MATRYCS

D5.1 | MATRYCS Analytics Building Services V1.0

WP5 – Analytics Building Service

December 2021

www.matrycs.eu

Modular Big Data Applications for Holistic Energy Services in Buildings

MATRYCS

Disclaimer

Copyright Message

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

Grant Agreement Number	101000158	Acronym	MATRYCS
Full Title	Modular Big Data Applications for Holistic Energy Services in Buildings		
Topic	LC-SC3-B4E-6-2020 | Big data for buildings		
Funding scheme	H2020- IA: Innovation Action		
Start Date	October 2020	Duration	36
Project URL	www.matrycs.eu		
Project Coordinator	ENG		
Deliverable	D5.1 - MATRYCS Analytics Building Services V1.0		
Work Package	WP5 - Analytics building services		
Delivery Month (DoA)	M13	Version	1.0
Actual Delivery Date	17/12/2021		
Nature	Other	Dissemination Level	Public
Lead Beneficiary	EURAC		
Authors	Daniele Antonucci, Manuel Mauro, Aaron Estrada, Pietro Zambelli [Eurac], Sofía Mulero Palencia, José Llamas, Víctor Serna, Roberto Sanz, Javier Antolín, Iván Ramos, Ana Vidal, Javier Cembranos, Laura Sanz, Gema Hernández Moral [CARTIF], Zoi Mylona, Nikolaos Sofias [HOLISTIC], Panagiotis Kapsalis, Haris Doukas, Vangelis Marinakis, Konstantinos Alexakis [NTUA], Leandro Lombardo, Marija Borisov [ENG], Carlonia Fortuna [COMSENSUS]		
Quality Reviewer(s):	Gema Hernández Moral, Sofía Mulero Palencia [CARTIF], Michal Staša [SEVEN]		
Keywords	Building Analytics, Energy Efficiency, Digital Twin, Geoclustering tool, Energy Conservation Measure (ECM), Building Automation Control (BAC), Technical Building Management (TBM), Measurement and Verification (M&V), IPMVP.		

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

Preface

MATRYCS focuses on addressing emerging challenges in big data management for buildings with an **open holistic solution** for Business to Business platforms, able to give a competitive solution to stakeholders operating in building sector and to open new market opportunities. **MATRYCS Modular Toolbox,** will realise a holistic, state-of-the-art AI-empowered framework for decision-support models, data analytics and visualisations for Digital Building Twins and real-life applications aiming to have significant impact on the building sector and its lifecycle, as it will have the ability to be utilised in a wide range of use cases under different perspectives:

- Monitoring and improvement of the energy performance of buildings - **MATRYCS-PERFORMANCE**
- Design facilitation and development of building infrastructure - **MATRYCS-DESIGN**
- Policy making support and policy impact assessment - **MATRYCS-POLICY**
- De-risking of investments in energy efficiency - **MATRYCS-FUND**

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

Who We Are

	Participant Name	Short Name	Country Code	Logo
1	ENGINEERING – INGEGNERIA INFORMATICA SPA	ENG	IT	ENGINEERING
2	NATIONAL TECHNICAL UNIVERSITY OF ATHENS	NTUA	GR	EPU
3	FUNDACION CARTIF	CARTIF	ES	CARTIF
4	RHEINISCH-WESTFAELISCHE TECHNISCHE HOCHSCHULE AACHEN	RWTH	DE	RWTH AACHEN UNIVERSITY
5	ACCADEMIA EUROPEA DI BOLZANO	EURAC	IT	eurac research
6	HOLISTIC IKE	HOLISTIC	GR	HOLISTIC
7	COMSENSUS, KOMUNIKACIJE IN SENZORIKA, DOO	COMSENSUS	SL	COM SENSUS
8	BLAGOVNO TRGOVINSKI CENTER DD	BTC	SL	BTC
9	PRZEDSIEBIORSTWO ROBOT ELEWACYJNYCHFASADA SP ZOO	FASADA	PL	FASADA
10	MIASTO GDYNIA	GDYNIA	PL	GDYNIA
11	COOPERNICO - COOPERATIVA DE DESENVOLVIMENTO SUSTENTAVEL CRL	COOPERNICO	PT	Coopérnico
12	ASM TERNI SPA	ASM	IT	ASM ASM Terni S.p.A.
13	VEOLIA SERVICIOS LECAM SOCIEDAD ANONIMA UNIPERSONAL	VEOLIA	ES	VEOLIA
14	ICLEI EUROPEAN SECRETARIAT GMBH (ICLEI EUROPASEKRETARIAT GMBH)	ICLEI	DE	ICLEI Local Governments for Sustainability
15	ENTE PUBLICO REGIONAL DE LA ENERGIA DE CASTILLA Y LEON	EREN	ES	EREN | Junta de Castilla y León
16	VIDES INVESTICIJU FONDS SIA	LEIF	LV	
17	COMITE EUROPEEN DE COORDINATION DE L'HABITAT SOCIAL AISBL	HOUSING EUROPE	BE	HOUSING EUROPE
18	SEVEN, THE ENERGY EFFICIENCY CENTER Z.U.	SEVEN	CZ	SEVEn

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

Contents

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

Figures

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

Tables

Abbreviations and Acronyms

Acronym	Description
BAC	Building Automation Control
DT	Digital Twin
DL	Deep Learning
DOM	Document Object Model
DT	Digital Twin
ECM	Energy Conservation Measure
EE	Energy Efficiency
FME	Feature Manipulation Engine
IPMVP	International Performance Measurement and Verification Protocol
KPIs	Key Performance Indicators
LSP	Large Scale Pilot
M&V	Measurement and Verification
ML	Machine Learning
OSM	Open Street Maps
TBM	Technical Building Management

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

Executive Summary

Deliverable 5.1 aims to describe the services that will be developed in MATRYCS identifying the types of development, input/output and application in different large-scale pilots (LSP). This deliverable is the first of a series of 3 documents that will update the status of the development of analytical services during the project (M13, M24, M31). The analytical services discussed in the document are related to (1) Analytics for energy performance - indoor condition evaluation and intelligent energy management; (2) Analytics for building systems and infrastructure; (3) Analytics for policy making and policy impact assessment on building level; (4) Analytics for building efficiency investments.

These Analytics Building Services, combined with the Digital Twins, will be exploited, in order to apply holistic energy services, categorized in 4 main groups, covering energy performance and indoor condition evaluation of buildings (MATRYCS-PERFORMANCE), building infrastructure (MATRYCS-DESIGN), policy making and policy impact assessment (MATRYCS-POLICY) and building efficiency investments (MATRYCS-FUND).

To make the document easier to read to the user, a summary of the services together with the relative pilots' scenarios are shown in the introduction and in the annex.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

1. Introduction

Building analytics can be powerful tools to assist in operating buildings, safely and efficiently. The result is tangible even at different scales, influencing the reduction of energy consumption and CO2 of neighbourhoods and cities.

With the increasing use of sensors in buildings and the resulting increase in data and information, the use of analytics to manage buildings and/or neighbourhoods/cities is critical.

Indeed, the data can be used to increase building efficiency, develop smarter systems, improve indoor comfort and air quality, provide better management of building systems and defining effective investments on buildings.

Moreover, with the events that have characterized the last two years, the interest in smart buildings has grown in particular because of the ability to guarantee technologies that can provide clean, reliable, and safe environments. Tenants are now looking for virus-safe buildings where they can live and manage their business. The use of big data combined with specific analytics makes it possible to verify and manage different information in one place, turning data into knowledge.

In the MATRYCS project, different analytics building services have been designed. The aim is to accommodate the demands of different pilots, covering a very wide range of possible demands from different types of users. Four main groups of services have been defined:

- MATRYCS – PERFORMANCE: It represents services able to manage all aspects related to energy performance and indoor conditions of buildings. Specifically:
  - **Energy Prediction (s1.1).**
  - **Building Automation Control (s1.2).**
  - **KPIs calculation (s1.3).**
  - **Technical Building Management (s1.4).**
  - **Optimization for network operations (s1.5).**
- MATRYCS – DESIGN: It focalizes in building infrastructure and how the design of infrastructures or their refurbishment can be supported. The main service is the ECM - evaluation with the related technologies catalogue.
  - **Technologies catalogues services to support the design and development of building infrastructure (s2.1).**
  - **ECM-based scenarios evaluation service (s2.2).**
- MATRYCS – POLICY: It provides services for policy making and policy impact assessment. The tools are:
  - **SECAPs decision-making support. (s3.1).**
  - **Energy Performance Certification harmonization and checking (s3.2).**
  - **National and EU policy impacts assessment and support (s3.3).**
- MATRYCS – FUND: in which analytics for de-risking investments in energy efficiency will be developed. The tools are:
  - **Services to support Energy Savings M&V towards improved Energy Performance Contracts (s4.1)**

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

- **Services to support the financing of the EE refurbishments (s4.2).**

Together with these groups, two other tools will be developed to cover fundamental aspects of building design and management. The two tools are **Digital Twins (s0.1)** and the **Geoclustering tool (s0.2).**

Figure 1 shows the pilots schema as defined in the proposal phase together with the four main service groups.

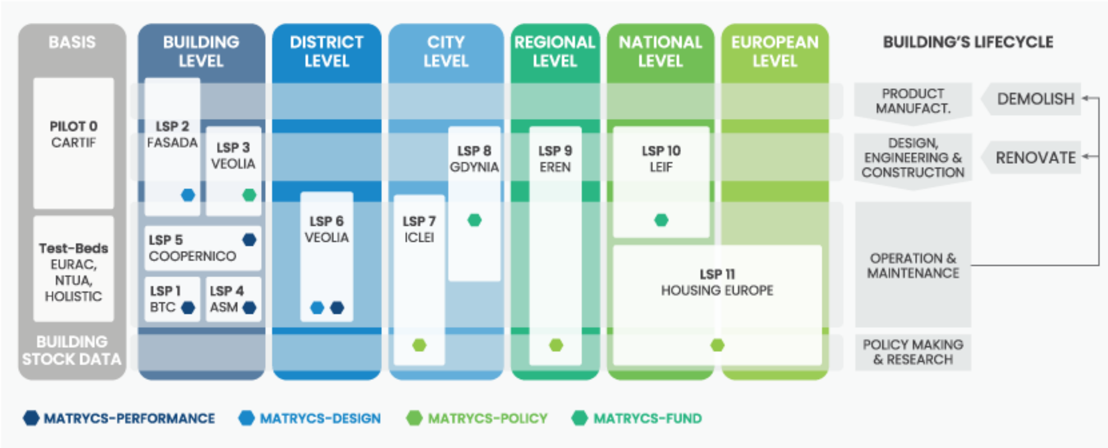

Figure 1 - MATRYCS Large scale Pilots classification and scope.

Furthermore, the following table (Table 10 of the Deliverable D6.1) shows the association of the case study with the related service.

Table 1 MATRYCS Services/Pilots' matrix

			MATRYCS Large Scale Pilots (LSPs)										
			1	2	3	4	5	6	7	8	9	10	11
MATRYCS SERVICES	s0.1	Digital twins	X	X				X		X	X		
	s0.2	Geoclustering								X			
	s1.1	Energy prediction	X			X	X	X					
	s1.2	Building Automation and Control (BAC)	X			X	X						
	s1.3	KPIs calculation								X	X		
	s1.4	Technical Building Management (TBM)	X				X						
	s1.5	Optimization for network operation				X		X					
	s2.1	Technologies catalogues					X						
	s2.2	ECM-based scenarios evaluation		X									
	s3.1	SECAPs decision-support							X				
	s3.2	EPC harmonization and compliance									X		
	s3.3	National and EU policy impacts assessment							X			X	X
	s4.1	M&V of energy savings			X			X					
	s4.2	Financing of EE refurbishments		X								X	

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

In order to have a more streamlined document the code of each service will be made available in the official repository of the project in github: https://github.com/MATRYCS

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

1.1 Purpose of this document

The document aims to provide exhaustive information about the analytics services for buildings have been developed and tested so far in MATRYCS and which will be improved over the course of the project by documenting them in subsequent versions of this deliverable.

Analytics are the point of contact between the user and the whole architecture of MATRYCS. Their role is fundamental because on the one hand it must offer the user the possibility of obtaining certain insights from the data provided and available within the platform, and on the other hand it is required a connection and synergy with all the MATRYCS components (developed in particular in WP3 and WP4) to ensure that the best use is made of the 5V characteristics of big data: Velocity, Variety ,Volume, Velocity and Value.

Evaluating the Figure 2 which represents the **D**ata**I**nformation**K**nowledge**W**isdom pyramid, it is clear that the part relating to information and knowledge on data can only be obtained through the use of suitable tools that with the use of machine learning techniques or specific KPIs provide the basis for the user to improve the performance of the building, generating energy savings, economic and CO_2 reduction.

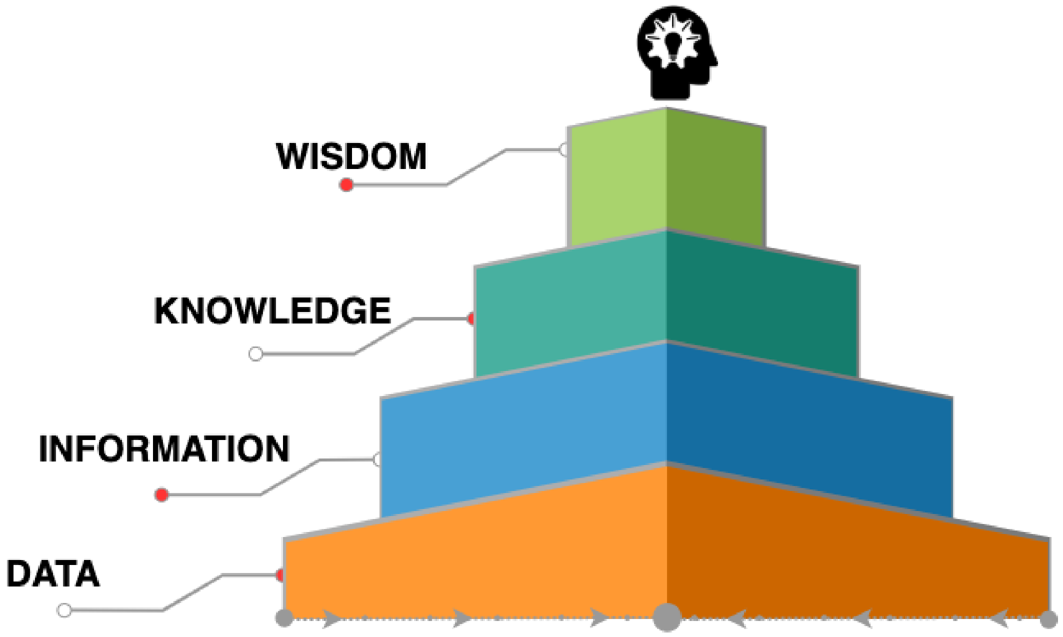

Figure 2 - From Data to Wisdom

1.2 Structure of the document

Each building analytical service described through 4 main sections:

- Tool description: This section will describe the tool in general, giving an overview of the service to be offered. The following sections will go into detail, defining the more technical aspects of the development and implementation of the tool, also taking into account the links with other WPs.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

- Input/output: each tool requires specific inputs to be defined in order to obtain correct results. In this section we will illustrate which inputs are used for each service, describing the parameters needed to ensure that the tool functions correctly and which outputs each service can generate.
- User experience: the aim of the tool is to be used by different stakeholders. In order to get a correct view of how the tool works, this section shows how an individual user can approach the service and what are the expected steps he/she has to take. By analyzing the user experience, it is possible to identify which points are necessary to ensure a user-friendly view of the platform.
- Application on MATRYCS project. This last section shows in a practical way how the tool has been used in the different MATRYCS large-scale pilots, providing possible generalizations of application in subsequent versions of the tools and deliverable.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2. Services development and application

2.1 MATRYCS – Digital Twin (s0.1)

2.1.1 Tool description

Digital Twin (DT) can be defined as a connected, digital representation of a physical building (district, city, region, etc.) and corresponding processes that are used to understand, predict and optimise performance in order to achieve a more cost-effective, straightforward and sustainable smart building. It brings together dynamic and static data from multiple sources in 2D/3D models and enables informed and effective decisions to be made. It provides real-time understanding of how a building is performing – enabling immediate adjustment to optimize efficiency and to provide data to improve the design and management of future buildings or districts.

There are several layer-based approaches to the architecture of a Digital Twin (DT). Starting with the core of the building's DT, three layers can be distinguished: a **physical layer** that corresponds to sensorisation (physical equipment that functions as data sources); a **data layer** in which all the available data would be concentrated, (such as the data repositories and the real-time information); and finally the **model layer**, where the reasoning, behaviour, simulation and prediction models would be located (Figure 3)

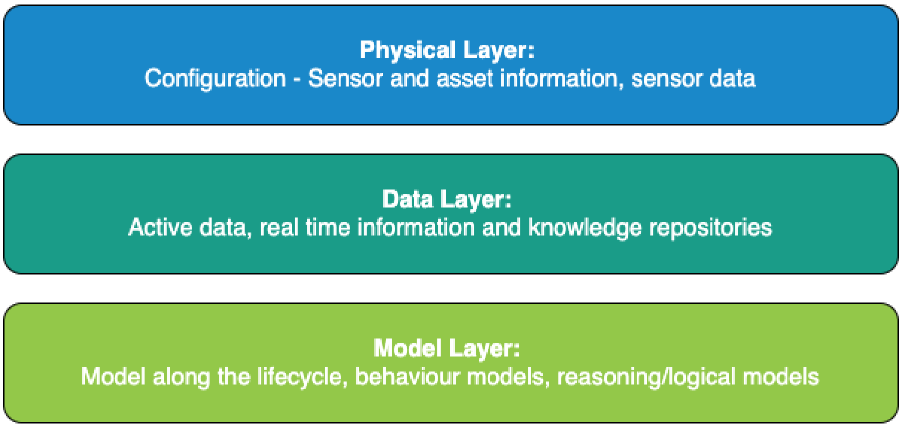

Figure 3 - Digital Twin schema

Other approaches expand the number of layers by adding a user interface layer or subdividing some of the previous layers (a cloud storage layer could be added in the data layer; an analysis layer could be included in the model layer; and the physical layer could also include communications, IoT). In the MATRYCS project, these three main layers will be used, and other layers will be included depending on the use cases and the specific requirements of each application.

This common model of DT will be taken under consideration based on three layers at three different scales (building, district and region / nation) since the approach is different in each case (e.g. file formats, services). These scales correspond to the main typologies of the MATRYCS use cases.

The main objectives of the development of the DT in MATRYCS are the following:

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

- The DT will compile the relevant information to be used at three different scales to support the implementation of energy services.
- The scales considered are related to the main standard to be used in each case: (1) Building level (IFC), (2) District / City level (CityGML) and (3) Regional /National scale (INSPIRE data); which should be combined with the corresponding real and monitored data in each scale.
- The DT will support the deployment of other services by integrating all relevant information at each scale, and automatically generating models based on other data sources if they do not exist.
- As an overarching service, the related business cases could be applicable in every intermediate scale and be related to the stakeholders that deploy their activity at that level. In principle, all LSPs could benefit from the DT service.

In MATRYCS project, DTs will be considered at three different scales: building, district and regional / national.

2.1.1.1 Digital Twin – Building level

At the building level scale, the following specific objectives are proposed:

- Create a coherent aggregation of data from multiple sources and link them to a digital model of the building to obtain an initial DT.
- Integrate the different analytics building services in the DT answering the needs specified in the pilots and captured in the user stories and the specifications.
- Modelling the elements and behavior of the entities taking part in the DT.
- Prepare/adapt the DT to the required inputs and outputs, the required predictions and behavioral analysis, as well as the features to be applied for the ML and DL definitions.

2.1.1.2 Digital Twin – District level

The Digital Twin at district level will provide relevant information on buildings at district to support the implementation of energy services. As a transversal tool, it will support the development of other services. At the end, the tool will allow users to:

- Visualize building situation in the city considering the geometry in a 3D view.
- Visualize the main building parameters for energy demand calculation: year of construction, building use, mean height, envelope parameters, theoretical energy system, etc.
- Compare building performance by means of Energy Performance Certificates (EPCs) if they are available in the area or default values of energy consumption according to climate parameters.

The definition of a 3D building scene at district level is important for example to define accurate profiles of the energy consumption including solar gains and shadowing effects by close buildings. The results displayed in the tool will be represented with different colours under standardized ranges for each variable.

The general concept of the tool architecture is presented in Figure 4. This architecture covers de integration of the recovered information at pilot scale in a CityGML file and the transformation to be represented in a user-friendly interface.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

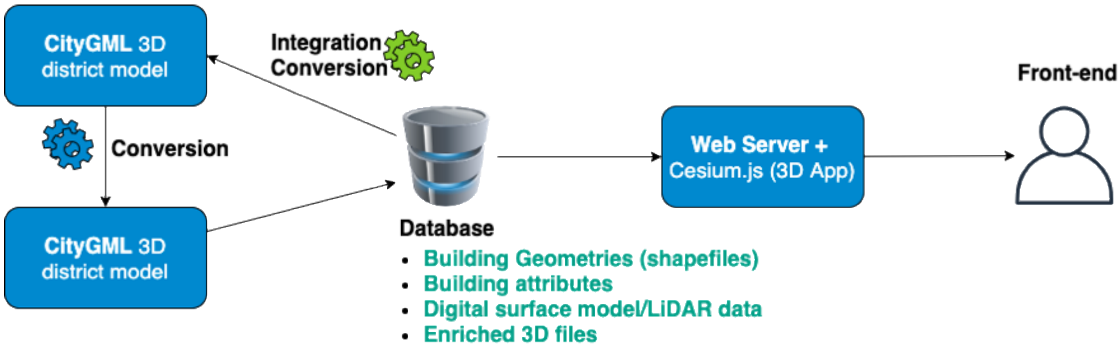

Figure 4 - Digital Twin tool architecture at district scale

- **Front-end**: The Front-end will be developed using Material UI components for Bootstrap and the 3D Javascript library Cesium. This library will provide a user-friendly 3D environment to visualize the location of each building in the district.
- **Back-end**: As a first approach, the Back-end will be developed in Python. Processing algorithms will be coded using python libraries for geo-located data and data transformation will be covered with Feature Manipulation Engine (FME) software. Alternatives to this software will also be studied. Flask framework will be use to ensure the communication between the enriched 3D files stored in the database and the Graphical User Interface.
- **Database**: The most suitable solution to store the data will be studied, which could be the use of a relation SQL database. In this sense, a PostgreSQL-PostGIS database that includes capabilities to work with geo-located data could be a good alternative.

2.1.1.3 Digital Twin – Regional level

The target of the DT at regional level is to create a coherent aggregation of data from multiple sources, in order to obtain an estimated version of the energy performance of the buildings in a given region.

Up to date, some potential hurdles have been detected for this topic:

- The incomplete data might be not uniformly distributed, for example, the Open Street Maps (OSM) data does not show information on the year of construction and use of the buildings, or there is a lack of relevant information in the cadaster data for some areas. Therefore, the analyses could offer results not 100% valid on certain cases.
- Some data sources might not be available in certain European countries, and the alternatives should be feasible.
- The large quantity of data will require the usage of Big Data methodologies in order to handle and store them. Therefore, it will be necessary to pay attention to data processing, so there would be relatively small files for general reports.

In the next stage these issues will be addressed:

- Analyze datasets: to decide whether cadastral data can be complemented with data taken from unprotected cadastral data;
- Decide the data model to be used for DT;
- Geometric part and data to add from the results.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.1.2 Input/output

2.1.2.1 Digital Twin – Building level

As a first approach to the development of the DT, the integration of sensor data and historical data in BIM models (using commonly accepted formats: IFC and RVT mainly) will be performed.

Sensor data can be integrated into Revit BIM models[1] with the Revit tool called Dynamo2. It allows to create predefined scripts and execute them from Revit, asking end users for a few necessary inputs only. The following script has been created to monitor the sensors included in the DT of the building. First, the IFC file can be imported into Revit. Then, the scripts in Dynamo are created and executed through the Dynamo Player, placed in the Revit interface, to execute the following commands programmed for DT monitoring tasks (Figure 5):

- To create a "sensor box" element.
- To define parameters that will represent the data in the "sensor box" element.
- Collect the sensor data directly from the excel/csv/other provided file and represent it into the previously defined parameters.
- To operate with the data. For instance, a schedule can be created and exported to PDF.
- To export the model to an IFC file with the new data inputs.

Figure 5 - Revit: Tool Dynamo data flow (script created for DT monitoring tasks)

[1] https://www.bimobject.com/en/product?sort=trending
[2] https://dynamobim.org/

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

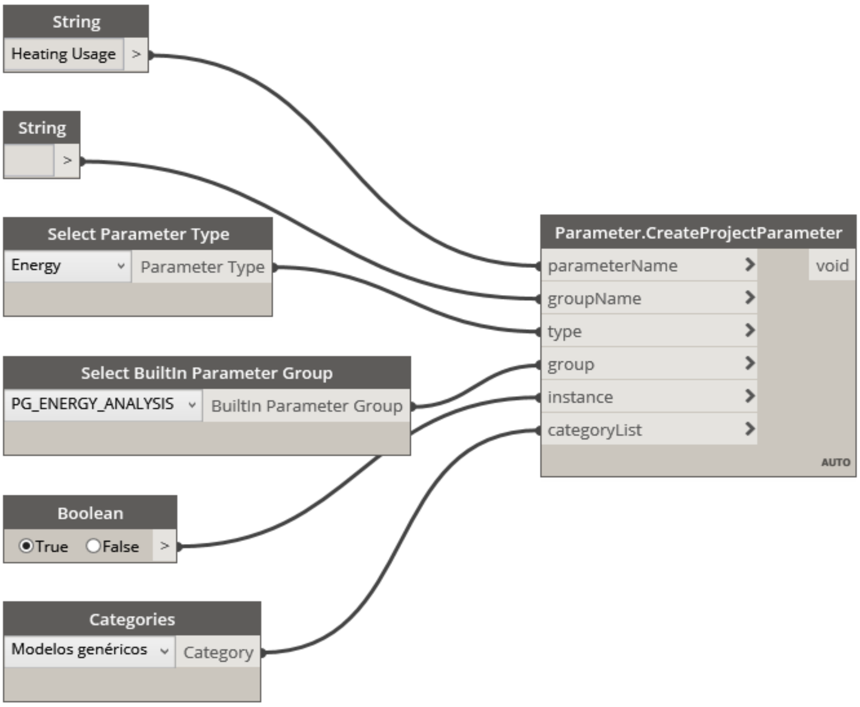

Figure 6 - Example of Revit/Dynamo scripts for integrating sensor data into BIM files

2.1.2.2 Digital Twin – District level

The necessary inputs for the service implementation will be the following:

- Buildings' geometries: they will be collected from cadasters, if available for the pilot. In case cadastral data are not available, building geometries will be obtained from OpenStreetMaps (OSM).
- Building attributes (e.g. year of construction, building use, number of floors, height) if available along with geometries.
- Other building parameters such as envelope values, glazing percentages or energy systems derived from EU databases (e.g. TABULA, BSO, Eurima)
- Digital Surface Model (DSM) derived from LiDAR[3] data or LiDAR point data to improve or obtain height values. If point LiDAR data is to be used, it needs to be pre-processed in order to identify points that were classified in a wrong class.
- Copernicus data from "Urban Atlas" in EU cities to identify changes and/or heights if LiDAR data is not available.

The outputs of the service will be the following:

- An engine to calculate building height based on different approaches.
- Transformation model to generate CityGML and 3D tiles generation.
- User-friendly visualization tool.

