Extended dEPCs applications toolkit v2

The D^2EPC project has received funding from the EU's Horizon 2020 research and innovation programme under grant agreement No 892984

Project Acronym:	D^2EPC
Project Full Title:	Next-generation Dynamic Digital EPCs for Enhanced Quality and User Awareness
Grant Agreement:	892984
Project Duration:	36 months (01/09/2020 – 31/08/2023)

DELIVERABLE D4.7
Extended dEPCs applications toolkit v2

Work Package:	WP4 Digital Platform for Dynamic EPCs Issuance and Enabled Applications
Task:	T4.3 Extended dEPCs applications toolkit
Document Status:	Final
File Name:	D^2EPC_D4.7_ Extended_dEPCs_applications_toolkit_v2_HYP
Due Date:	31.08.2023
Submission Date:	31.08.2023
Lead Beneficiary:	HYP

Dissemintation Level

Public	☐
Confidential, only for members of the Consortium (including the Commission Services)	☒

Authors List

Leading Author				
First Name		Last Name	Beneficiary	Contact e-mail
Thanos		Kalamaris	HYP	t.kalamaris@hypertech.gr
Co-Author(s)				
#	First Name	Last Name	Beneficiary	Contact e-mail
1	Stefanos	Makris	HYP	s.makris@hypertech.gr
2	Pavlos	Kakarantzas	HYP	p.kakarantzas@hypertech.gr
3	Giorgos	Fiotakis	HYP	g.fiotakis@hypertech.gr
4	Nikolas	Spiliopoulos	HYP	n.spiliopoulos@hyperteck.gr
5	Stelios	Choumas	HYP	s.choumas@hypertech.gr
6	Christos	Malavazos	HYP	c.malavazos@hypertech.gr
7	Lefteris	Kalonaris	HYP	l.kalonaris@hypertech.gr
8	Panagiotis	Klonis	CERTH	pklonis@iti.gr
9	Konstantinos	Chatzintinos	CERTH	kchatzi@iti.gr
10	Gerfried	Cebrat	SEnerCon	gerfried.cebrat@senercon.de

Reviewers List

Reviewers			
First Name	Last Name	Beneficiary	Contact e-mail
Paris	Fokaides	FRC	eng.fp@frederick.ac.cy
Phoebe-Zoe	Georgali	FRC	res.gp@frederick.ac.cy
Naghmeh	Altmann-Mavaddat	AEA	naghmeh.altmann@energyagency.at

Version History

v	Author	Date	Brief Description
1.1	**T. Kalamaris**	**08.06.2023**	**First draft based on the previous version**
1.2	**T. Kalamaris**	**07.07.2023**	**Document draft ready, sent to task partners for feedback collection**
1.3	**T. Kalamaris**	**28.07.2023**	**Updated information based on input from partners**
1.5	**T. Kalamaris**	**04.08.2023**	**Document ready for peer review**
1.7	**T. Kalamaris**	**21.08.2023**	**Document updated based on the peer review comments**
2.0	**T. Kalamaris**	**31/08/2023**	**Version ready for submission**

Legal Disclaimer

The D^2EPC project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 892984. The sole responsibility for the content of this publication lies with the authors. It does not necessarily reflect the opinion of the European Climate, Infrastructure and Environment Executive Agency (CINEA) or the European Commission (EC). CINEA or the EC are not responsible for any use that may be made of the information contained therein.

Copyright

Executive Summary

This report is the second version of the Extended dEPCs Applications Toolkit. D4.7 v2 delivers a detailed description of the D^2EPC toolbox that provides a series of invaluable services to the end-user. The respective modules correspond to the “Energy Performance Benchmarking” and the “Energy Performance Verification and Credibility” and are thoroughly discussed within the deliverable.

D4.7 initially presents a conceptual overview of the application’s toolkit to be integrated into the D^2EPC architecture along with the literature research which preceded the design of components. Then, the report dives into the sub-modules that constitute the two main components of the toolkit, providing insights into their overall operation and functionalities to be delivered to the end-users and other stakeholders.

Starting from the Energy Performance Benchmarking, the module comprises two separate sub-modules, the Classification and Benchmarking Tool. The Classification Tool is equipped with the appropriate modern data mining algorithms able to extract information from operational and temporal building characteristics. Its main focus is to generate distinct classes for the categorisation of the buildings under study. On the other hand, the Benchmarking Tool steps on the generated classes and materialises the benchmarking service. This service contributes to the evaluation of the building performance by establishing a set of performance criteria – built upon the comparison of buildings- to be satisfied by future EPC users.

The large number of installed devices and the huge volumes of generated data raise the need for specific applications to check the operation of IoT equipment and the quality of extracted data. For this reason, the D^2EPC Applications Toolkit includes the Energy Performance Verification & Credibility module which is equipped with the appropriate software to undertake the device monitoring and data quality assessment. The Network Monitoring Tool keeps track of the operational status of deployed devices and delivers a network representation to the end-user via a Credibility User Interface (UI). Meanwhile, the Data Quality Tool performs tailored checks on the collected data to ensure their adequacy prior to their utilisation from other project components.

Beyond the components’ description, the report includes the finalised versions of the modules along with their technical design and presents validation results, part of the works conducted in the context of T4.3.1 and T4.3.2.

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

Term	Description
API	Application Programming Interface
DQ	Data Quality
CSV	Comma Separated Value
DWD	Deutscher Wetterdienst
EPC	Energy Performance Certificate
EPB	Energy Performance Benchmarking
EPG	Energy Performance Gap
EPVC	Energy Performance Verification and Credibility
GDPR	General Data Protection Regulation
GEG	Gebäudeenergiegesetz
HVAC	Heating, Ventilation and Air-Conditioning
IT	Information Technology
ISO	International Organisation for Standardisation
IoT	Internet of Things
KPI	Key Performance Indicator
PHP	Hypertext Preprocessor
SRI	Smart Readiness Indicator
GWP	Global Warming Potential
RES	Renewable Energy Sources
IML	Information Management Layer
DT	Digital Twin
SVG	Scalable Vector Graphics
UI	User Interface

1 Introduction

1.1 Scope and objectives of the deliverable

The main focus of the deliverable is to provide an overview of the D^2EPC Extended dEPCs Applications Toolkit and report the works conducted as of M26 of D^2EPC. More specifically, D4.7 describes the operation and properties of two separate tools constituting the Applications Toolkit. The "Energy Performance Benchmarking" is responsible for the categorisation of buildings on various traits and the "Energy Performance Verification and Credibility" undertakes the monitoring of the installed IoT Devices in the project's pilots and performs tailored checks to the extracted datasets to ensure the overall data quality.

The report aims to provide insights into the research performed towards the designing of the relevant components, the definition of their integrated functionalities and the overall development status up to this stage of the project.

1.2 Structure of the deliverable

The content of D4.7 is structured into 5 chapters to provide a holistic view of the Extended dEPCs Applications Toolkit. A brief description of the sections and their respective content is presented below:

- **Chapter 2 - Insights on the Extended dEPCs Applications Toolkit**: This chapter includes a detailed description of the modules residing in the Applications Toolkit. It begins with a general view of the toolkit and the literature research that preceded the implementation of the tools. It further dives into the functionalities of the sub-modules that materialise the desired services.
- **Chapter 3 - Extended dEPCs Applications Toolkit - Technical Design and Implementation**: This chapter delivers the interactions of the Extended dEPCs Applications Toolkit modules with system components of the D^2EPC system architecture along with the operations carried out. It further provides technical aspects of the components' development and data flows.
- **Chapter 4 - Extended dEPCs Applications Toolkit Validatio**n: This chapter includes the works conducted towards the validation of the incorporated components as of M26 of D^2EPC.
- **Chapter 5 – Conclusions**: This chapter wraps up the included information and provides insights on the second version of the deliverable expected in M36.

1.3 Relation to Other Tasks and Deliverables

D^2EPC work structure is broken down into eight building blocks in total, the Work Packages (WP), which are further segmented into tasks that undertake specific pieces of work. Task 4.3 is part of D^2EPC's fourth WP, which implements the project's digital platform as well as the enabled applications. The components that were delivered within the task have been initially described in the project's architecture under WP1 and specifically T1.4. Both modules of the Applications Toolkit (EPB and EPVC) are highly engaged in data acquisition and provision. As a result, T4.3 is closely related to the activities of T2.5 (in which the project's Common Information Model is designed and delivered), T3.1 (i.e., the task responsible for the definition of the Information Management Layer), and T3.3 (which delivers the D^2EPC Digital Twin component) for the static and dynamic data acquisition and streaming of the assessed information. Other interactions of the toolkit's modules concern the project's calculation engine (T4.1), which takes as input the collected data from the pilots to calculate various quantities (e.g., performance indicators) and the alerting system (T4.2), which aims to inform the user about the building's energy performance or warn the user for various inconveniencies (e.g., data quality). Finally, provided that the integrated services are delivered to the end user via a web application, T4.3 is highly related to the works of T4.4 which undertakes the design and development of the main D^2EPC web platform.

2 Insights on the Extended dEPCs Applications Toolkit

2.1 Overview of the Applications Toolkit

D^2EPC aims to deliver a holistic platform that comprises a plethora of services and applications that provide insights to end-user in regards to the building's operation and performance. The "Extended dEPCs Applications" corresponds to a toolkit that includes the appropriate software components which activate two distinct functionalities of the D^2EPC framework:

- The categorization and, hence, the comparison between buildings of similar characteristics delivered by the Building Energy Performance Benchmarking (EPB) module
- The validation of the dynamic data elements entered as input for the issuance of the dynamic EPC undertaken by the Energy Performance Verification and Credibility (EPVC) module.

In Figure 1, a conceptual diagram of the components that constitute the D^2EPC system is presented.

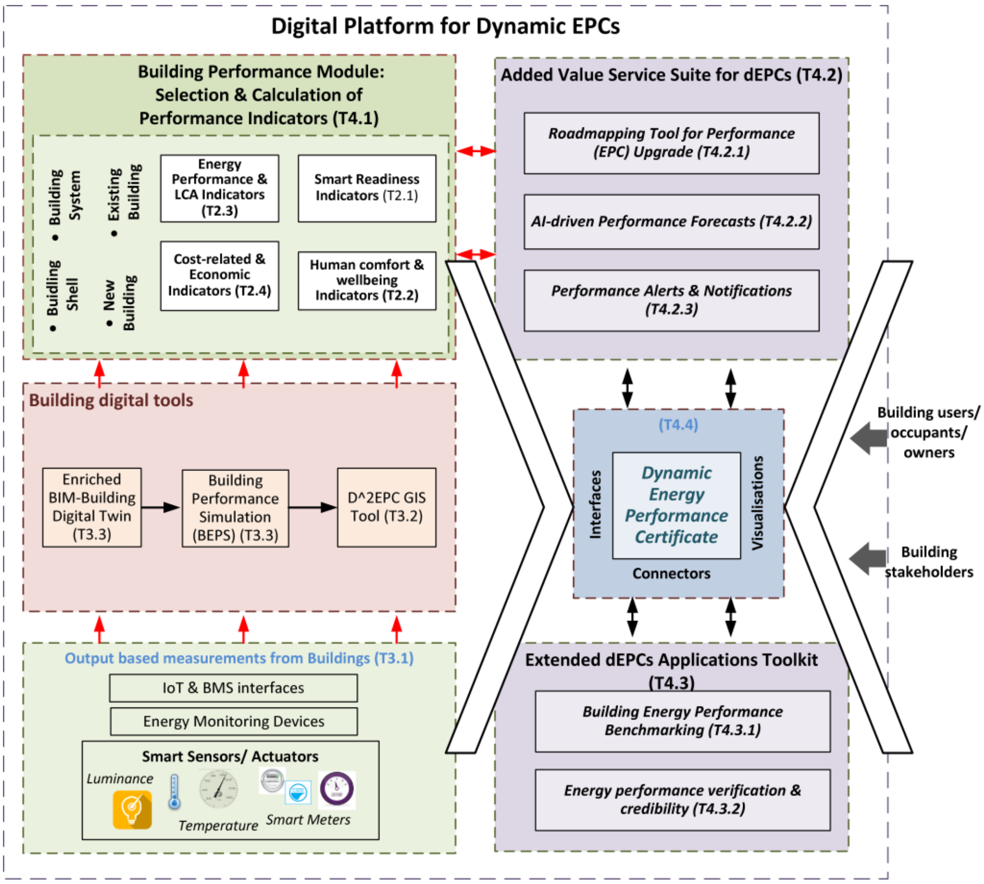

Figure 1. The Extended dEPC Applications Toolkit within D^2EPC [1]

2.2 Literature Research

The design process of the modules residing in the Extended dEPCs Applications Toolkit has been based on a wide desk of research performed on data quality and network monitoring tools, as well as classification and benchmarking processes. The results of the research are presented in the next sub-sections.

2.2.1 Research in Energy Performance Benchmarking

2.2.1.1 Current Scene in Building Benchmarking

Benchmarking is the enduring comparison of processes, to close the gap to the best-performing entity systematically [2]. So, the task is to define “best performing” and quantify a value to be compared to. One possibility is to work with quantiles and compare the existing values with the 1% or 10% quantile. However, it can also be decided to compare the energy demand of a building complying with the latest building code and to the average [3].

There is criticism about the efficiency of asset-based EPC, pointing to prebound and rebound effects. It was found that ‘there is very little difference in actual average consumption for households across the EPC spectrum’ [4]. Based on the results iof ‘The Energy Performance Gap (EPG)’ [4], the Energy Performance Gap (EPG) - i.e., the deviation between the predicted energy consumption during the design stage of a building and the measured consumption during the actual operation- was inversely correlated with energy efficiency. More specifically, higher energy use (or higher EPG) was associated with energy-efficient buildings, while inefficient buildings presented lower energy use (lower EPG). The main factor that causes these discrepancies has been identified as the occupant’s behaviour which plays a determining role in the building’s actual consumption. Many studies have identifed ‘rebound’ efects, where a lower effective price of heating encourages increased energy use by the occupants.

To eliminate this type of influence in the building classification, the operational rating should be considered, which involves actual building measurements. However, the tenants’ behaviour is not always homogeneous, provided that a multitude of factors (e.g., thermal comfort, insufficient building insulation etc.) might trigger increased energy usage. For this reason, input from more than one normalized metric should be considered (e.g., as per the SRI framework, Human Comfort etc.)

There are two options for EPC rating. One uses absolute values and the other uses those of reference buildings for defining EPC classes. Thus, there is embedded benchmarking in both approaches, while in the second one, more relevance is given and D^2EPC is centered around. Furthermore, it is possible to have a twofold approach [5, 6, 7] as shown in Figure 2.

1. Compare to the energy demand of a building complying with the latest standards.
2. Compare to the average of the existing stock.

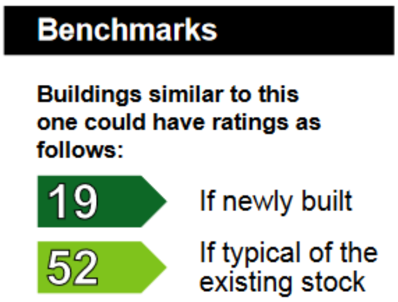

Figure 2 Benchmark as in UK’s EPC [3]

2.2.1.2 Research in Benchmarking Approaches

In contrast to conventional statistical or simulation models, data mining techniques are combined with building data characteristics to address the fundamental issues with building energy benchmarking. Sensitivity analysis is taken as a feature selection problem and building grouping is achieved via clustering. Numerous data mining techniques were implemented and conducted in four stages: data cleaning and statistical analysis, sensitivity analysis, building energy classification and model performance evaluation. To determine which building and energy system factors have the greatest impact on energy use, the feature selection algorithm RF-RFE is used and identification of comparable buildings is made using the k-prototypes clustering algorithm.

