Test and Evaluation Report
Deliverable 7.4

POLicy Data Exploitation & Re-use

POLDER

PROJECT IDENTIFIER	POLDER
PROJECT TITLE	POLicy Data Exploitation & Re-use

DOCUMENT VERSION	1.0
PLANNED DELIVERY DATE	M43 (June 2022)
ACTUAL DELIVERY DATE	M43
DOCUMENT TITLE	Test and Evaluation Report
WORK PACKAGE	WP7
DOCUMENT TYPE	Report – Word

ABSTRACT	The objective of this report is to test the architecture proposed in deliverable 7.2 and evaluate the results of the Partner Polder Ecosystem. The results are validated according the key components of the system proposed. A good testing strategy is implemented in order to minimize risks and using the results for redesigning the design faults. This strategy can guaranty a successful software before is used by end users.
KEYWORDS	Tests, results, evaluation, performance

FUNCTION	DATE	ENTITY
Author	07/04/2022	ACCURO
Editor	15/04/2022	MANTIS
	22/04/2022	FORTEARGE
	22/04/2022	ACD
	22/04/2022	ARD
	15/04/2022	MANTIS
	09/06/2022	ACCURO
	22/06/2022	BEIA
	28/06/2022	NOMMON
	30/06/2022	QUOBIS
Reviewer		

PARTNERS CONTRIBUTIONS RECORD

#	Partners	Contributor on Phase 1	Date of Contribution	Contributor on Phase 2	Date of Contribution 2
1	ACCURO			X	07/04/2022
2	Acd			X	22/04/2022
3	ARD GROUP			X	22/04/2022
4	BEIA			X	
5	ForteArGe			X	22/04/2022
6	FCC			X	27/04/2022
7	Keypro Oy				
8	Mantis Software			X	15/04/2022
9	NOMMON			X	28/06/2022
10	Quobis			X	30/06/2022
11	SIS				
12	Vexiza				
13	Wealize				

CHANGES RECORD

VERSION	DATE	ENTITY	DESCRIPTION OF CHANGES
0.1	07/04/2022	Accuro	Document template creation, and document structure and guidelines.
0.2	15/04/2022	MANTIS	The empirical results for social media monitoring and analytics system were added to the document.
0.3	22/04/2022	Forteargе, ACD, ARD, Mantis	Added inputs from Turkish consortium in Sections 2, 3, 4 and 4.2.4.
0.4	27/04/2022	FCC	Added information related to data acquisition adapters and OCB
0.5	09/06/2022	ACCURO	Added Accuro's tests and results in §2.1, 3.2 and 4.2, and common use case with VEXIZA in §5
0.6	22/06/2022	BEIA	Contributed to sections 2.1 and 3.2
0.7	28/06/2022	NOMMON	Contributed to sections 2.1, 3.2, 4.2 and 5
0.8	29/06/2022	ACCURO	Document compiling, and final editing and formatting
0.9	30/06/2022	QUOBIS	Contributed to sections 2.1
1.0	29/06/2022	ACCURO	Document compiling, and final editing and formatting

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATION	MEANING	DESCRIPTION
A*	a graph traversal and path search algorithm	
APIs	Application Programming Interfaces	

ABBREVIATION	MEANING	DESCRIPTION
CitySDK	A service development kit for cities and developers.	
CKAN	Comprehensive Knowledge Archive Network	Open source data portal software
CNN	Convolutional Neural Network	
FIWARE	Future Internet Core Platform	A middleware platform, driven by the European Commission, for the development and global deployment of applications for Future Internet.
KPIs	Key Performance Indicators	
NoSQL	Not only SQL	
OPM	Occupancy Presence Model	
RNN	Recurrent Neural Network	
RVO	Reciprocal Velocity Obstacles	
SCIS	Smart Cities Information System	A project funded by EU under H2020 programme
Socrata	A software-as-a-service platform for open data publishing and visualization.	
SQL	Structured Query Language	

1. INTRODUCTION

1.1. DOCUMENT OBJECTIVES AND SCOPE

The purpose of this document is to compile the tests that have been conducted to evaluate the behavior of the POLDER platform in different aspects and to analyze the results obtained from these tests in order to make adjustments in the configuration to offer the desired service to the end users.

The evaluation and validation of the platform is performed from different points of view to address the greatest number of issues that may affect the end user experience, either directly or indirectly.

1.2. DOCUMENT STRUCTURE

In order to approach testing from different perspectives, this document is structured following the platform architecture, composed of 5 different layers as shown in Figure 1.