[3] https://pro.arcgis.com/en/pro-app/latest/help/data/las-dataset/what-is-lidar-.htm

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.1.2.3 Digital Twin – Regional level

The Digital Twin at regional level service will use information from public data sources. The sources identified are cadastral information and information from Open Street Maps. The information used will be manly the following: geometric information, gross floor area, number of floors above ground, number of floors below ground year of construction, information on the use of the building, number of dwellings. It is important to note that the geometric information needs to be georeferenced.

The service will use the common data model as a base and will complete the information following this data model. The output of the tool will be a Digital Twin at regional level for a region will all the information gathered.

This service will be used as the basis for the calculation in other services: mainly s1.3 (KPIs calculation) and s3.2 (EPC harmonization and compliance). These services will use the building information contained in the DT at the regional level in order to extract the necessary parameters to carry out the calculation and estimations.

But in addition, this DT will be fed by the results of these services and other sources, adding processed information (energy demand estimations, consumption, etc.)

All the necessary data for the service to operate can be retrieved from MATRYCS' databases directly or by using the APIs (cadaster API or OSM API) in the future if necessary.

2.1.3 User experience

2.1.3.1 Digital Twin – Building level

For the first approach, using Dynamo tool, the scripts created can be executed directly with the Dynamo Player, which is placed inside Revit interface: Manage tab -> Visual Programming panel -> Dynamo Player. It provides a default directory where to see the status of the scripts and allows end users to launch a script, provide inputs and edit the script in Dynamo, if necessary.

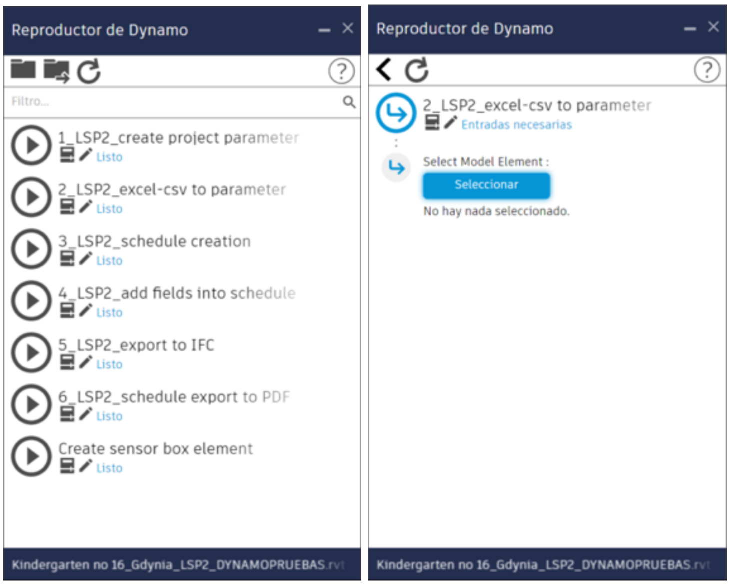

Figure 7 - Dynamo Player

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

This approach implies using a commercial software, hence, the use and programming of several open-source BIM tools will be evaluated in the next steps. The tools that are considered in the first instance that could be useful are: Xbim Xplorer[4] and BIMvision[5] (Figure 8) and Xeokit.

The Xbim[6] toolkit is a .NET open-source software development BIM toolkit that supports the BuildingSmart Data Model (also known as the Industry Foundation Classes IFC).

Xbim allows .NET developers to read, create and view Building Information (BIM) Models in the IFC format. There is full support for geometric, topological operations and visualization. In addition, Xbim supports bi-directional translation between IFC and COBie formats. Core libraries for data manipulation are all written in C#, the core of the geometry engine is written in C++.

Xbim also provides XbimWebUI, an open-source Typescript/Javascript library that can be used for web presentation of BIM models. It uses WebGL and is independent of any third party WebGL framework. XbimWebUI, being based on web technologies, has great flexibility by allowing the display of IFC models in all types of web browsers.

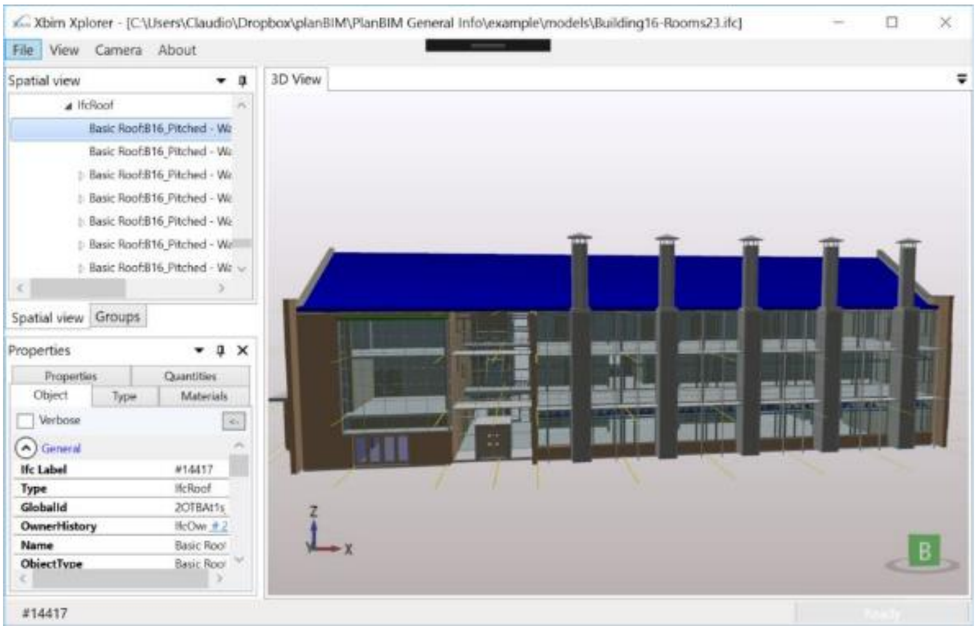

Figure 8 - Xbim explorer sample interface

BIMvision is a freeware IFC model viewer available only for Windows Operating Systems (Figure 8). It allows to view the virtual models coming from CAD systems such as Revit[7], Archicad[8], BricsCAD BIM[9], Advance[10], DDS-CAD[11], Tekla[12], Nemetschek VectorWorks[13] and others without the need to have commercial licenses for these systems or to have a viewer from each provider in particular. BIMvision offers visualization of BIM models created in IFC format 2×3 and 4.0.

[4] https://docs.xbim.net/downloads/xbimxplorer.html
[5] https://bimvision.eu/
[6] https://docs.xbim.net/
[7] https://www.autodesk.com/products/revit/overview
[8] https://graphisoft.com/solutions/archicad
[9] https://www.bricsys.com/en-intl/bricscad-bim
[10] https://www.graitec.com/advance-bim-designers-2020-is-here/
[11] https://www.dds-cad.net/
[12] https://www.tekla.com/
[13] https://www.vectorworks.net/en-GB

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

It supports a plugin system that can be extended by other plugins, in order to develop a specific data visualization or processing extension. The plugin SDK is written in C++ and depends on the the .NET framework API. Unlike other solutions, they do not provide a web-based SDK, which makes this a platform dependent application.

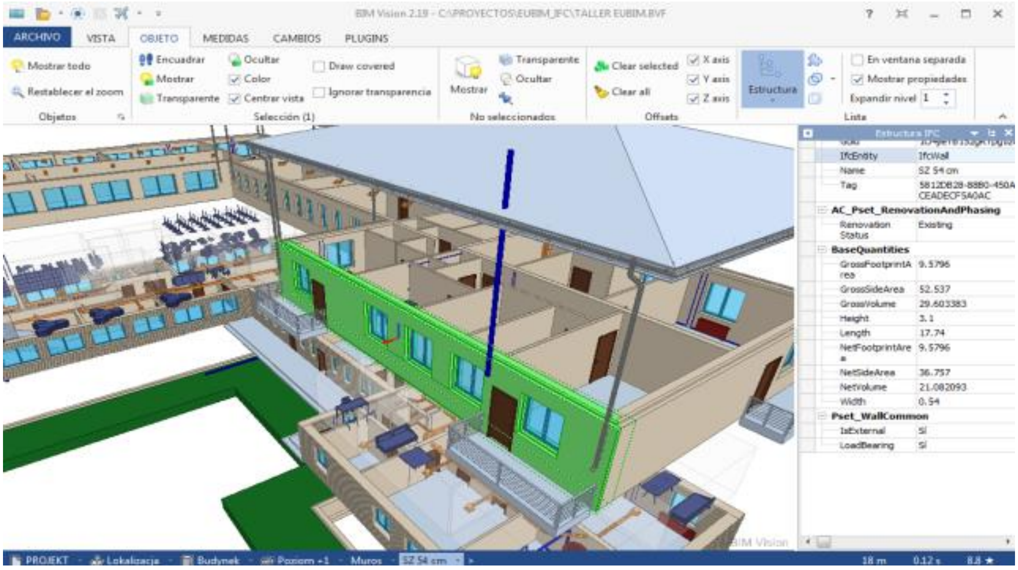

Figure 9 - BIM vision sample interface

Finally, Xeokit-bim-viewer from Xeolabs is a 2D/3D BIM viewer that runs in the browser. The company involved in the development of this tool also provides an open-source JavaScript software development kit for working with high-detail, full-precision BIM models. This feature allows users to view huge and complex models in all major browsers, including mobile.

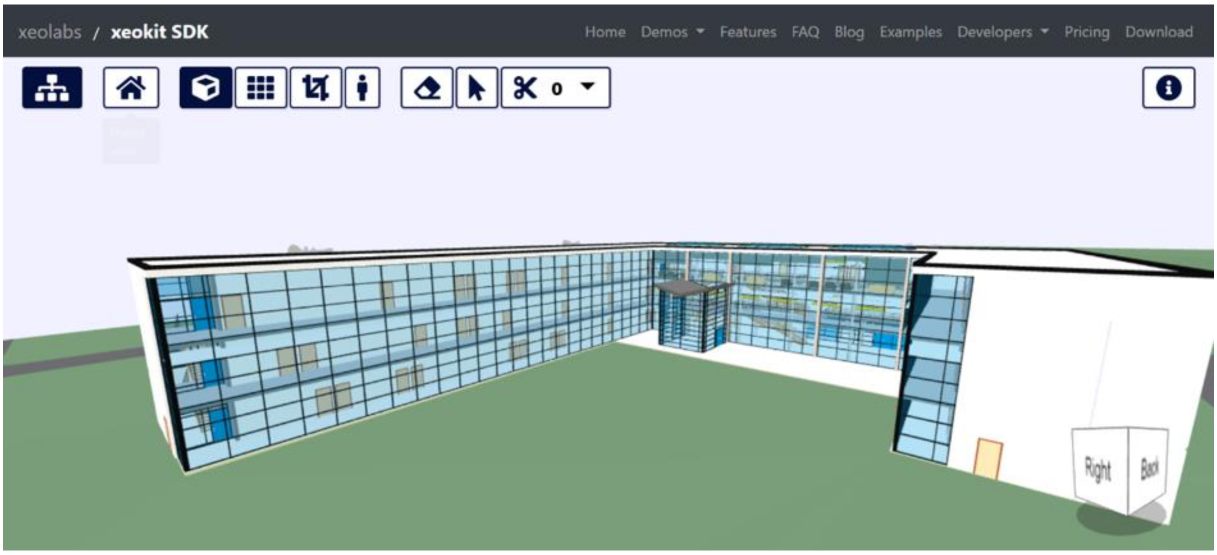

Figure 10 - BIM vision ion Xeolabs

Windows applications, such as BimVision or .NET application-based tools, such as those that rely on the Xbim .NET library, are by nature dependent on a specific type of architecture and operating system.

The use of web-based tools, such as XbimWebUI and Xeokit SDK, provides greater flexibility since, as already mentioned, they work in all major browsers and thus, eliminating dependence on a specific type

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

of architecture and operating system. Preliminary testing has been carried out with both tools and, even though both tools are very strong in terms of behavior in terms of number of features and performance, XbimWebUI will be used in the initial proof of concept.

It is considered that the use of these tools could complement the visualization envisaged in the MATRYCS Framework without relying on commercial software such as Revit or ArchiCAD. In any case, the final decision to use these tools will depend on the results obtained in the first iteration of the development and the ease of integration in the MATRYCS platform.

2.1.3.2 Digital Twin – District level

The service will be use by means of a web-based interface in which CityGML models will be included. Using a browser, the user will be able to select the area of interest in which 3D models are available. The following steps could describe the user experience:

- Enter the service.
- Choose the CityGML model to be evaluated.
- Select and compare features from different buildings.

Figure 11, presents a graphic view of a 3D scene similar to the one proposed in this service.

Figure 11 - 3D scene represented in Cesium

2.1.3.3 Digital Twin – Regional level

The Digital Twin at the regional level aims to display information about the buildings in a region. The information that can be shown is less specific that for DT at other levels, so mainly information about the geometry of the building and some other general parameters. Besides, if this service has been fed by other service as s1.3 or s3.2 it could show estimation information (energy demand and consumption, CO_2 emissions, etc.).

The utility of this energy DT could be tested mainly in decision-making processes, where it would be necessary to have a clear idea about the uses and consumption of energy, its geographical distribution, and what would happen if an external factor modified the situation, at the regional level.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

The Digital Twin at regional level visualization could be generated using Leaflet (javascript library), creating a web page. This web page could be visualized form one standard browser. More functionalities could be added to the tool, but more complex tools related to this DBT will be developed at regional level for s1.3 and s3.2.

The following figures show a possible view of the DBT at the regional level.

Figure 12 - Digital Twin at regional level

Figure 13 - Digital Twin al regional level showing demand information (estimate)

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.1.4 Application on MATRYCS Pilots

2.1.4.1 Digital Twin – Building level

The use cases addressed at this scale are:

- Use Case 1 [s0.1 DT] (Digital Twin creation).
- Use Case 2 [s0.1 DT] (Services integration in Digital Twin).

The DT will be tested and validated in **LSP 01** (BUILDING OPERATION: Facility and resources fingerprinting for Efficiency and Optimal Balancing of Energy Vectors [BTC]) and **LSP 02 (**BUILDING REFURBISHMENT: Sustainable building assessment and optimization of refurbishment options [FASADA]) pilots. For this first approach, LSP02 is used, whose data available are:

- Kindergarten Revit file.
- Kindergarten IFC file.
- Energy consumption data (per month, twenty months stored).
- Temperature and humidity sensor data in 4 minutes frequency (two years stored).

	time	TAir	RH	tmw1	tmw2	tmw4	tmw5	tmw3	tmc	TR
0	05/07/2018 12:22	29,8	38	23,5	23,3	22,3	22,9	22,9	22,6	23.565.178.521.944.100
1	05/07/2018 12:26	29,8	38	23,5	23,4	22,3	22,9	23,3	22,8	23.715.850.035.112.700
2	05/07/2018 12:30	29,8	38	23	23,3	22,3	22,6	22,8	22,8	23.501.516.459.248.100
3	05/07/2018 12:35	29,9	38	23,4	23,4	22,3	23	24	22,5	23.920.454.231.476.700
4	05/07/2018 12:39	29,9	38	23,4	23,3	22,3	22,8	22,3	22,4	2.335.057.241.851.160
5	05/07/2018 12:43	29,9	38	23,8	23,6	22,5	23,3	23,1	22,9	23.810.560.735.115.600
6	05/07/2018 12:47	30	38	23,5	23,7	22,3	22,6	22,1	22,5	23.322.651.871.149.600
7	05/07/2018 12:52	29,9	38	23,2	23,5	22,3	22,4	22,1	22,2	23.245.755.509.933.400
8	05/07/2018 16:08	29,4	39	22,2	22,4	21,9	21,9	21,3	21,6	2.259.790.801.214.880
9	05/07/2018 16:12	29,4	39	22,1	22,3	21,9	21,9	21,3	21,6	22.583.661.518.106.700
10	05/07/2018 17:18	29,3	40	21,9	22,1	21,9	21,8	21,1	21,7	2.248.950.295.672.130
11	05/07/2018 17:44	29,3	40	22	22,2	21,7	21,8	21,2	21,4	22.440.008.979.649.200
12	05/07/2018 19:43	29,2	40	21,8	22	21,9	21,7	21,1	21,4	22.441.219.816.078.200
13	06/07/2018 0:09	29	40	21,4	21,6	21,7	21,5	21	21,2	22.243.255.076.423.500
14	06/07/2018 2:51	28,8	40	21,3	21,5	21,5	21,3	21	21	22.118.727.865.082.600
15	06/07/2018 3:34	28,8	40	21,3	21,5	21,4	21,3	20,9	21	2.205.185.616.685.540
16	07/07/2018 4:05	28,8	40	21,2	21,3	21,6	21,3	20,7	21,2	22.051.544.453.843.600
17	07/07/2018 4:09	28,8	40	21,2	21,4	21,4	21,4	21	20,9	22.072.682.166.731.200
18	07/07/2018 11:09	28,9	40	21,6	21,8	21,6	21,4	20,9	21,2	22.193.172.648.205.800
19	07/07/2018 11:43	28,9	40	21,6	21,8	21,5	21,3	20,9	21,2	22.145.481.763.712.600
20	07/07/2018 11:47	28,9	40	21,6	21,8	21,5	21,4	21	21,2	22.187.912.187.873.500
21	07/07/2018 13:15	29	40	22	22,6	21,6	21,7	21,1	21,4	2.240.849.482.226.570

Figure 14 - Example of sensor data available in LSP 02.

2.1.4.2 Digital Twin – District level

This use case at district level will be developed in the LSP6. The datasets necessary to build the DT at district level is the **LSP6** DISTRICT HEATING NETWORK: Energy demand prediction to design and develop DHN and optimize the operation [VEOLIA] are:

- Cadaster buildings, if available for the pilot. Building attributes are also useful in characterizing demand estimates.
- Buildings from OpenStreetMaps (OSM), if cadastral data are not available.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

- LiDAR[14] data to improve height calculation. These data need to be pre-processed to identify the data that were classified in a wrong class.
- Copernicus data from "Urban Atlas" to identify changes and/or heights if LiDAR data are not available.

2.1.4.3 Digital Twin – Regional level

The use cases will be centered in the **LSP9**:

- Use Case 1: EPCs data error checker and validator.
- Use Case 2: Visualization of estimated EPC for the s3.2 service, as well as the Use Case 1 and 2 for the s0.1 service considering Digital Twin (Services integration in Digital Twin).

Currently the following sets of data are under process:

- **Spanish cadastre:**
  - Relevant data for Castilla y León region have been collected (2248x4 files: BU, Bupart, CP and CZ files - 43.02GB).
  - In the process of loading to the platform.
  - There is an online service of unprotected data, but up to date, the information cannot be downloaded in batch.
- **OpenStreetMaps:**
  - Maps that can be consulted online.
  - Maps that can be consulted online.
  - Possibility to download all the data (directly or using API).
- **Regional Spanish EPC:**
  - Public EPCs offered by Castilla y León. Available in web.
  - EPC XML repository of Castilla y León. This information is not public (it belongs to the regional administration). Currently, the process of collecting it is taken place. This corresponds to around 99.000 files (17,8 GB).

[14] https://pro.arcgis.com/en/pro-app/latest/help/data/las-dataset/what-is-lidar-.htm

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.2 MATRYCS – Energy Prediction (s1.1)

The energy prediction service (s1.1) is part of MATRCYS PERFORMANCE services. The tool developed aims to support energy analysis and improve building performance through better management of the controls that characterise the building/plant system using predictive algorithms. In particular, the energy consumption of buildings or the production of photovoltaic systems will be predicted. In this first version, 3 datasets from 3 different pilots have been used:

- LSP4 – ASM for data referring to photovoltaic production and electricity consumption.
- LSP5 – COOPERNICO for data related to photovoltaic consumption.
- LSP6 – VEOLIA for data related to heating energy consumption.

2.2.1 Tool description

Energy prediction service – s1.1, offers predictions of various energy sources. It uses time series of these energy sources, in order to predict the next time steps' value in a modular way, providing the suitable algorithms and models depending on the use case and the predicting target variable. It is also configurable, by taking under consideration the variables needed to be used as inputs and decide the optimal algorithms to be used in each case. The energy prediction service can be used as a stand-alone service, where the results can be visualised or can be used as a sub-service, providing input to more advanced applications, which would require a predicted value of an energy source. The service is being developed in Python programming language and utilises several ML/DL libraries such as Scikit Learn, Tensorflow and Keras.

The service architecture is shown below:

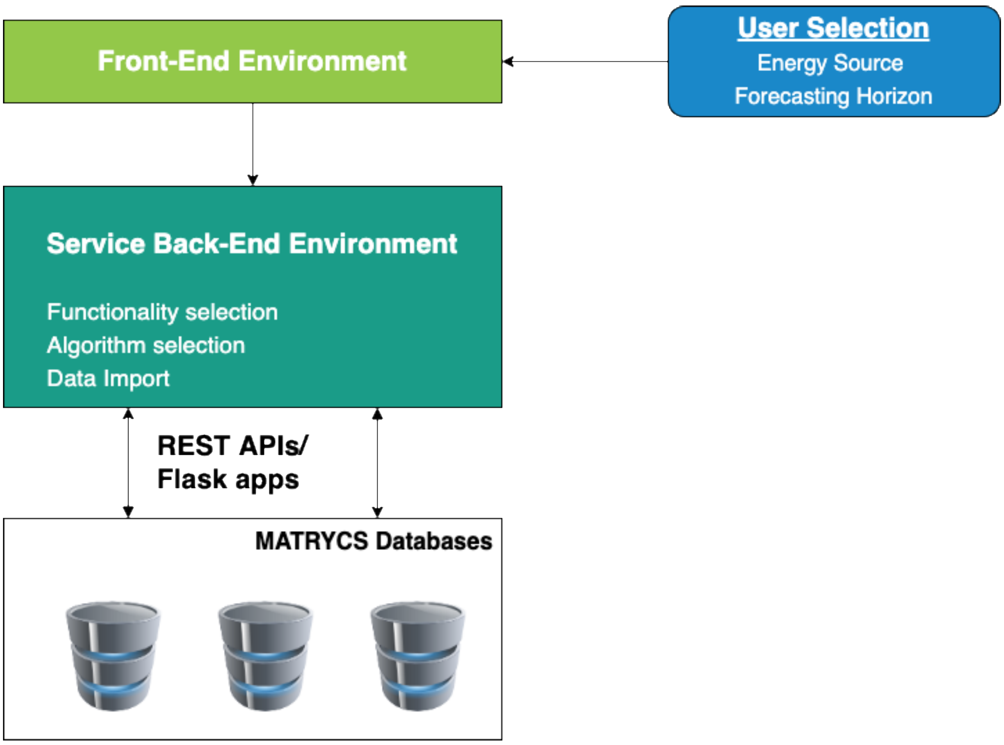

Figure 15 - Architecture of the Energy Prediction service (s.1.1)

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

The service requires to select the energy source that will be predicted, as well as the forecasting horizon. These selections will be available in the front-end environment of the service. The back-end environment is responsible to determine the functionalities needed to perform the prediction. These functionalities are the selection of the optimal algorithm, based on the selected energy source and forecasting horizon, and the import of the relevant data. The service is designed to provide results with several different combinations of available datasets (mainly classifying the weather data availability), by taking into consideration that the data needed to create an energy prediction module can be varied. The data import functionality, is connecting with the MATRYCS databases of WP4 with REST APIs and simulates the connection with a weather API. At this level of progress, weather data are available through specific CSV files to demonstrate the proof of concept. In later stages of MATRYCS project, the connection of the weather API will be incorporated into the service in order to automatically retrieve the required weather data.

The back-end operation of the service is presented in the Figure 16 below:

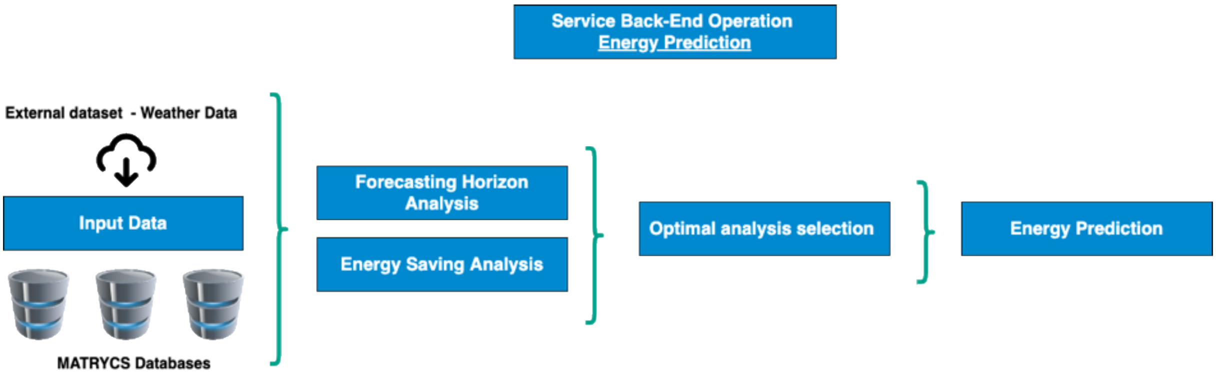

Figure 16 - Back-End operation of Energy Prediction Service (s1.1)

2.2.2 Input/output

The service will take as input time series data of energy sources and weather data (either predictions or historical data depending on the implementation). This would be in either consumption or production side and considers various sources, such as PV production, heating production and energy consumption of an end-user (building, EV, etc.). It will also use as input weather information, either historical data or forecasted weather, depending on the implementation. Based on the application, the service will decide which algorithm should be used from the developed ML and DL models.

The output of the service will be the forecasts of the respective timeseries. Then there will be two types of outputs, the first would be visualising the forecast to the end-user or use the forecast as an input for another service. In the case of an end user, visualisation tools will be used to present the results in graphs and charts and in the second case, through REST APIs the results will be delivered to the relevant applications.

All data needed for the service to run, can be retrieved from MATRYCS' databases directly or will be given as an input from the user (in later stages of MATRYCS). In the case of user input, the relevant pre-processing must take place in order to configure the inputs based on the models' architecture.

The Figure 17 presents the inputs and outputs of service s1.1, and the main back-end operations.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

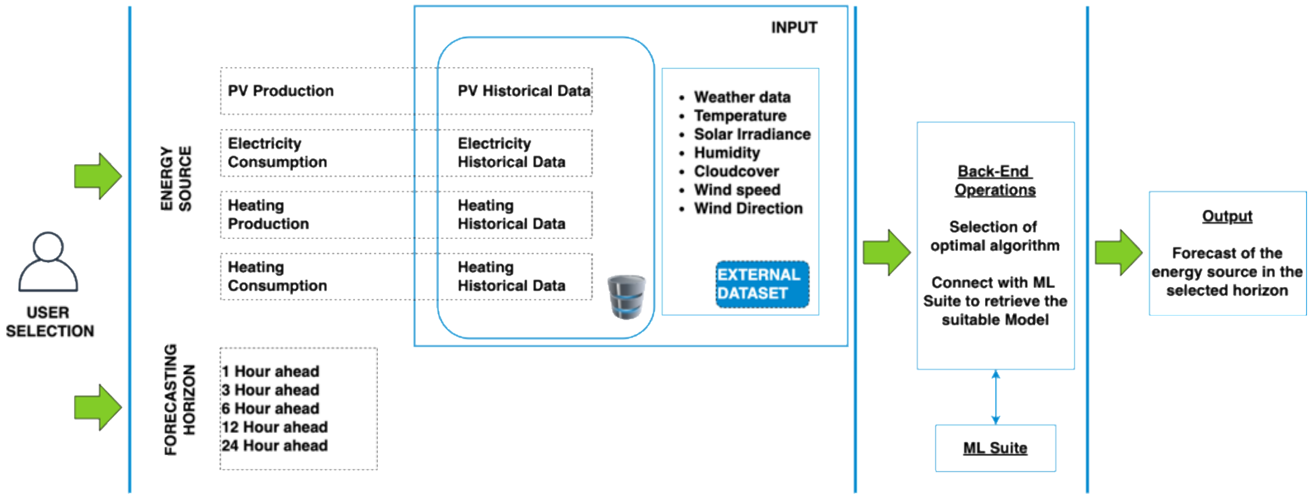

Figure 17 - Inputs/outputs of the Energy Prediction Service (s1.1)

2.2.3 User experience

In the case of the service used as standalone service, the following steps describe the user experience:

- Enter the platform
- Choose the energy source that will be predicted
- Select to upload data (in later stages of the project once user input is incorporated)
- Select the forecasting horizon of the prediction
- Output (visualizations)

In the case in which the service is used a sub-service for advanced applications, the user is not directly involved. The relevant back-end operations of the applications will call Energy Prediction service as needed and will use the outcome based on its functionalities. Hence, there is not going to be a graphical representation to the user.