Another approach is presented in 'An Energy Performance Benchmarking of office buildings: A Data Mining Approach' [5], where data mining techniques are used to create an energy performance benchmark for office buildings. The K-prototypes method is employed to classify buildings and a wrapper model based on regression analysis for feature selection. The key concept is to group the structures containing mixed-type data (both categorical and numerical) and construct benchmarks within each group based on the relative importance of each structure, resulting in eight different kinds of energy benchmarks for office buildings, one for each cluster and validated using Adjusted R-squared.

In 'Evaluating the Energy Readiness of National Building Stocks Through Benchmarking' [6], a benchmarking approach is chosen that takes the least amount of input data. A 10-100 scale was added as a rating system directly proportional to energy efficiency. The benchmarking results are organized in a table called "benchmarking table". As a higher cut-off, they utilized an EPC value of 500 kWh/(m^2a), which corresponds to class H for all the building clusters taken into consideration. Any EPC that was higher than this one was removed from the calculations for the fit distribution because it was deemed to be an out-of-scale outlier. For fitting distributions and construction of the benchmarking tables, R programming language was used, which is not demanding in computational resources and is simple to implement in spreadsheet-like tools as well.

Finally, in 'Development of whole-building energy performance models as benchmarks for retrofit projects' [7] a systematic development process of whole-building energy models as performance benchmarks for retrofit projects is presented. Regression methods of energy benchmarking used rely on statistical models developed by using an existing data set (from utility data and on-site measurements), to find potential links between a dependent variable (e.g., utility bill) and some independent variables (e.g., weather data, occupancy patterns, operational schedules). The second goal of this approach is to characterize dominant factors or significant determinants of energy use (via sensitivity analysis). The use of this model explained the correlations between climate-adjusted energy use intensities and a few explanatory building characteristics such as building age, occupancy patterns and schedules, HVAC (heating, ventilation, air conditioning) type, and lighting equipment and controls.

2.2.2 Research in Network Monitoring Tools

Generally, the term network monitoring corresponds to a system that includes the appropriate software (and hardware) components that allow for tracking a variety of aspects related to the network and its operation. Such aspects can be the traffic, utilisation of bandwidth, or the status of several devices or sub-systems that constitute the overall network [8]. A network monitoring tool enables the timely detection of connection failures/disruptions, device malfunctions, or any other type of issue that widely occurs in data flow and might lead to system delays or even shutdowns.

In today's IT systems, the utilization of a network monitoring tool offers a variety of benefits, guaranteeing a smooth, accurate and faultless information exchange among software components [9]. More specifically:

- A monitoring tool enables the network administration to forestall the probable outages that may happen in a high-complexity network. It provides the necessary visibility for the early locating of an outage, prior to the generation of any bottlenecks to the system.
- In cases when the system is down, network monitoring significantly facilitates problem-solving as it provides invaluable information all the way to the bottom of the issue. Live network maps and performance metrics are utilized to save as much time as possible for successful network crisis management.
- A well-organised network monitoring yields notable added-value for the IT personnel and the enterprise as a whole. Immediate access to the source of the issue saves significant amounts of troubleshooting and, as a result, this staff time is freed up for the completion of day-to-day tasks and projects or further business development. Ultimately, this time-saving is translated to a speedy return-of-investment of the purchased tool.
- In the age of information, an increasing number of devices (e.g., IoT) are being deployed to all types of commercial, industrial and residential buildings. Moreover, modern services require internet-based communication to ensure remote access. Consequently, the IT environments are growing in size and the networks are gradually becoming more complex. Reliable monitoring tools are considered mandatory, as the systems are rendered susceptible to performance fluctuations from increased complexity and internet dependency.
- The network monitoring tools can also act as a stepstone for the first level of system security as they provide insight into its normal operation. In cases when all devices are up and running and no spikes in traffic levels are recorded, a baseline behavior can be inferred for the standard performance of the network. Any deviation from the baseline operation could be an indication of a security threat.
- The majority of network monitoring tools grant access to historical data (e.g., log files). Those can heavily contribute to the assessment of the installed equipment regarding its performance. As an example, trends analysis can be performed in order to ensure the current infrastructure meets the business needs or, alternatively, updates should be introduced in the system.

In IoT systems that include massive information flow among a large number of deployed devices and cloud infrastructure, the integration of a network monitoring concept is extremely valuable. An IoT system comprises various appliances utilized in both industries or residences. These appliances can be metering, sensing or actuating instruments that do not differ in essence from a conventional computer. To get maximum value out of them, they need to operate seamlessly at peak performance for the vast majority of the time. Based on the IoT system's nature, it is quite unrealistic to allocate human resources for constant checking on the operation of numerous devices. An automated monitoring tool can undertake this procedure and alert the involved users when a malfunction or communication disruption occurs within the network. From then on, certified personnel can act upon the issue via remote access.

Nowadays, there is a plethora of tools tailored to monitor IoT networks. Based on the findings on the Best IoT Device Monitoring Tools from Comparitech [10] tools such as Domotz[1], splunk[2], Datadog IoT Monitoring[3], senseye[4], skyspark, [5]TeamViewer IoT[6], AWS IoT Device Management are highly recommended due to the user-friendliness and functionalities they offer. However, these services are not open-source and expect a paid membership to activate (apart from a short free trial period).

[1] https://www.domotz.com/?utm_source=ctech&utm_medium=review&utm_campaign=iot-device-monitoring
[2] https://www.splunk.com/en_us/data-insider/iot-monitoring.html?301=/en_us/iot.html
[3] https://www.datadoghq.com/dg/monitor/iot/
[4] https://www.senseye.io/
[5] https://skyfoundry.com/
[6] https://www.teamviewer.com/en/iot/

2.2.3 Research in Data Quality

The term data quality has numerous definitions. In general, quality data are considered those that can be utilised for various intended uses (operation, decisions, planning) provided that they correctly represent the real world in regards to the physical quantity they measure.
According to ISO 8000-8 [11], data quality is classified into three fundamental categories. The **syntactic** quality, the **semantic** quality and the **pragmatic** quality. The first category concerns the level of data conformity to the specified syntax. The second one describes how well data are matched to the content they represent. Lastly, the third category refers to the extent to which the recorded data are fitting, relevant and valuable for the objective they were initially recorded for. In Figure 3[7], the three previously mentioned categories are presented in tabular form, along with the main goal, properties and measures associated with each category.

Semiotic Level	Goal	Properties	Measures
Syntactic Data Quality	Consistency	Syntax, rules, symbols	e.g. % of inconsistencies in data
Semantic Data Quality	Comprehensiveness and accuracy	Completeness, meaningfulness, correctness, unambiguous	e.g.% of errors in data sample
Pragmatic Data Quality	Usability and usefulness	Timeliness, conciseness, accessibility, reputability, understood	e.g. time of update, user surveys, influence on decision-making and outcomes

Figure 3. Fundamental categories of data quality

ISO 25012 [12] delivers a general data quality model applicable to any information system which stores data in a structured way. The standard introduces several dimensions describing the overall quality of a dataset segmented into two main classes, inherent and system-dependent data quality. Inherent data quality concerns the intrinsic potential of data to meet specified requirements and is expressed via the said dimensions. More specifically:

- **Accuracy** which concerns whether the data values recorded for an object or a metric are correct
- **Completeness** which addresses the level of comprehensiveness of a dataset
- **Consistency** which examines whether the same data kept at different places within a system are identical
- **Credibility** (or Validity) which covers the level of trustfulness of a dataset
- **Timeliness** which corresponds to the speed of data dissemination

Figure 4 sums up the aforementioned dimensions and presents indicative examples per case.

[7] http://wiki.doing-projects.org/index.php/Data_Quality_Management

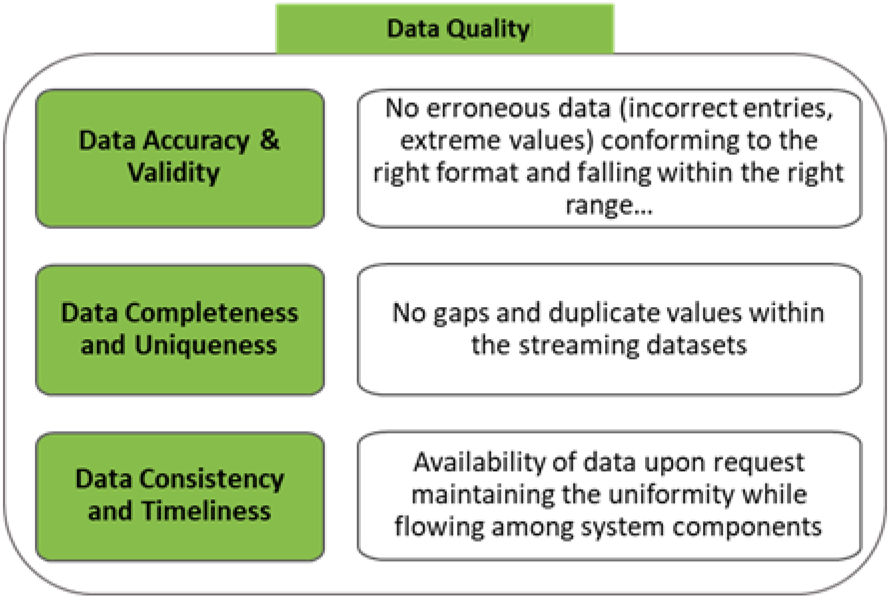

Figure 4. Inherent Data Quality Dimensions

The other data quality category mentioned in ISO 25012 is system-dependent. It is related to the level of achieved quality through a computer system under certain conditions. The respective dimensions are:

- **Availability** which is related to sufficient storage for the extracted data and accessibility by authorized users.
- **Portability** which is the ability of data to be recorded, updated, or removed from one system to another
- **Recoverability** which regards a system's ability to restore data that have been lost, accidentally deleted, corrupted, or made inaccessible

In the present day, there are numerous propriety and open-source data quality tools available. The main purpose of such tools is to contribute to the overall data management of an enterprise by delivering specific functionalities. Some significant functionalities are mentioned below:

- Perform checks in the data flows to identify and remove errors, typos, or redundancies in an automated way
- Apply cleansing algorithms for the detection and elimination of any extreme values present in the datasets (i.e., outliers)
- Maintain logfiles that include details on identified errors (and the corresponding timestamps) within the streaming datasets
- Calculate specialized performance indicators and allow for the monitoring of the systems' progression in regards to the quality of the recorded data
- Provide visualizations to the end-user towards facilitating the data management process.

2.2.3.1 Data Quality in IoT

Regarding the Internet of Things, IoT systems are typically prone to low-quality data due to various causes. Indicatively, the multiple data sources, vast data volume, data type variations (structured, unstructured, high-dimensional etc.), and the growing scalability of IoT systems significantly increase the probability of recording bad-quality data.

In 'Data Quality Management in the Internet of Things' [13], a three-layered structure is utilised to showcase the DQ problems and challenges in IoT. In a high-level representation, a conventional IoT system comprises the device layer, the processing layer and the application layer. The device layer includes the metering, sensing and actuating IoT devices deployed in the physical world (buildings). The network layer is responsible for the acquisition of IoT data with the utilisation of several specialised protocols, allowing for interoperable communication. Finally, the application layer contains several applications that undertake the processing, analysis, storage and streaming of the collected datasets to other software components that deliver services to the end-users. Based on the feedback of many researchers dealing with IoT data quality, bad data problems can be generated in every abovementioned layer. Faulty installations, incorrect placements, vandalisms (device layer), constraints in network resources, connectivity issues (network layer) and data processing (application layer) can be sources of invalid, incomplete or inconsistent data.

The most common data errors occurring in sensing/metering devices correspond to anomalies, missing values, constant values, stuck-at-zeros, noise and drifts. Typical anomalies are the extreme values within a dataset (outliers) that normally stem from temporary device malfunctions. Meanwhile, constant values and stuck-at-zeros are commonly correlated with malfunctioning or disconnected devices. Missing and noise values can be generated by unstable wireless networks, power failures or environmental interference (blockages, walls, weather etc.). Lastly, the drifts are readings that deviate from the true value over time due to the degradation of sensing material which is an irreversible chemical reaction.

Figure 5 presents a series of affecting factors, indicative examples and error types occurring in each layer of an IoT structure.

Layer	Affecting Factors	Examples	Error Types
Perception layer	• Sensors • Environment • Security • Privacy • Network	• Battery problems • Precision limitation • Mechanical failures • Bad weather • Device upgrades • Unstable network • Non-encrypted	• Missing value [66] • Incorrect value
Network layer	• Network • Environment • Security • Privacy	• Unstable network • Bad weather • Security attacks	• Missing value • Incorrect value
Application layer	• Streaming processing • Security • Privacy	• Manually errors • Obsolete schema definition • Streaming operators	• Wrong schema definition • Misplaced value • Broken join relationship • Misplaced column values • Missing record

Figure 5. Data Quality Threats per layer

2.2.3.2 Data Quality in EPC Issuance

As dictated by metrology (as mentioned in 'Basic Principles of Engineering Metrology' [14] the science that studies measurement), a meaningful measurement needs to be extracted based on common standards that deliver the approved methodology to be followed for different setups. In Europe, the measuring instruments directive MID defines the regulative framework. Within the directive 2004/22/EG for metering energy flows MI-002, M-003 and MI-004 are relevant.

For the issuance of an operational EPC, beyond the monitoring of the data flows and credibility of measured physical entities, there is a need for validating the data in terms of EPC input. Provided that a multitude of different metering devices are engaged in the EPC calculations, the acquired data are prone to generating meaningless or erroneous results.

The errors from metering setups can be segmented into avoidable, systematic and random. Avoidable errors are those related to the time stamp of the measurement, encoding analog information, interpreting strings with floating point data or any other error generally occurring when transmitting data electronically. Systematic errors are the result of an improperly calibrated apparatus that introduces the same one-directional bias into all of the measurements. Lastly, random errors are difficult to detect and lay in the working principle of the meter, possibly exposed to random environmental influences.

The D^2EPC-relevant metering setups are the heat, gas and electricity meters utilised for consumption measuring and billing. Regarding heat meters, the calculation of the resulting energy involves temperature and mass flow measurements. The respective measurement errors might originate from a detached or wrongly-positioned sensor, signal digitization or calibration. For the gas meter, only the volume flow is measured without any measurement-based corrections for temperature or pressure. Furthermore, another influencing factor is the gas quality, which also affects the correctness of the result if given as energy. Lastly, for the electric meters, it is important having enough measurements per small intervals, if current and voltage are out of phase.

Generally speaking, the following problems may occur when meter data is taken on a monthly basis:

- The meter may have low power (smart gas meter, heat meter), so the measurements are inaccurate
- The communication channel might have problems
- The meter might have been exchanged (and the old value before exchange and values of the new meter before installation have not been recorded properly)

2.3 D^2EPC Energy Performance Benchmarking Module

Within T4.3.1, a classification mechanism is delivered to benchmark the buildings under study based on configuration data and KPI results. This mechanism steps on data-mining techniques to enable the categorization of buildings, taking into account metrics linked to the building's operation and human-centric features. The main purpose of the module is to act as a basis for informing future dynamic EPC users whether they meet the set performance criteria or not and which paths should be followed for performance improvements.

To classify buildings, a proper dataset was constructed in order to test the most suitable clustering algorithms. The dataset consists of building metadata, including primary space use labels and the other essential building characteristics as described in section 2.3.1 and data related to the SRI framework, Operational Rating Class and contributing factors and Asset Rating Class.

To continue with, clustering results are analyzed, and a general report is composed after post-processing for each one of the clusters. Apart from computing building categories, the classification mechanism is utilized for functionality conducted by the other benchmarking sub-components as well.