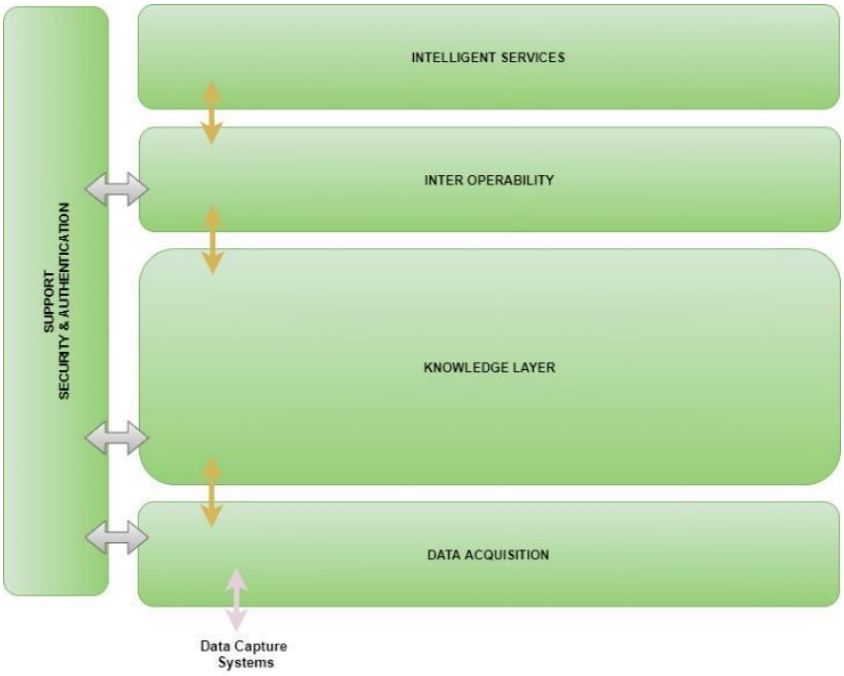


Figure 1. POLDER's Open Urban Platform Architecture - Level 0 schema

Considering this, the following structure is proposed to be followed.

- Data acquisition layer tests
  - Data acquisition
    - Organization
      - Tests
      - Results

- Data adapters
  - Tests
  - Results
- Knowledge layer tests
  - Orion Context Broker
    - Tests
    - Results
  - Advanced analytics
    - Organization
      - Objective
      - Tests
      - Conclusions
- Interoperability layer tests
  - Platform dashboard
    - Tests
    - Conclusions
  - Dashboards for data visualization and analysis
    - Organization
      - Tests
      - Conclusions
- Intelligent services layer tests
  - Scenario
    - Objectives
    - Tests
    - Results
- Security layer tests
  - Tests
  - Conclusions

2. DATA ACQUISITION LAYER TESTS

As stated in *D6.3 Reference Capabilities*, the data acquisition layer implements the necessary mechanisms to integrate information from different data sources in order to meet the requirements of the target environment.

Some specific adapters should be implemented according to the devices and systems to be integrated into the platform and the required protocols to gather information from the original data sources.

These adapters can be divided into two groups: adapters developed by data providers to store data in databases or send data to data analysis systems, and adapters needed to send data to the context broker.

2.1. DATA ACQUISITION

2.1.1. ACCURO

In order to capture the information in terms of video or photos, an IP camera has been selected. The model of this camera is the Trendnet camera model TV-IP1319PI.

Figure 2 shows the camera model.

Figure 2. IP camera model TV-IP1319PI

This camera must generate a fluid video from 8MP to 4K UHD with a resolution of 3840 x 2160 at 15 fps, which allows a range of 30 meters.

In order to carry out the configuration, a PoE (Power over Ethernet) device has been used, which is used to power the device and use an RJ45 network cable.

The types of compression and resolution that the camera has can be shown in the following Table 1.

Table 1. Compressions and resolutions of the camera

• **Compression**	• H.265+ / H.265 • H.264+ / H.264 • MJPEG
• **Resolution**	• Main stream: H.265+/H.265/H.264+/H.264: 3840 x 2160, up to 15 fps • Secondary stream: H.265/H.264/MJPEG: 640 x 360, up to 30 fps • Third flow: H.265/H.264: 1280 x 720, up to 30 fps

As can be seen in Table 1, high resolution settings from 3840 x 2160 to 640 x 360 can be used using the H.265 and H.264 protocols.

It is also very important to know the communication protocols that this camera has in order to configure the ports which will be used to have access remotely. Table 2 shows the communication protocols where the camera can be configured.

Table 2. Network protocols

• **Network protocols**	• IPv4, IPv6, UDP, TCP/IP • DHCP, NTP, DDNS, SMTP, FTP, LLTD • SNMP v1, v2c, v3, QoS/DSCP, IEEE 802.1X • NFS, SMB/CIFS • HTTP, HTTPS • PPPoE • UPnP, RTSP, RTP

As can be seen in Table 2, the most important protocols that can be used are the HTTP, HTTPS, RTSP and RTP protocols, which can be configured to be able to select the port and thus be able to access the camera.