The following figure presents a preliminary sitemap of Energy Prediction service as a standalone service.

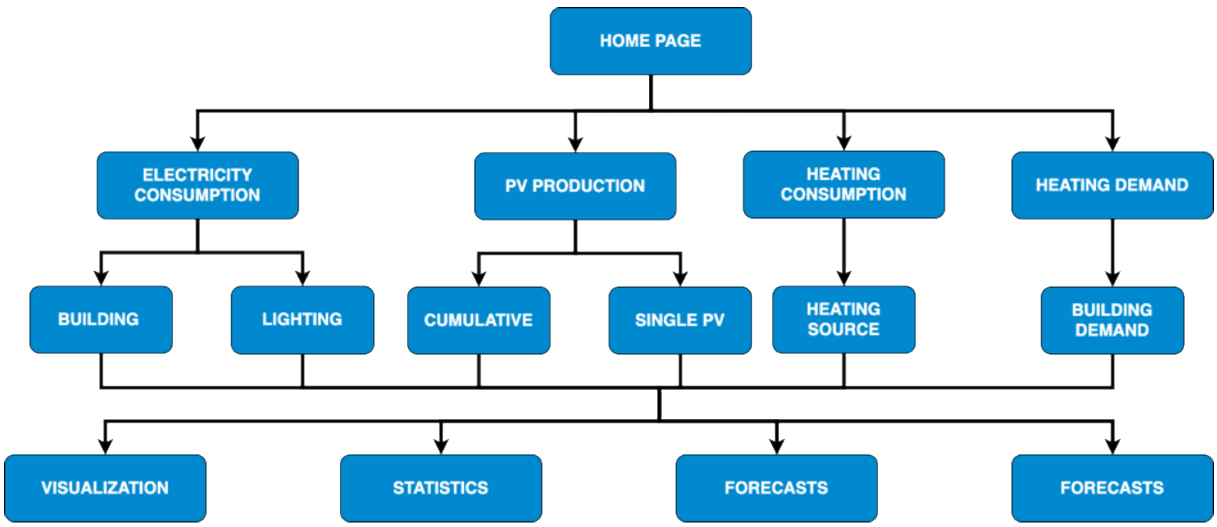

Figure 18 - Sitemap of Energy Prediction service (s1.1)

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

An initial mock-up of the service is presented in the figure below. The scenario presented is for the case of PV production, under LSP5. Several PV plants can be selected, and the prediction service can provide a visual outcome of the forecasted results. The different horizon forecasts are also presented and the values of each prediction will be displayed on screen on a user-friendly way.

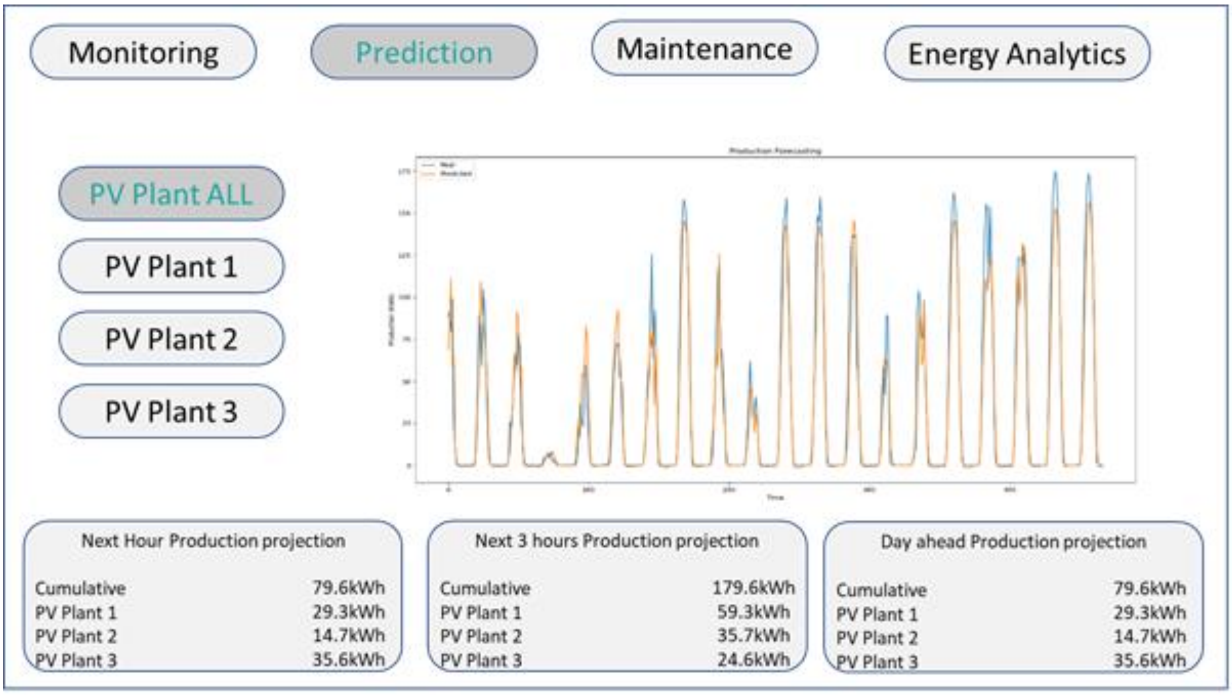

Figure 19 - Example of data visualization from the Energy Prediction Services (s1.1)

2.2.4 Application on MATRYCS Pilots

The use cases will be focused on **LSP4 - ASM**:

- Use Case 1 (UC04_01: Optimization of electrical distribution system operation)
- Use Case 2 (UC04_02: Action plans for performance improvements) and

LSP5 – COOPERNICO:

- Use Case 1 (LSP5-> UC05_01: Predictions about carbon footprint)
- Use Case 2(UC05_02: Identification of PV performance issues, UC05_03)
- Use Case 3(UC05_03: Optimization of self-consumption of PV system) and

LSP6 – VEOLIA:

- Use Case 1 (UC06_04: Operation of DH Network)

Energy Prediction service will be applied to three LSPs of MATRYCS directly as a service. The application to each one is described below:

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

- **LSP4 – ASM**: Energy prediction refers to RES production and energy consumption. Two different cases are considered for the pilot. The first case is the RES production, which involves PV plants. The available data in order to make energy predictions, are the historical PV production data, weather data (temperature, solar irradiance, humidity), the plant's technical and operational information. The second case refers to a building's electricity and heating consumption, in which the input considered is the electricity consumption of the building's operation systems.
- **LSP5 – COOPERNICO**: Energy Prediction service in LSP5 refers to PV production. The available data consists of several solar plant's information, such as PV production, electrical measurements, nominal power, nominal capacity, general information of solar plants (number of modules, PVs efficiency, size, location) and weather data. Multiple models have been developed and tested in order to predict the future PV production, with sets of inputs and forecasting horizons. Transfer learning techniques have been also investigated and tested to be incorporated in the service when data are not sufficient enough to create unique predictions.
- **LSP6 – VEOLIA:** Energy prediction service in LSP6 refers to heating energy prediction. The heating energy production from the energy generators (boilers) is being predicted and the heating load demands from the consumer's perspective is under development. The inputs consist of power generation, weather information from the energy generators side, and temperature measurements and weather information from the consumer's side. Currently the design of the service has made, and initial tests with the data sample is in progress.

Energy prediction service will also be used as sub-service to other services, which require the predicted energy, namely s1.2 – BAC Automation & Control and s1.5 – Operations for Network Optimisation.

The application to MATRYCS LSPs that are built upon service s1.1 will be enhanced with advanced functionalities in order to provide an integrated overview based on the Use Cases of each pilot. These functionalities will incorporate the data that are available, the requirements and the needs of each LSP, and will take under consideration the potential limitation and constraints. Overall, the tool that is envisioned to be offered to the LSPs which will use the s1.1 – Energy Prediction service is presented below. The basic functionalities will cover the Energy Prediction as described in the previous paragraph, the Visualisation module, the Energy Analytics Module and the Anomalies Module.

The following figure summarises the functionalities and applications on MATRYCS pilots, which service s1.1 will offer.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

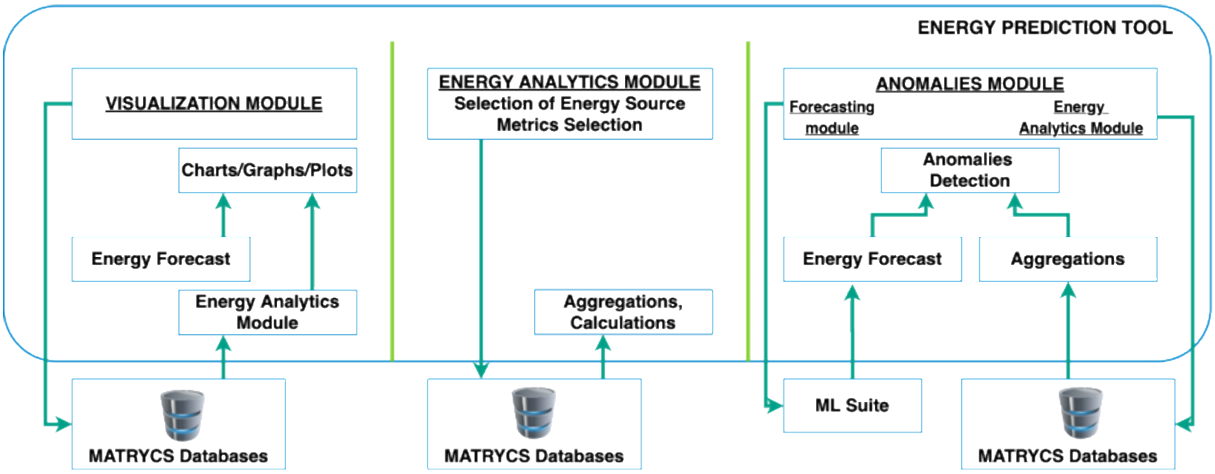

Figure 20 - Application functionalities of Energy Prediction service (s1.1)

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.3 MATRYCS – Building Automation and Control (BAC) (s1.2)

2.3.1 Tool description

Building Automation and Control service-s1.2 offers a centralized and automating framework in order to assess a building's heating, ventilation, air conditioning, lighting and sub-systems, as well as the cumulative energy consumption of the building. The monitoring of the building's performance will be available as it will provide information to the user regarding the building's performance on several energy efficiency indicators. Then, control actions of the building's different areas, possible action plans and measures which improve the energy efficiency of the building will be supported.

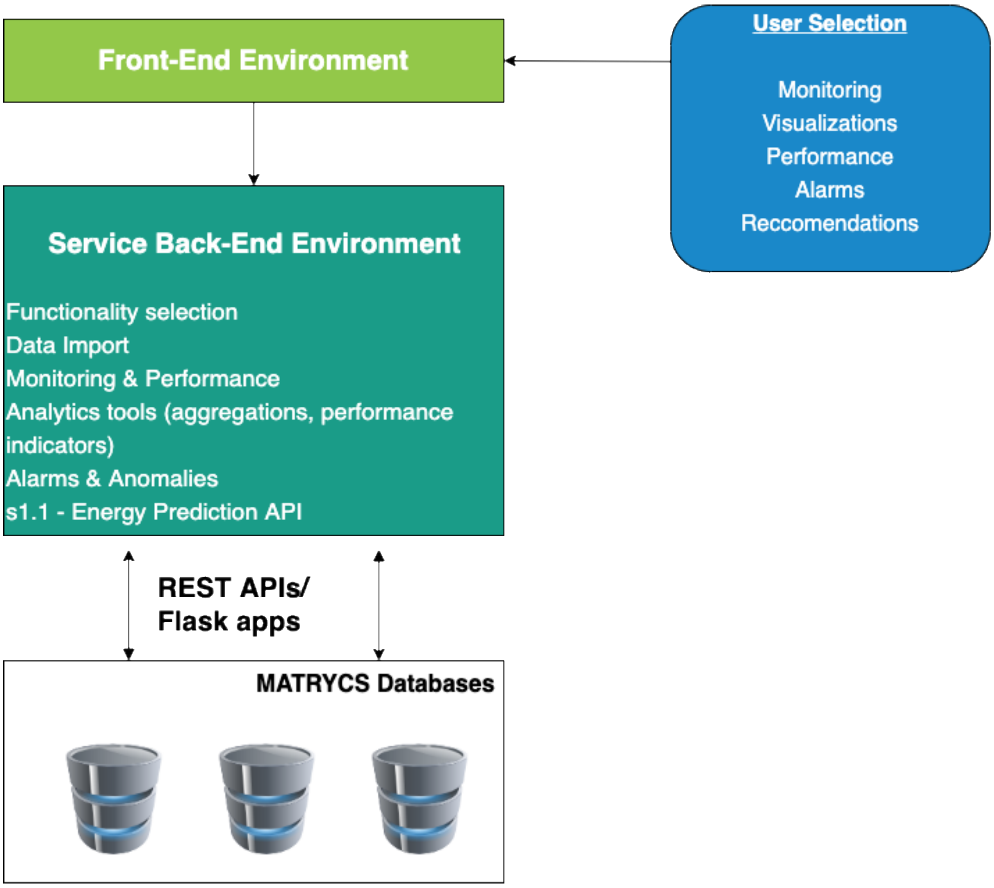

Figure 21 - Workflow of service s1.2

The service is being implemented in Python programming language, and several AI modules are incorporated in order to cluster the different type of areas of the buildings, their use and offer insights. The service will be enriched with advanced visualizations in order to depict the different areas of the building. Specific benchmarks and performance indicators of the building are calculated and presented and the possible control and action plans can be proposed. The service will also use as a sub-service, results of s1.1-Energy Prediction, to assess the impacts of the proposed actions.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.3.2 Input/output

The service uses as input real-time and historical data from buildings, such as heating and cooling system, indoor temperature, ventilation system, lighting consumption, electricity consumption will be used if available. Building characteristics are also considered (area, type of building, use of building, occupancy, etc.) in order to analyze the building's performance. Weather data are also used, in order to assess the impact of the control actions to the performance indicators. The service can use the results of service s1.1 – Energy Prediction as a sub-service, when assessing the future heating and electricity consumption.

The output of the service is the performance of the building's systems and sub-systems based on the performance indicators. The ability to visualize through graphs, plots and charts the analysis of the building's performance will be offered as output, in order to provide the overview of the energy efficiency of the building and offer insights regarding recommended control actions on different areas of the building. Alarms can be triggered on special events (such as, high consumption in a certain day) or in the case of a high value of inefficiency based on the performance indicators.

The data will be retrieved from MATRYCS databases and will be used to create the building's profile and monitor performance. The below figure presents the inputs and outputs of service s1.2, and the main backend operations.

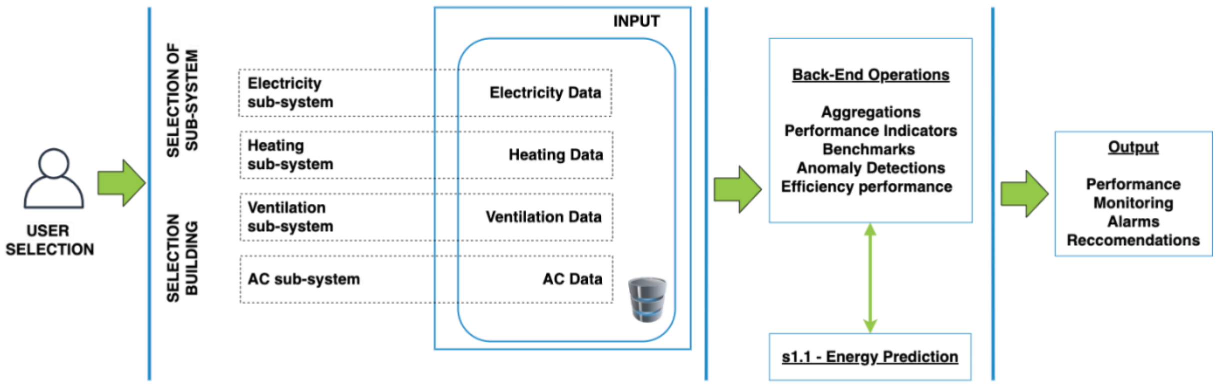

Figure 22 - Input and output of the service s1.2

2.3.3 User experience

The steps that describe the user experience are presented in the following flowchart:

- Enter the platform
- Choose type of building
- Choose building characteristics
- Choose areas of interest (electricity consumption, HVAC efficiency)
- Output

An initial mock-up of the service is presented in the figure below. The scenario presented is for the case of building performance monitoring, under LSP1. The user can select a particular area of the building (different zones/rooms) or the whole building in order to visualise the results. Then, the option to select different subsystems (electricity system, heating system, AC system) of the building is also offered. The

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

real-time performance of a subsystem is also offered, in ordered to provide real-time system level control monitoring. The energy performance indicators are calculated and presented per area/room and provide insights of the energy efficiency of the different subsystems or for the whole building. The possibility to offer prediction services in the subsystems is also supported, using the relevant energy source of the building and service s1.1-Energy Prediction as a sub-service.

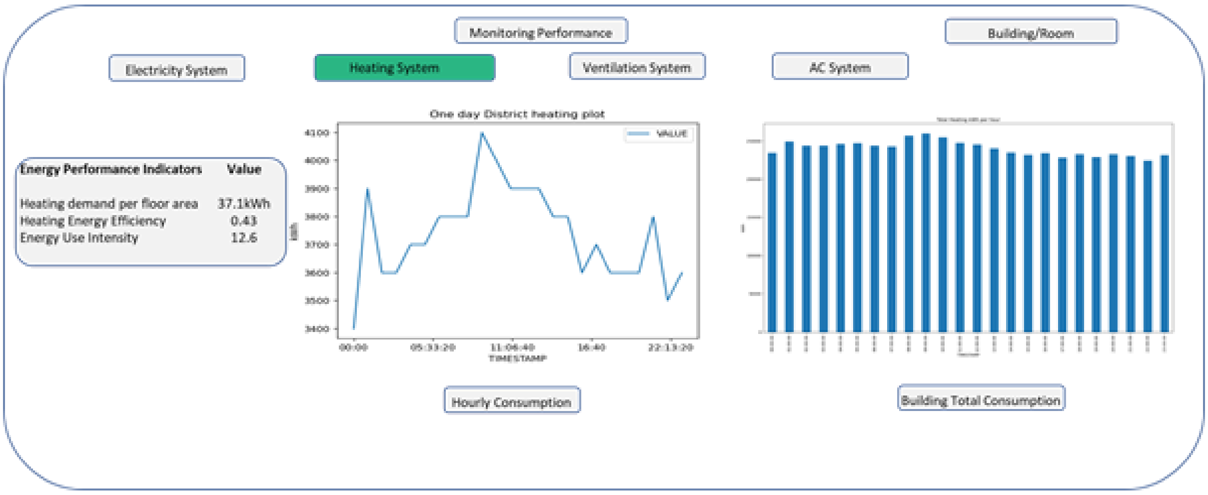

Figure 23 - Example of the information provided to the user by the front-end of service s1.2

2.3.4 Application on MATRYCS Pilots

The use cases will be focused on

LSP1 - BTC:

- Use Case 1 (UC01_02: Action plans for performance improvements) and

LSP4 – ASM:

- Use Case 1 (UC04_02: Action plans for performance improvements) and

LSP5 – COOPERNICO:

- Use Case 1 (UC05_03: Optimization of self-consumption of PV system)

The application to each one is described below:

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

- **LSP1 – BTC:** Building Automation and Controls service in LSP1 refers to the monitoring of selected buildings and several areas of the buildings. In these areas, the overall energy efficiency is calculated based on the type of buildings and action plans for performance improvements are proposed. There are three types of buildings considered for this LSP, a multi-purpose building where offices, commercial stores and leisure shops are all accommodated in one building, a warehouse which is used for storage and logistic purposes including cold storage facilities and a water park. As of now, the multi-purpose building has been studied as the data availability is sufficient to perform the service. Building characteristics, such as the size of the building, the total and per floor area, the topology of the HVAC systems and the building's configuration is provided. The historical electricity and heating consumption per floor and room is provided along with weather measurements. The scenario for LSP1, is monitoring the building condition in real time and providing the relevant energy performance indicators for selected areas of the building. The identified indicators so far, are the Total Energy Use, Energy Use Intensity, Electrical Load Factor, Lighting Power Density, Daylight Effectiveness and HVAC Energy Efficiency. Using service s1.1-Energy prediction as a sub-service, the future electricity and heating demand is also predicted, in order to assess the impact of action plans for performance improvements by using the benchmark values of the performance indicators.
- **LSP4 – ASM:** Building Automation and Controls service in LSP4, is similar to the case of LSP1. One building is available for the application of the service in LSP4, in which the electricity demand and the HVAC system characteristics are provided. More specifically two boiler systems, two heat pumps and three AHUs (location, rated power, operation mode) are provided. The Building Automation and Control service will benchmark their performance and use the energy efficiency indicators in order to provide the overview of the system and propose actions to improve the performance. The identified indicators so far that can be used for this scenario are, Total Energy use, Electrical Load Factor, Energy Use Intensity, Energy Performance Coefficient (considering the on-site renewable power generation). The monitor of building refers to the monitoring of selected buildings and several areas of the buildings. In these areas, the overall energy efficiency is calculated based on the type of buildings and action plans for performance improvements are proposed. Using service s1.1-Energy prediction as a sub-service, the future electricity and heating demand is also predicted, in order to assess the impact of action plans for performance improvements by using the benchmark values of the performance indicators.
- **LSP5 – COOPERNICO:** The application of the Building Automation Control on this pilot is under evaluation. The general approach is to use the same development done in the other LSP4 and LSP1 to manage electrical system with photovoltaic integration.

The following figure summarizes the functionalities and applications on MATRYCS pilots, which service s1.2 will offer.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

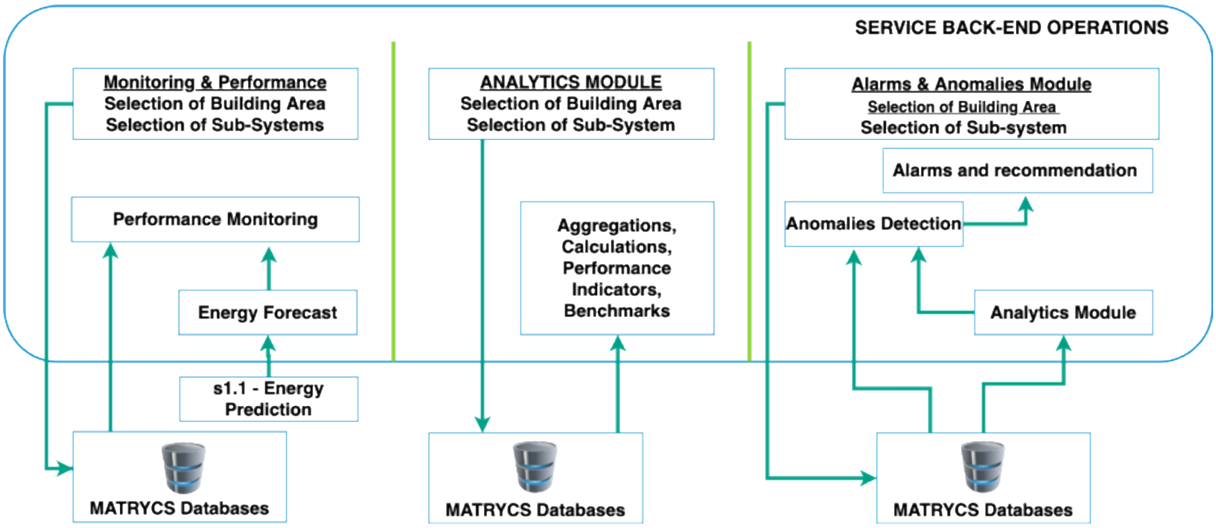

Figure 24 - Application functionalities of the Building Automation Control service (s1.5)

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.4 MATRYCS – KPIS Calculation (s1.3)

2.4.1 Tool description

KPIs calculation service (s1.3) aims to provide estimations of energy demand and consumption based on the generation of simplified models. Other indicators (social, environmental and economic…) will also be calculated.

For the energy estimation, the idea is to generate the model using the Digital Twin at the regional level and other sources (mainly public information on building typologies such as Building Stock Observatory, Tabula Episcope, Spanish Building Code and climate data). Concerning energy demand and consumption estimates, these KPIs are being modelled using machine/deep learning techniques, specifically artificial recurrent neural network (RNN) architecture such as long-term short-term memory (LSTM). Long Short-Term Memory networks are a type of recurrent neural network capable of learning order dependence and recognizing patterns in data sequences. What differentiates RNNs and LSTMs from other neural networks is that they take time and sequence into account, they have a temporal dimension. Therefore, LSTMs are well-suited for classifying, processing and predicting time series given time lags of unknown duration. In this case, data from consumption sensors, weather data and occupancy data will be used to generate an energy demand prediction model that can be used to make estimates and the results can also be exploited by other related services. Once this model is filled, the service will use algorithms based on EPBD in order to estimate some values for these buildings: energy demand and consumption and CO_2 emissions.

The information calculated for each building can be aggregated at block, district, municipality, region or province level. The information can be shown to the user in a georeferenced way or in the form of a report.

The potential of the tool is that can provide an estimate of some energy-related indicators for a large number of buildings using only public information sources.

The main challenges to achieve the objectives are:

- It is difficult to create a digital twin as basis without accurate information: OSM has no information about height and year of construction.
- TABULA/EPISCOPE has different information depending of the country, so it is difficult to create something working at European level
- Obtain enough data to develop robust and accurate prediction models for energy demand and consumption

2.4.2 Input/output

The inputs used by the tool are the following:

- Basic information for the generation of the DT at regional level:
    - Spanish cadastre
    - OpenStreetMaps
- Information about Building typologies:

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

  - TABULA/EPISCOPE
  - Building Stock Observatory
  - Spanish Building Code
- Historical data (energy consumptions, weather data,...)

It is important to note that the Spanish cadastre of the region of Castilla y León has 2249 files (corresponding to its 2249 municipalities) and information on around 1.490.000 buildings. In this sense, it is paramount to work with big data techniques. The DT at regional level that will be used as the basis for this service, will be built by filling the common data model defined for MATRYCS WP3.

The expected output for the tool is a Digital Twin at Regional Level with estimated information for energy demand and consumption and, if possible, CO_2 emissions.

2.4.3 User experience

This service aims to mainly help the Regional Energy Efficiency planner, in order to calculate the consumption (and other KPIs) of the building sector in the region, detailing the levels of energy performance by provinces, cities and neighbourhoods. This service will complement the estimates calculated by s3.2 service (Support Energy Performance Certificate harmonization and checking).

The information calculated for this tool will be shown in two complementary ways: (1) in georeferenced mode; and (2) in report mode.