2.3.1 Benchmarking Dataset Creation

In the pursuit of crafting a robust Benchmarking tool, the fundamental prerequisite demanded the assembly of a comprehensive and data-rich dataset. In this regard, an exploration of the foundational D^EPC backbone tools was undertaken, encompassing the Smart Readiness Indicators framework, the Operational Rating and the Asset rating. This exploration resulted in the strategic data selection to staff the dataset. By adding a broad spectrum of pertinent building characteristics, the dataset came into its final form and includes 1000 theoretical building cases apart from the D^2EPC pilot sites. All data were organized in distinct columns and each row captures and details essential facets of a building case. The production of useful energy performance insights and statistics could contribute to decision-making processes regarding building structural retrofitting and building performance enhancement. The following table presents the building characteristics and energy performance indicators that synthesize the dataset:

DATASET COLUMNS		
Heating(SRI Domain)	SRI Score	Total GWP(kg CO_2/m^2)
Cooling(SRI Domain)	SRI EPC Class	Total GWP category
Ventilation(SRI Domain)	Asset Rating EPC Class	Country
Lighting(SRI Domain)	Operational Rating EPC Class	European Region
DHW(SRI Domain)	Cooling(Operational Rating Annual Primary Energy)	Building Category
Electricity(SRI Domain)	Heating(Operational Rating Annual Primary Energy)	Primary usage
Dynamic Envelope(SRI Domain)	Electrical Appliances & Lighting(Operational Rating Annual Primary Energy)	Construction Year
Monitoring & Control(SRI Domain)	On-site RES(Operational Rating Annual Primary Energy)	Construction Decade
Electric Vehicle Charging(SRI Domain)	On-site RES Label	Building Area
Building Area Label	-	-

2.3.2 Benchmarking Cases

As previously stated, the Energy Performance Benchmarking (EPB) module is structured around two distinct sub-components: the building classification component and the fact that facilitates the extraction of statistics, coupled with a comprehensive comparison mechanism for buildings, encompassing specific metrics and options. Consequently, the EPB cases are discretely categorized according to their respective functionalities.

This partitioning encapsulates the modular nature of the EPB system, portraying its multi-function nature. The building classification component serves as the foundation for categorization, while the other sub-component empowers users with the capability to extract insightful statistics and engage comparisons. By embracing both of the aforementioned components, the EPB module emerges as a tool in the pursuit of informed decision-making and robust evaluation of energy performance across a broad spectrum of building scenarios.

2.3.2.1 Classification Tool Case

The efficacy of the classification tool relies on the clustering methodology employed. Specifically, the utilization algorithm assumes a pivotal role in partitioning the diverse building cases into distinct categories. The determination of the optimal number of categories hinges upon an informed evaluation based on the silhouette score.

To facilitate the comprehension of the differences of the computed building categories and subsequently the classification of buildings, the tool generates a proper output that acts as category descriptions. This output is designed to endow the end user with valuable insights into the differentiating characteristics of each category.

Moreover, drawing upon the capabilities of the classification tool, exploiting the distinct building attributes, yields the opportunity to extract supplementary statistical insights. An attempt has been undertaken to extract valuable information related to the discrepancy in the European region and building primary usage across the computed building categories. Furthermore, an analogous approach is adopted for elucidating the distribution of building classifications concerning the different construction decades. In addition, the distributions of Operational Rating EPC, Asset Rating EPC and SRI EPC are presented.

The cases can be seen numbered below:

1. European Region discrepancy
2. Primary Usage discrepancy
3. Building arrangement per construction decade
4. Operational Rating EPC discrepancy
5. Asset Rating EPC discrepancy
6. SRI EPC discrepancy

The classification tool as it is presented in the next sections, will be utilized again in comparison cases where a building under study is compared to other buildings regarding specific metrics given secondary benchmarking options.

2.3.2.2 Building repository Overview statistics

The ensuing batch of cases furnishes an insightful panorama pertaining to Energy Performance Certificates and the comprehensive measure of total global warming potential. By harnessing the reservoir of building data encapsulated within the D^2EPC framework, comprising both the encompassing dataset and pilot sites, end users are able to gain a holistic perspective on energy performance across a spectrum of building attributes.

Augmenting the user experience in the D^2EPC platform, a selection of distinct modes is made available, effectively catering to the diverse nature of the D^2EPC framework’s EPCs and total global warming potential. Furthermore, the users have the choice to select one or more secondary benchmarking options according to their preferences. The secondary options are related to the building's primary usage and geographic parameters, including the European region and country, the epoch of construction and the expanse of the building area.

2.3.2.3 Comparison cases for a building under study

The final category of benchmarking cases centers on comprehensive comparisons encompassing specific metrics and Energy Performance Certificates for a designated building under scrutiny such that the end user is empowered with the capability to find out and evaluate the energy performance of a particular building of interest. Importantly, this evaluation extends beyond buildings solely within the same category, embracing a broader spectrum of comparisons across different categories.

The pivotal means to achieve such evaluations resides in the utilization of the classification mechanism in combination with the repository of buildings. In scenarios where the building under study is compared to buildings of the same category, a selective process is conducted and solely, buildings within that category are curated after the category of the building under study is determined via a fitting process. On the other hand, when comparisons extend across various categories, buildings are chosen based on characteristics aligned with the end user's preferences.

In both scenarios, buildings are grouped respectively and precise calculations are executed to yield the desired output. The output of comparisons comes in plot or textual form and is presented in the D^2EPC platform at the Benchmarking tab.

2.4 D^2EPC Energy Performance Verification & Credibility Module

D^2EPC project aims at delivering a new-age performance certificate focusing not only on the building's expended energy but also on other aspects related to its sustainability, indoor ambient conditions, smartness and overall operation. To achieve this, a vast amount of information needs to be collected, analysed and circulated among a multitude of software components. Therefore, every piece of data requires to be gathered and distributed, guaranteeing their syntactic, semantic and pragmatic quality provided that credible information is mandatory for EPCs. The issuance of an energy performance certificate includes contracting; thus, high-quality data empowers the mutual trust between involved parties.

The overall information flowing from the project's demonstration cases to the system's modules can be segmented into two main types of elements. The static elements represent the building's configuration (location, usage, size, typology, materials etc.), and the dynamic elements correspond to the data recorded by the deployed sensing and metering devices (i.e., energy and environmental metrics).

Within T4.3.2, a component is delivered to undertake the verification process for the entirety of dynamic data captured by the IoT infrastructure in the D^2EPC pilots and guarantee the reliability of the data collection. The Energy Performance Verification and Credibility (EPVC) module is designed to constantly monitor the status of deployed devices and check specific data quality features in an automatic and continuous manner.

EPVC is a composite component that comprises two separate sub-components: the Network Monitoring Tool and the Data Quality Tool. The former is responsible for receiving and analysing the operational status of the IoT devices installed locally at the pilot sites. The latter is responsible for verifying the qualitative and quantitative reliability of the collected data, defining their suitability to be used by other project components. Lastly, another sub-module, the Verification & Credibility UI, is delivered to notify the end-user in the D^2EPC platform. Equipment malfunctions, communication disruptions at the installed IoT devices, as well as identified problems in the collected data are presented via alerts.

2.4.1 D^2EPC Network Monitoring Tool

According to the research mentioned in 2.2.2, as well as the project requirements described in D1.7, a list of desirable characteristics (Figure 6) has been compiled indicating where the focus for the implementation of the D^2EPC Network Monitoring Tool is gathered. The most important functionalities to be integrated concern the alerting mechanism, the representation of the devices residing in the network and the scalability of the tool.

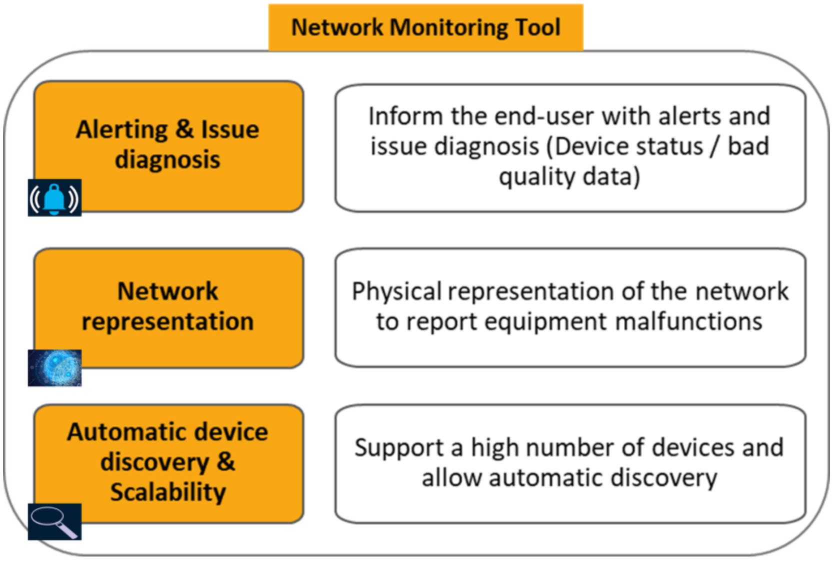

Figure 6. D^2EPC Network Monitoring Tool desired functionalities

Hypertech has already developed and integrated into its solution a web-based monitoring tool delivering a representation of the deployed devices and the IoT gateways to the user. The purpose of this tool is to provide the necessary information to the pilot partners (access is only provided to authorized persons) in regards to the issues that might occur in the deployed IoT devices. The tool is frequently updated and presents the status of the entire IoT equipment per prosumer (i.e., pilot site). Therefore, the personnel in charge per pilot is able to monitor the installed equipment, identify any connection losses and swiftly act towards restoring the communication per device. It is worth noting that Hypertech's monitoring tool only reports information about the status of the devices and not the quality of the collected data. The respective functionality is integrated into the Data Quality sub-component of the EPVC, described in the next sub-section.

In Figure 7, the main dashboard of the monitoring tool is provided. Starting from the left, a list of the available prosumers (per pilot) is presented, along with the overall status of the corresponding network and the number of devices that remain active. In the upper middle and right part of the dashboard, the percentages of active IoT gateways and devices in total are offered via pie charts. Finally, an interactive map is also delivered to the user, indicating the location per prosumer (in a GDPR-compliant manner).

Each user can further dive into details related to the status of the network of interest by choosing a distinct prosumer ID. Figure 8 presents the prosumer-specific dashboard, unveiling the functionalities of the tool. In the upper left, the user can be informed about the IoT gateway status as well as the percentage of active devices deployed in the respective pilot. Below, a status history notifies the user about the latest status updates in the network and the (indicative) time they occurred. Lastly, the dashboard offers a detailed list of the deployed devices per zone and provides information regarding their status and latest updates.

In general, three different cases can be identified:

i. The IoT gateway and the deployed IoT devices are indicated with the green color. In this case, everything is up and running.
ii. The IoT gateway is indicated with the red color. In this case, the communication is completely lost with the deployed IoT devices in the pilot and immediate actions need to be taken in order to restore the connection.
iii. The IoT gateway is indicated with green color while some IoT devices are indicated with red color. In this case, a separate issue per the OFFLINE device needs to be identified and distinct actions should be considered.

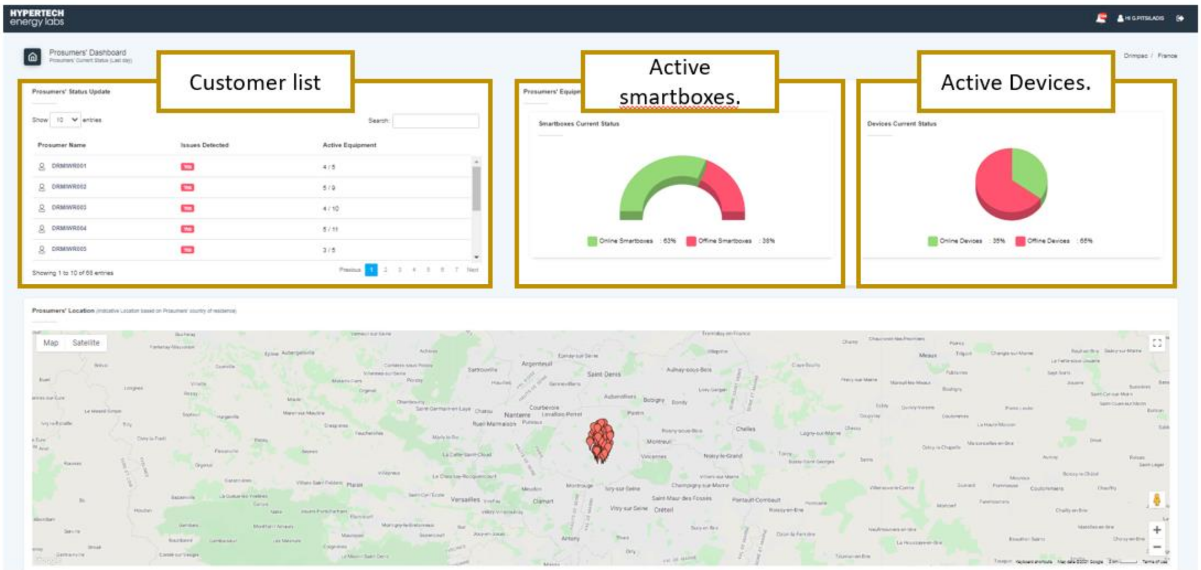

Figure 7. D^2EPC monitoring tool main dashboard

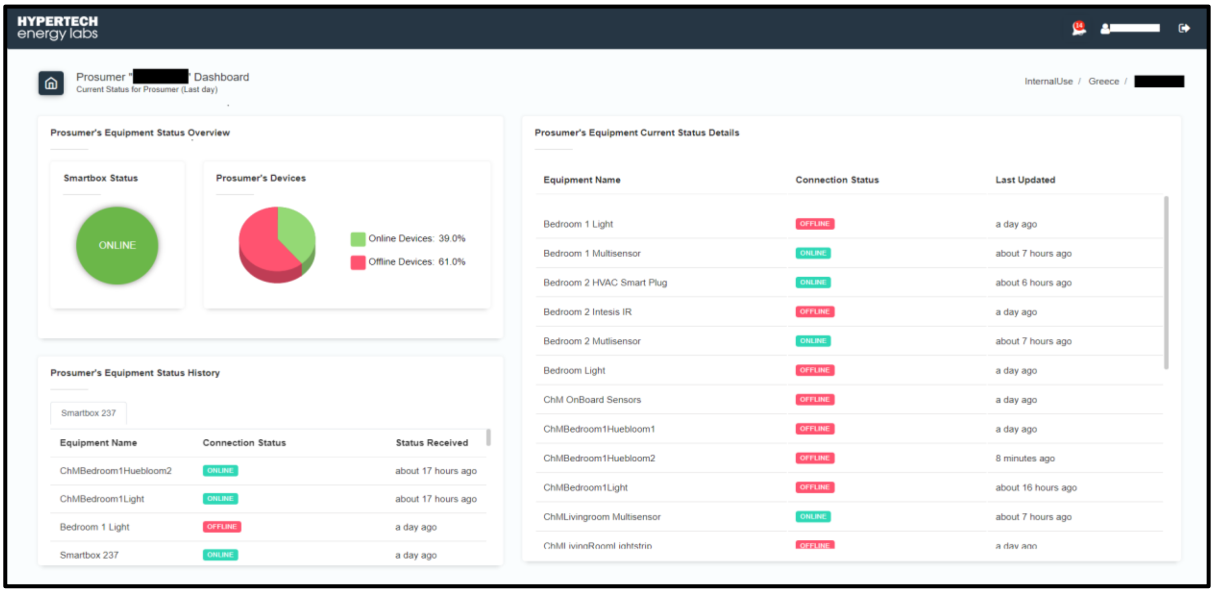

Figure 8. Prosumer/pilot site dashboard

2.4.2 D^2EPC Data Quality Tool

D^2EPC's work package 2 includes the works conducted toward the definition of the project's key performance indicators. Meanwhile, Working Group 2 has been assembled to deliver the framework for the building's operational rating. The outcome of WP2 and WG2 has led to a set of of dynamic data requirements expected to be satisfied for the overall project rollout. The main purpose is to obtain the appropriate information for the calculation of specific performance indicators related to the building's operation and indoor conditions. To satisfy the abovementioned requirements, IoT equipment has been deployed in the project's demonstration cases to deliver various measurements of energy consumption and ambient conditions metrics.

