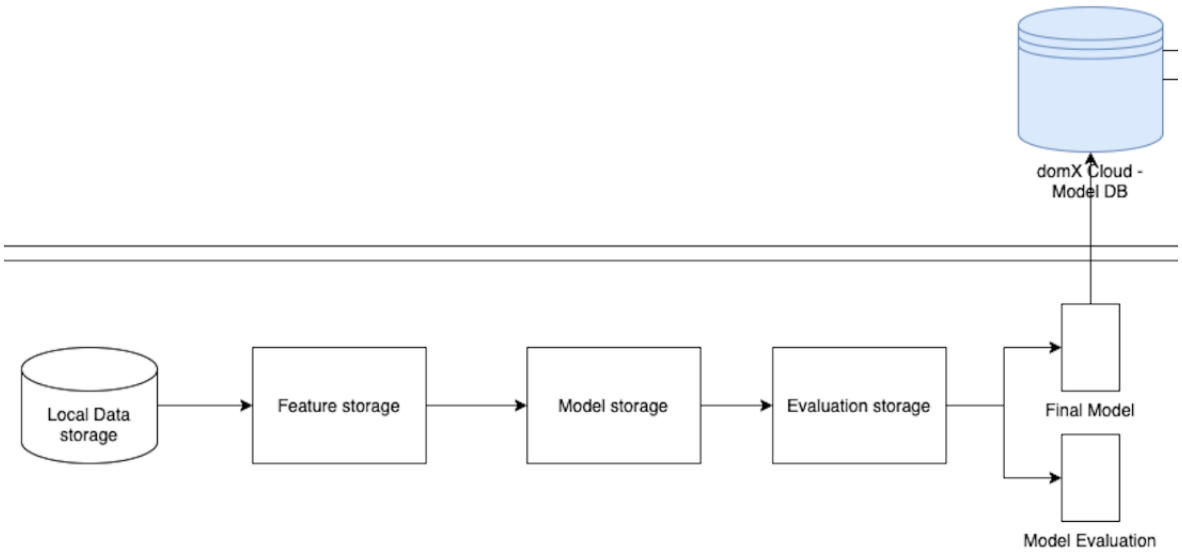
Type	Data collection Process

Name	*Public datasets collection*
Documentation	

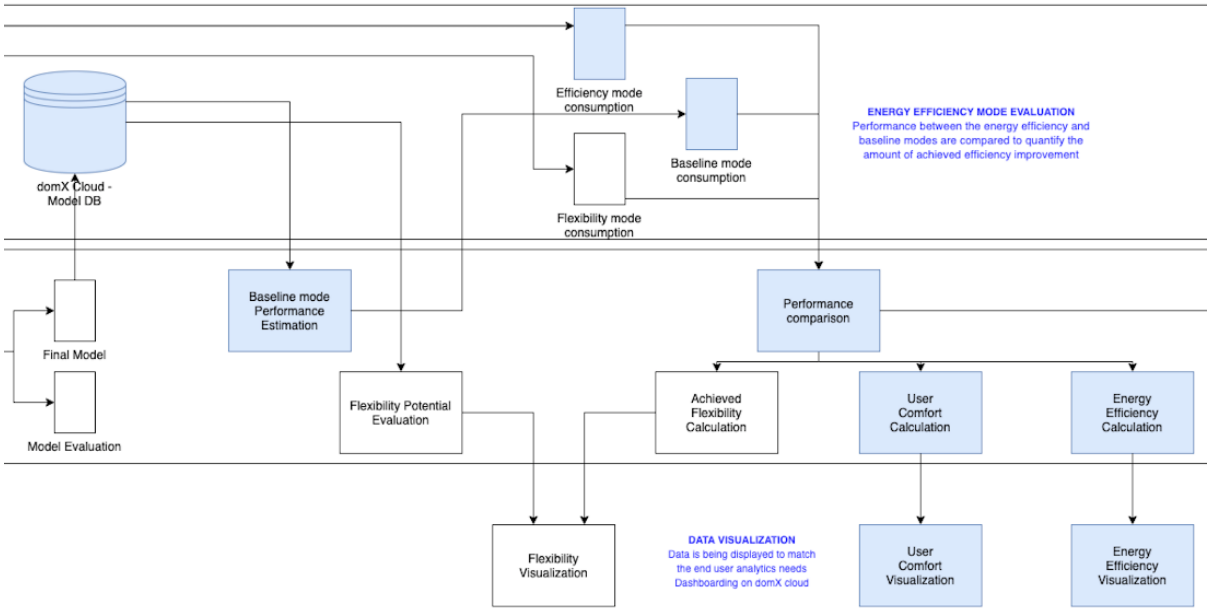
<Process > **[ID:<BC6-UC15-004>]**

Type	Data harmonization Process
Name	*Harmonization of collected data sets and storage in domx Cloud SQL DB and Big Data*
Documentation	