For the georeferenced mode, the user can retrieve the desired information indicating the expected level of aggregation. The user could browse the maps and click for check more specific information. For example, at the building level the user could see all the information contained in the DT at the regional level for this building. In other scales (district, block...) the user could see information on: the area of buildings, number of buildings contained and the estimated information in an aggregated way (aggregated taking into account the area of the buildings). The following figure shows the appearance of the GUI for the georeferenced mode of the service.

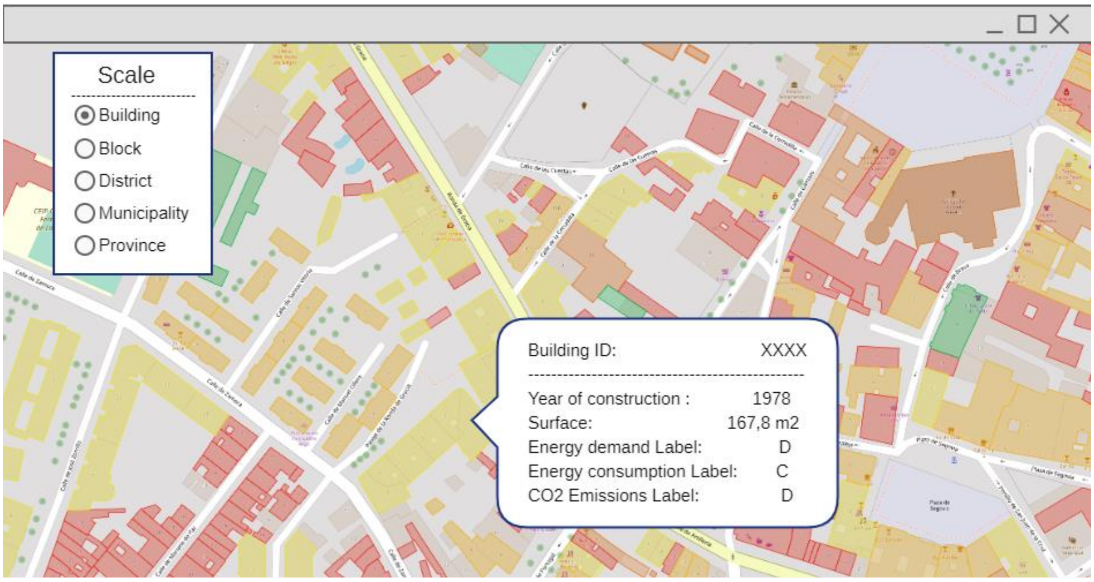

Figure 25 - GUI of Georeferenced mode of the services s3.1 together with s3.2

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

For the report mode, the user selects the level of aggregation, the area of interest (region, province, municipality…) the desired indicator and other information, as shown in the following figure:

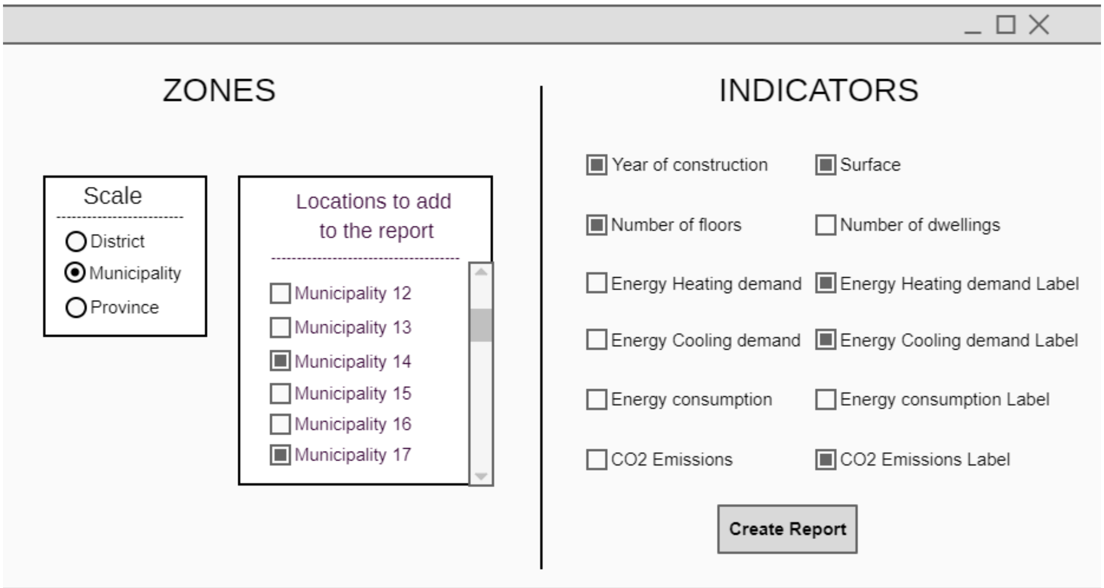

Figure 26 - Zones and Indicators definition

Once the parameters are selected, the service creates a report with the information that has been previously calculated in form of graphs and/or tables. This report could be displayed on a screen or stored as a file. An example can be seen in the following image:

Figure 27 - Example of report for the service s1.3

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.4.4 Application on MATRYCS Pilots.

The use cases will be focused on

LSP8 - GDYNA:

- Use Case 1 (UC08_01: Geo_municipality service) and

LSP9 –EREN:

- Use Case 1 (UC09_02: Visualization of estimated EPC) and

The application to each one is described below:

- In **LSP8 - GDYNIA**, the service will provide information for the municipality of Gdynia municipality, that is, at the city level. Depending on the available data, the corresponding LSTM networks will be trained to generate predictive models of the variables of interest considered.
- In **LSP9 – EREN**, s1.3 will be complemented by service s3.2. Both services will show estimations for the values of energy demand, consumption and CO2 emissions, using different approaches such as Machine learning/Deep learning techniques (LSTMs recurrent neural networks). This LSP will use information for the entire region Castilla y León region.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.5 MATRYCS – Technical Building Management (s1.4)

2.5.1 Tool description

The objective of the s1.4 Technical Building Management systems (TBM) services is to offer support for preventive and predictive maintenance of systems at the building level, through technical functionalities focused on tasks such as the detection of failures or anomalies. Among those features is the ability to identify temperature sensors that are not working properly.

The service will be able to compare the actual measurement with the modeled behavior. This could be extended to monitor the behavior of other relevant sensors included in the building management system.

Two main blocks of tasks must be performed to develop this service:

- The first approach to train the model from historical data from a specific sensor (data mining process)
- Second approach to continuous monitor the data generated by the previous sensor (TBM1).

Technologies / tools used to develop the service:

- Predictive model will be obtained at laboratory level. On demand.
- TBM1 → To detect events and to generate alarms. Developed in R and with proprietary license.

2.5.2 Input/output

The inputs to the service are the following:

- Historical monitoring information of the temperature sensor (time series of temperature measurements) to obtain predictive models. The more historical data, the better. To be developed in R and no licensed to final user.
- Continuous monitoring information (frequency to be defined) and performed predictive model for specific temperature sensor.

The outputs of the service are the following:

- Prediction model for a specific time series.
- Alarms generation (to be sent by email or similar) and incident record in a log file (database).

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

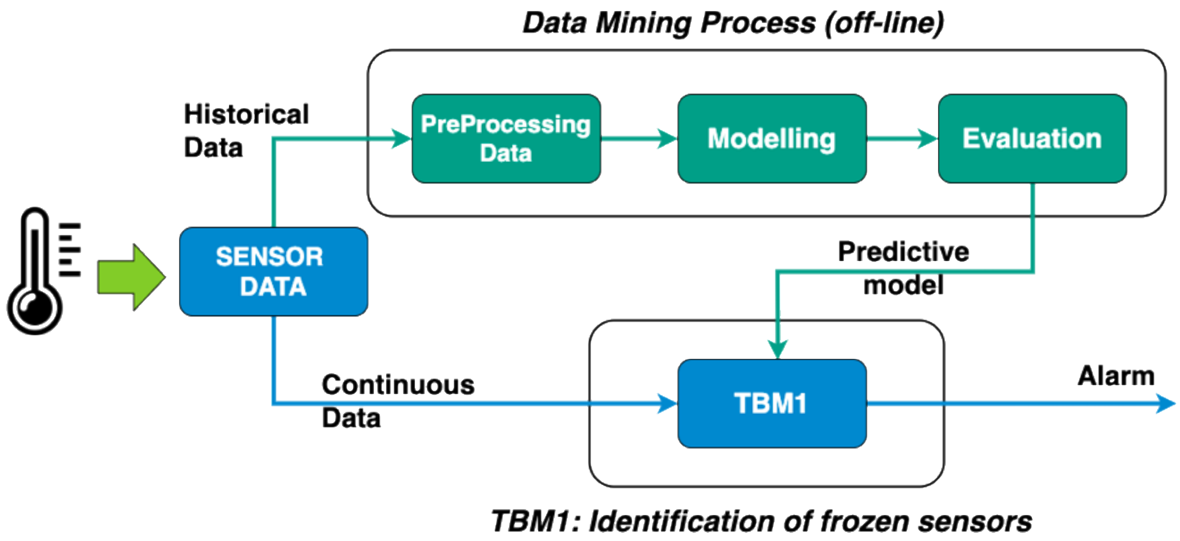

Figure 28 - TBM – Data process and model identification

All the data necessary for the execution of the service will be retrieved directly from the MATRYCS databases. At an initial stage, it can be provided in the form of flat files. For the calculation of models, common R processing libraries will be used, to allow a possible later integration.

2.5.3 User experience

The steps that describe the user experience process, are listed below:

- Enter the platform
- Selection of action plan to improve system performance
- Feature selection and scheduled plan
- The system provides details about undesirable situations detected
- Alarms will be generated (by email or similar), creation of a database of incidents for later consultation.

2.5.4 Application on MATRYCS Pilots.

The use cases will be focused on **LSP1 - BTC**:

- Use Case 1 (UC01_01: Action plans for preventive maintenance) and

LSP5 – COOPERNICO:

- Use Case 1 (UC05_02: Identification of PV performance issues)

the application to each one is described here under:

- **LSP1 – BTC:** The data available are from various sensors of temperature of Atlantis Water Park located in BTC City (Ljubljana, Slovenian). The service will be focused on the identification of temperature sensors that do not work, comparing the actual measurements with the modelled behavior. This could be extended to monitor other relevant sensors included in the building management system.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

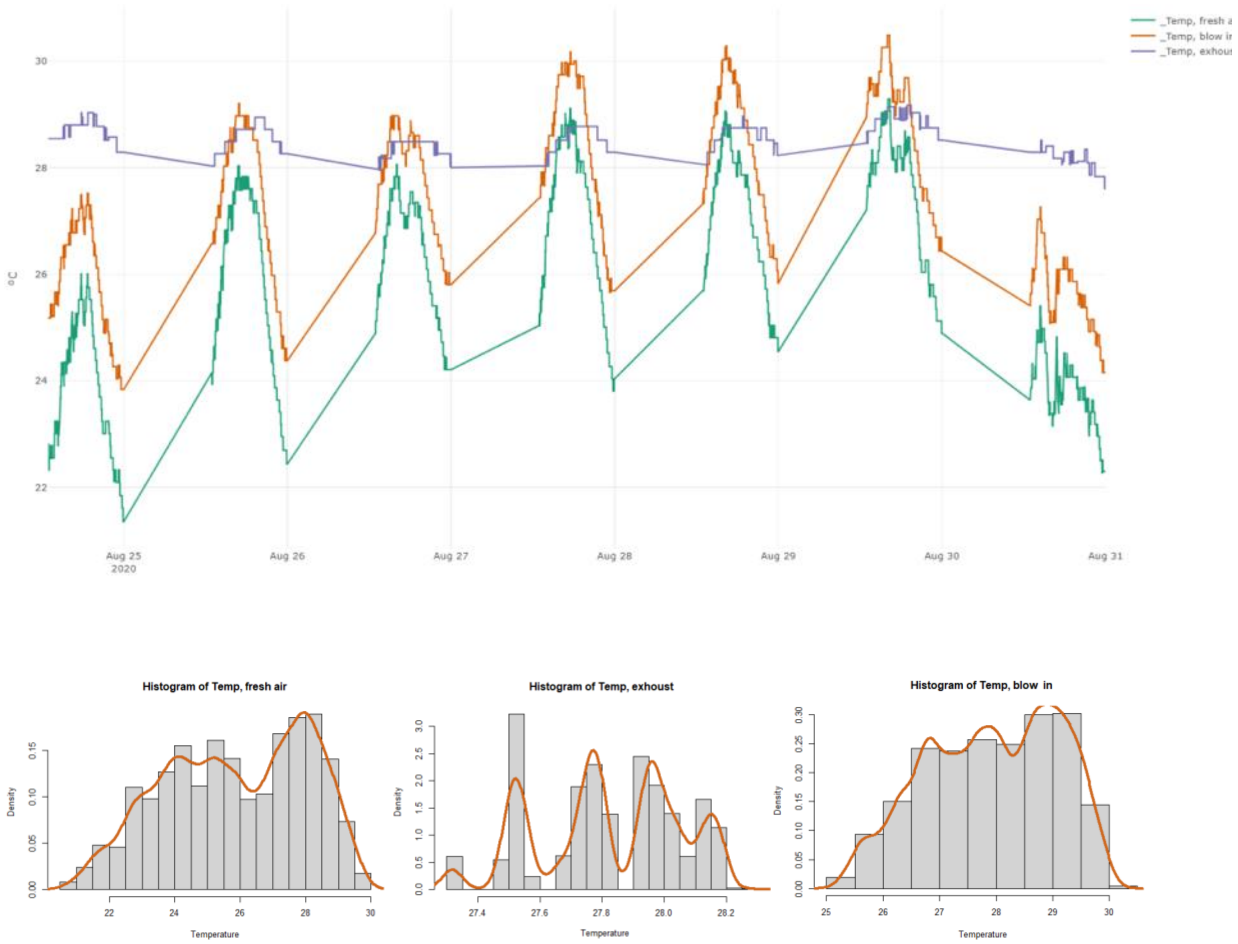

Figure 29 - Output for the service s1.4 on LSP1 data

This service may be applicable to other pilots and other types of sensors.

- **LSP5 – COOPERNICO:** The objective of the service in this pilot would be focused on identifying when a photovoltaic system is not working properly. Its application is still under analysis, conditioned by the amount and variety of data available in the pilot, and the possible overlaps with the s1.1 Energy Prediction service.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.6 MATRYCS – Optimization for network operations (s1.5)

2.6.1 Tool description

Service s1.5 – Optimisation for Network Operations offers support in different type of network and capabilities for operations optimisation. The optimisation process consists of several functionalities and processes based on the network's characteristics. In MATRYCS two types of networks are considered under two LSPs, a District Heating Network (DHN) and a local electricity network, each with its distinctive requirements. Service s1.5 will provide the analysis of the networks' critical nodes and their performance. As the optimisation process of a network, and specifically considering the structure of both the networks to be used in MATRYCS, can be performed in several areas (generation side, demand side, network nodes and flows), the services that are developed under s1.5, is intended to serve the use cases and the improvement of the functions of these networks. To this end, the analysis of the demand and supply side and the specific way which the network can be optimised is designed. In the figure below, the architecture of the service is presented.

The developed application will be based on optimization techniques and heuristic algorithms, with the scope to achieve energy matching between production sources and energy consumers. These algorithms could be generalised in order to include many cases (1 source of production vs multiple consumers, multiple production sources vs 1 consumer, multiple producers vs multiple consumers etc.) Prioritisation techniques will also be considered in order to schedule the energy supply. Cases where energy storage option exists (for example batteries) will also be considered. The tool will also utilise visualisation for better understanding of the produced results.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

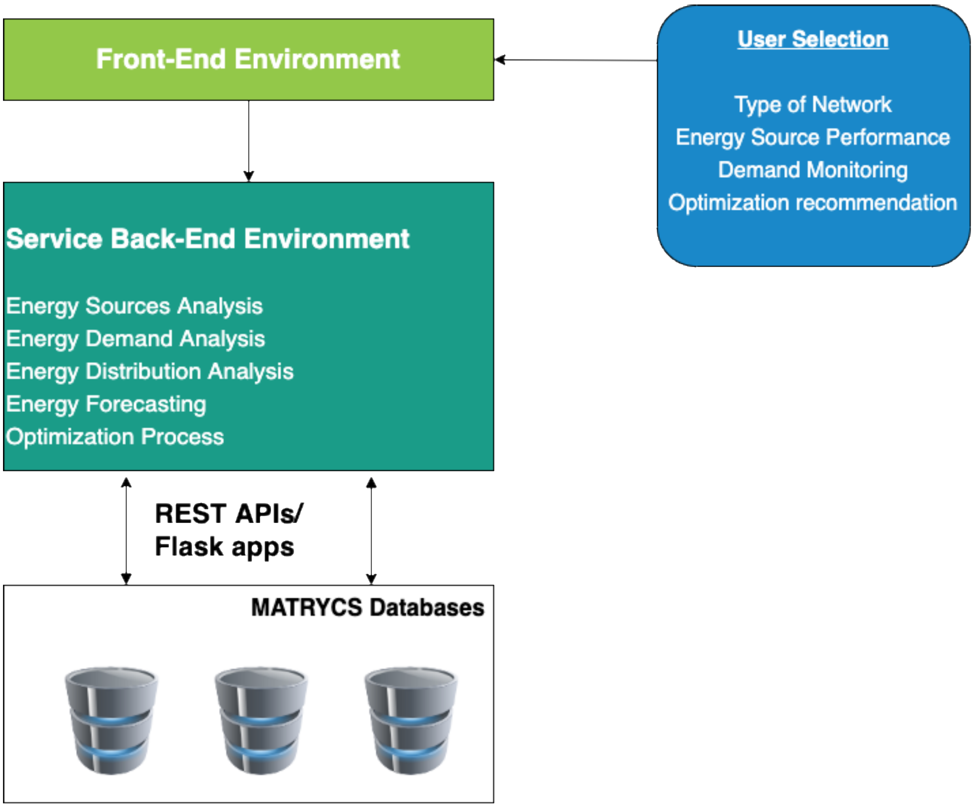

Figure 30 - Workflow of service s1.5

The selection of the network triggers the suitable analysis for network optimisation. This is done by taking into consideration the supply and the demand side in the network, which behave in different ways. Service s1.1 – Energy Prediction is also used in this service, as a subservice in order to analyse the future performance of the networks and evaluate the optimisation processes.

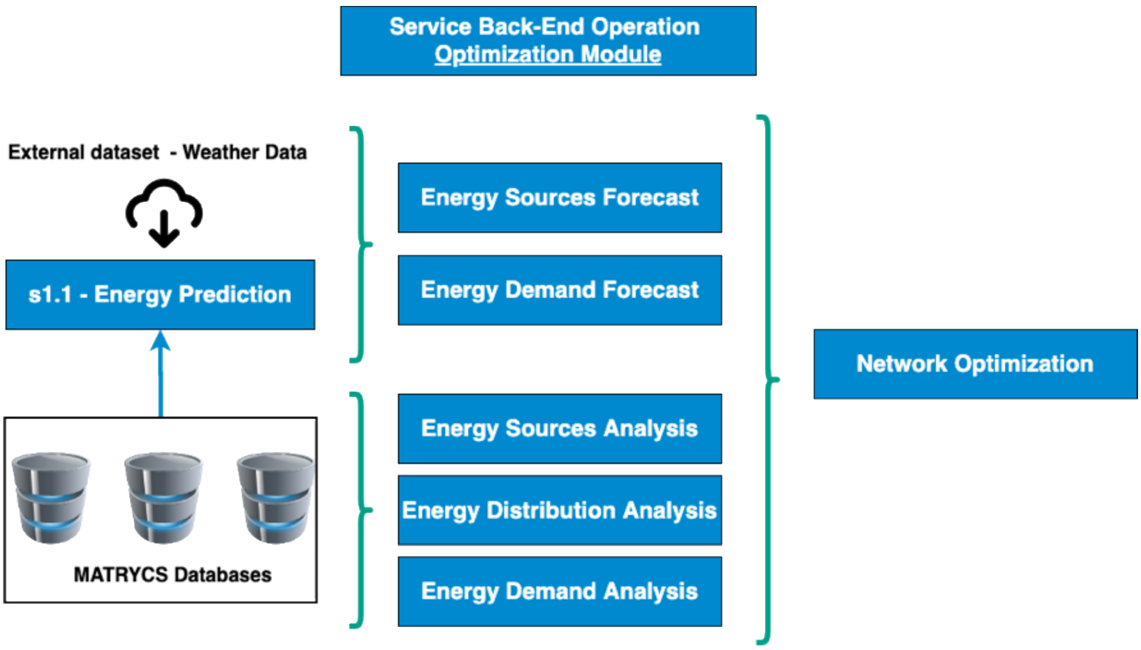

Figure 31 - Back-end operation of the service s1.5

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.6.2 Input/output

The service for each network will take as input, the energy sources both in supply and demand side. The inputs will be in time series data and are used to simulate the network operation in several conditions. The performance analysis will showcase how the network responds to different demands and how the supply side responds to the demand. Performance indicators are designed in order to depict points of interest in each network that can lead into optimisation processes. The performance indicators can also be used to alarm of possible anomalies, which may result in operational inefficiencies and overall network downtime. These supporting elements then are combined in order to provide the key points where the optimisation can take place, on simulated scenarios based on historical performance. Service s1.1 – Energy Prediction, will also be used after the simulation process of each network is completed. This will allow, to perform specific optimisation techniques in critical network areas as identified by the simulating process.

The main output of the service will be the suggestion on how to allocate supply side to the different consumers or how to prioritize energy production sources in a grid. This will be offered both on a historical level via insights based on the performance indicators and a future operation (day-ahead) level. The day-ahead level will leverage service's s1.1 – Energy Prediction results, in order to use the results of the simulated scenarios, for future network performance. It will also offer monitoring capabilities of the supply side and of the energy demand end users.

The figure below presents the inputs and outputs of service s1.5, and the main backend operations.

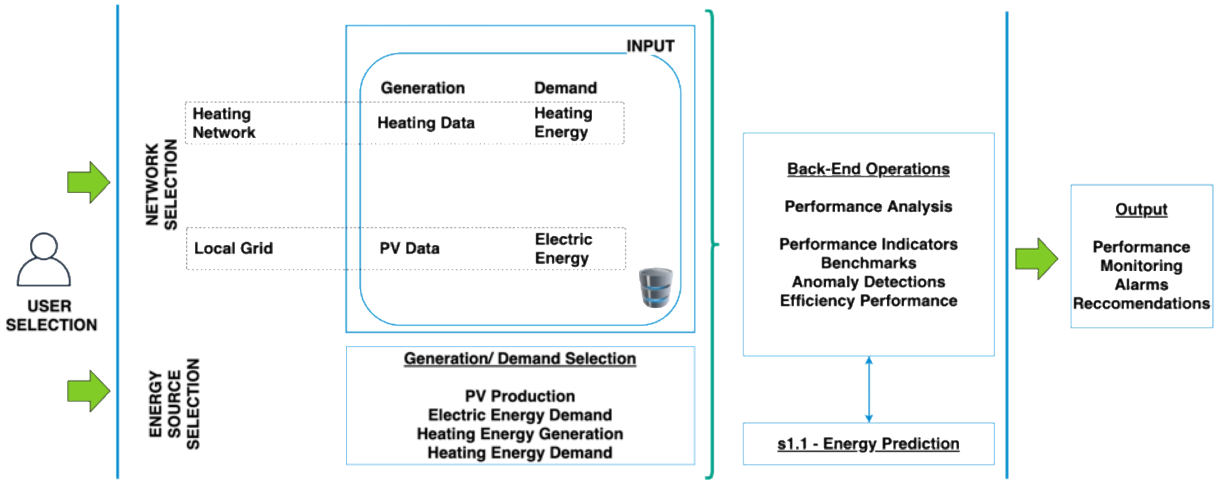

Figure 32 - Input and output of the service s1.5

2.6.3 User experience

The user experience of s1.5 refers mainly to a network operator. The ability to select the two examined networks (heating and electricity) will be offered in order for the service to provide the relevant functionalities. The user actions are presented in the flowchart below.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

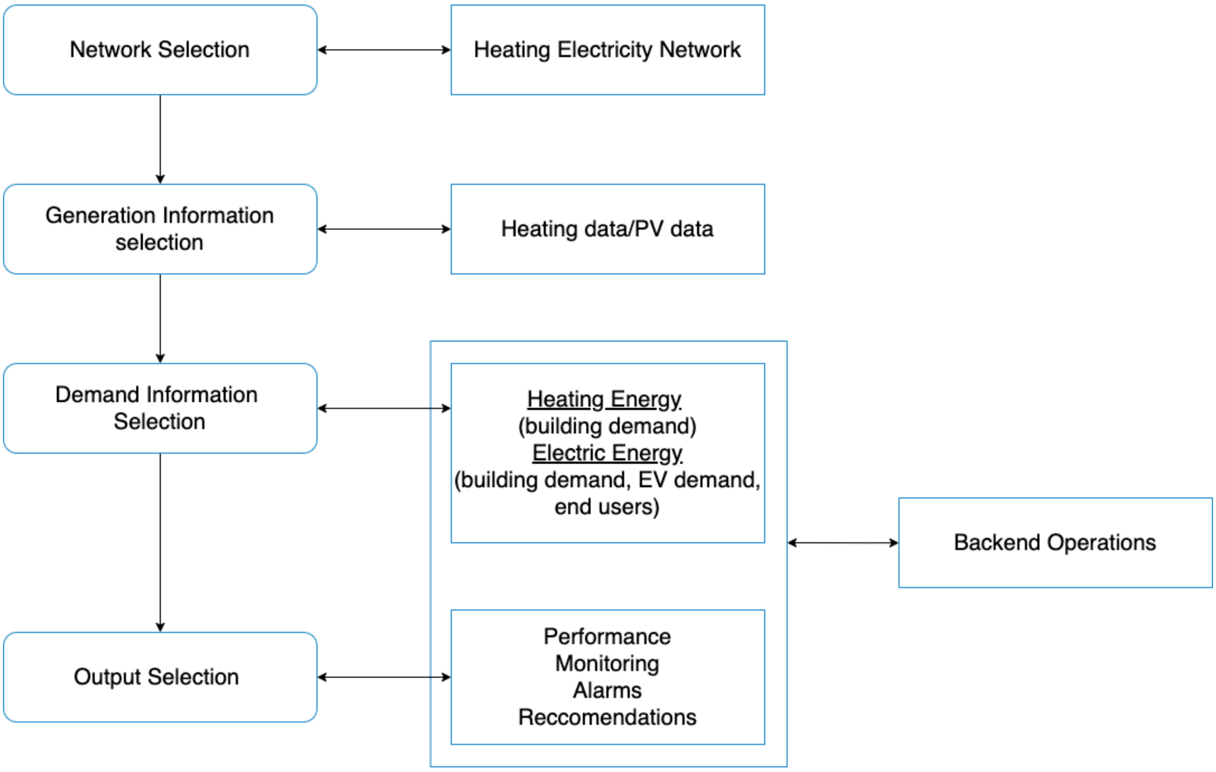

Figure 33 - Sitemap of Optimisation for network operations (s1.5)

The main steps that describe the process from the electricity network perspective as a demonstration, are the following.

- The user will enter the service interface.
- The user will give as input the grid topology by entering information of PV production sources (PV, wind energy etc).
- The user will give as input the information of the energy demand, by selecting the consuming points (grid, storage/batteries, buildings, others).
- The output for the user will be the visualisation of the functionalities that are designed, such as the historical performance of the generation and demand sources, and the scheduling plan for the network optimisation which has been generated by the respective optimisation algorithms according to the case.