The large number of installed devices increases significantly the overall system's complexity. Each piece of deployed equipment needs to be registered along with a series of attributes regarding the corresponding building, space, and device type (static configuration). Furthermore, the measurements from IoT devices (usually more than one type per device) are recorded in a time series format in order to maintain the sequence of events (dynamic data). Therefore, a robust IoT system should undertake the streaming of the combined static and dynamic elements.

D^2EPC's Information Management Layer (IML) component steps on the well-established Hypertech's solution, which has been designed to satisfy the IoT system requirements towards a secure, speedy and scalable information flow. IML is composed of both software and hardware components. In brief, the three main IML parts correspond to: i) the Wireless Sensor Network (WSN), which includes the deployed IoT sensors and meters in the pilot buildings, ii) the IoT Gateway, which is a raspberry-based computing device that collects the data captured by the WSN and iii) the IML Cloud which gathers and processes the acquired information from the IoT Gateways and further streams it to other system components. Concerning the D^2EPC case, the majority of demonstration cases are already equipped with rich IoT infrastructure covering a large portion of the D^2EPC requirements. Consequently, the IML Cloud is adjusted to interface with the respective web platforms that collect pilot data, without the usage of IoT Gateways.

IML is based on a concrete data model specially designed to ensure the syntactic, semantic and pragmatic quality of data. Thus, it manages to extract and distribute the time series datasets in a consistent, comprehensive and timely manner. However, communication disruptions (e.g., power outages, internet disconnections) or other malfunctions are highly likely to occur. The overall quality of the collected data is heavily susceptible to any type of operation disturbances. Data gaps or extreme values render a dataset inadequate to be used as input into calculation methodologies. For this reason, Hypertech's IML solution is equipped with state-of-the-art algorithms able to detect and treat discontinuities and inaccuracies (i.e., outlier detection and treatment). In ANNEX C of D3.4 [15], the integrated algorithms are presented in detail in regards to the utilised cleansing and imputation techniques.

Under the works of D^2EPC, the IML component is enriched with a brand-new Data Quality Tool, which extends the data quality assessment. Beyond the cleansing techniques, additional validation is applied to the streaming datasets to deliver even more reliable data for the issuance of the dynamic EPC.

The tool is designed on the basis of specific applied checks (adapted to the nature of the project's metrics), which evaluate various features relevant to the type, shape and range of the collected data. More specifically:

- **expected_characteristics**: Based on the type of communication (Restfull or event-based) with each of the pilot sights, certain characteristics of the incoming dataset (index, number of columns, number of rows, temporal granularity and start-end period) are expected per data source (i.e., sensors/meters). This check evaluates whether the characteristics coincide with the anticipated ones.

- **max_data_gap**: The extracted datasets from the pilot cases are examined on a 24-hour basis. This check identifies the maximum period in the past 24 hours during which no data were recorded by the device.
- **only_positive_values**: As dictated by the nature of the D^2EPC-relevant metrics (energy, power, temperature, humidity, illuminance, air quality), no negative values should be delivered by the IoT measuring devices. A negative value indicates a device malfunction or faulty installation that needs to be addressed as soon as it is observed.
- **how_many_zeros**: This check has been integrated in order to highlight the amount of zeros existing within the dataset, as a large amount of zeros could be an indication for data inconvenience for the calculation of D^2EPC quantities.
- **is_monotonic**: Apart from the positivity of the acquired data values, there are metrics that describe cumulative quantities (e.g., the cumulative energy consumption). In this case, the values are expected to grow sequentially, meaning that the previous measurement should always be smaller than the current one.
- **Is_occupied_space**: In cases when access to occupancy data of a space is granted, several measurements can be further evaluated based on the metric type. Indicatively, the power consumption of an unoccupied space should be maintained at baseload levels, provided that no person is present to trigger any energy-consuming actions. If there is no access to occupancy data, the building's occupancy is inferred -where feasible- by the respective operation schedule based on the building's typology (commercial pilots). The '**is_occupied_space**' has been integrated into the '**Is_outside_range'** check, provided that occupancy affects the expected range of measurements.
- **Is_outside_range**: Although the integrated cleansing algorithms in the IML manage to eliminate the majority of the inaccuracies detected within the data, some values -escaping detection- can still be considered as erroneous. For example, the relative humidity has a range of values between 0% and 100%. A very high (or low) measurement of relative humidity within this range might not be treated as an outlier. Nevertheless, such humid (or dry) indoor conditions are quite unrealistic to occur in real-life scenarios. To tackle this type of drawbacks, a baseline value per metric is generated based on maximum and minimum values identified in previously accepted measurements. The determination of each baseline focuses on examining each value and its relevance with the rest of the dataset, taking into consideration seasonal recency that most likely affects the majority of ambient conditions datasets. From then on, each time a dataset is evaluated in the Data Quality Tool, the received values are compared to the respective baseline values towards reporting their validity. Other predefined checks in specific metrics have also been introduced. The range of sensor measurements per metric - extracted from the device data sheets- has been taken into consideration for the baseline. Lastly, in the CO_2 case, all assessed measurements are expected to surpass the minimum outdoor CO_2 concentration limit of 300ppm.

As described in the previous section, Hypertech's solution includes a Network Monitoring Tool that utilises the operation status of each device and delivers a physical representation of the corresponding network to the end-user. However, within D^2EPC, the datasets extracted from the already-existing IoT equipment in the pilots (via interfacing) do not include the operation status of the respective devices as an attribute. In order to derive an indication of the device status, some extra checks are integrated into the Data Quality Tool presented below:

- **is_stuck_device**: This function analyses a chunk of data and calculates the number of consecutive values. If the values are not altered for a specific amount of time, then the sensor/meter is considered stuck
- **is_dead_device**: As an analogue to the previous function, 'is_dead_device' calculates the number of consecutive missing values reported from a device and infers its status based on a predefined time interval.

2.4.3 D^2EPC Data Verification KPIs

The data verification process which takes place in the EPVC Data Quality Tool provides the necessary information for the calculation of specific Key Performance Indicators. These KPIs enable the monitoring of the IoT Networks progression and the assessment of the overall data quality with the passage of time. In Table 1, the definitions and calculation processes of the indicators that have been considered for integration in the D^2EPC data verification framework are presented in tabular form.

Table 1. D^2EPC Data Verification KPIs

Indicator	Definition	How to calculate
Percentage of Empty Values	Empty values indicate information is missing from a data set.	Count the number of data gaps within a data set and divide it by the total number of timestamps in the period of interest.
Deviation from the baseline value	Data values are compared with the respective baseline values defined from validated past data.	A baseline is generated based on previous trusted data. Then, the percentage of deviation between the actual and baseline values is calculated.
Amount of Dark Data	How much information is unusable due to data quality problems?	Look at how much of your data has data quality problems.

2.4.4 D^2EPC Credibility UI

The materialisation of the D^2EPC Credibility UI has been based both on Hypertech's Network Monitoring solution and the Data Quality tool developed to serve the project needs. Results from the checks performed on the collected datasets in fixed 24-hour intervals are expected to be reported and delivered to the end-user. To address this requirement, the Network Monitoring tool was properly modified to interface with the Data Quality tool and include the D^2EPC pilots to its scope. As the device status is not part of D^2EPC's data model, the applied data quality checks (2.4.2) are properly configured to infer the device statuses which are further communicated to the network monitoring tool. On the device-level, 'OFFLINE' is considered a sensor or meter which has not sent any values in any of its included metrics. For example, a multisensor -measuring temperature and humidity- is deemed OFFLINE if it didn't send either temperature or humidity measurements in the past 24 hours. In any other case, the device is characterised as 'ONLINE'.

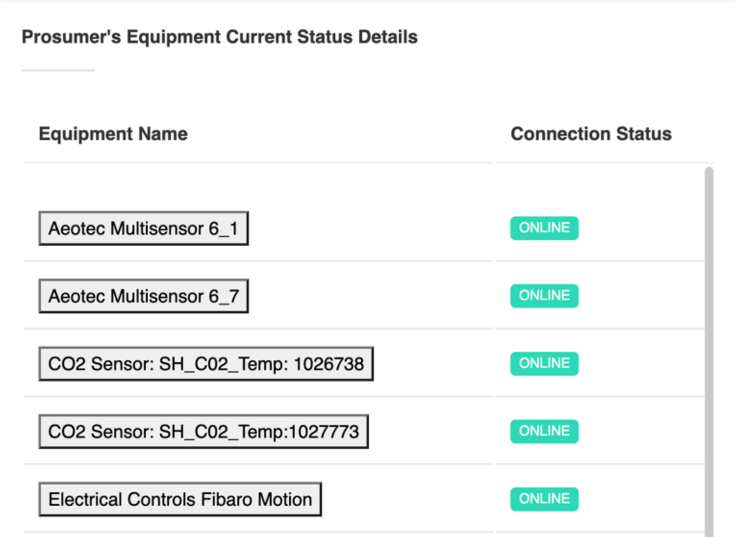

Figure 9. List of IoT devices and their status for the nZEB smarthouse

One of the main newly-added functionalities of the updated monitoring tool is the inclusion of the report generated by the Data Quality tool. End-users are able to select one of the meters/sensors from the list of available devices and receive information for all the corresponding metrics extracted from the selected device. The report includes dataset characteristics, issues and overall quality performance product of the quality checks analysed in section 2.4.2 and the calculated KPIs of section 2.4.3. Figure 10 presents the results of a data quality report extracted from a temperature-humidity sensor.

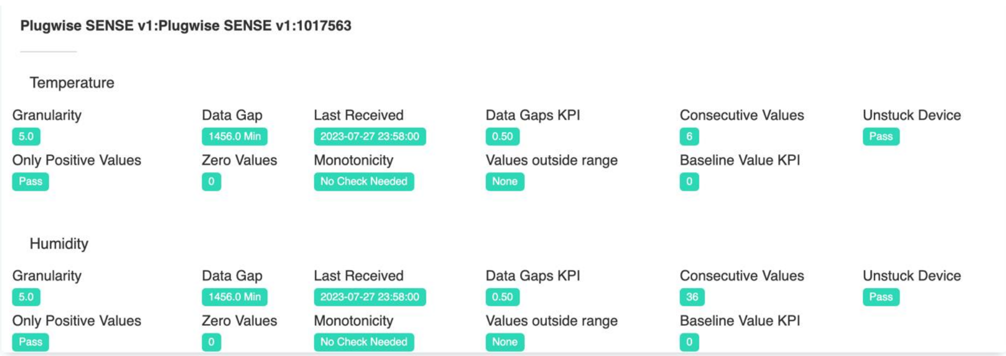

Figure 10. Data Quality report per metric in a temperature-humidity sensor

On the prosumer-level, due to the nature of the IML's data model and Hypertech's solution overall, each pilot prosumer is identified through an IoT gateway, which collects all measurements from all IoT devices and streams them to the IML cloud. When the IoT gateway has lost connection, all communications are stopped with the building's IoT network and the prosumer is characterised as 'OFFLINE'. Provided that D^2EPC pilots are already equipped with IoT infrastructures without having IoT gateways compatible with the IML, a 'virtual' IoT gateway has been attributed to each pilot partner to replicate the methodology followed in Hypertech's solution. As a result, a D^2EPC pilot 'prosumer' is characterised as 'OFFLINE' when the virtual IoT gateway has lost connection i.e., all the associated devices did not send data for the last 24 hours. Figure 11 presents the outcome of the applied modifications -from an operational and aesthetic scope- on HYP's network monitoring tool.

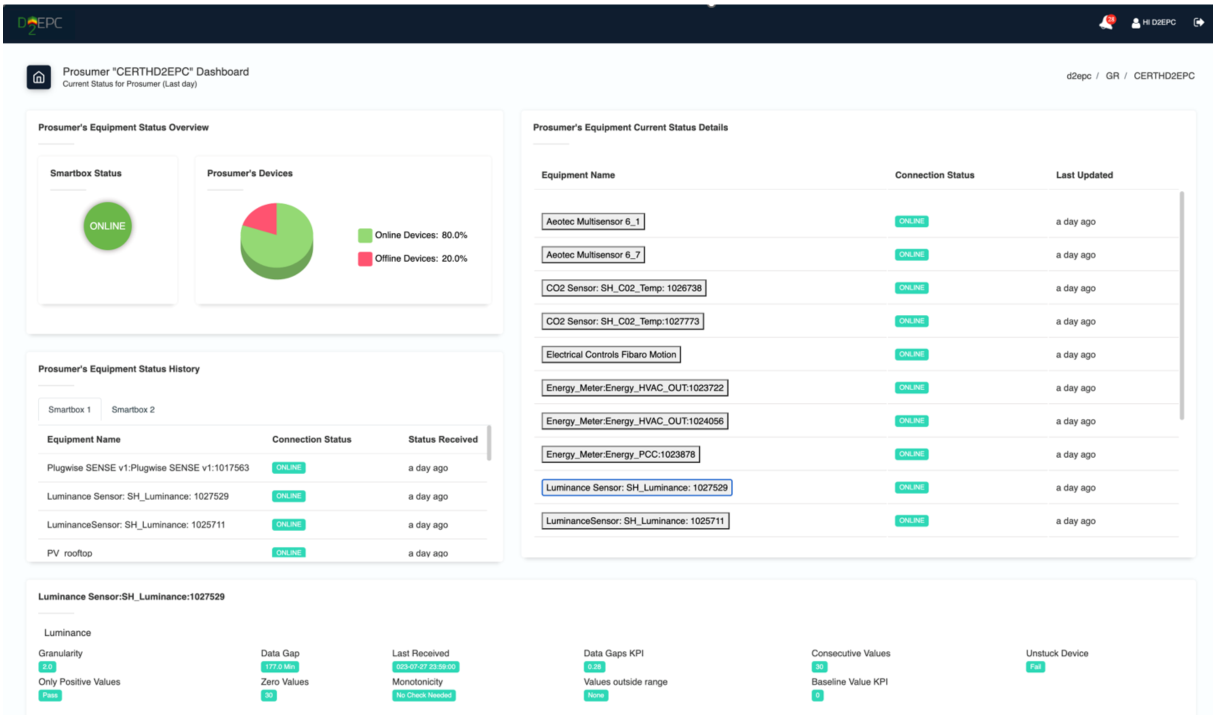

Figure 11. D^2EPC Network Monitoring Tool enhanced interface

2.4.5 Data Verification Approach in D^2EPC

The overall D^2EPC methodology on data verification corresponds to a multidimensional approach incorporating different stages of data quality assessment and involving different system components. The finalised static and dynamic data input in the D^2EPC platform undergoes multiple checks, securing the adequacy of information and validity of results in regards to the project's calculated quantities. Beyond the IML and EPVC components contributing to the dynamic data validation, other components of the D^2EPC systems architecture are also involved in the overall data verification process. Inside DT [16] resides the Input Data Validator. Its main purpose is to confirm the correctness of static elements of data about to be stored in the project's repositories. Those elements are inextricably linked with user input which is a common factor for error generation in a database. The Validator performs tailored checks on the data types and detects any missing element required as input for the calculations of either the asset or the operational-based EPC ratings. The module ultimately delivers a report containing all the fields ineligible to pass the validation and prompts the user to provide any missing piece of information directly through the project's web platform. Finally, regarding the dynamic data elements, the Input Data Validator ensures that every single measurement acquired from the IML (after the EPVC checks) is mapped to a unique provided ID that corresponds to a specific sensor or meter installed in one of the D^2EPC pilots.