Figure 3 shows the configurations of the HTTP, HTTPS and RTSP ports to be able to select the port according to the convenience of the company. For example, this example shows port 80 for the HTTP port, port 554 for the RTSP port, and port 443 for the HTTPS port.

Figure 3. HTTP, RTSP, and HTTPS port settings

Once the IP of the camera is configured and we know the port to which it is going to be accessed, the next thing is to create a user with a password to have a simple access control and in this way have the correct line of code that includes all these parameters. The line of code for RTSP would look like this:

cv2.VideoCapture('rtsp://admin:camara123@128.0.50.218:554/CH001.sd')

A camera object can be created using Open CV through cv2.VideoCapture(). In this way this object will have the information of the frames per second that are being sent from the camera.

2.1.1.1. TESTS

The information is obtained remotely through the IP camera and these frame per second are analyzed by means of a trained detector. This trained detector detects different objects with a level of confidence that can be adjusted depending on the level of accuracy required.

Figure 4 shows an example in real time when the information of the camera is processed, and a trained algorithm is used to detect different objects such as people, bags, etc.

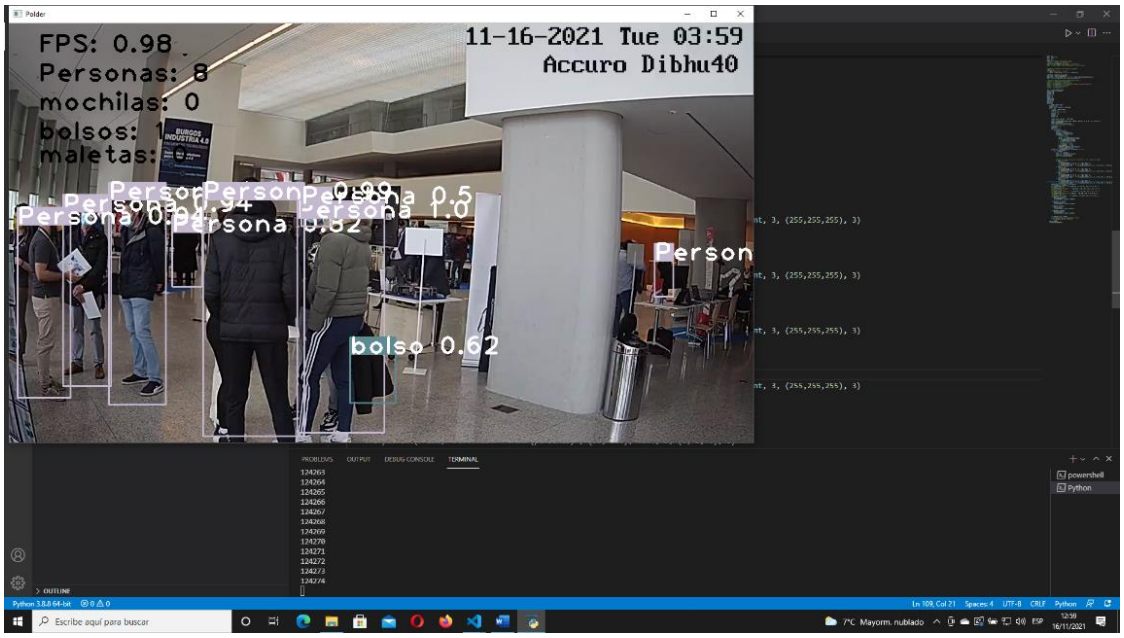


Figure 4 Data acquisition and object detection

2.1.1.2. RESULTS

The results show a good quality video streaming and a good detection of objects that are sending to the database located remotely in a server from Accuro. The lower part of the image from Figure 4 shows the number of identifications for each object detected that is sending to the database while the image from the camera is processed in real time. The boxing boxes represents the detection of different objects. Therefore, the data acquisition has been performed successfully.

2.1.2. ACD

2.1.2.1. TESTS

2.1.2.1.1. ENERGY MONITORING

The DQN algorithm is tested on different time interval dataset shown in figures below. The reward function results are observed for each dataset. The Best results are achieved by conducting on 1 hour interval dataset. The results are discussed in section 2.1.2.2 in detail.