An initial mock-up of the service, based on the electricity network is presented below under LSP4. The user can select to view the overview of the network based on its topology. Information regarding the real time performance can be viewed along with the next hour's projections (based on service s1.1 – Energy Prediction incorporation). The ability to show detailed information regarding previous performance of the generation sources and demand sources is also offered. The optimisation of the network performance is offered in a separate tab, where it will demonstrate the optimisation module of previous network states and provide a scheduling plan for future operational activities.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

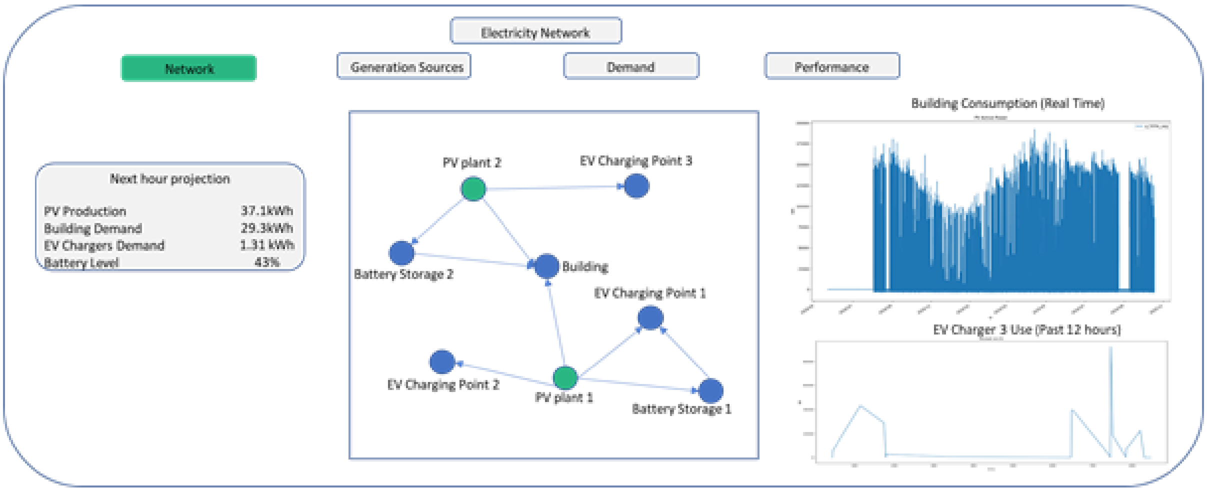

Figure 34 - Example of the information provided to the user by the front-end of service s1.5

2.6.4 Application on MATRYCS Pilots.

The use cases will be focused on **LSP4 - ASM**:

- Use Case 1 (UC04_01: Optimization of electrical distribution system operation) and

LSP6 – VEOLIA:

- Use Case 1 (UC06_04: Operation of DH Network)

the application to each one is described here under:

- **LSP4 – ASM**, the use case consists of energy production (supply side) by PVs, which has to be transmitted to several end-users (demand side). The end-users identified are a building electricity needs (lighting, HVAC), battery storage and EV chargers. For this use case, the use of the PV production in the network is examined and the optimisation process will be directed in using efficiently the produced energy towards the load demands, and optimally planning the operation of the network for minimising unused produced energy from RES.
- **LSP6 – VEOLIA**, the use case consists of heating generation from biomass and gas fuelled boilers (supply side) and buildings heating energy consumption (demand side). In this use case, the network optimisation is directed towards prioritising the use of biomass energy source instead of gas energy source. The DHN will be scheduled optimally by analysing the load demands, in order to provide an operational schedule which will allow optimising the operational hours of both energy sources, in favour of the biomass use.

The service will offer extended functionalities, in order to integrate all the relevant information of the generation and demand sources and provide an overall overview of the network as well. The following figure summarises the functionalities and applications on MATRYCS pilots, which service s1.5 will offer.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

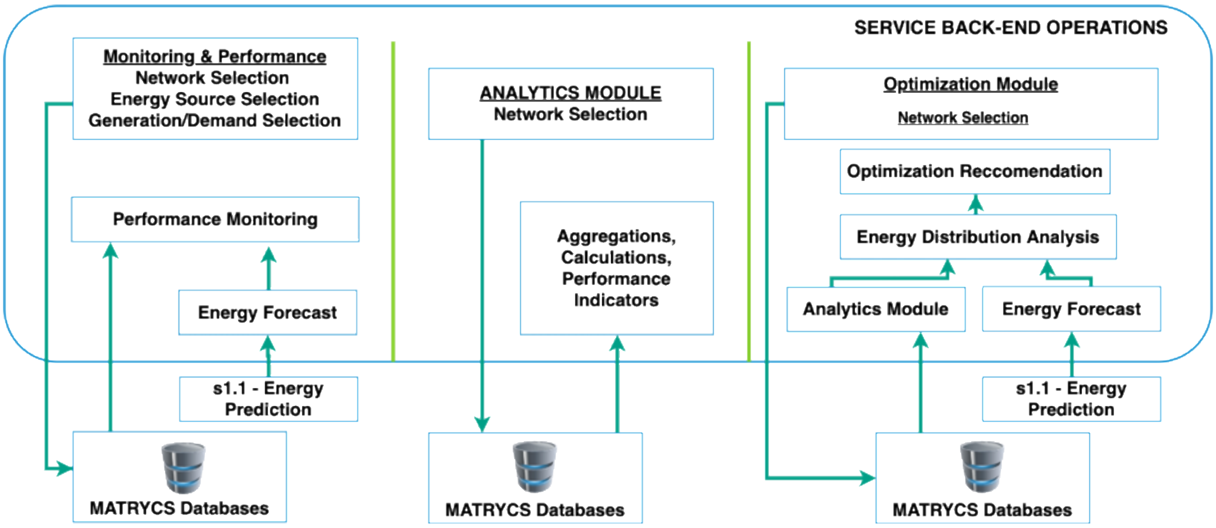

Figure 35 - Application of the s1.5 – Back-End Operations

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.7 MATRYCS – Technologies catalogues services to support the design and development of building infrastructure (s2.1)

The design or renewal of buildings is concerned with several aspects such as functionality, aesthetics, safety, operation costs and has to take into account criteria such as costs and regulations. Increasingly technology is playing a part in this process not only in the computer aided design and renders but also with energy computations and simulations as a function of building features including the technical proprieties of the materials and technology used for the end result. A catalogue of energy conservation measures (ECM) to renovate buildings that complies with the guidelines established by the recent European and global directives has been recently proposed[15]. They define an (ECM) as any kind of technological resource employed to improve the energy performance of a building. They also classify ECMs in four different groups of: (1) Passive, (2) Active, (3) Renewable Energy Source (RES), and (4) Control ECMs. According to the respective publication,

Passive ECMs aim to significantly reduce the energy needed to heat and cool a building independently of the energy and the equipment that will be chosen to heat or cool the building. This is realized by replacing the current windows, improving thermal insolation, reducing the air leakage, etc.

Active ECMs. These measures involve the replacement of the HVAC energy supply components by others with increased performance.

Renewable Energy Source ECMs include the use of sustainable energy sources such as wind, sun, water and geothermal.

Control ECMs refer to measures that achieve an optimised performance of energy supply technologies, making a smart use of the remaining fossil fuels and rationalizing their use, while still achieving comfortable living conditions. Some of these are thermostat control and advanced controls also multidimensional control related to weather, sequencing, air flow control, etc.

The OptEEmAL H2020 project has developed an ECM Catalogue[16] that includes materials and device information for some of the ECM groups above. The catalogue connects to the Building Energy Performance (BEP) simulation models in an automated way. One of the components still missing is a simulation technology that is able to help perform energy optimization using green energy from the PV panels and also enable simulations and computations with the size of the PV panels, that also depends on the location.

2.7.1 Tool description

The tool performs PV estimation that is able to inform the design and dimensioning of an integrated building energy provisioning system that includes green energy generators. Based on metadata, a mathematical model and site-specific parameters, the tool forecasts the produced energy for that specific building.

[15] https://www.sciencedirect.com/science/article/abs/pii/S2352710219313646

[16] https://github.com/arc-lasalle/OptEEmAL-ECMCatalogue

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.7.2 Input/output

The tool expects as input PV specific data such as Location (Lat/Long), Tilt of the panels, Orientation and Nominal power and outputs daily estimates with 15 min granularity as depicted in Figure 36. The default configuration with 15 minutes granularity can also be tuned if needed.

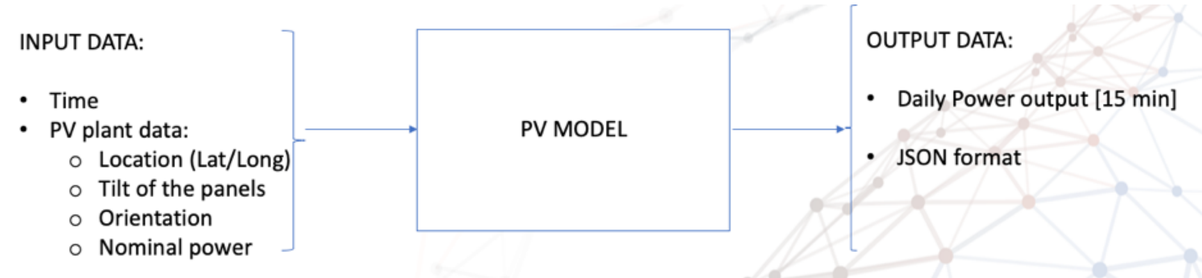

Figure 36 Input to and output from the estimation tool.

2.7.3 User experience

The user can interact with an existing catalogue such as the ECM Catalogue or directly with the estimation tool. In the first case, it is assumed that the estimation tool is integrated with the catalogue that provide APIs for ingesting the Weather data, visual user interface for inserting the PV data. The outcome can be integrated in the catalog overall UI possibly aggregated or even not directly shown to the user but only used to inform the final design suggestions of the building. Alternatively, estimates can be displayed in a graph as shown in Figure 37.

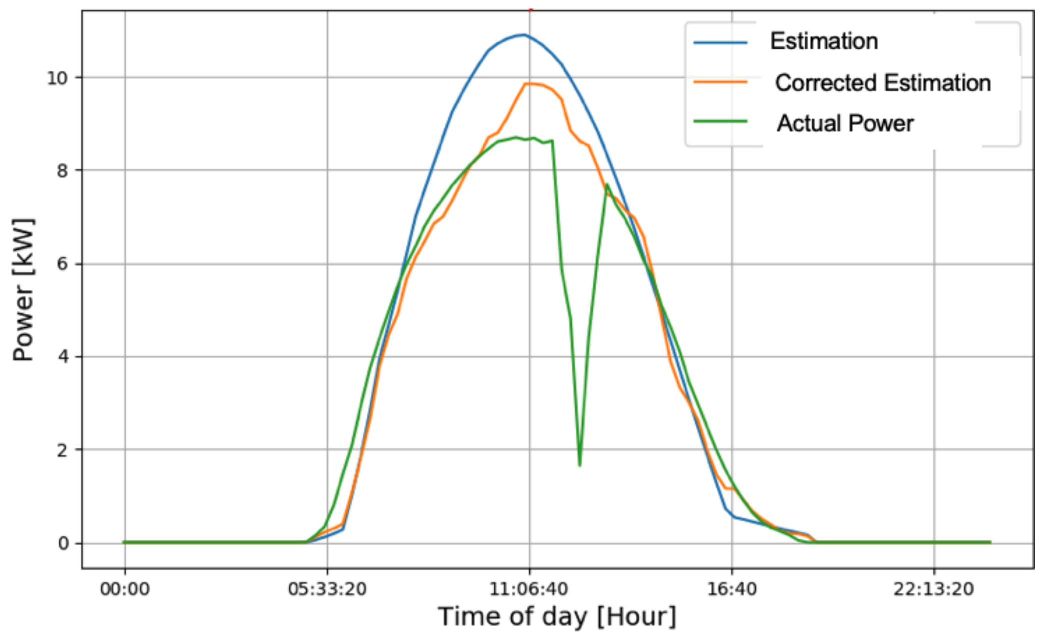

Figure 37 One day PV forecast.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.7.4 Application on MATRYCS Pilots.

The use cases will be focused on LSP5-> UC05_04 and **LSP1 – BTC.** We will show energy saving, CO2 and costs trade-offs based on various design options for buildings. Different measures will be evaluated from different perspective, therefore the objective function will cover not only the energy, CO2 savings but also the investment costs. The EMC of all 4 groups will be evaluated:

– Passive: improving Façade insulation, changing the glazing.

– Renewable: PV installation

- Active: changing of the existing HVAC systems

– Control: optimizing utilization of RES energy and cost optimization

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.8 MATRYCS – ECM-based scenarios evaluation service (s2.2)

2.8.1 Tool description

Increasing the energy efficiency of existing buildings is a priority from three different points of view: the environment, the security of energy supply and the quality of life of citizens. The renovation of existing buildings has a particularly high potential. It is therefore logical, from an economic and operational point of view, to implement appropriate energy saving measures when a building is renovated. The ECM tool developed in MATRYCS is based on a wider study and development carried out by EURAC together with the Province of Bolzano aimed at a simplified performance evaluation of the building with the purpose of defining the priority of energy upgrading interventions in a stock of buildings.

The planning process implemented in the calculation tool has several steps:

- **Detailed audit of the building stock**: collection of information relating to the actual state of the building-installation system of each building (geometric and structural aspects, final energy uses and inefficiencies of the envelope components and systems)
- **Simplified feasibility studies**: definition of typical interventions and assessment of the impact on buildings. Investment costs and costs for ordinary and extraordinary maintenance of buildings are considered in the business plan.
- **Calculation of key indicators**: a series of key indicators are calculated, covering various aspects, i.e. cost-effectiveness of the intervention and reduction of CO2 emissions, which also represent the criteria for assigning priorities for intervention.
- **Prioritization of interventions:** A fundamental pillar of the tool's methodology for evaluating the cost-effectiveness of a rehabilitation intervention is the search for synergies with extraordinary maintenance. The synergies of an energy upgrading intervention are quantified through the concept of "avoided cost" compared to an extraordinary maintenance intervention.

2.8.2 Input/output

The "Audit" section (Figure 38) collects information on the actual state of the building-plant system regarding individual buildings. The data entry in this section is aimed at the knowledge of the final energy uses, at the identification and analysis of possible inefficiencies and energy criticalities of the building and of the present systems. The compilation of the building card involves a series of operations consisting of the survey and analysis of data relating to the building-plant system in standard operating conditions (geometric-dimensional data, thermophysical data of the components of the building envelope, performance of the plant system, etc.). This information is essential to proceed in the analysis and economic evaluations of energy requalification of the building.

13 macro categories (Figure 39) were identified in each of which different information from the building is requested.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

The tool, based on the data given by the user, provides an assessment of what costs and savings can be obtained by implementing a specific measure of energy efficiency of the building. Specifically, in the case of intervention with insulation of the envelope, the data required are:

- Building typology
- Roof type
- Roof surface
- Unit cost of insulation
- Roof transmittance pre intervention
- Post intervention roof transmittance
- Climate data
- Unit thermal energy cost
- Thermal energy unitary emissions
- Seasonal global average efficiency
- Roofing incentive (default value = 0)
- Avoided cost coverage (default value = 0)
- Thermal energy unit cost
- Specific thermal energy emissions

In the case of **insulation of opaque walls**. The inputs are:

- Building type
- Opaque walls surface
- Insulation unit cost
- Pre-intervention wall transmittance
- Post intervention wall transmittance
- Weather data
- Thermal energy unitary cost
- Emissions unitary thermal energy
- Seasonal global average efficiency
- Walls incentive (default value = 0)
- Avoided cost walls (default value = 0)

other interventions taken into account are: **roof insulation, replacement of windows, change of heat generator, new lighting technologies**.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

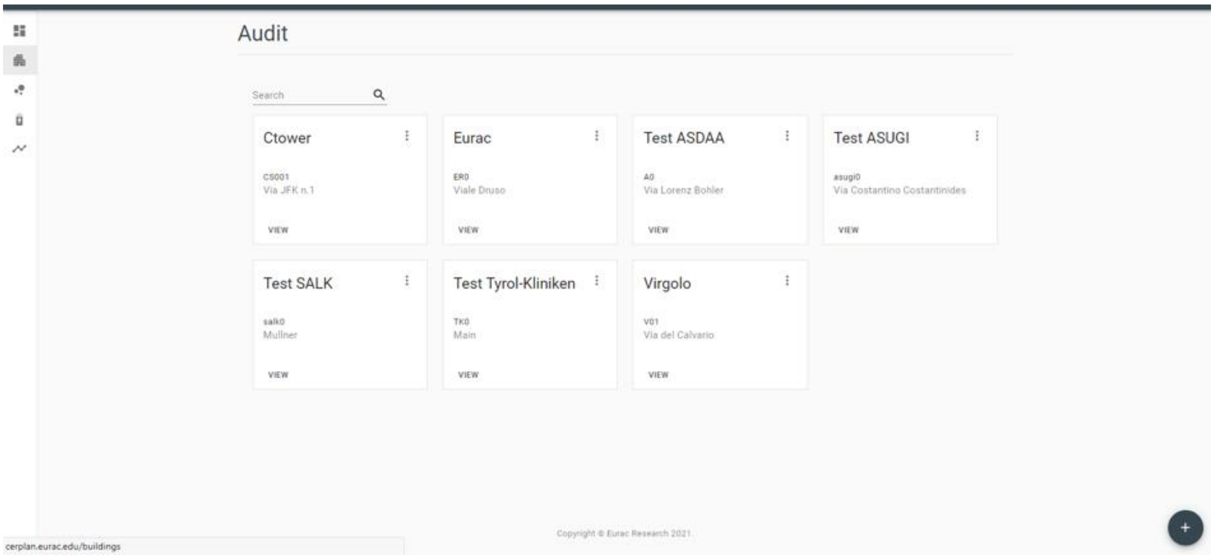

Figure 38 - Audit Section ECM tool

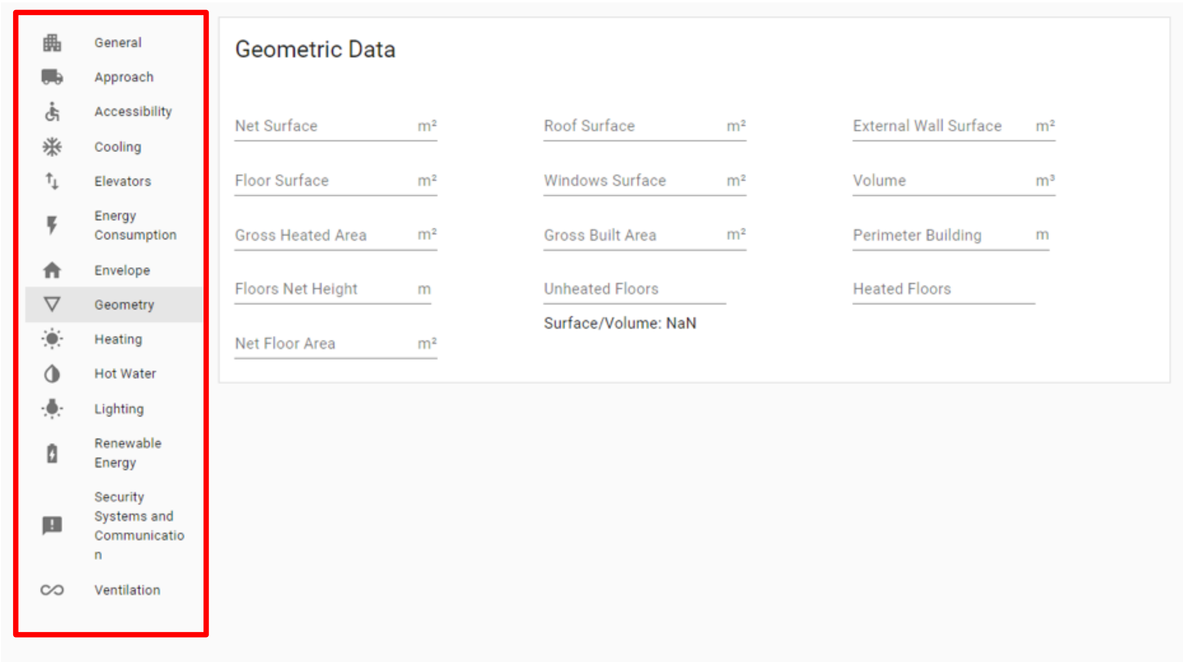

Figure 39 - Inputs for the ECM tool

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

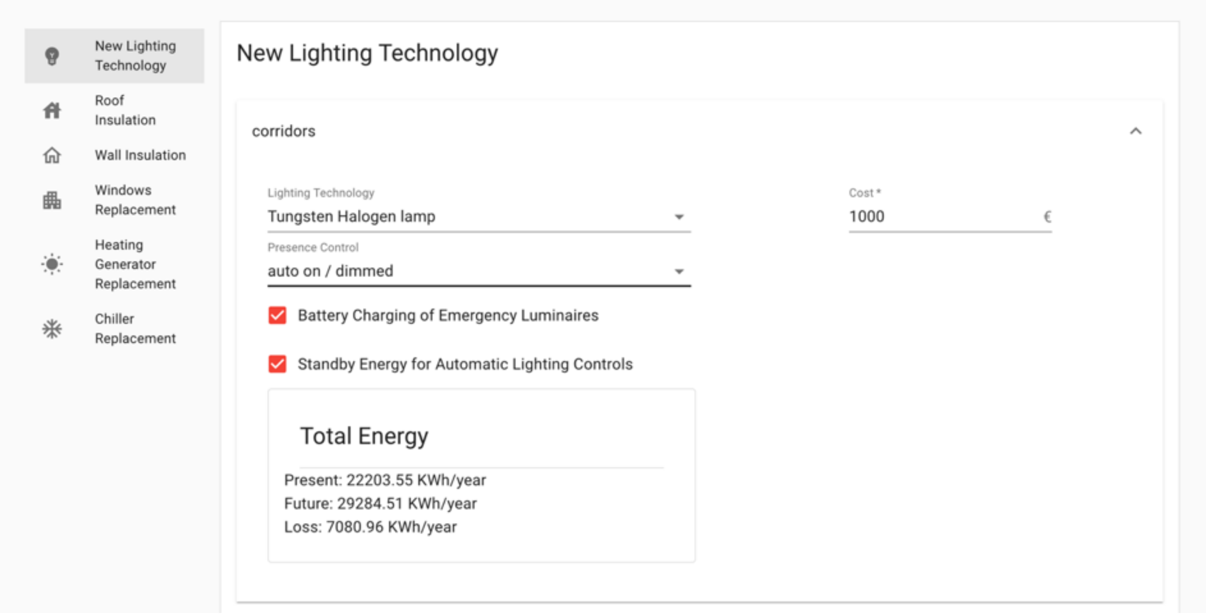

Figure 40 - Output for the ECM tool.

In this first version the tool runs locally within the Eurac servers. The tool will then be connected directly with MATRYCS governance, from which the data useful for analysis will be taken (Figure 41).

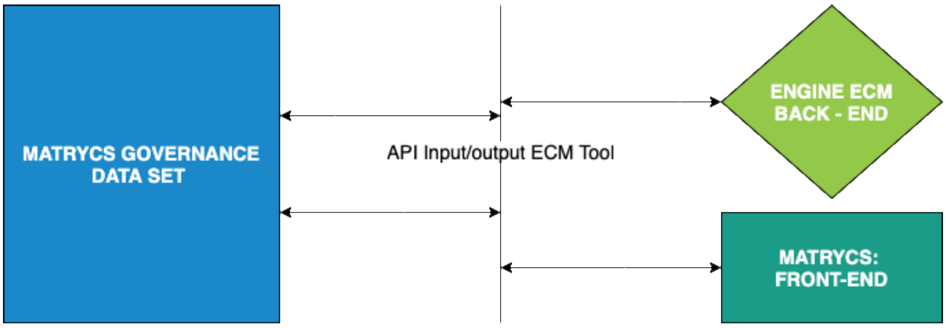

Figure 41 - ECM tool connection to MATRYCS GOVERNANCE

2.8.3 User experience

By entering the application, the user will be able to fill in all the forms related to the current geometric and performance characteristics of the building. Each user can connect as many buildings to his account as he wants. The tool will ask for various information as indicated in the previous paragraph.

The calculation of improvement interventions to be implemented is based on the number of information provided by the user. To obtain a ranking of the interventions to be implemented it will be necessary to provide the required inputs for each ECM. Following the calculation, the user will be able to modify the inputs at any time. The platform will be navigable through dedicated tabs whose access will be allowed through a sidebar on the left side of the screen (Figure 42).

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

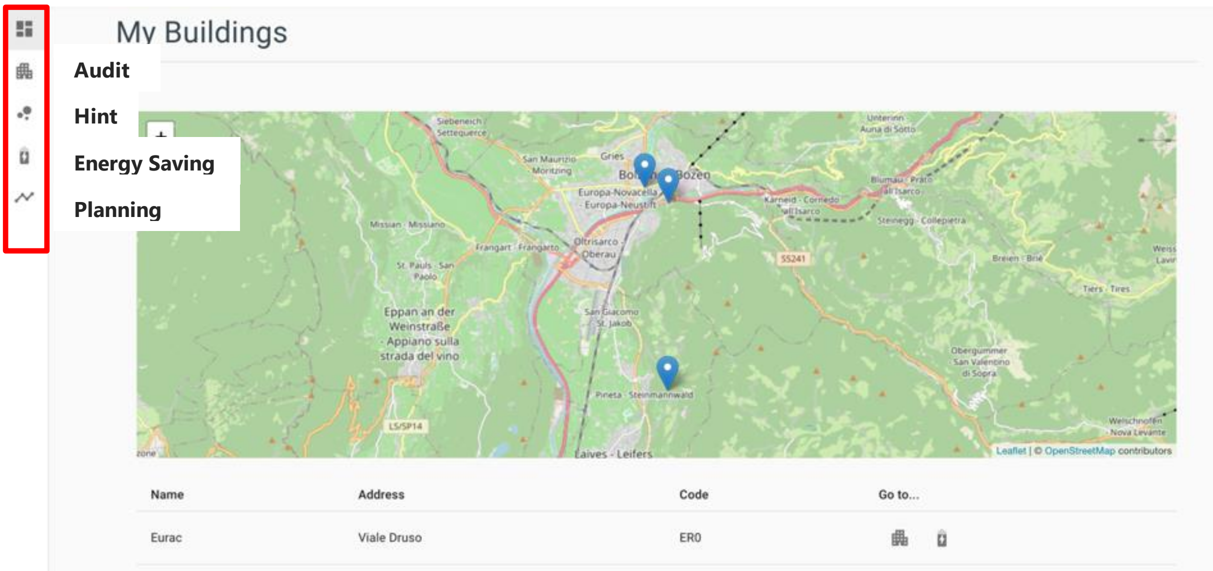

Figure 42 - ECM page selection

- The "**Audit" page** allows the user to enter all data related to the building, plants and related consumption. The data entered are static and refer to annual values.
- The "**Hint" page** suggests the actions to be taken following the calculation made for each type of ECM.
- The "**Energy Saving**" page shows the calculated values for each ECM giving the possibility to the user to choose eventual new technologies to be used providing a value of obtainable savings.
- The "**Planning**" page allows the user to enter data about the available budget to understand what interventions are feasible. Where possible, possible government incentives are also evaluated.

2.8.4 Application on MATRYCS Pilots

The use cases will be focused on **LSP2 - FASADA**:

- Use Case 1: UC02_01: Building evaluation of ECM / Building refurbishment

The case study consists of a kindergarten where facade renovations have been performed. A first energy audit for an ECM has been carried out, further information is needed and being acquired.