3 Extended dEPCs Applications Toolkit – Technical Design and Implementation

3.1 Building Energy Performance Benchmarking - Technical Outline

In this task, a classification mechanism for buildings was developed as a sub-component of the benchmarking tool, which in combination with the sub-component responsible for statistics production and comparison making, it facilitates the monitoring of the overall energy performance of buildings under study. Furthermore, it extracts useful insights and comparison results towards contributing to users' decision-making process. The decision-making concerns any course of action in order to bring improvements to buildings and maintain or improve occupants' comfort, health and well-being while achieving essential goals concerning energy consumption. The results produced refer to normalized metrics as per the SRI framework, Operational and Asset Rating tools. [1]

3.1.1 Building Energy Performance Benchmarking within D^2EPC's System Architecture

The Building Energy Performance Benchmarking (EPB) module comprises the necessary sub-components that realise the building's classification and enable its comparison with other buildings in regards to specified metrics. The first sub-component corresponds to the Classification Tool, which includes the algorithms responsible for data mining based on various traits regarding both infrastructure and temporal characteristics. The second sub-component residing in the EPB is the benchmarking tool which steps on the defined classes (from the classification tool) to evaluate the performance of different buildings based on determined reference values.

EPB communicates with the appropriate components within D^2EPC to extract the necessary information and deliver the calculated results. More specifically, it interfaces with the Digital Twin to retrieve the building information as well as the project's common Repository to obtain KPIs, weightings and reference values. Finally, the component provides the benchmarking results to the D^2EPC Web Platform while any updated reference values are stored in the project's Calculation Engine. Figure 12 presents a functional diagram that includes the entirety of the EPB's interactions.

D^2EPC end-users (i.e., Building Services/Material Industry, Suppliers, Engineers, Building designers, Facility Managers, ESCOs) request from the Web Platform Asset Rating Data benchmarked visualization and the request is transmitted to the Building Energy Performance Benchmarking Tool. The Building Energy Performance Benchmarking tool requests operational-based energy performance data from the D^2EPC Repository and performs the correlation between the assets/systems and the energy performance. The correlation result is sent for visualisation through the Web Platform Figure 13 Building Energy Performance Benchmarking Sequence Diagram (Figure 13).

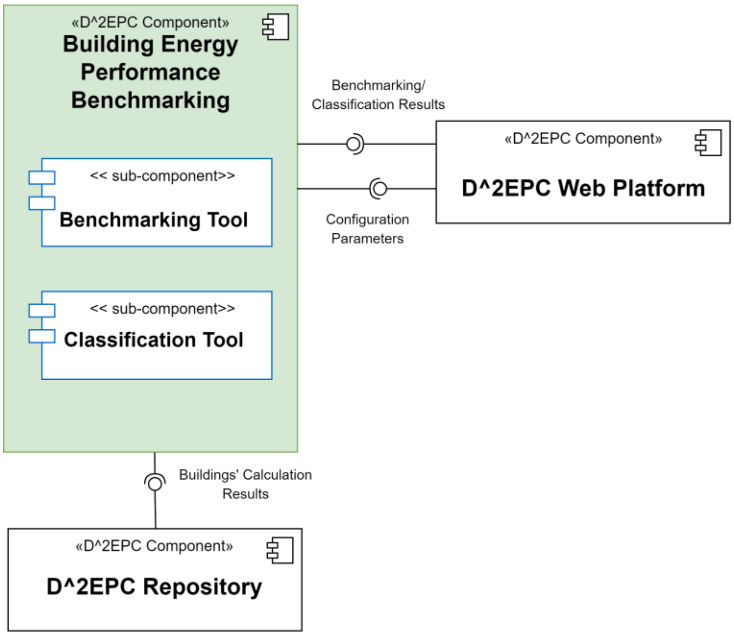

Figure 12 Building Energy Performance Benchmarking Functional Diagram [1]

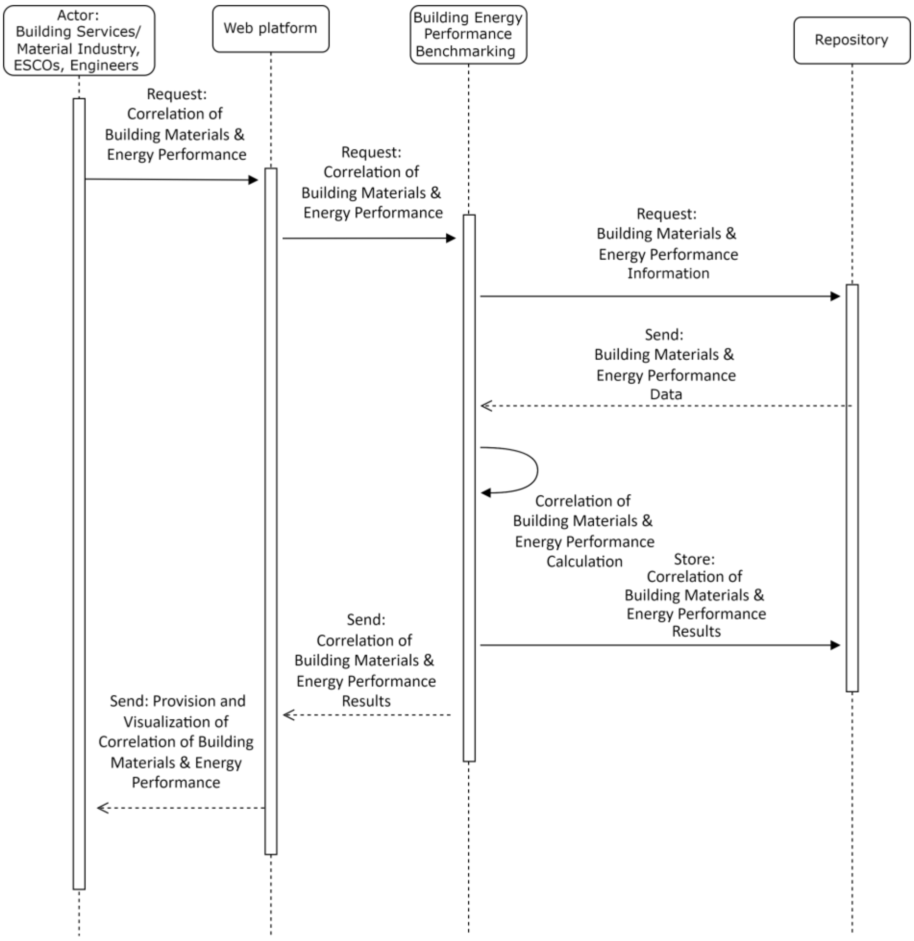

Figure 13 Building Energy Performance Benchmarking Sequence Diagram [1]

3.1.2 Building Energy Performance Benchmarking Technology stack and Implementation

The tools and modules used for the development of the Classification and Benchmarking Tool (developed from scratch within D^2EPC), along with their respective versions, are presented in Table 4.

Table 2. Benchmarking Tool Technology Stack

Tool	Version	Libraries		
		Name	Version	Usage
Python	3.10	pandas	2.0.2	Data Manipulation
		scikit-learn	1.2.2	Perform Clustering
		matplotlib	3.7.1	Results validation in graphical format
		numpy	1.25.0	Mathematical Operations
		tabulate	0.9.0	Data validation after data manipulation

The APIs described in Table 3 contain the information exchange between the Benchmarking Tool and other D^2EPC system components. The input data corresponds to the information retrieval from the project's common repository.

The output data regarding the benchmarking results streamed to the D^2EPC Web platform to be delivered to the end-users.

Table 3. Benchmarking Tool Interfaces

Data		From	Payload Format	Communication Protocol	API (method)
Input	Pilot Buildings Information and Benchmarking Options	D^2EPC repository	JSON	HTTPS	REST (POST)
	Dataset of 1000 Theoretical Building Cases	Docker filesystem	JSON	-	-

Data		To	Payload Format	Communication Protocol	API (method)
Output	Benchmarking Results	UI/UX	JSON	HTTPS	REST (POST)

3.2 Energy Performance Verification and Credibility – Technical Outline

For the materialisation of the previously mentioned EPVC subcomponents, existing software has been utilised (i.e., HYP's Network Monitoring tool) new applications have been developed from scratch (i.e., Data Quality tool), and interfaces with the appropriate data sources have been established (i.e., IML, DT, Alerting Mechanism). In the following sub-sections, the technical outline of the EPVC is described in brief.

3.2.1 Energy Performance Verification and Credibility within D^2EPC's System Architecture

EPVC communicates with the Information Management Layer, which gathers all data extracted from the D^2EPC pilots. After a series of checks performed on the datasets, EPVC streams the processed information to the D^2EPC BIM-based Digital Twin (DT), which corresponds to a virtual representation of the project's demonstration cases. If the data quality is not acceptable, an alert is generated and sent to the Performance Alerts and Notification module, which is responsible for connecting and pushing notifications to the D^2EPC platform. In parallel, EPVC forwards the credibility alerts and other related information to the Credibility UI to be presented to the end-user

In Figure 14, a functional diagram is presented, highlighting the interactions between the EPVC and other D^2EPC components as described in the project's architecture. Figure 15 provides a sequence diagram describing the operations carried out by EPVC. [1]

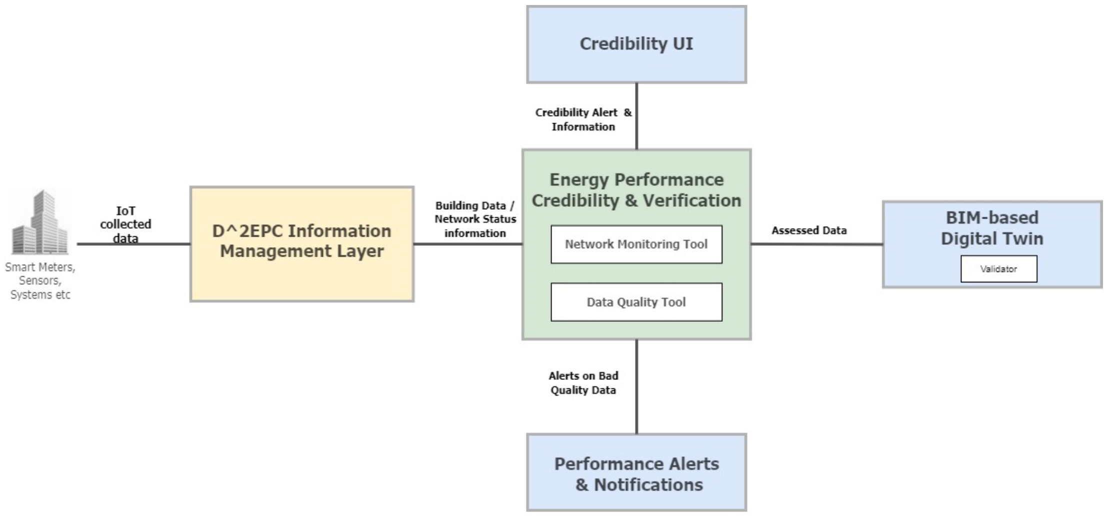

Figure 14. Energy Performance Verification & Credibility Functional Diagram.

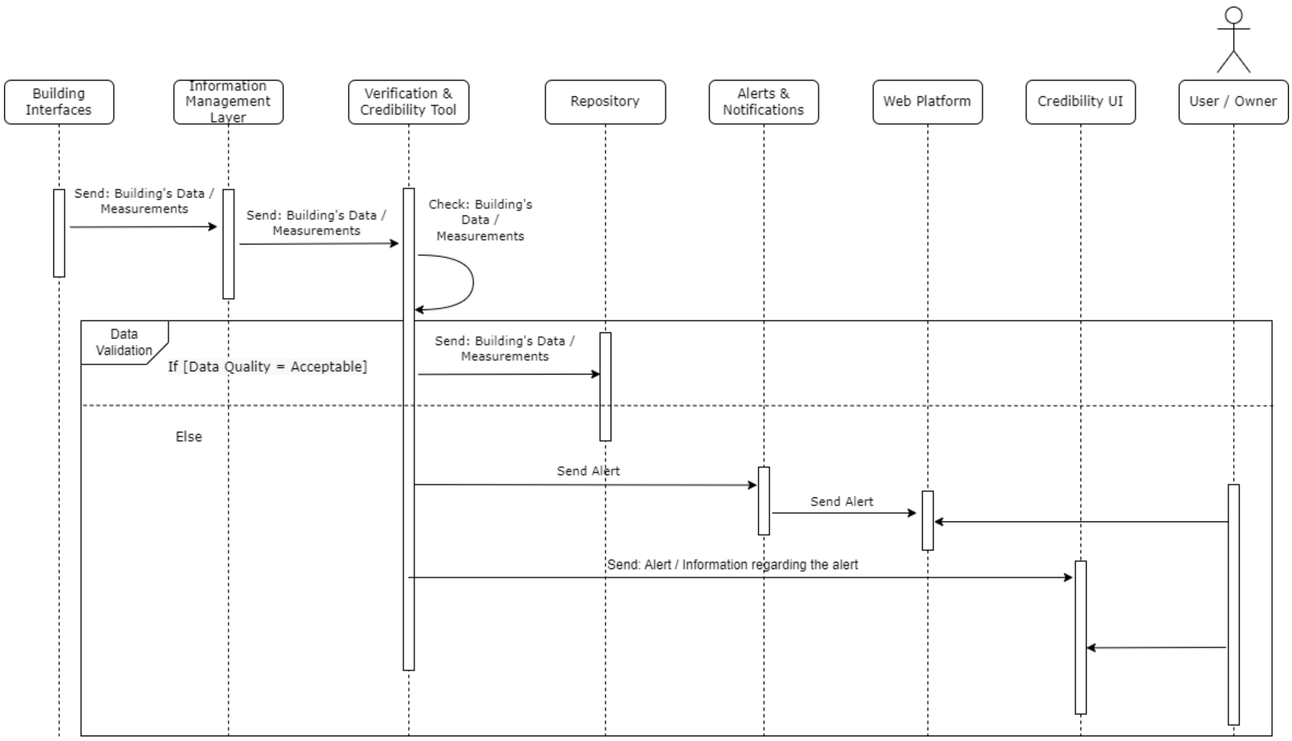

Figure 15. UC2.1 Sequence Diagram.

3.2.2 Energy Performance Verification and Credibility Technology Stack and Implementation

The tools and modules that were used for the development of the Data Quality Tool (developed from scratch within D^2EPC), along with their respective versions, are presented in Table 4.

Table 4. Data Quality Tool Technology Stack

Tool	Version	Libraries		
		Name	Version	Usage
Python	3.8	Pandas	1.4.3	Data manipulation
		Numpy	1.20.3	Mathematical operations
		Bullwark	0.6.1	Python's Data Quality Framework
		Datetime	3.2	Datetime manipulation

The APIs described in Table 5 contain the information exchange between the Data Quality Tool and other components. The input data corresponds to the information retrieval from the IML data sources.

The output data regard the information that is streamed to the Common Repository as well as the Performance Alerts & Notifications tool.