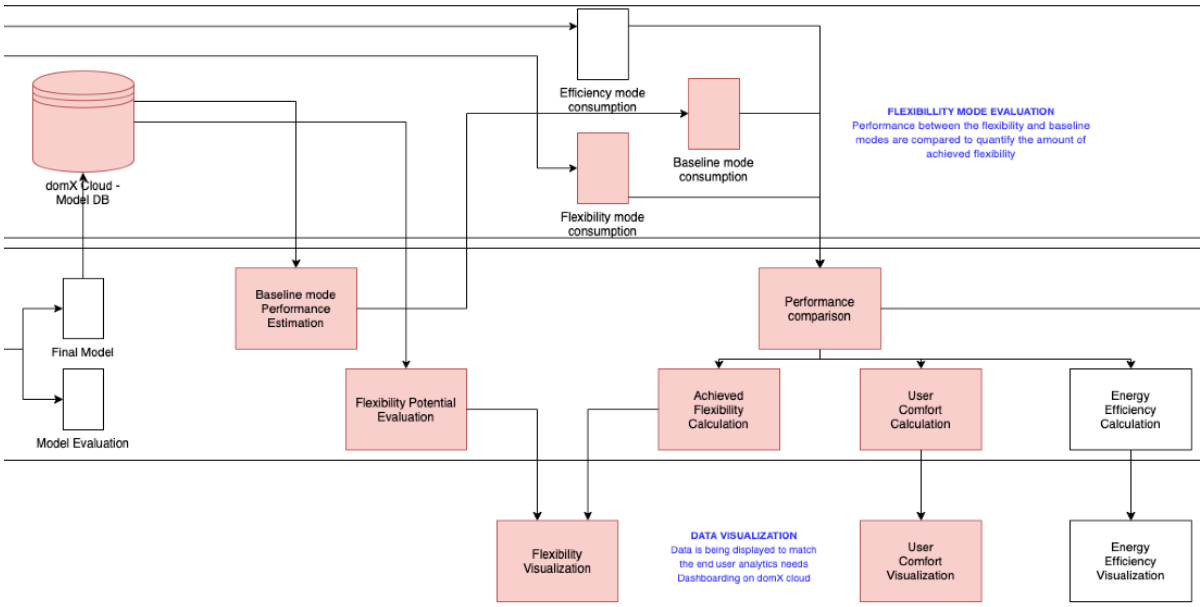
AI toolbox



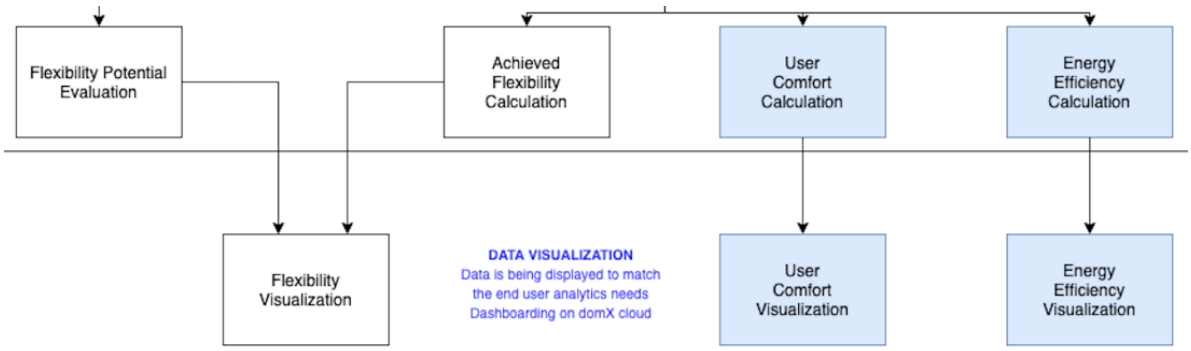
Energy Efficiency evaluation

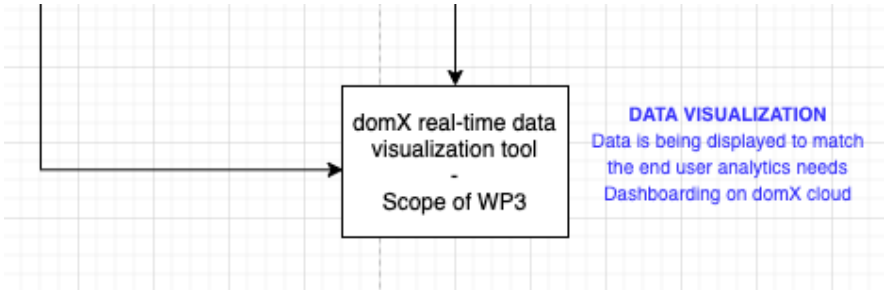


Flexibility evaluation

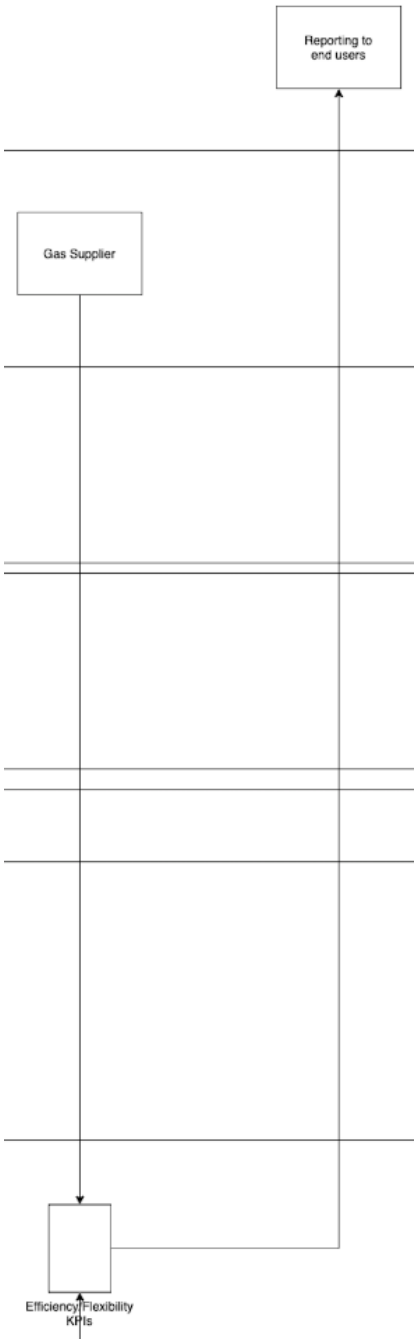


Data Visualization





Data Reporting



VI. REFERENCES AND INTERNET LINKS

Web platform

[1] Energis Cloud owned by Energis and rendered accessible to the members of the consortium through Helexia : http://www. Energis.cloud

[2] Engie Connect platform : The pre-existing IoT- based building automation platform used by Engie for facility management and energy management . In UC11-12-13, it will be used to interact with the equipment installed in the buildings (e.g. read the state of the devices, collect data and send control commands) via API.http://www.engie.gr/engieconnect/