The list of use case according to deliverable D2.1 are: UC02_01: Building evaluation of ECM / Building refurbishment

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.9 MATRYCS – Support SECAPs decision-making (s3.1)

2.9.1 Tool description

Service s3.1 – Support Sustainable Energy and Climate Action Plan (SECAP) decision making, is a service which aims to enhance decision making at city level, regarding policymaking for climate adaptation and mitigation actions. The service will be offered through an integrated tool, which includes regional administration data, visualised information and offer insights and access to policy related data. The service will make use of aggregated data gathered through questionnaires that local governments fill and will be consisted of several functionalities, aiming to support activities when local governments are preparing the SECAPs, evaluating policies and identifying measures and actions to be taken for further mitigation reasons.

The service is designed to be used from a policymaker perspective of a local government. The functionalities that will be offered, are being designed under the scope of the level of information and analysis, which are required to make informed decisions when creating the SECAPs.

The main functionalities to be included consists of data visualisation, comparisons between cities based on similarities algorithms, matching techniques and recommendation proposals. KPIs and metrics based on several categories will be used in order to define the cities performance in climate adaptation and facilitation of mitigation techniques. The metrics are based on the cities answers and concerns categories such as gas emissions, climate hazards identified, RES penetration, waste management and climate mitigation actions among others. The tool will extract information from qualitative and quantitative data from the questionnaires to create aggregations and reasoning rules so as to find connections between implemented actions and offer insights.

Brief descriptions of the functionalities are presented below.

- Data Visualisation: Gathered data from the questionnaires will be displayed for easy information access.
- Benchmark information: By aggregating all the gathered information from the cities' answers, benchmarks are created for the different categories and provide the progress of each city.
- Metrics information: An analysis of the questionnaires is accumulated into several different KPIs. This process enables the overall performance of a local government in climate adaptation and can offer the areas that needs improvement.
- Recommendations: Based on the metrics information and the benchmark definitions, recommendation and measures proposals can become available by identifying similar actions taken from cities facing analogous problems.
- Insights offering: Based on the questions answered, the analysis of the different categories can provide insights as to where local governments should focus.

The internal modules of the service will treat the data and make the necessary actions in order for the service's applications to run. These modules are developed as Python scripts and are summarised below.

- Data formatting: The various data, both qualitative and quantitative, are clustered accordingly, to represent the cities answers and enable data analytics.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

- Data analysis: Based on reasoning rules and clustering of the several categories, the data are analysed in order to provide insights of cities performance.
- Querying: As the service will provide a wide range of visualisation of the data to support decision making, several scripts are created to access the data for display.
- Metrics creation: An analysis of both the unique and aggregated cities data, is performed in order to create different metrics based on the questionnaires' categories. The metrics can provide an overview of how the cities are performing in relation to each other.
- Cities matching: Internal modules that find similarities between cities, which can provide recommendations of climate adaptation actions based on similar performance.

2.9.2 Input/output

Data inputs are gathered from a set of questionnaires, which include various types of information (qualitative and quantitiative data). 14 different categories of cities' information are reported, such as Governance & Data Management, Climate Hazards & Vulnerabilities, City-wide emissions, Adaptation, Emissions Reduction, etc, and each of the category includes numerous related questions. The questionnaires are collected on annual level and are in the form of CSV files.

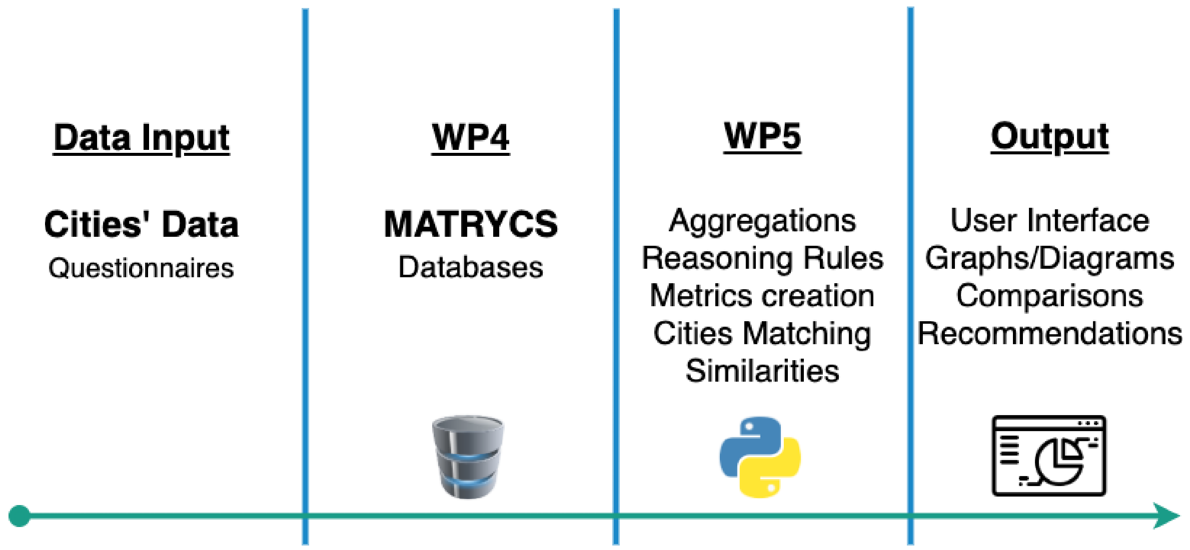

Figure 43 - Workflow of the service s3.1

The output of the service will be a web application, which will display the envisioned functionalities to the end-users. The service is designed in a modular way, allowing easy integration of additional functionalities, if needed. The main goal of the service is to provide easy-to-use functionalities for policymakers, by gathering information from various data sources and provide insights for climate adaptation actions.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.9.3 User experience

The user experience for service s3.1, is mainly from the perspective of a policymaker of a local government, i.e., city view. However, it could also provide functionalities in later stage for a more general audience and interested parties, where the aggregated information of the cities can be displayed. This will allow different user experiences and consequently the type of actions that a user can perform during the service use. The user experience from the perspective of the city view is described below.

- Entering the service. The user will enter the service after the authentication process. The authentication process will be required for the city view, as it will allow the functionalities that are designed for each city.
- Selection of the different categories. The categories that provide both general and detailed information regarding the city information will be available for selection.
- Functionalities selection. The functionalities of the service will be either provided directly as visualised information or can be selected as standalone choices. For example, for a selected category (Renewable Energy Sources) the performance of a city can be shown directly and the suggested measures in order to improve could be selected in order to gather the insights and provide support.

An initial mock-up of how the service can be displayed to the user as part of the city view is presented below.

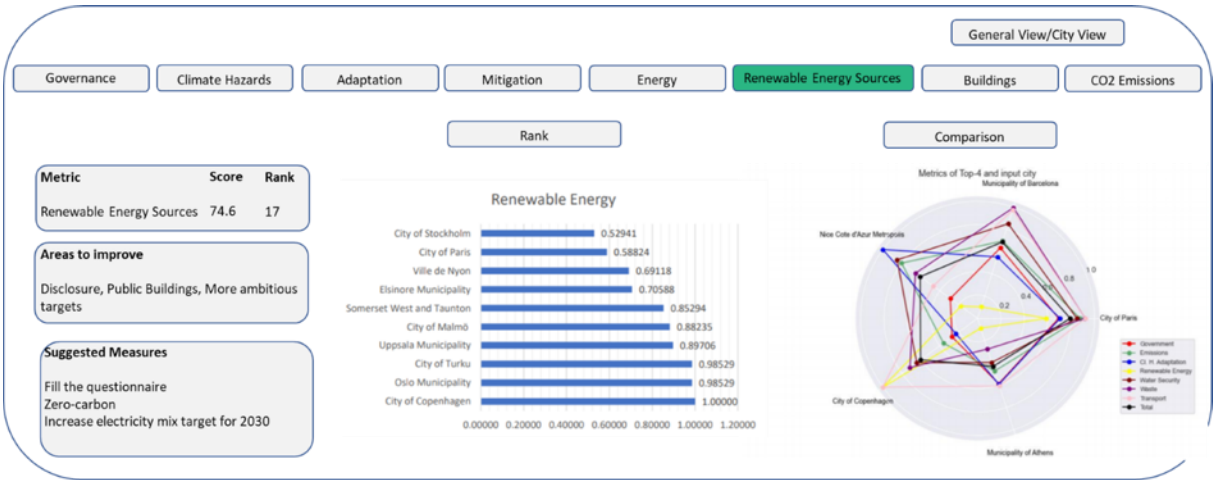

Figure 44 - Example of information provided by the Front-End of the service s3.1

2.9.4 Application on MATRYCS Pilots.

The use cases will be focused on **LSP7 - ICLEI**:

- Use Case 1 (UC07_01: Policy-makers decision-making support) and
- Use Case 2 (UC09_02, UC07_02: Building's investment prioritization)

The service 3.3 is designed in a modular way that derives directly from the use cases and needs of the pilot. The service is designated to be used as a supporting tool to ICLEI's activities, thus, to be used for

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

local governments' activities towards tackling climate change. At this initial stage, it is created primarily as a visualisation tool with extended functionalities to support local policy makers. However, due to its modular design, extended applications that could integrate external databases, applications for reports and information extraction, will be evaluated for integration in later stages of the project.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.10 MATRYCS – Services to support Energy Performance Certificate harmonization and checking (s3.2)

2.10.1 Tool description

The EPC harmonization and compliance service (s3.2) has the general objective of working on the next generation of energy performance assessment and certification. For this, the service will work with existing EPCs (Energy Performance Certificates) and will extract and process all the information from them in order to use this information as knowledge.

The target group for this service are designers, policy makers, citizen and owners, but the main beneficiaries will be policy makers.

For s3.2, the Digital Twin at the regional level will be used as the basis for generating the data model used by the service. The main source to generate this digital twin will be the cadastre.

The main actions to be carried out by the service are the following: (1) harmonization of the complete EPCs from a Region or City, (2) detection of errors or inconsistences in EPCs from a Region and (3) generation of estimates for some indicator in a region.

For the harmonization some data models will be analysed in order to select the best one for the project.

To detect errors or inconsistencies, one or more EPCs can be checked for anomalies. First, the data from the EPC will be compared with that coming from the digital twin. After, that the service performs different verifications, e.g., checking if some parameters are within the expected ranges, for example insulation thickness, windows air chamber, "U" values, thermal bridges, facility energy efficiency, air renewal flow rate... The service generates a report (on screen or in a file) with the possible errors detected.

To generate the estimate, the idea is to extract valuable information from the registered buildings EPCs, to extrapolate results and make better decisions. To do this, first some typologies will be extracted (supervised or unsupervised), and after that the typologies will be applied to other buildings in order to estimate energy demand and consumption and CO2 emissions.

2.10.2 Input/output

The inputs needed by the tool are the following:

  - **Spanish cadastre** (Open Street Map could be also used)
- **Regional Spanish EPC:**
  - **Public EPCs** offered by Castilla Y Leon (CyL) administration. Available in web.
  - **EPC XML repository** of CyL. This information is not public (it belongs to the regional administration). Around 99.000 files (17,8 GB)
- **Data tables from EPC regional register of Castilla y León region →** https://analisis.datosabiertos.jcyl.es/explore/?sort=modified
- **LiDAR** in order to check the potential of LiDAR to correct errors in the Spanish cadaster

The outputs of the tool are different depending on the specific use case of the service. One output is the harmonized EPCs. Another output consists in a report with the anomalies or inconsistencies

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

encountered in one specific EPC. Finally, the last result is the estimate generated by the service for the energy demand and consumption and CO_2 emissions (the values or the label, to be decided).

All the data required for the execution of the service could be retrieved directly from the MATRYCS databases or provided as user inputs (in one use case the user can upload a new EPC for verification) For the calculation of models, Python libraries already included in the WP4 ML models and processing module could be used, to allow possible integration in later stages.

2.10.3 User experience

Specific users could benefit from the tool are the following: Energy Performance Certification (EPC) issuers, EPC's Regional Registry Management Service, regional Energy Efficiency Strategy planners and buildings buyers and sellers.

Entering the service. The user will enter the service after the authentication process. The authentication process will be required for the city view, as it will allow the functionalities that are designed for each city.

The user will indicate the EPC that she/he wants to check. There are two possibilities: (1) check a new EPC; and (2) check EPCs previously stored in the database.

For the first possibility the user uploads a new EPC in xml format. The interface for this step is shown in the following picture:

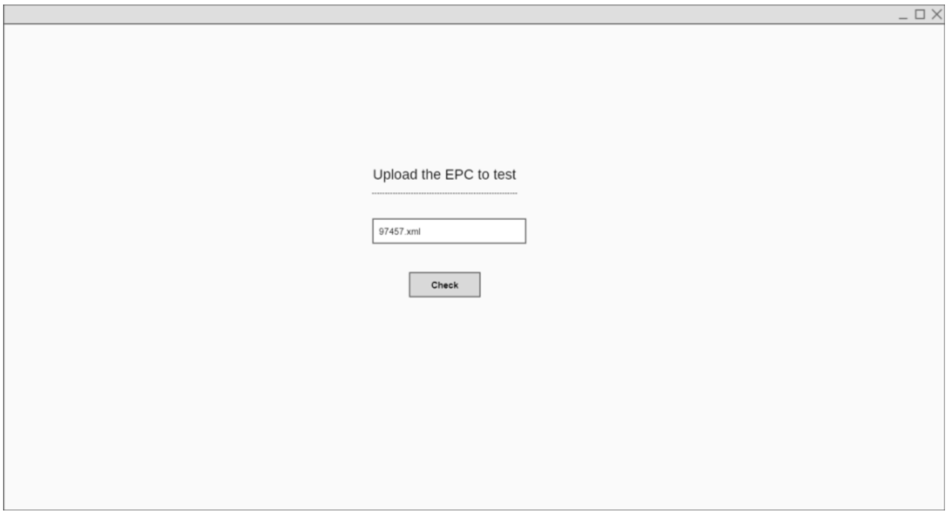

Figure 45 - EPC tool - file import

For the other possibility, the user can select the EPC or EPCs that she/he wants to check from a list. The list is based on the list of EPC XML files contained in the database. There is also a filter to find files faster. The interface can be seen here:

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

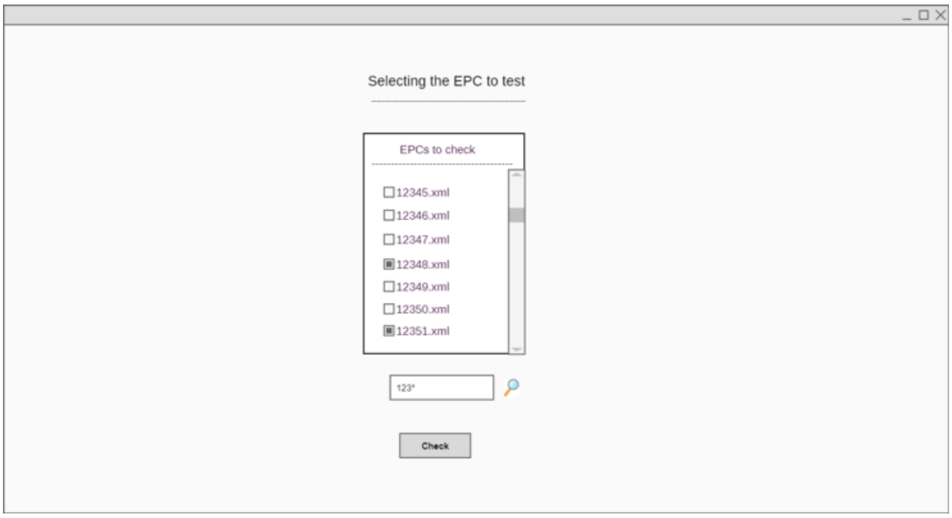

Figure 46 - EPC selection for testing

Once the button "check" is clicked the user can see the check results on the screen. Also, the user can download the report.

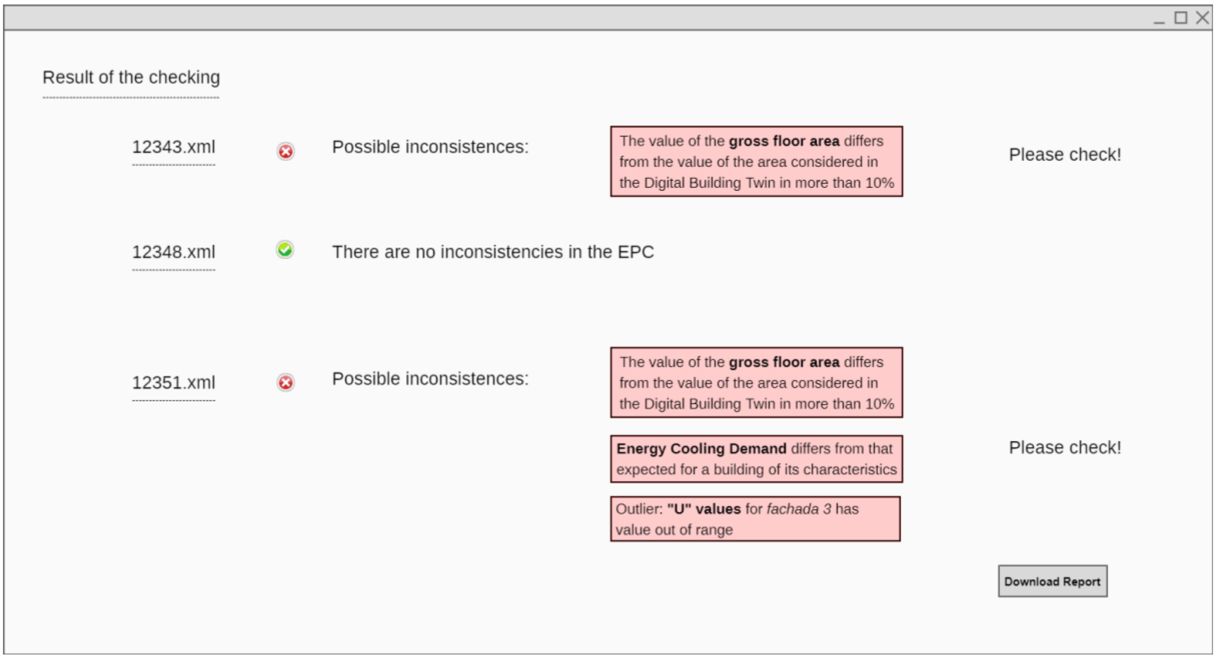

Figure 47 - Result of EPC tool testing

In the case of the information related to energy estimated by the tool, the approach is similar to the one explained in section 2.4.3, so please, refer to this section to verify this specific approach.

2.10.4 Application on MATRYCS Pilots.

The use cases will be focused on **LSP9 - EREN**:

- Use Case 1 (UC09_01, EPCs data error checker and validator) and
- Use Case 2 (UC09_02, Visualization of estimated EPC)

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

The main outcome of the EPCs data error checker and validator will be a relevant tool that will enable EPC issuers to offer trustful calculations on energy consumption and savings, and also EPC registers to verify registered EPCs, by detecting potential errors (thus increasing EPCs datasets quality).

The main outcome of the application focused on the visualization of the estimated EPCs will be a relevant tool that will allow energy efficiency planners to obtain valuable information from the registered building EPCs, as support to make better decisions in the planning stages and to provide information of estimated EPCs to different actors.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.11 MATRYCS – Services to support national and EU policy impacts assessments (s3.3).

2.11.1 Tool description

Service 3.3 – "National and EU policy impacts assessment and support" will offer insights for data-driven policy making and impact assessment for energy-efficient buildings. The service will prioritise investments on energy efficiency interventions based on historical performance and through portfolio optimisation. By taking into consideration already implemented energy efficiency projects, it enables the identification of the most efficient way for fund allocation, considering the respective possible risks as well.

It will combine information on the building level, by analysing energy efficiency both on performance level and in alignment with sustainable practices with relevant national and/or EU policies. The approach for this service is two-fold; an analysis of a selected set of buildings, in which renovation actions have taken place, and the integration of the associated policies that would allow measuring and evaluating the linked policy objectives.

The analysis of the buildings refers to the energy efficiency actions that have been implemented and in order to elicit the EE portfolio optimization. This will enable policymakers in national and European level to draw conclusions from previous actions, identify the suitable ones along with the optimal fund allocation towards energy efficiency measures in buildings. The analysis of the policies refers to the specification of their objectives towards energy-efficient buildings. Upon the specification, they can be linked with the analysis of the buildings and the proposals of the portfolio analysis, thus allowing to assess their impact.

2.11.2 Input/output

The service will use as inputs the relevant data based on the two-fold approach analysis.

For the energy efficiency analysis of the buildings, the inputs consist of the building characteristics and energy performance information. The building characteristics include information such as the location of the buildings, number of dwellings, year of construction, energy classification. Previously financed energy efficiency projects and information regarding the cost, energy savings achieved from building interventions and refurbishments and financing, can be utilised by incorporating data insights and experience of already implemented projects and their results. Expert opinion can also be included as a determining factor, when calculating the quantification of risk levels for energy efficiency building interventions. Such risks can include several indicators; political, regulatory, social or economic. The analysis will be performed in selected buildings of the housing sector, in order induct it in a larger scale (national or/and European).
Upon calculation of the optimal portfolio analysis, the assessment of the impact in the relevant national/EU policies will follow. The relevant policies to be analysed have been preliminary defined as followed, but not limited.

- European Green Deal
- European Industrial Strategy

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

- Blueprint initiative in the construction sector
- Next Generation EU
- Recovery and Resilience Plans
- Renovation Wave for Europe
- Affordable Housing Initiative (100 renovated Lighthouse Districts)
- European Pillar of Social Rights
- Cohesion Policy 2021-2027
- Just Transition Mechanism
- Citizens' initiative "Housing for all"
- New European Bauhaus
- Housing Partnership of the Urban Agenda for the EU
- EPBD

2.11.3 User experience

The service is at the design level at this stage. More information regarding the user experience will be provided, once the details will be clarified.

2.11.4 Application on MATRYCS Pilots

The use cases will be focused on **LSP11 – HOUSING EUROPE**:

- Use Case 1 (LSP11-> UC11_01: Portfolio optimization)

Service s3.3 will focus on the housing sector, by providing the portfolio analysis in selected buildings of interest. The results upon identification of the optimal portfolio, aims to allow the assessment of several EU policies towards the housing sector that the pilot leader is operating and providing insights for future energy efficiency interventions in the building sector.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.12 MATRYCS – Services to support energy savings M&V towards improved Energy Performance Contracts (s4.1)

2.12.1 Tool description

The objective of this service is to develop a tool aiming at measure and verify the energy and economic savings (with the possibility of emission savings) achieved after the implementation of Energy Conservation Measures (ECMs) in a building following the IPMVP (International Performance Measurement and Verification Protocol).

The service will be able from one side to create a Measurement and Verification plan following the IPMVP Option C. Option C is the one suitable for complete facilities with availability of data before (which is known as the Baseline period) and after the implementation of the energy efficiency measures (Reporting period). In addition, the service will be able to generate reports showing energy, economic and emission savings after the implementation of ECMs.

The software/tools that will be used for the development of the service are described below for the three main components: Front-end, Back-end and database.

- **Front-end (GUI):**

  This is the part of the service that interacts with the user and must be user-friendly, easy to understand (information included/shown by the service) and web-based. So far, the technology to develop this part has not been decided but should be based in a server/client architecture and offer a UX/UI useful for the user who will use the service. In addition, the interface must cover all the possibilities in the use of the service, allowing the users to define the project, upload data, select the independent and dependent variables, execute the model creation, calculate savings, etc. In short, it must follow the definition of the usage scenarios and use cases identified in the project, as previously described in Deliverable 2.1.

- **Back-end** (Data import and export, Data analysis, Statistical engine, savings calculation, reporting):

  The heart of the service, which is in charge of the data analysis and mathematical model calculation for the definition of the behaviour of the building (using the available data from the Baseline period), is being developed in Python. In addition, the savings calculation algorithms and other data connectors and utilities for uploading and downloading data from the service to the storage database will also be developed using Python.

- **Database** (store raw data from pilots):

  Many possibilities for storage database are currently being explored. The most suitable for now are relational SQL databases. The amount of data inserted into the database will basically be one year of data for the baseline period (before the ECMs implementation) and one year of data for the reporting period (after the ECMs implementation). Different storage frequencies could be available (15 minutes, 1 hour, etc.), but it is not expected to have a very large amount of data (most of the

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

times, building data are recorded not very frequently, and the number of different dependent and independent variables, is relatively low), so, it is not necessary to have Big Data.

The picture below shows the architecture of the service and the connection among the previously described modules. Note that the graphic only intends to represent a very easy to understand service workflow and how data are transferred from one part to another and how the user interacts with the service, but specific steps and definitions on how to store the data into the database, data format and other questions are not included in it.

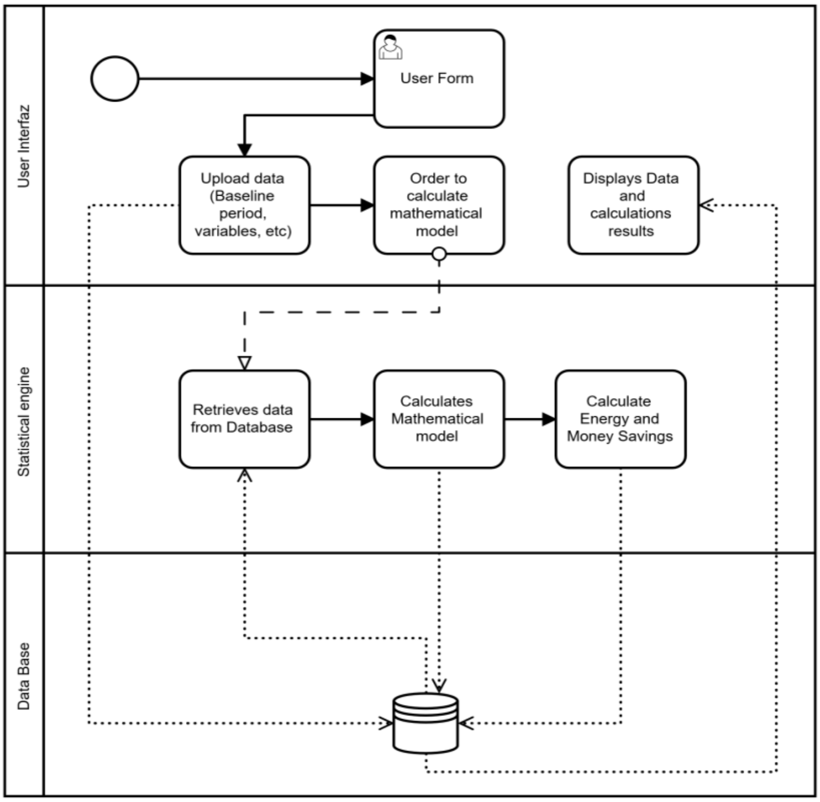

Figure 48 - Relation among modules

2.12.2 Input/output

The inputs to the service are the following:

- Metadata to define the project (location, type of building, ECMs, etc.).
- Energy consumption raw data (electricity, natural gas, etc.) before and after the implementation of the ECMs.
- Independent variables raw data. The main independent variable will be weather conditions data (temperature, radiation, etc.) before and after the implementation of the ECMs. Other independent variables such as occupancy could have also impact on the building's energy consumption.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

- Static factors (conditioned surface, set-point temperature, etc.) before the implementation of the ECMs and any changes that occurred with respect to the reference conditions.
- Energy prices.
- Emission factors.

The outputs of the service are as follows:

- M&V plan for the specific building under study. Optimal energy mathematical model.
- Energy savings report.
- Economic savings report.
- Emissions savings report.

All the data necessary for the execution of the service could be retrieved directly from the MATRYCS databases, or else, be provided as user inputs (although it should be noted that some of the inputs will necessarily be indicated by the user). For the calculation of models, Python libraries already included in the WP4 ML models and processing module will be used, to allow possible integration in later stages.

2.12.3 User experience

The main users of the tool will be experienced energy users such as ESCOs, energy facilities managers, energy auditors, etc.