Table 5. Data Quality Tool Interfaces

Data		From	Payload Format	Communication Protocol	API (method)
Input	Energy Consumption and Ambient Condition Metrics	Information Management Layer	JSON	HTTPS	REST (POST)
Data		**To**	**Payload Format**	**Communication Protocol**	**API (method)**
Output	Energy Consumption and Ambient Condition Metrics	D^2EPC Common Repository	JSON	HTTPS	REST (POST)

	Alerts	Performance Alerts & Notifications	JSON	HTTPS	REST (POST)

4 Extended dEPCs Applications Toolkit Validation

4.1 Validation of the Performance Benchmarking

4.1.1.1 Classification Tool

As elaborated in section 2.3.3.1, the classification tool generates output to provide users with a comprehensive understanding of the distinct building categories formulated through the clustering algorithm. This output, presented in a tabular format, is prominently featured on the principal benchmarking page within the D^2EPC platform. The presentation closely resembles the illustrative figure depicted below:

	heating	cooling	ventilation	lighting	dhw	electricity	dynamic_envelope	ev_charging	monitoring_control	sri_score
count	358	358	358	358	358	358	358	358	358	358
mean	44.3067	49.7873	44.6861	41.5659	41.8337	55.2539	43.5172	55.5964	60.8734	48.6023
std	25.3408	17.0831	26.833	22.1494	29.5194	32.4935	25.2952	33.386	33.0505	9.56765
min	0	9.95817	0	0.369425	0	1.21338	1.37134	0.425856	2.75551	31.1626
25%	21.6016	36.2512	24.1729	24.8261	18.3675	15.6629	23.6414	15.673	16.0706	37.5119
50%	41.6797	50.0628	32.9617	33.1118	25.676	71.0707	32.6173	67.0543	76.8027	52.954
75%	69.1826	62.9488	77.499	64.073	77.665	82.5685	72.2889	87.1303	85.758	56.3144
max	100	97.6745	100	94.9295	95.9415	96.6359	100	99.5318	99.8048	63.3272

	building_area	op_cooling	op_heating	op_el_app_lighting	on_site_res	total_gwp(kg CO2/m^2)
count	358	358	358	358	358	358
mean	8296.02	685.385	866.232	1647.34	1638.13	2081.87
std	949.388	126.611	87.0543	223.783	207.369	298.743
min	6673	453	700	1201	1202	1448
25%	7497.25	576.25	790	1475.75	1516.25	1801
50%	8251	689.5	879	1665	1641	2128.5
75%	9116.25	790.5	938	1840.75	1806.25	2373.75
max	9992	899	999	1996	1996	2500

Figure 16. Classification Tool output for a single building category

The above table provides users with valuable insight into a building category and endows them with the capability to make comparisons with other potential building categories.

Furthermore, in order to extract supplementary insights for the buildings classified in each category, the dataset buildings are grouped as per their category, and the following output in graphical form is calculated:

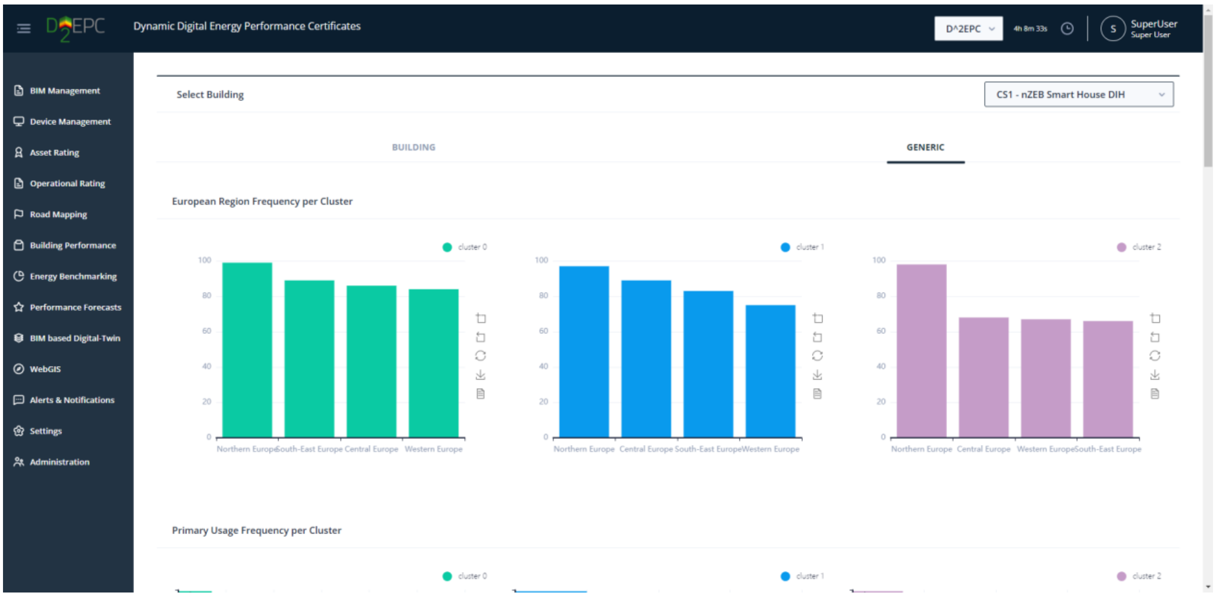

Figure 17. European Region Frequency per Cluster

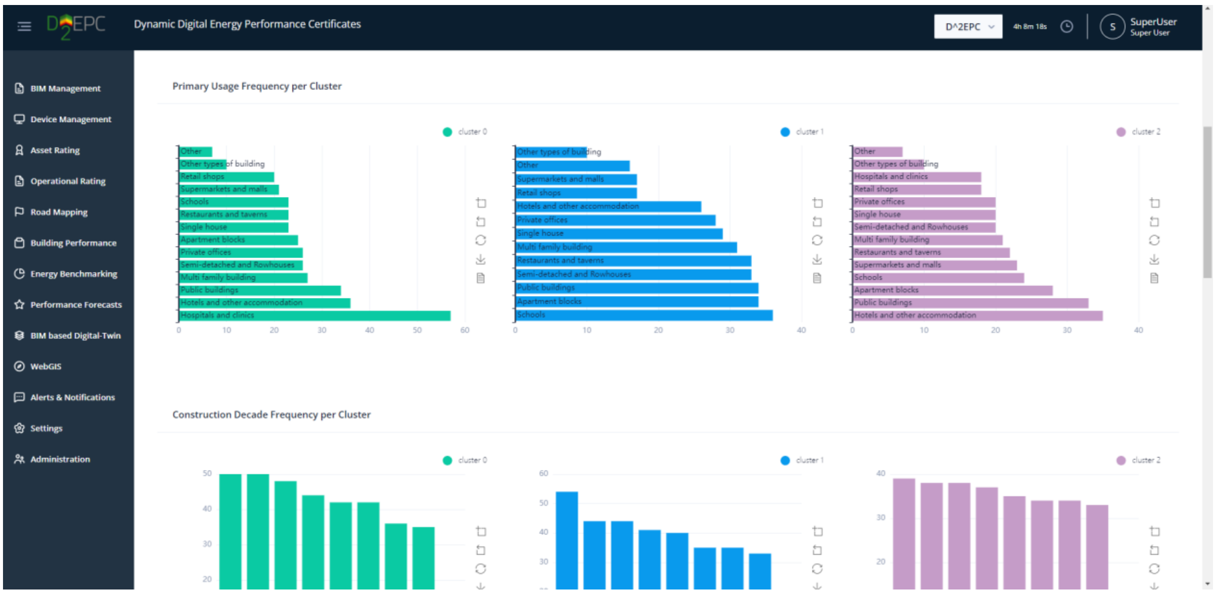

Figure 18. Primary Usage Frequency per Cluster

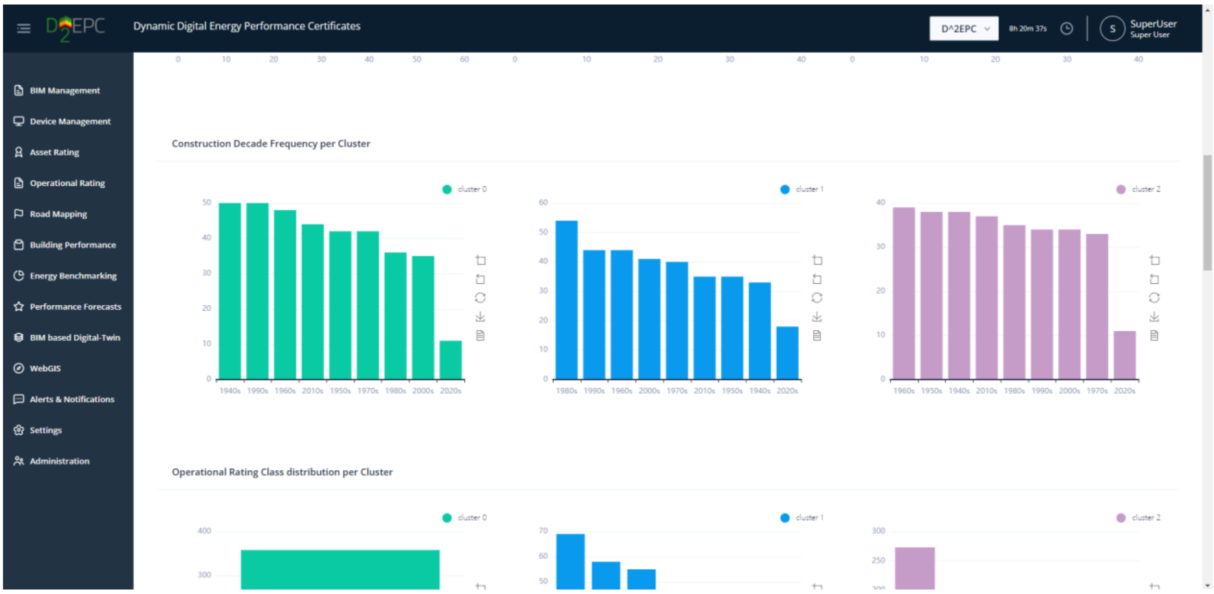

Figure 19. Construction Decade Frequency per Cluster

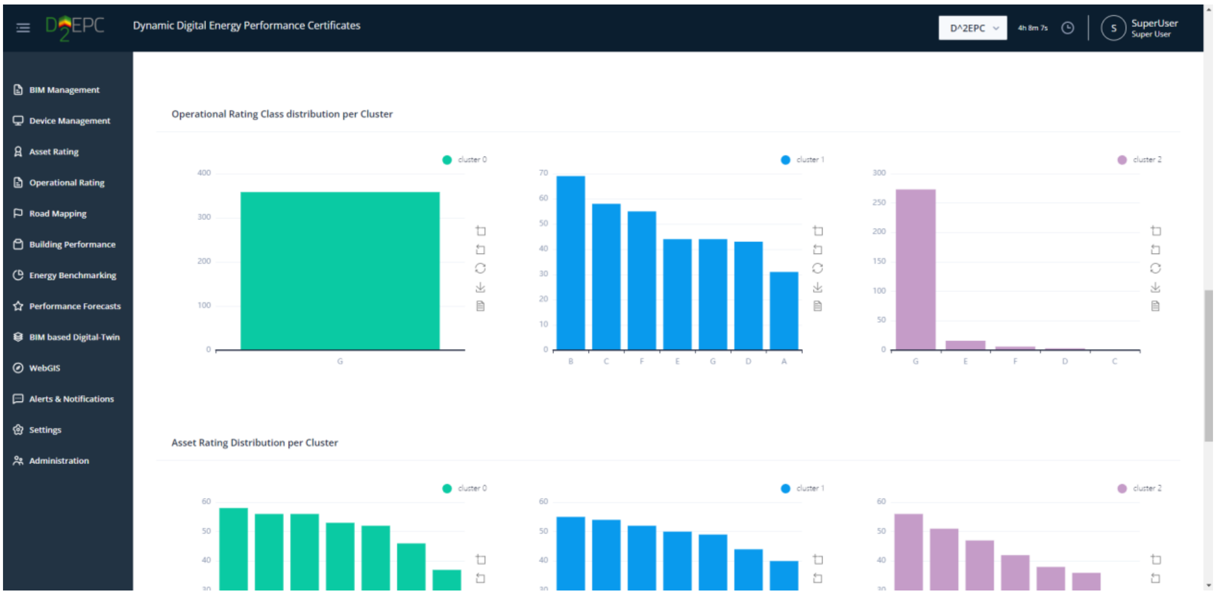

Figure 20. Operational Rating Class distribution per Cluster)

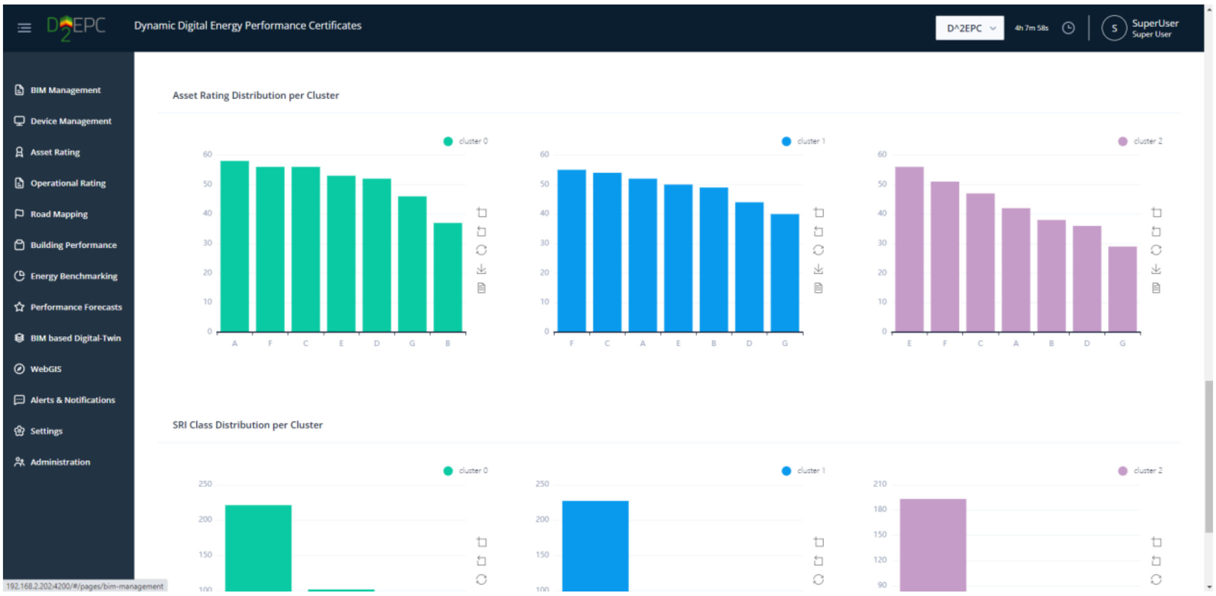

Figure 21. Asset Rating Class distribution per Cluster

Figure 22. SRI Class distribution per Cluster

The above figures provide insight regarding building characteristics combined with energy performance and allow users to understand any correlations between the different energy performance metrics and a wide spectrum of building characteristics.

4.1.1.2 Building repository overview statistics

This category of benchmarking cases provides users with an avenue to comprehensively explore the entirety of the building repository in search of energy performance insights. This dynamic tool empowers users with the capability to selectively engage secondary options, which in turn pertain to specific building characteristics. By employing these secondary options, users can readily obtain an energy performance overview encompassing Operational, Asset, and SRI ratings.

The following figure depicts the Operational, Asset and SRI Classes frequency for buildings that match 3 secondary options related to the construction decade, the European region and primary usage. The values of these three secondary parameters are as follows:

- Construction Decade: "1990s"
- European Region: "Northern Europe"
- Primary Usage: "Restaurants and Taverns"

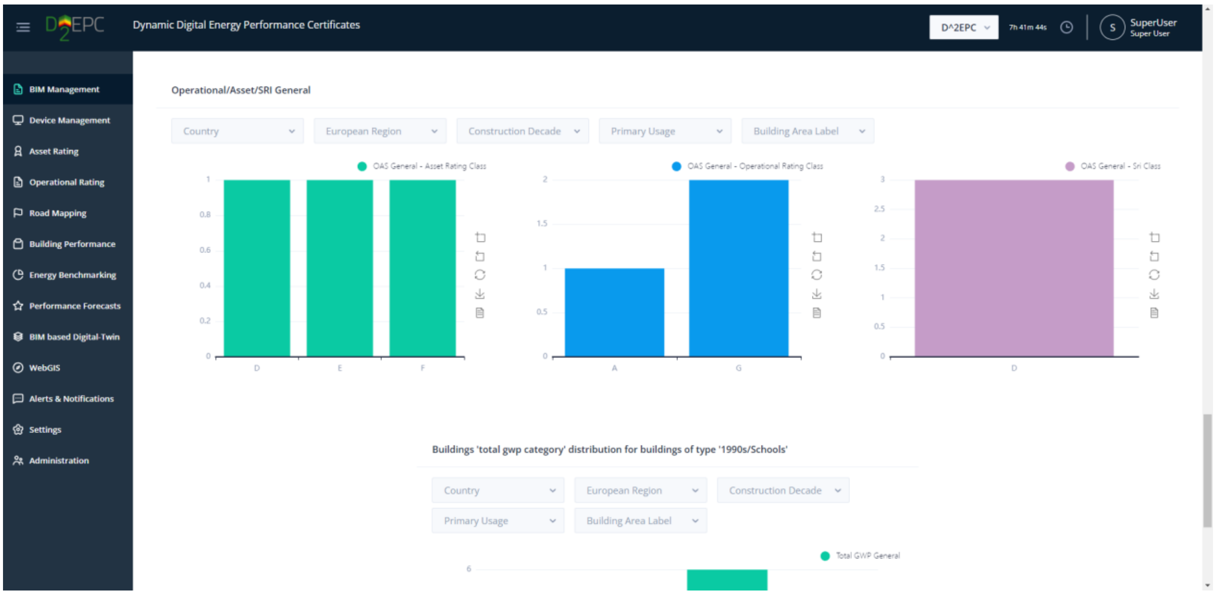

Figure 23. Operational, Asset, and SRI Class distribution for buildings matching certain user-defined criteria

As can be observed in the above Figure 23 the user may select among five secondary options. For each option, a drop-down menu is available for the user to select the desired value of each parameter.