The service follows the IPMVP workflow to measure and verify savings after the renovation/installation of energy efficiency measures in buildings. This means that first it is necessary to define the M&V plan with which to estimate the savings associated with the renovated building in a second step.

M&V plan creation (Before the intervention)

Generally speaking, after authentication in the MATRYCS service, the user must create a new project and define the general data of the project (type of building, location, ECMs, etc.) through a questionnaire/form. Then the user should upload the raw reference data (energy consumption, independent variables and static factors) to the service. The reference data could be uploaded to the tool in different ways such as CSV files, XLSX files, JSON, etc. Once the tool obtains the user data, the tool runs, analyse and processes the data through the statistical engine and provides the user with the optimal mathematical model of the reference scenario to predict savings (one mathematical model for each type of fuel). The mathematical model will be shown to the user through a virtual interface with the possibility of being printed, for example, in PDF format. The user has the opportunity to stop here and obtain the M&V plan if the only information available is the reference data.

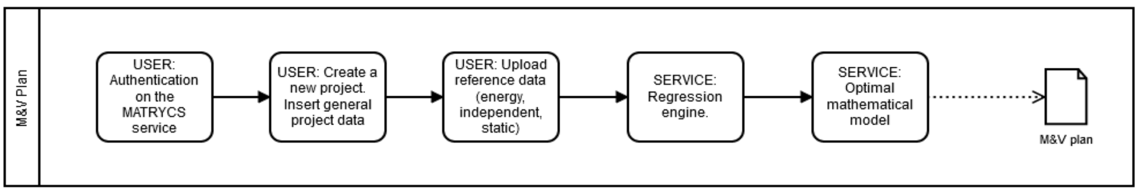

Figure 49 - Basic workflow for the model calculation

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

<u>Savings report generation (After the intervention once the ECMs have been implemented in the building)</u>

In this second step and if the user provides the tool with the raw data after the retrofitting process (energy consumption, independent variables, changes in the static factors, energy prices and emission factors) for at least one year after the ECMs intervention, the tool will run again using the mathematical model generated in the previous step and the savings calculation algorithms that provide the user with the corresponding reports. Again, the reporting data could be uploaded in the tool in different ways (CSV, XLSX, JSON, etc.). These reports will be displayed directly in the service interface with the possibility of being printed, for example, in PDF format and will obtain the M&V plan (including the mathematical model) and the savings. Depending on the information provided to the tool, the tool will provide only the energy savings report or also the economic and emission savings reports. These savings reports can be calculated periodically (e.g. savings can be calculated each year).

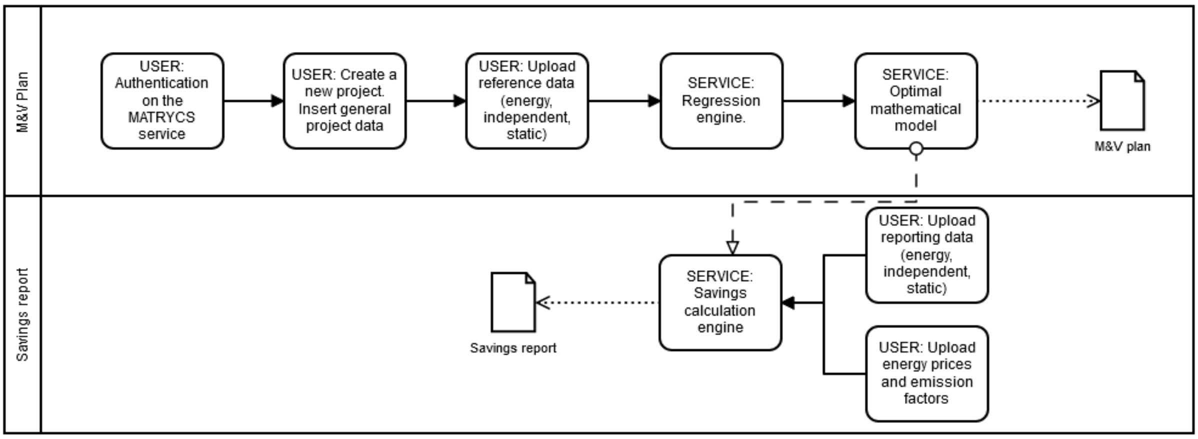

Figure 50 - Basic workflow for the savings calculation

Here below an initial mock-up on how the service will be presented to the user is shown.

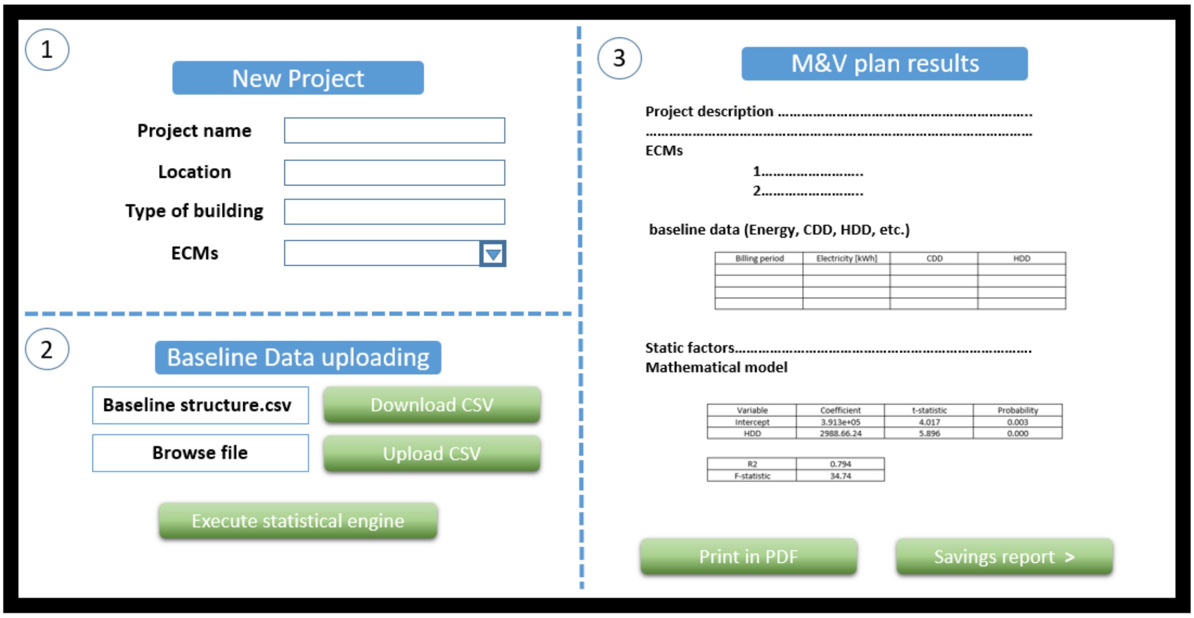

Figure 51 - M&V plan creation

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

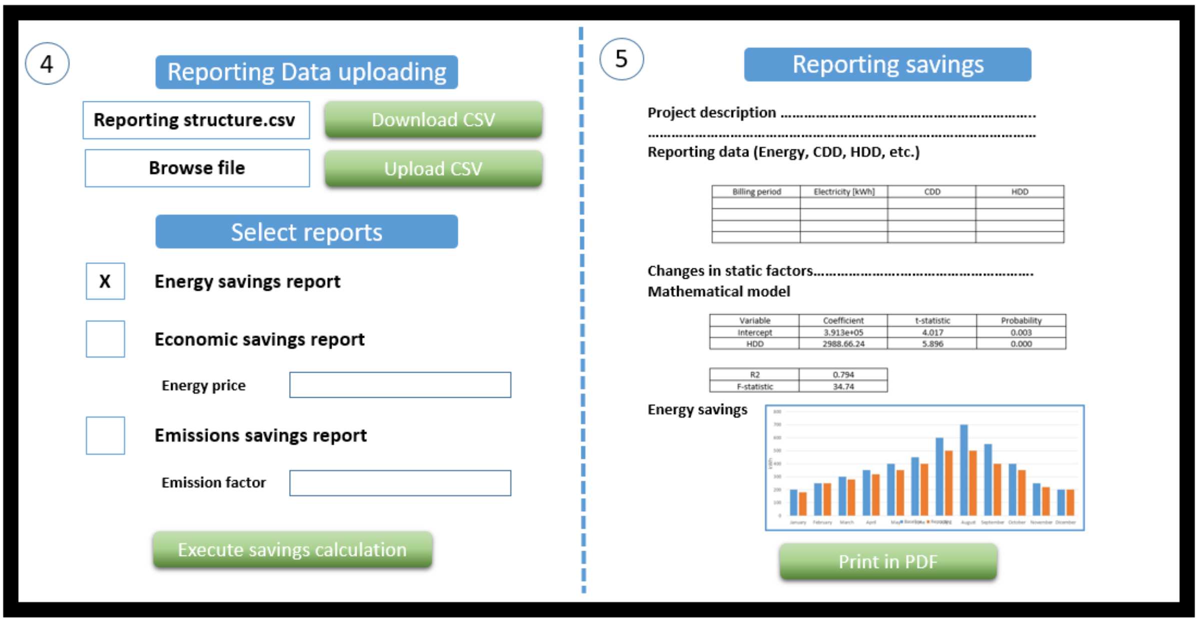

Figure 52 - Savings report generation

2.12.4 Application on MATRYCS Pilots

The pilots in which the tool will be tested and validated are the ones of VEOLIA (LSP03 – CP Arca 5 and LSP06 - Torrelago).

The data available from both Large-Scale Pilots that will be used for the purposes of the tool are the following:

- 15-minute frequency energy consumption data (e.g. boilers gas consumption, thermal production, etc.).
- 15-minute frequency weather conditions data (e.g. outdoor temperature).
- Static data.
- Energy prices.
- Emission factors.

The energy consumption data and the weather conditions information of the reference building will allow obtaining the adjusted mathematical model with which to estimate the energy savings once the data of the reporting period (after the implementation of the ECMs) are completely available (at least one year of data).

The static data will allow to check if the building has suffered any variation with respect to the reference situation (e.g. change in the conditioned area of the building, change in the use of any area of the building, etc.).

On the other hand, energy prices and emission factors, will allow estimating not only the savings in terms of kWh, but also in terms of € and kgCO2.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.13 MATRYCS – Services to support the financing of EE refurbishment (s4.2)

2.13.1 Tool description

The purpose of the tool is to provide economic indicators to investors, funders, and users to support them in the decision-making process. The tool foresees a close collaboration with the H2020, EENvest and TripleA projects. Currently, considering the development that is carried out within these projects it was decided to develop a platform that takes advantage of the services developed by EEnvest in a simplified form and to provide these results in the MATRYCS project. This approach allows on the one hand to validate a the EEnvest software funded by the European Commission and on the other hand to increase the use and users without having to develop a new tool from scratch.

The process involves the development of a flask app that uses the API provided by the EENvest software to send the data provided by the MATRYCS case study dataset and get back useful information for the creation of KPIs and a dedicated report.

The information and indicators that can be obtained depend on the data flow shown in the Figure 53

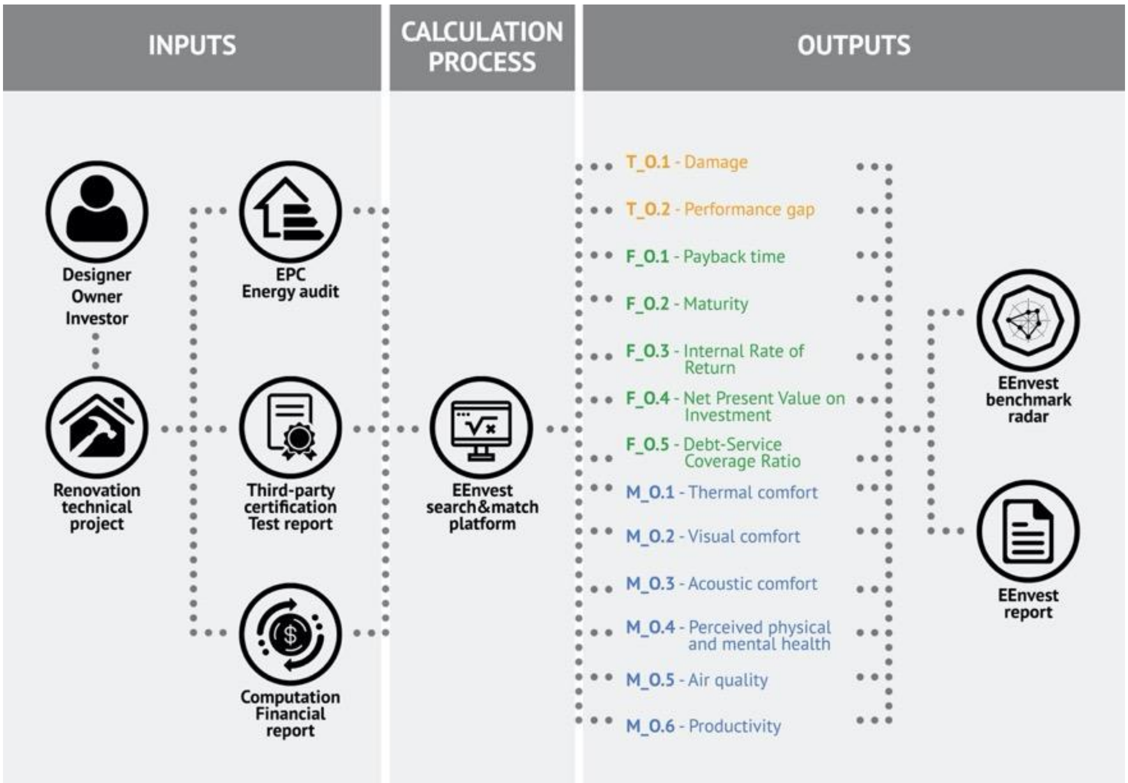

Figure 53 - Input from EENvest Platform[17]

The calculation of indicators depends on a well-defined list of inputs, based on which the results will be provided. The tool that will be developed in MATRYCS will try to simplify the work generated by the other two projects, providing an easier reading of the results.

[17] EEnvest 2021-06-09, Presentation GA #833112, Polimi

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.13.2 Input/output

The tool will follow the diagram shown in Figure 54the application will connect directly to MATRYCS GOVERNANCE to acquire all the necessary data. Such data will be sent to the EEnvest platform through the use of a REST API. the result will be fronted to the MATRYCS user in a dedicated front-end where all the necessary indexes and information will be displayed.

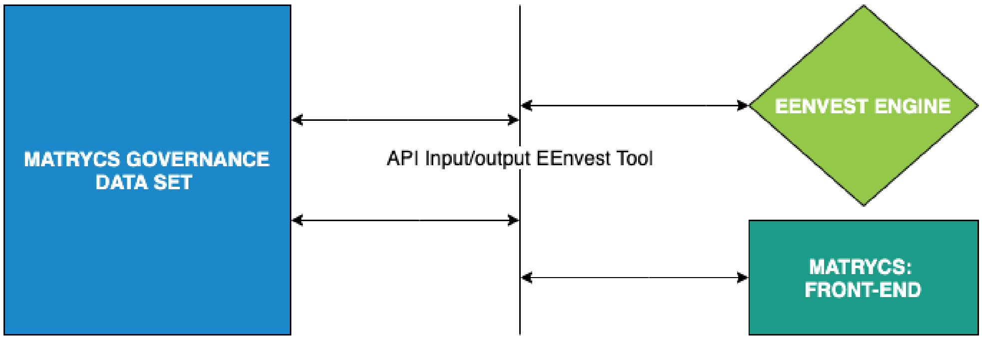

Figure 54 - Connection to MATRYCS GOVERNANCE

2.13.3 User experience

The service is at the design level at this stage. More information regarding the user experience will be provided later when all the components characterizing the tool will be defined and developed.

2.13.4 Application on MATRYCS Pilots

The use cases will be focused on **LSP10 – LEIF**:

- LSP10-> UC10_01: Risk assessment of funding in energy efficiency investments

The pilot aims to reduce the uncertainty linked to EE investments, which can be attributed to the lack of relevant skills and ability to assess investments. It will strengthen debt and equity financing of EE projects. The pilot leader is LEIF the only institution in Latvia that has reliable data on investments in EE and actual performance of investments in terms of energy savings. The pilot concept will demonstrate MATRYCS framework through cross-domain integration of a variety of heterogeneous historical and live data on financial performance, underlying EE impact of the investments, through historical extensive smart meters data integration.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.14 MATRYCS – Geo-clustering services as support to the Big Data Vision (s.0.2)

2.14.1 Tool description

The Geo-clustering is an intuitive tool that a wide range of actors in the building sector can take advantage of it. The main objective of the tool is to allow the comparison, analysis, and evaluation of building stock performance using cluster techniques. The tool can be tailored depending on the user's need and specific interest, the energy manager can compare the performance of buildings located in a same or different area as well as knowing what could be a possible saving in the refurbishment of a building applying specific action, building designer can it to figure out the best technology for a certain specific use and climate, researchers can analyse the impact of specific building features on energy consumption or to assess the relation between building envelop, mechanical system typology and thermal comfort.

Furthermore, Geo-clustered data are extremely valuable to highlight trans-national similarities and provide useful information to policy makers and legislative bodies for more tailored policy decisions.

The geo-cluster tool enables any platform user to:

- visualise the energy and comfort information of your buildings;
- compare buildings' performance through the use of well-defined Key Performance Indicators (KPIs);
- benchmark your building against others in the same category, that is having similar functionalities or with a specific building of your own portfolio; ;
- cluster buildings according to the selected KPIs and get the resulting clusters on a map.

The definition of the KPIs is extremely important for the analysis of the dataset and to identify meaningful output. The KPIs are used both for comparing buildings and for clustering.

This latter is a process that groups buildings in a set of meaningful classes, called clusters. A cluster is a collection of data objects that are "similar" to another one that can be treated collectively as one group.

In the MATRYCS clustering the buildings are clustered in groups having similar features. In addition to selecting the building type, the user, depending on his needs, chooses the relevant KPI among those identified during the project. The resulting clusters are represented with different colours and range of values in the map.

Currently two unsupervised machine learning algorithms have been used: K-Means algorithm and HDBSCAN.

In general, a cluster represents a collection of aggravated points having common similarities. The key point is to define the exact number of centroids in the dataset (k). The centroid is the central point of the cluster identifying the quantity. Each data point is assigned to each of the clusters by reducing the sum of the in-cluster squares. the K-means algorithm identifies k number of centroids, and then assigns each data point to the closest cluster, keeping the centroids as small as possible.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

HDBSCAN is a density-based clustering algorithm that constructs a cluster hierarchy tree and then uses a specific stability measure to extract flat clusters from the tree18.

Unlike K-Means, HDBSCAN and in general density-based methods algorithm work well even when the data isn't clean and the cluster are widely shaped.

2.14.1.1 Tool architecture

The tool architecture is shown in Figure 55.

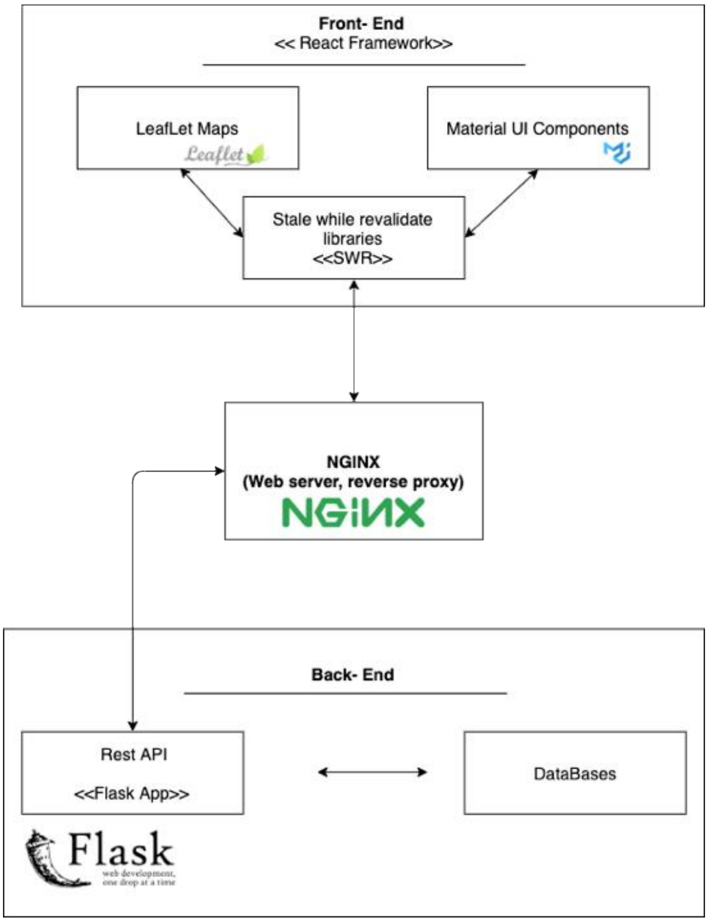

Figure 55 - Geo-clustering tool architecture (on localhost)

The front-end has been developed in a way to make it independent and able to integrate in a larger system. By doing so we have delegated functionalities such as authentication and cookies management to subsystems external to the tool, assuming a valid session before access to the tool is granted.

The interested subsystem created consists of two parts: the front-end and the back-end, which in turn is composed by the database and the application managing the logic of the whole tool.

In general, there are 3 main components that manage the state and re-rendering processes of its children: Map, Sidebar and Bottom bar. Furthermore, to avoid unnecessary re-render processes triggered by an update of the state of a particular higher order component, we use the Context API, so

[18] A hybrid approach to Hierarchical Density-based Cluster Selection – Claudia Malzer, Marcus Baum – 2020 IEEE International COnmference on Multisesnor Fusion and Integration for Intelligent Systems (MFI), Karlsruhe, Germany, pp 223-228

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

that only specific interested elements of the DOM are updated whenever an event is triggered. Further additions to the architecture would be implemented following the pattern designed in the figure 49.

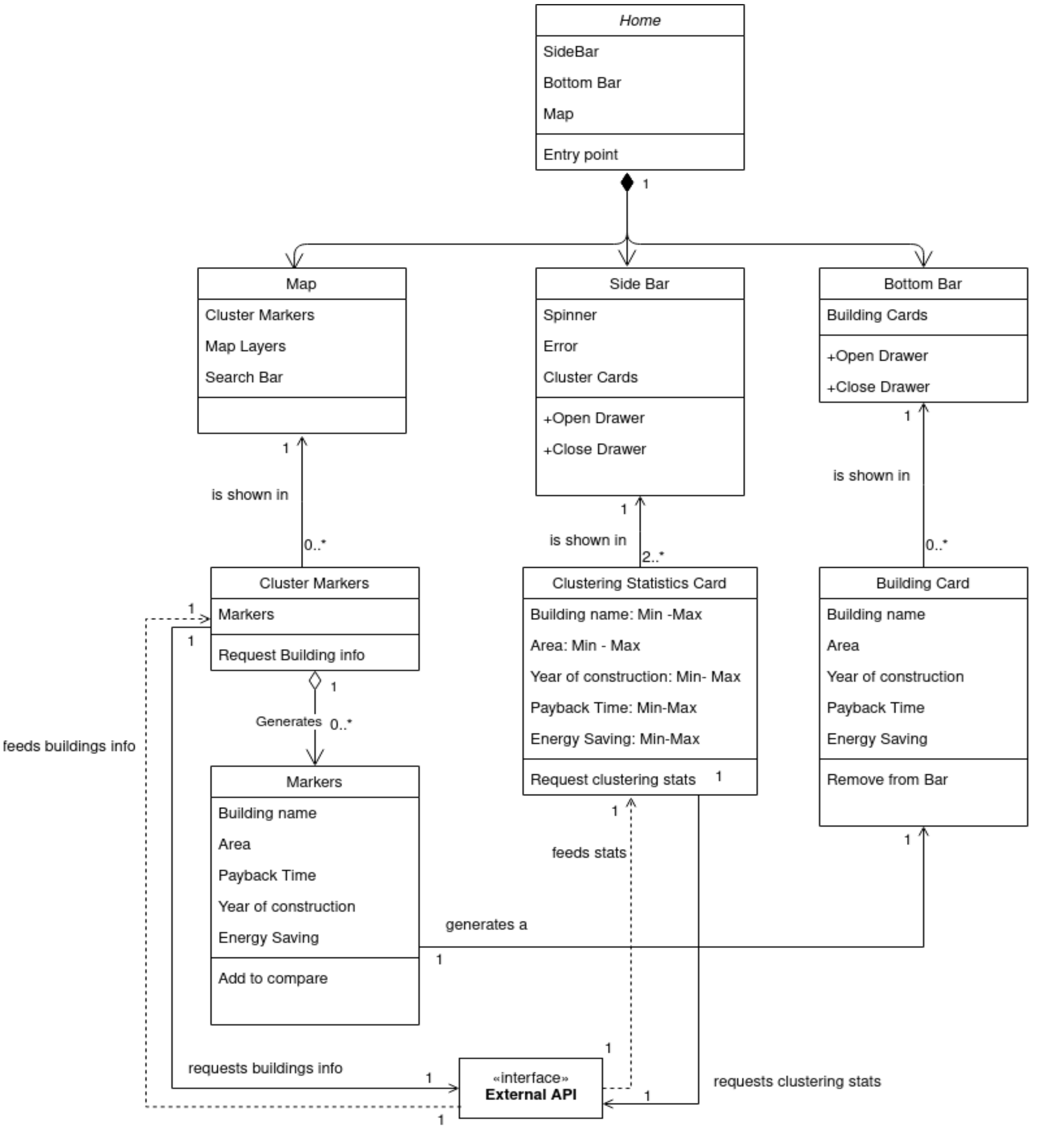

Figure 56 - Additions to the geoclustering architecture in the application to the LEIF dataset

2.14.1.1.1 Front-end

The front-end of the tool has been developed considering different aspects such as performance, usability, and presentation. To reach these aims different frameworks have been used:

- **Next.js**: The front end is a single page application created with Next.js (https://nextjs.org/), a React-based framework designed to support different types of optimizations such as client-side redirection, image optimization, server-side rendering, and more. This framework is especially well adapted for the design of lightweight applications that are capable of running even on low computing-power devices. In particular, this tool presents several features such as component hydration, image optimization and the use of dynamic data fetching.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

- **Material-UI:** We have integrated into the project the Material-UI framework for graphics and user interaction. This allows the creation of a user-friendly site with a highly recognizable style, facilitating the interactions with the different features of the page. Moreover, the utilization of this framework has accelerated the development of the application by providing readily available but highly customizable react components.
- **Leaflet.js:** As a graphic support to geo-localize the different points of interest in the application we have resorted to the open-source project Leaflet.js, which provides an interface for the visualization and interaction of map layers. This framework is a mature project actively supported by a big community of developers that offers a wide range of features, which have provided a solid base for the development of the front-end.

2.14.1.1.2 Back-end

In this first phase the tool is developed for operation on a local machine. The components that characterize the back-end of the geoclustering tool are:

- **Server:** the EURAC server is used to test the tool in this development phase. In fase di produzione il tool sarà spostato nell'infrastruttura EGI.
- **Database**: Data provided by the case studies in xls and csv formats were transformed into RDF files of type *.ttl* using the "Brick schema" ontology. The resulting file IS currently within a local database (GraphDB). In the production phase the MATRYCS structure will be used, in particular the reasoning engine connected to the NEO4J - NEOSMANTICs database where the dataset will be saved.