Following the same methodology, users can similarly delve into the global warming potential metric for buildings aligning with their specified criteria. Within the dataset, each building is accompanied by a dedicated column denoting its corresponding global warming potential. In a bid to streamline the insights generation process, distinct global warming potential categories were derived, named after value ranges. The following figure depicts the output for buildings matching the criteria set below:

- Construction Decade: 1990s
- Primary Usage: "Schools"

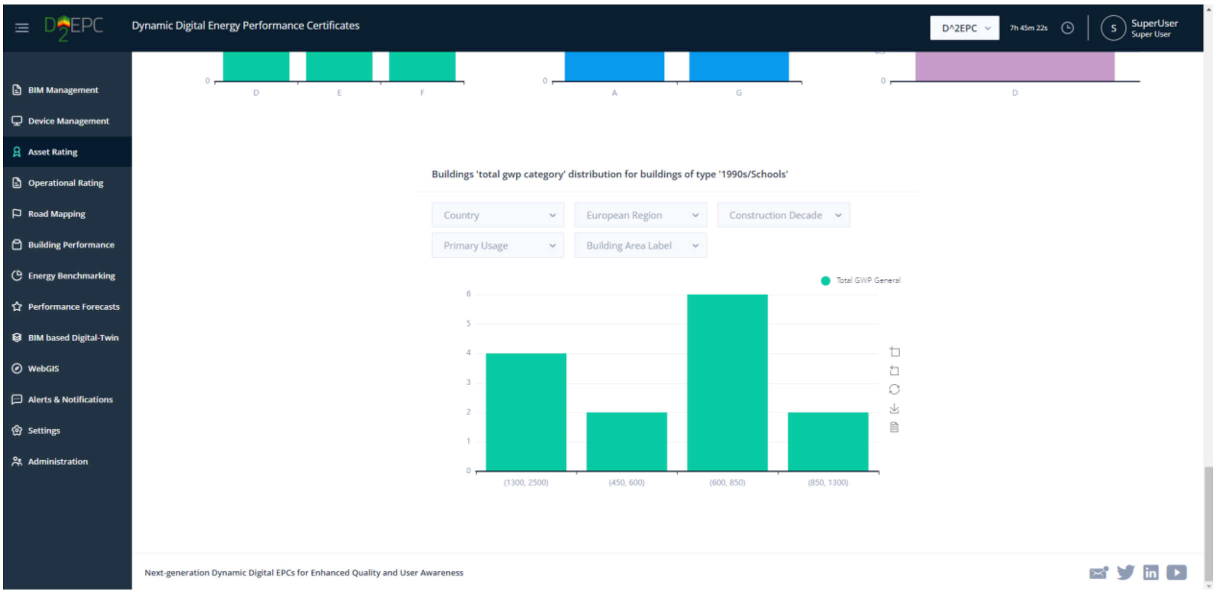

Figure 24. Global warming Category distribution for buildings matching certain user-defined criteria

4.1.1.3 Building under study-based cases

Concluding the array of benchmarking cases are those centered around a specific building undergoing analysis. These encompass seven distinct scenarios. The initial quartet of cases demands the utilization of the classification mechanism, chosen for its efficacy in grouping the building into a category that emerged from the classification process. Here, comparisons are concluded exclusively with buildings that share the same category.

The subsequent set of cases, numbering three, pivots towards a broader scope. These cases involve comparisons between the building under study and the entirety of the D^2EPC building repository. This approach facilitates a panoramic understanding of the building repository.

The results of the first four cases involving the classification mechanism are as follows:

The initial case pertains to a SRI domain-specific comparison of the building under study with the mean values associated with buildings under the same category. The second case centers around Operational Rating, focusing on the integral attributes that contribute to the computation of the Operational Rating Class. The third case is centered on the percentile comparison of the Asset Rating class and the fourth case is devoted to evaluating the global warming potential of the building.

The results of the aforementioned cases manifest in either graphical or textual formats and these results are detailed below. It is worth noting that for experimentation purposes, the results are produced with the utilization of the CERTH pilot site as a building under study.

For the first case, the results are both in graphical and textual formats as follows:

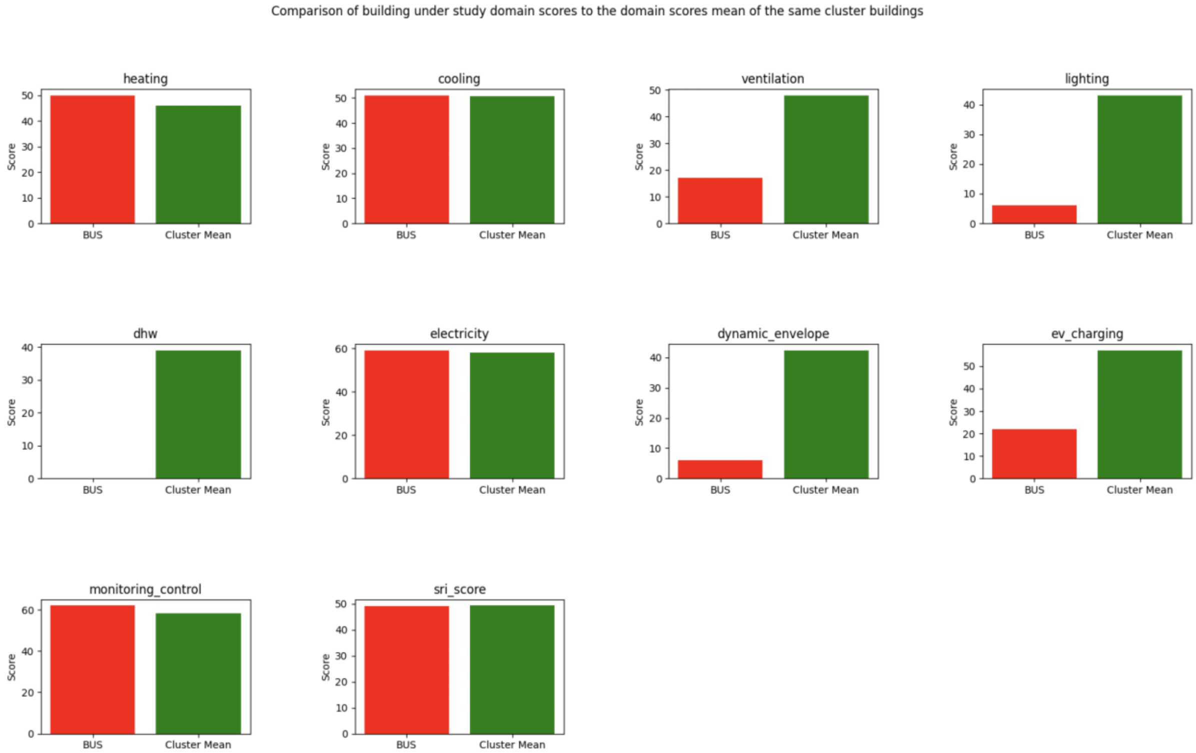

Figure 25. – SRI Domains building category based

For this case, apart from the graphical results, percentile comparisons for the SRI domain scores are presented in textual form, as can be observed below:

Table 6. SRI domain scores percentile comparisons

PERCENTILE COMPARISON RESULTS
BUS **Heating** sub-domain is greater than **62.15%** of the buildings in its assigned category
BUS **Cooling** sub-domain is greater than **53.91%** of the buildings in its assigned category
BUS **Ventilation** sub-domain is greater than **27.06%** of the buildings in its assigned category
BUS **Lighting** sub-domain is greater than **6.97%** of the buildings in its assigned category
BUS **DHW** sub-domain is greater than **0.0%** of the buildings in its assigned category
BUS **Electricity** sub-domain is greater than **47.99%** of the buildings in its assigned category
BUS **Dynamic_Envelope** sub-domain is greater than **1.05%** of the buildings in its assigned category
BUS **Ev_Charging** sub-domain is greater than **36.15%** of the building in its assigned category
BUS **Monitoring_&_Control** sub-domain is greater than **32.55%** of the building in its assigned category
BUS **SRI_Score** sub-domain is greater than **34.88%** of the building in its assigned category

The next case is Operational Rating based, and the results come in graphical and textual formats as well:

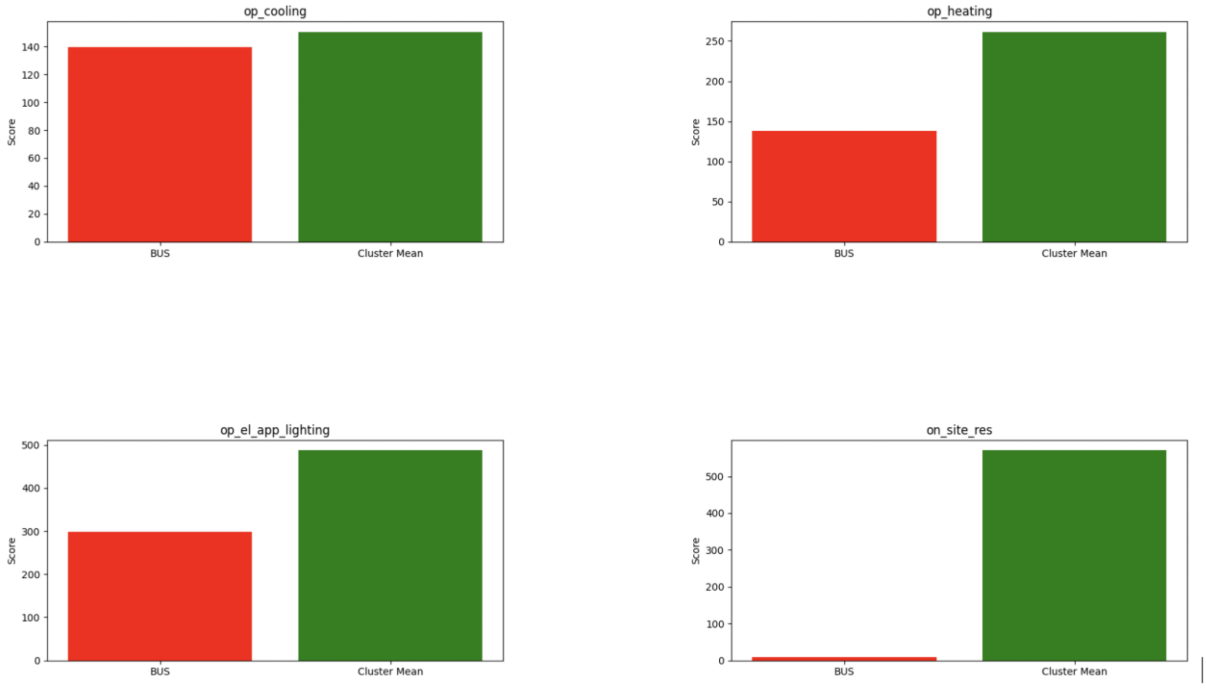

Figure 26. – Operational Rating contributing factors comparison

Table 7. Operational Rating contributing factors percentile comparisons

O.R. RELATIVE INFO PERCENTILE COMPARISON RESULTS
BUS **op_cooling** yearly primary energy is greater than **19.87%** of the buildings in its assigned category
BUS **op_heating** yearly primary energy is greater than **22.41%** of the buildings in its assigned category
BUS **op_el_app_lighting** yearly primary energy is greater than **2.53%** of the buildings in its assigned category
BUS **on_site_res(kWh/m2)** is greater than **34.03%** of the buildings in its assigned category

The next case refers to Asset Rating class percentile comparison as the results come in textual form as follows:

Table 8. Asset Rating Class percentile comparison

O.R. RELATIVE INFO PERCENTILE COMPARISON RESULTS
BUS Asset Rating Class: A
BUS **Asset Rating Class** is greater than **0.0%** and lower or equal to **100.0%** of buildings' Asset Rating Class of the same category

Finally, the last of the cases related to building category is about total global warming potential, where the building value is compared to the mean value of the buildings under the same category, and results come in graphical format. Furthermore, a percentile comparison comes in textual format as follows:

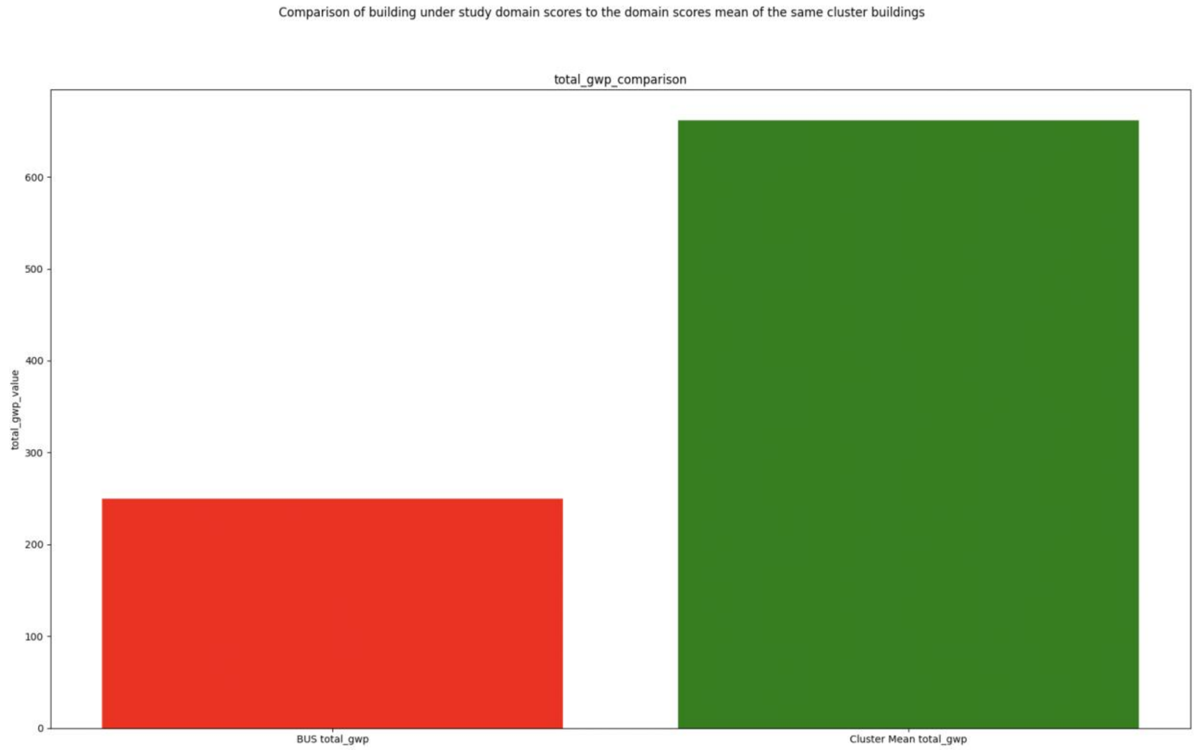

Figure 27. – Total Global Warming Potential Comparison

Table 9. Total Global Warming Percentile Comparison

TOTAL GLOBAL WARMING PERCENTILE COMPARISON
BUS Total Global Warming Potential Category (100, 250):
BUS total_gwp_category is higher in rating than 100.0% and lower or equal in rating than 0.0% of buildings under the same category.

The benchmarking cases are concluded with the last 3 cases regarding a building under study and refer to a comparison among buildings from the entire building repository.

The first case is about comparing SRI domain scores with mean values of buildings that match certain criteria. The following results are the product of a comparison among the CERTH pilot site and buildings with the same Country, Primary Usage, and Construction Decade from the whole building repository.

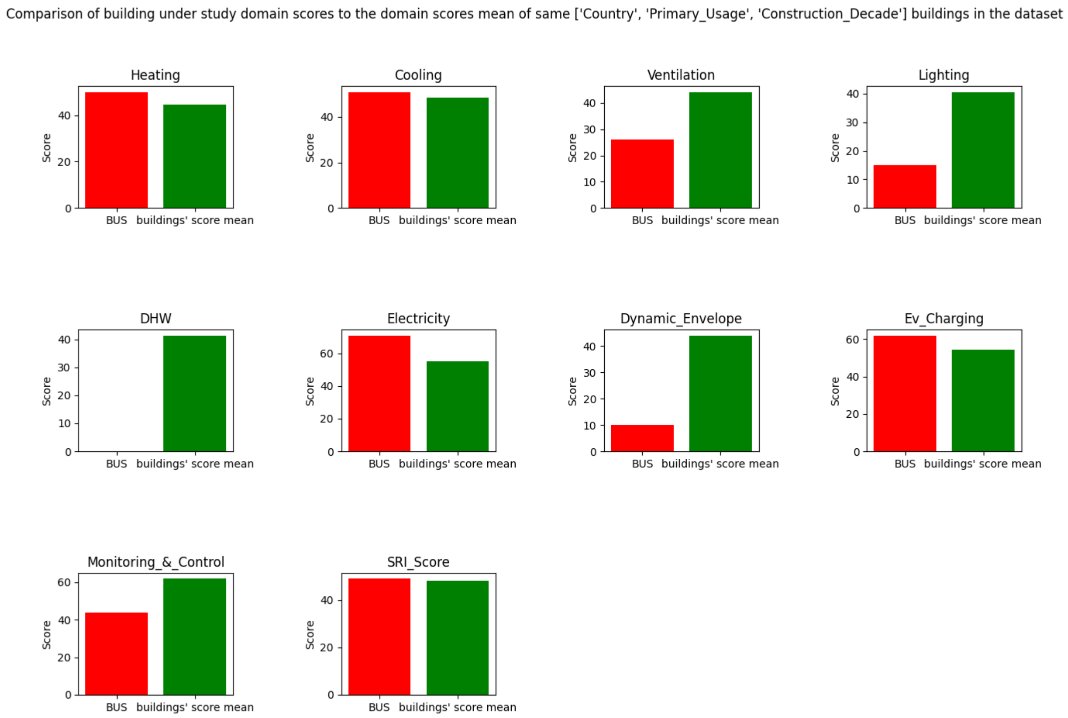

Figure 28. – SRI Domain scores comparison among building under study and buildings matching certain criteria

Second to last comes the case regarding percentile comparisons of Operational Rating Class, Asset Rating Class, and SRI Class among the CERTH pilot site and buildings with the same Country, Building Area Range, and Construction Decade. Results come in textual format as follows:

Table 10. Operation Rating, Asset Rating and SRI Classes percentile comparisons

PERCENTILE COMPARISON RESULTS
BUS **Operational Rating EPC** is greater than **89.92%** of the buildings with the same Country, Building Area Category, and Construction Decade
BUS **Asset Rating EPC** is greater than **85.27%** of the buildings with the same Country, Building Area Category, Construction Decade
BUS **SRI EPC** is greater than **9.30%** of the buildings with the same Country, Building Area Category, Construction Decade

To conclude the benchmarking cases, the last one is the comparison of the buildings under study global warming potential to buildings matching certain user-defined criteria. The following results concern the comparison among the CERTH pilot site and buildings with the same Construction Decade, Primary Usage and Building Area Range as the CERTH pilot. The results come in textual format as follows:

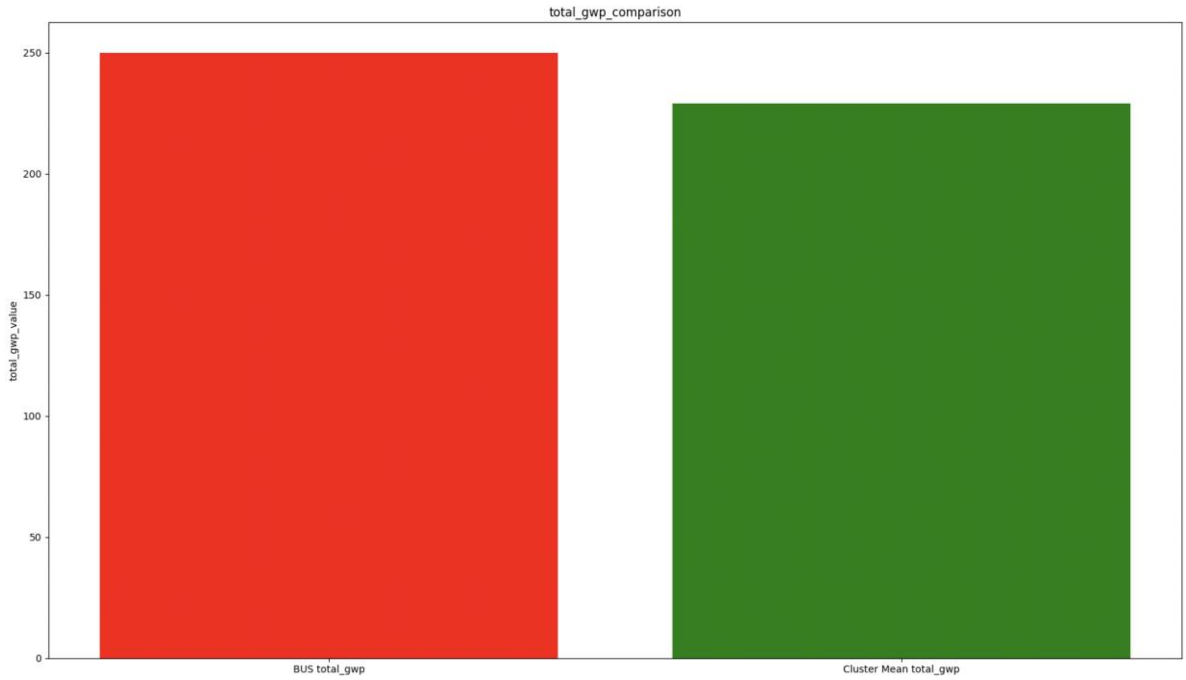

Figure 29. – Total Global Warming Potential value comparison among buildings under study and mean value of buildings matching certain criteria

Table 11. Total Global Warming Percentile Comparison

TOTAL GLOBAL WARMING PERCENTILE COMPARISON
BUS total global warming category is rated higher than 0.0% of the buildings with the same construction decade, primary usage and building area range
BUS total_gwp_category is lower or equal in rating than 100.0% of buildings with the same construction decade, primary usage and building area range

4.2 Validation of the Data Quality Tool

The Data Quality Tool has been implemented and tested on data input originating from the D^2EPC pilot cases. The applied tests simulated the actual data provision that was later realised between the pilots and Hyp's IML component under the works of D^2EPC. More specifically, data acquired by CERTH's and FRC's Restful APIs have been segmented into consecutive chunks of energy/power consumption and ambient conditions measurements captured by the metering and sensing infrastructure in the two buildings. Subsequently, the data quality checks have been successively performed on the incoming data. It is worth noting that a long-term quality assessment has been carried out -on historical data dating back almost a year- in order to collect sufficient information towards the calculation of the data verification KPIs. Results for one year of data acquired from nZEB and FRC are presented in Table 12.

Table 12. Verification KPI results from three different spaces in FRC university

Data Verification KPI	nZEB		FRC					
	Office		Floor		Ampthitheatre		Canteen	
	Temperature	Humidity	Temperature	Humidity	Temperature	Humidity	Temperature	Humidity
Deviation from the Baseline Value	0%	0%	0%	0%	0%	0%	0%	0%
Percentage of Empty Values	50%	50%	7%	7%	<1%	<1%	17%	17%
Amount of Dark Data	50%	50%	7%	7%	<1%	<1%	17%	17%

Based on the results in Table 12, it is observed that all values extracted from both pilots were beneath the expected range, rendering all measurements suitable for the KPI calculations. Large amounts of missing data have been reported for the nZEB temperature and humidity dataset. These were derived from a multisensor measuring both metrics; thus, the missing info is attributed to miscommunications of the device (was later confirmed with colleagues from the nZEB pilot). Contrary to the smarthouse, a small portion of data was missing in the FRC spaces. Finally, the indicator 'Amount of Dark Data' is calculated as a sum of the previously calculated indicators, and provided the fact that no deviations from the baseline value have been observed, the 'Amount of Dark Data' coincides with the 'Percentage of Empty Values'.

5 Conclusions

D4.7 is the second version of the T4.3 deliverable and provides insights on the works conducted towards delivering the design and implementation of the modules residing in the Extended dEPCs Applications Toolkit. Starting from the literature review, the document highlights the current status of benchmarking approaches and further focuses on the functionalities of modern network monitoring tools as well as the data quality assessment as performed in the present day.

Regarding the Energy Performance Benchmarking module, the report presents the main operation and properties of the included Classification and Benchmarking Tools, as described in the project's architecture. Initially, special reference is made to the data input of the clustering algorithms, which consists of building metadata (i.e., primary space use) as well as attributes built upon the SRI domains. The produced clusters act as a basis for the building categorisation. More specifically, each of the D^2EPC demonstration cases is expected to be ascribed to a cluster to enable its comparison with buildings of similar specifications. Furthermore, the building under study is compared to a reference building -with the same primary space use- and a report is generated demonstrating the distance from an ideal overall energy performance.

Concerning the Energy Performance Verification and Credibility module, the two tools materialising the desired services are the Network Monitoring and Data Quality Tools. The Network Monitoring Tool steps on Hypertech's solution, which has already developed a scalable and secure web application that provides a physical representation of the deployed IoT devices along with their operation status. Meanwhile, the Data Quality Tool steps on open-source tools and modules to perform tailored checks on energy metering and ambient conditions data. A set of key performance indicators is further incorporated into the EPVC to enable the monitoring of the system's progression towards the overall quality of provisioned data.

This version of the deliverable D4.7 follows the maturity of the implementation of interacting components and presents the finalised versions of the modules constituting the Extended dEPCs Applications Toolkit. In the EPB case, the selected data input and benchmarking processes are thoroughly described. Moreover, indicative examples of comparative results are presented in the validation section based on the utilised visualisation methods. In the EPVC case, any developments or modifications imposed on HYP's monitoring tool to match the project's specifications and needs are delivered. In addition, the outcome provided to the user via the Credibility UI based on the operational status of the deployed IoT devices and the data quality assessment is included.

Finally, the second report is enriched (chapter 3) both for by EPB and EPVC with i) functional diagrams describing the interacting components, ii) sequence diagrams highlighting the operations carried out, and iii) the overall technology stack used (tools, libraries).

References

[1] CERTH, GSH, DMO, HYP, AEA, SEC, FRC, "D^2EPC Framework Archtecture and specifications v3," 2019.

[2] SpringerGabler, "Gabler Wirtschaftslexikon," 2018. [Online]. Available: https://wirtschaftslexikon.gabler.de/definition/benchmarking-29988. [Accessed 21 10 2022].

[3] A40 Tradepark, "Energy Performance Certificate," 2020. [Online]. Available: https://a40tradepark.com/downloads/unit-1-millbuck-epc-c58.pdf.

[4] B. Coyne and E. Denny, "Mind the Energy Performance Gap: testing the accuracy of building Energy Performance Certificates in Ireland," 17 7 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8550629/. [Accessed 21 10 2022].

[5] C. Alvarez, L. Motta and L. da Silva, "An Energy Performance Benchmarking of office buildings: A Data Mining Approach," in *IEEE International Smart Cities Conference (ISC2)*, 2020 .

[6] A. Ferrantelli, J. Belikov, E. Petlenkov, M. Thalfeldt and J. Kurnitski, "Evaluating the Energy Readiness of National Building Stocks Through Benchmarking," 2022.

[7] O. T. Karaguzel and K. P. Lam, "Development of whole-building energy performance models as benchmarks for retrofit projects," in *Winter Simulation Conference (WSC)*, 2011.

[8] "https://www.cisco.com/c/en/us/solutions/automation/what-is-network-monitoring.html," Cisco. [Online]. [Accessed 10 2022].

[9] "https://www.helpsystems.com/blog/top-benefits-network-monitoring," 10 2022. [Online].

[10] "https://www.comparitech.com/net-admin/iot-device-monitoring/," 10 2022. [Online].

[11] I. 1. 4, "ISO 8000-2: Data Quality," 2020.

[12] I. J. 1. 7, "ISO/IEC 25012 Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model," 2008.

[13] Zhang, L., Jeong, D. and Lee, S., "Data Quality Management in the Internet of Things," 2021.

[14] A. Abu-Zarifa, "Basic Principles of Engieering Metrology," 2013. [Online]. Available: http://site.iugaza.edu.ps/aabuzarifa/files/METRO20152_CH1.pdf. [Accessed 15 10 2022].

[15] T. Kalamaris, T. Papapolyzos, G. Giourgotopoulos, D. Ioannidis, P. Chatzipanagiotidou, S. Koltsios, N. Katsaros, G. Cebrat and M. Susnik, "D3.4 - IoT Platforms & Interfaces v2," 2022.

[16] CERTH, isZEB, KTU,GSH,CLEO, SEC, DMO,HYP, "Buildings digital twin for EPCs issuancce," 2022.

[17] Österrreichisches Institut für Bautechnik, "OIB Richtlinie 06," 2015. [Online]. Available: https://www.oib.or.at/sites/default/files/richtlinie_6_26.03.15.pdf.

[18] Bundestag, "Gesetz zur Einsparung von Energie und zur Nutzung erneuerbarer Energien zur Wärme- und Kälteerzeugung in gebäuden (Gebäudeeneregiegesetz GEG)," 8 8 2020. [Online]. Available: https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl#__bgbl__%2F%2F*%5B%40attr_id%3D%27bgbl120s1728.pdf%27%5D__1665841262911.

[19] F. Keimeyer, T. Kenkman, P. Hennig, S. Jank and S. Metzger, "Informative und transparente Heizkostenabrechnung als Beitrag für den Klimaschutz," 2015. [Online]. Available:

https://www.umweltbundesamt.de/sites/default/files/medien/479/publikationen/climate_change_01_2016_informative_und_transparente_heizkostenabrechnung_als_beitrag_fuer_den_klimaschutz.pdf. [Accessed 15 10 2022].

[20] DWD, "Datensatzbeschreibung Aktuelle monatliche Gradtage nach VD3807 für Deutschland," 2019. [Online]. Available: https://opendata.dwd.de/climate_environment/CDC/derived_germany/techn/monthly/heating_degreedays/hdd_3807/recent/BESCHREIBUNG_derivgermany_techn_monthly_heating_degreedays_hdd_3807_recent_de.pdf.

[21] J. Smith, "inteGRIDy sample reference," *EU,* pp. 1-2, 2016.

[22] O. Pasichnyi and et.al., "Energy performance certificates — New opportunities for data-enabled urban energy policy instruments?," 2019. [Online]. Available: https://www.researchgate.net/publication/330359991_Energy_performance_certificates_-_New_opportunities_for_data-enabled_urban_energy_policy_instruments.