The connection between the Back-end and the front -end is provided via a dedicated REST API. Currently a swagger API has been developed (Figure 57):

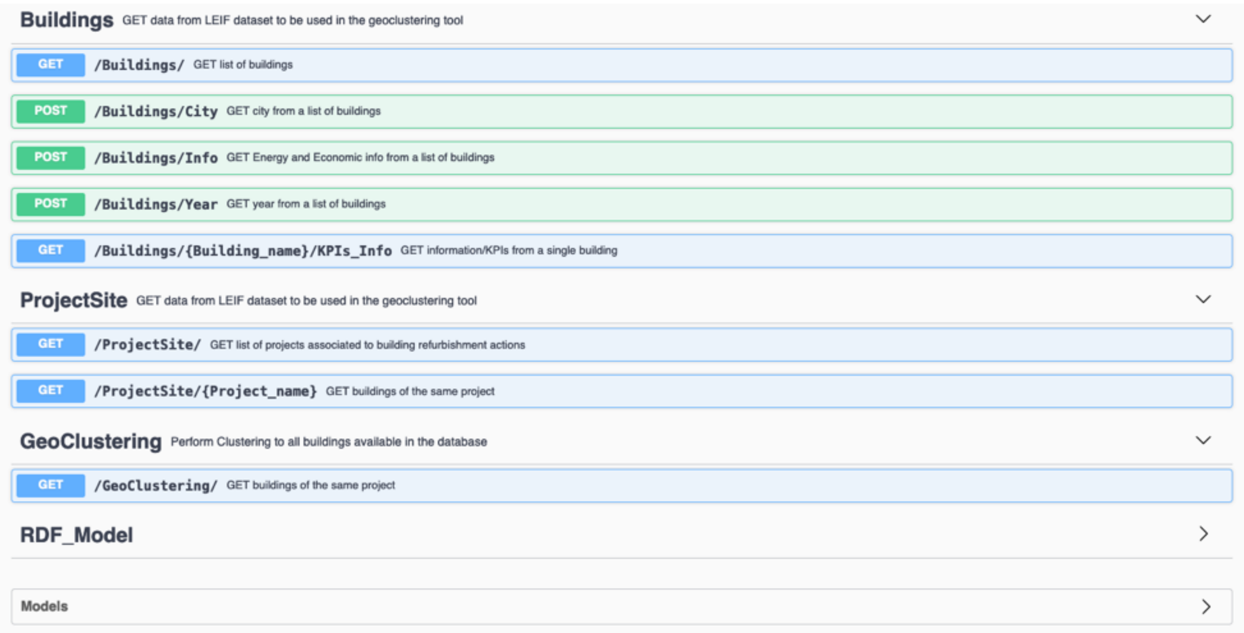

Figure 57 - SWAGGER UI for Geo-clustering API

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

2.14.2 Input/output

The geo-clustering tool needs both data and metadata to perform the analysis. The metadata is static information of the datasets that does not change over the time. The metadata file is generated following the data model defined in WP3. It is an RDF file (.ttl) that use the Brick ontology as semantic elements definition.

The latter describes precisely and unambiguously the different elements that characterize the building and its systems. The peculiarity is that it identifies different entities by means of a glossary or dictionary and links them to each other using relationships.

Figure 58 shows the connection between the Geo-clustering tool, the MATRYCS- PROCESSING, and the MATRYCS- GOVERNANCE. The RDF file is stored in NEO4J database using a specific plugin (neosemantics- n10s). A Retrieval Engine queries the database and expose the results in a dedicated rest API.

The clustering models are defined using BENTOML, that is a high-performance framework for serving, managing, and deploying machine learning models. BENTOML queries the Retrieval Engine to get the data needed to generate the clustering model. The model development and evaluation are performed in the serving framework, the results are exposed in a dedicated API, to be used by the Geo-clustering tool.

The metadata, used only for visualization and comparison of data between different buildings, are directly retrieved from the reasoning engine.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

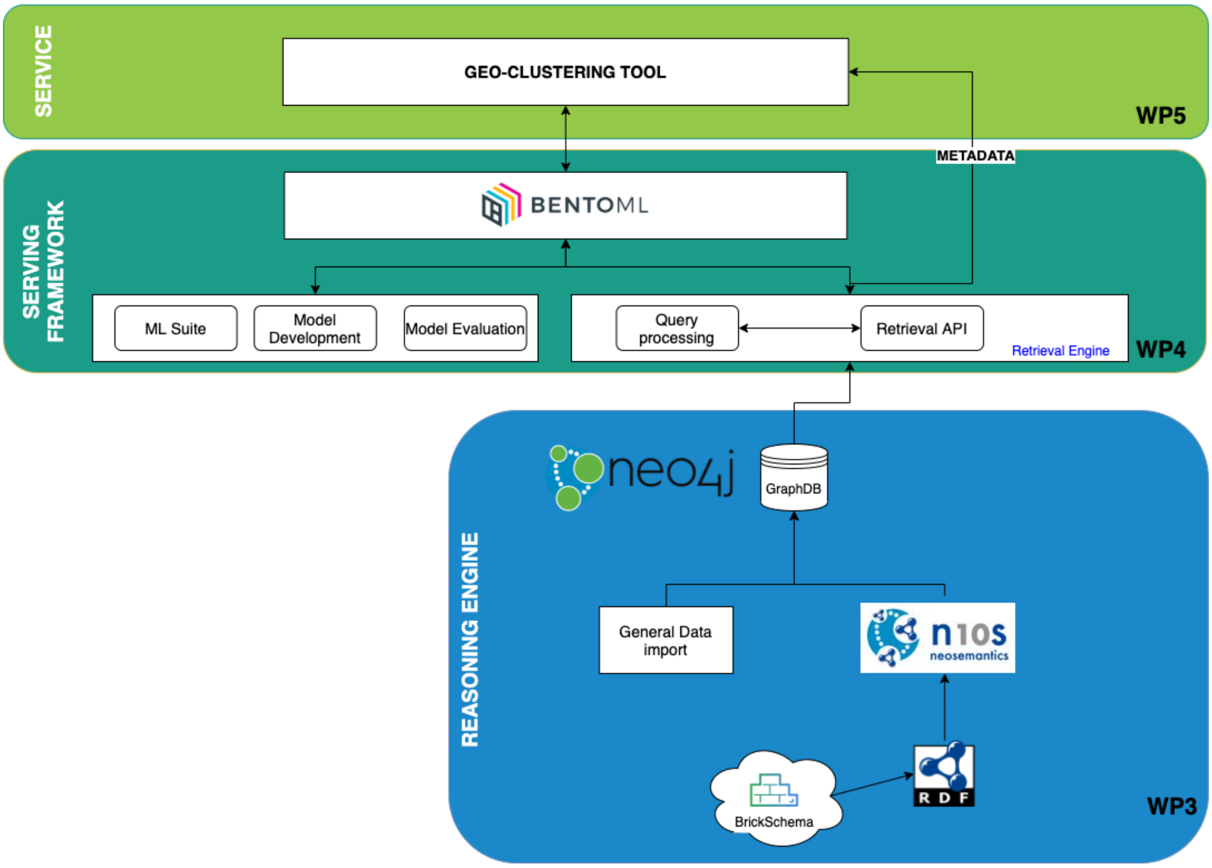

Figure 58 - Connection of the Geo-clustering tool with MATRYCS-GOVERNANCE (WP3) and MATRYCS PROCESSING (WP4)

2.14.3 User experience

Figure 59 illustrates the user approach to the geo-clustering service. After authentication to the MATRYCS analytics service platform, the user can either load the data in MATRYCS through a dedicated tool or, if already done, he can choose the service to be used (in this case the geo-clustering tool).

The tool allows two types of use: Data analysis and Data simulation. In data analysis, the previously loaded data is analysed in three ways: Data benchmarking, Data visualization and Clustering. In the latter, the resulting clusters provide information about lower and upper limits of some indicators, such as the energy saving, the Indoor Environmental Quality (IEQ) and economic analysis.

The user can also simulate in which cluster his building belongs by providing a list of necessary information. In this case, the tool places the building in the cluster it belongs to, to which the limit values of the indicators are associated.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

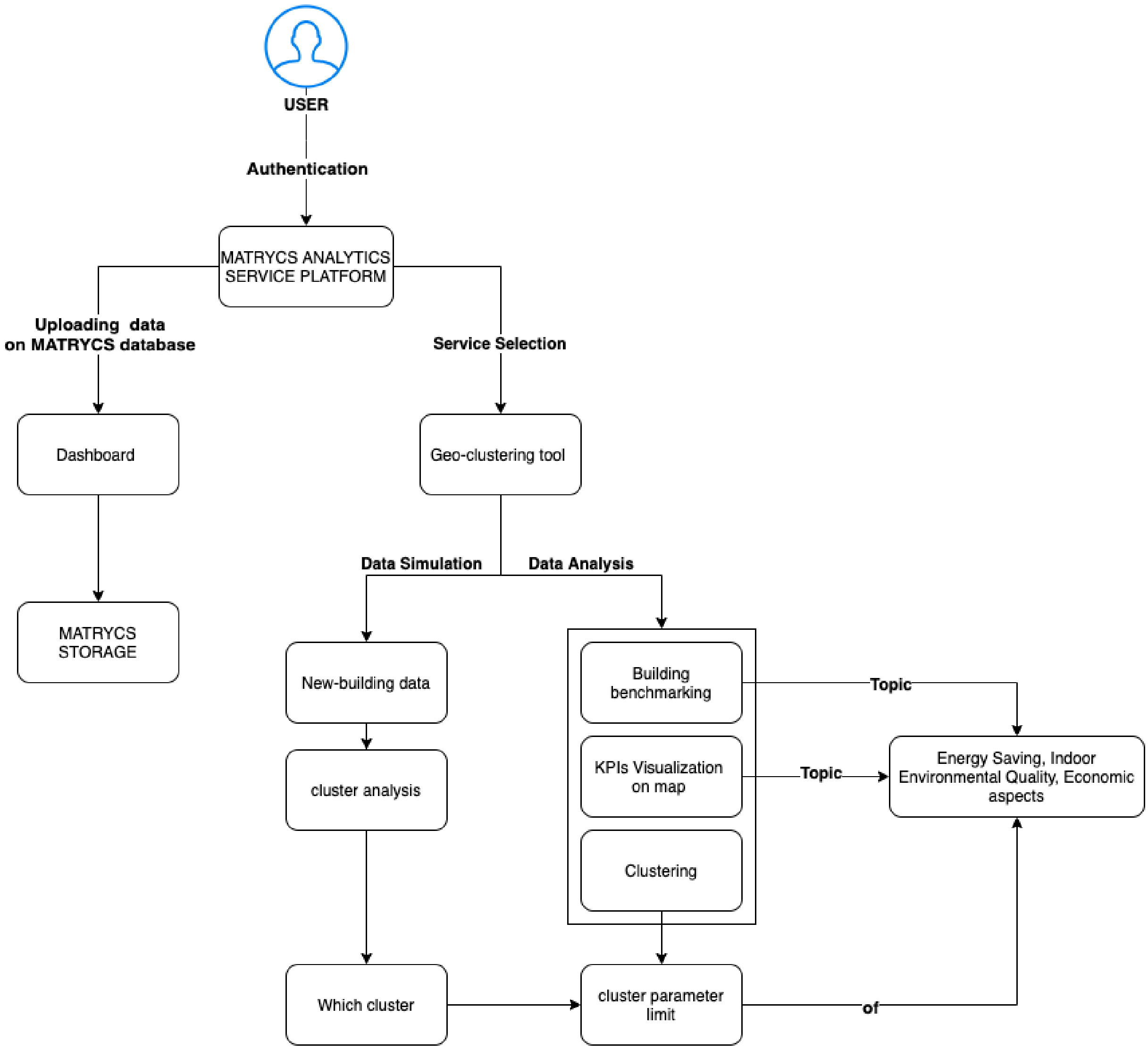

Figure 59 - User workflow for the service s0.2

2.14.4 Application on MATRYCS Pilots

Due to the lack of data of the LSP8 – GDYNA (UC08_01: Geo_municipality service), to validate the Geo-clustering tool the **LSP10 (LEIF)** dataset was used has indicated as possible application (Figure 65). This dataset provides a list of values related to buildings feature, energy consumption of a building stock, weather and economic data.

Moreover, the dataset has several information about the energy consumption of the building before and after the application of some Energy Conservation Measures (ECM).

The inputs are:

1) Year of the building
2) Type of building
3) Number od building per renovation project
4) Heating area of the building
5) Renovation actions
6) Town and County

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

7) Total energy consumption before and after renovation [kWh/m2]
8) Total heating consumption before and after renovation [kWh/m2]
9) DHW consumption before and after renovation [kWh/m2],
10) Electricity consumption before and after renovation [kWh/m2]
11) Cost of intervention
12) evaluation of CO2 reduction,
13) monthly average data of external temperature for different years and weather station

We identified the following KPIs to be used in the analysis:

- Energy saving (Normalized data)
- Cost saving
- Simple Pay Back Time of investment

The buildings have been clustered using the following parameters: Area, Year of construction, Heating consumption before Renovation, Renovation Cost, Action applied (Energy Conservation Measure - ECM), Heating Degree Days before the application of the ECM (Figure 60).

The data before being clustered were filtered by city and building type. The user can simulate the possible savings that could be obtained in the redevelopment of a building on the basis of the data provided by the dataset, through the specifications that each cluster has (Figure 62). For example, minimum and maximum value of energy savings, economic, and return on investment.

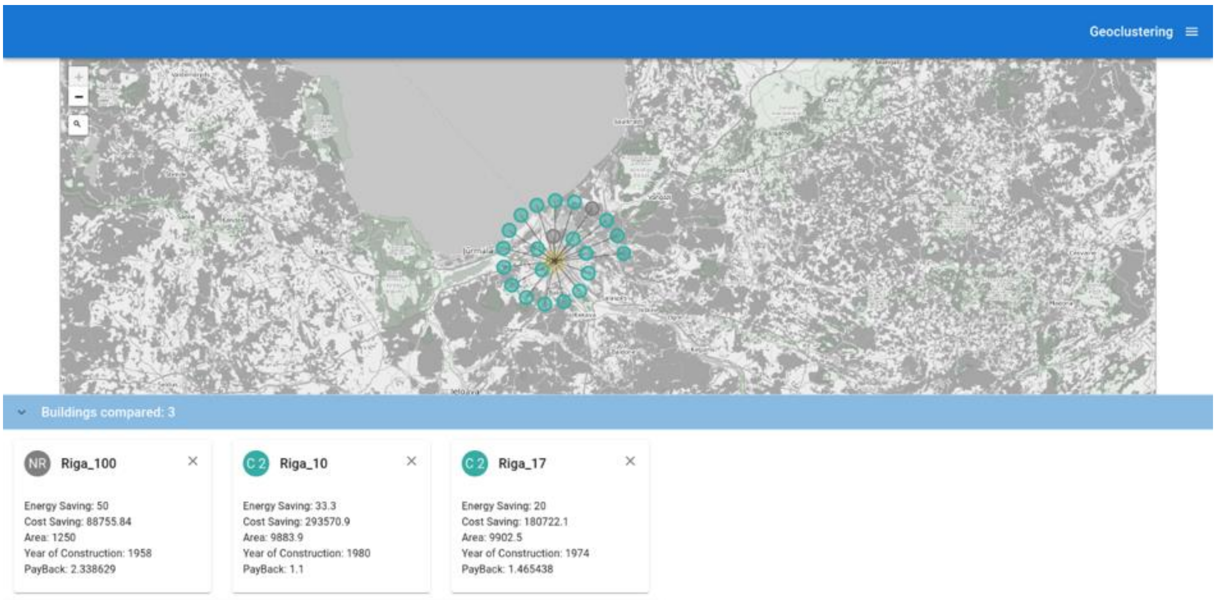

Figure 60 - Building clustering

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

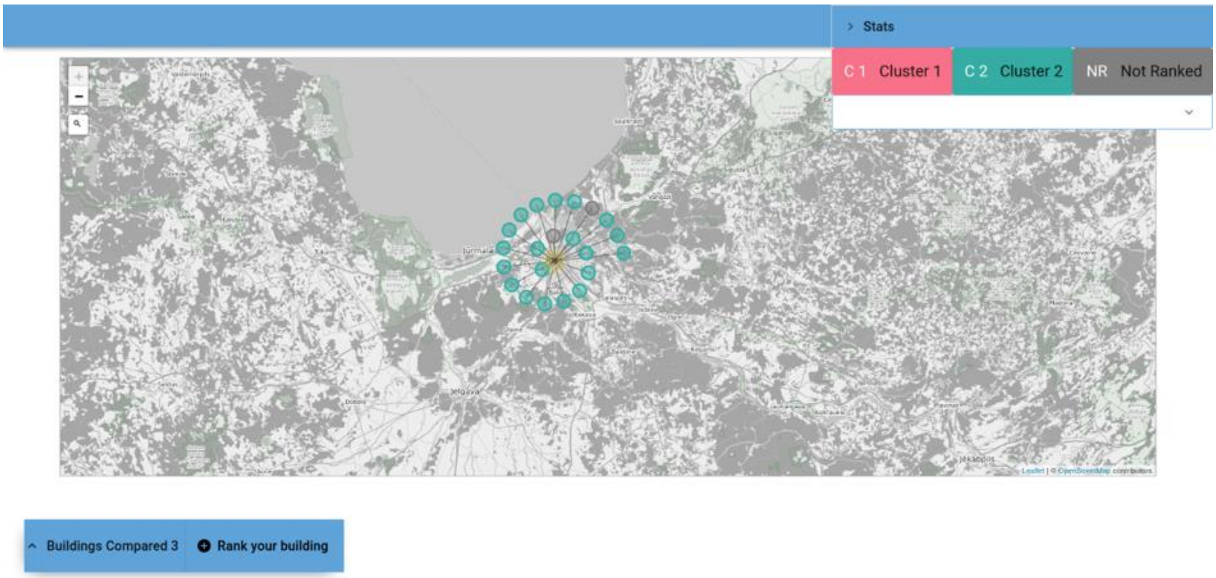

Figure 61 - Geo-clustering interface

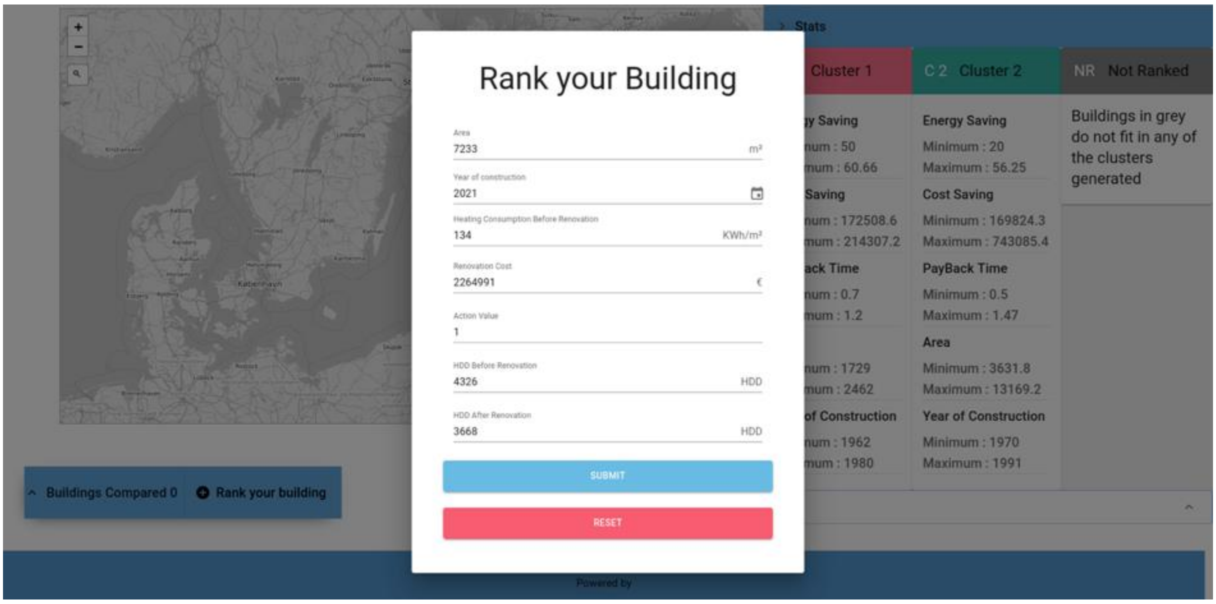

Figure 62 - Rank the new building

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

3. MATRYCS – Building Services Integration

The **Analytics Building Applications** together with **Visualizations and Reports Engine (WP4)** and **Digital Twins** are the three pillars of the Software as a Service in the MATRYCS architecture. The integration of these services with the MATRYCS processing and governance (PaaS) is summarized in the Figure 63

The security framework will handle the authentication and access of different users to a common platform where all services will be available. The security framework will guarantee the possibility of using different analytical services according to the data provided and complying with the EC regulations on Data Protection (GDPR).

Each single service will use the data or analysis of MATRYCS processing and governance in three main ways:

- Direct access to building database
- Connection to specific frameworks where different machine learning algorithms can run generating data and analysis to be shown in a dedicated front-end.
- Using the reasoning engine.

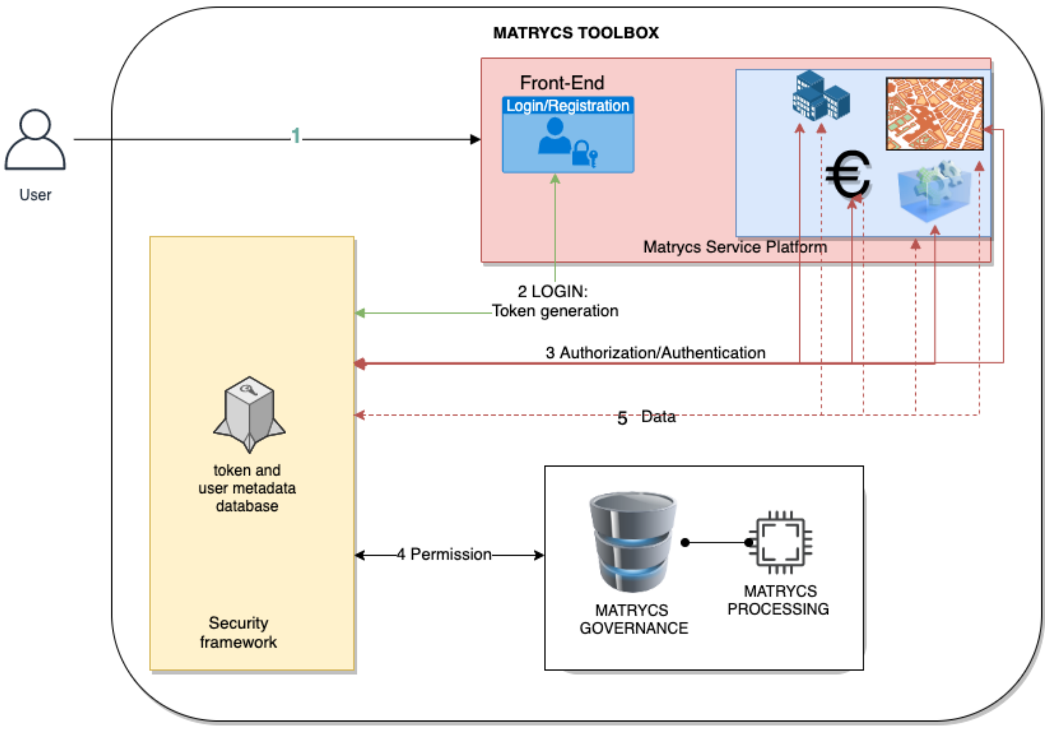

Figure 63 - MATRYCS Toolbox connection

According to the type of service one of the three types of connection will be set-up. For example, considering the development of the geoclustering tool, the application of a clustering algorithm will be referred by a specific serving framework (BENTO ML) running inside the MATRYCS processing. In all

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

cases an API will be provided to manage the exchange of information between the analytical services and the MATRYCS processing or governance.

3.1 MATRYCS service dashboard

Technical partners will be in charge to deploy the services. The latter will be developed to run independently as a docker container (virtual machine is also an optional under evaluation).

A container contains everything needed to run a small piece of software. Each container includes all the code, its dependencies and even the operating system itself. This allows applications to run almost anywhere: a desktop computer, a traditional IT infrastructure or the cloud. (Figure 64Figure 64).

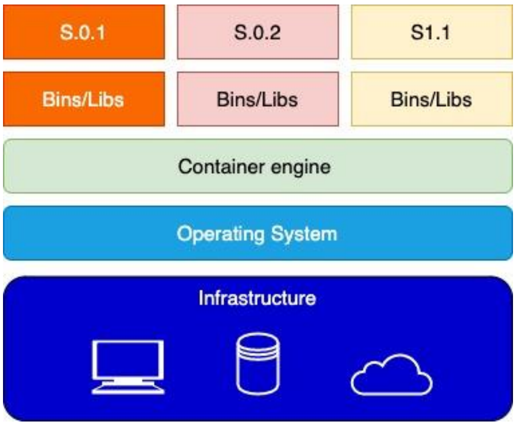

Figure 64 - General Container infrastructure

Each application provides a specific front-end to the user that he can use to perform his analysis. In order to facilitate the choice of all services and to manage account security a general platform with a dedicated front-end will be developed. The user after login will be able to access the chosen service based on the data provided, as shown in Figure 63.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

4. Conclusion

This document aims to show the first development related to the services defined in MATRYCS. The need to store data from the different pilots, analyze them and define how each service must act is an operation that requires time and several steps.

In particular, every single service needs to connect to WP3 for the data storage to be used. In fact some services may need a direct connection with the database without having to go through MATRYCS processing (WP4). The latter is instead very useful for all those cases where the application of algorithms or data analysis systems is essential. Examples are the services s0.2, s1.1, s1.2, s.4.1, etc. which apply algorithms for prediction, clustering and post-processing of data whose task is delegated to WP4.

Fundamental will be then the connection with WP6 where it will be carried out a verification of the results obtained in WP5, requesting feedback to the different pilots in order to continue to improve the service developed and make it usable to external applications.

Next steps are the continuous development of the application, the connection with the database, testing of MATRYCS processing features and implementation of the security framework.

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158

5. Annex

Following the work done in the deliverable D2.1. Hereunder, the list of all scenarios and services associated to each LSP can be found. The definition of each scenario can be visualized in the D2.1. The final application of each service in one or more LSPs will be observed in WP5 (Analytics Building Services), through the development of the services, and tracked as well in WP6 (Deployment and Validation in Real Environments) and reported in corresponding deliverables.

- ✓ Service included in the pilot
- ? Service proposed to be added to the pilot
- X Service included in the pilot but not communicated by service provider
- ? Service proposed to be added in the pilot but not communicated by service provider
- Service not included in the pilot
- ✓ Service included in the proposal, and proposed for addressing the User Scenarios
- ✓ Service proposed to be added after the Sessions with Pilots, proposed for addressing the User Scenarios
- X Service included in the proposal, and not proposed for addressing the User Scenarios

		Services in the Pilot (and matched with User Stories)														Usage Scenarios
PILOT		s0.1	s0.2	s1.1	s1.2	s1.5	s1.3	s1.4	s2.1	s2.2	s3.1	s3.2	s3.3	s4.1	s4.2	No.
LSP1	BTC	✓			✓		✓	✓		✓				✓	✓	20
LSP2	FASADA	✓					X		✓	✓				✓	✓	3
LSP3	VEOLIA						✓	✓						✓		19
LSP4	ASM			✓	✓	✓	?	?						✓		10
LSP5	COOPERNICO		?	✓	✓		✓	?	✓					✓		16
LSP6	VEOLIA	X	?	X		X	✓	?	X	X						18
LSP7	ICLEI		?				?				✓		X			12
LSP8	GDYNIA	X					✓			X					X	
LSP9	EREN	✓	✓				✓		✓	✓		✓				12
LSP10	LEIF		✓				?		✓	✓			✓	✓	✓	18
LSP11	HOUSING EUROPE												✓			3

Figure 65 - The final number of pilots' usage scenarios and services matrix

The MATRYCS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no.101000158