Time	Total	Electricity_cost	sensor0	sensor1	sensor2	sensor3
01.05.2018-00:00:00	0.475966634	0.139	3.33E-05	0.031666667	0.007316667	0.43695
01.05.2018-00:05:00	0.478933333	0.139	0	0.03175	0.007233333	0.43995
01.05.2018-00:10:00	0.483616699	0.139	1.67E-05	0.031783333	0.007033333	0.444783333
01.05.2018-00:15:00	0.482600033	0.139	1.67E-05	0.031783333	0.006966667	0.443833333
01.05.2018-00:20:00	0.346716666	0.139	0.000133333	0.03175	0.00705	0.307783333
01.05.2018-00:25:00	0.259499999	0.139	0.000283333	0.031733333	0.007033333	0.22045
01.05.2018-00:30:00	0.299166666	0.139	0.000283333	0.031833333	0.007	0.26005
01.05.2018-00:35:00	0.3112	0.139	0.00025	0.03185	0.007033333	0.272066667
01.05.2018-00:40:00	0.309233334	0.139	0.000183333	0.031866667	0.007116667	0.270066667
01.05.2018-00:45:00	0.308983366	0.139	1.67E-05	0.0319	0.007033333	0.270033333
01.05.2018-00:50:00	0.298716667	0.139	5.00E-05	0.031816667	0.007033333	0.259816667
01.05.2018-00:55:00	0.2958333	0.139	3.33E-05	0.031733333	0.007066667	0.257
01.05.2018-01:00:00	0.29595	0.139	5.00E-05	0.031683333	0.007116667	0.2571
01.05.2018-01:05:00	0.293066667	0.139	0	0.031716667	0.00725	0.2541
01.05.2018-01:10:00	0.292783367	0.139	1.67E-05	0.031566667	0.007266667	0.253933333
01.05.2018-01:15:00	0.292716667	0.139	0	0.031066667	0.007116667	0.254533333
01.05.2018-01:20:00	0.293550033	0.139	1.67E-05	0.030966667	0.007033333	0.255533333
01.05.2018-01:25:00	0.350900001	0.139	0	0.088066667	0.007016667	0.255816667
01.05.2018-01:30:00	0.394266667	0.139	0	0.113516667	0.024716667	0.256033333
01.05.2018-01:35:00	0.390333367	0.139	1.67E-05	0.031666667	0.101783333	0.256866667
01.05.2018-01:40:00	0.383516666	0.139	0	0.03175	0.097183333	0.254583333
01.05.2018-01:45:00	0.389183366	0.139	1.67E-05	0.031783333	0.100833333	0.25655
01.05.2018-01:50:00	0.39695	0.139	0	0.031783333	0.10915	0.256016667
01.05.2018-01:55:00	0.4010167	0.139	1.67E-05	0.03175	0.113066667	0.256183333
01.05.2018-02:00:00	0.4013167	0.139	1.67E-05	0.031733333	0.114266667	0.2553

Figure 5. 5 minutes interval

Time	Total	Electricity_cost	sensor0	sensor1	sensor2	sensor3
01.05.2018-00:00:00	0.475966634	0.139	3.33E-05	0.031666667	0.007316667	0.43695
01.05.2018-01:00:00	0.29595	0.139	5E-05	0.031683333	0.007116667	0.2571
01.05.2018-02:00:00	0.4013167	0.139	1.67E-05	0.031733333	0.114266667	0.2553
01.05.2018-03:00:00	0.253833367	0.139	1.67E-05	0.088066667	0.12215	0.0436
01.05.2018-04:00:00	0.19865	0.139	0.000266667	0.031816667	0.123283333	0.043283333
01.05.2018-05:00:00	0.198850033	0.139	1.67E-05	0.031783333	0.124433333	0.042616667
01.05.2018-06:00:00	0.198383367	0.139	6.67E-05	0.031066667	0.12405	0.0432
01.05.2018-07:00:00	0.199016667	0.139	0	0.031866667	0.123666667	0.043483333
01.05.2018-08:00:00	0.198116666	0.148	0	0.03175	0.123233333	0.043133333
01.05.2018-09:00:00	0.083083333	0.148	5E-05	0.031716667	0.007733333	0.043583333
01.05.2018-10:00:00	0.082266633	0.148	3.33E-05	0.031833333	0.00685	0.04355
01.05.2018-11:00:00	0.163983367	0.148	6.67E-05	0.113516667	0.007066667	0.043333333
01.05.2018-12:00:00	0.082233333	0.148	0	0.031733333	0.007416667	0.043083333
01.05.2018-13:00:00	0.082183366	0.148	6.67E-05	0.03175	0.007083333	0.043283333
01.05.2018-14:00:00	0.092850034	0.148	1.67E-05	0.030966667	0.007216667	0.05465
01.05.2018-15:00:00	0.117816667	0.148	0	0.0319	0.0073	0.078616667
01.05.2018-16:00:00	0.1035	0.073	0	0.031783333	0.00725	0.064466667
01.05.2018-17:00:00	0.101799967	0.073	3.33E-05	0.031566667	0.0069	0.0633
01.05.2018-18:00:00	0.08275	0.073	0.000233333	0.03185	0.006866667	0.0438
01.05.2018-19:00:00	0.082366667	0.073	0.000216667	0.031666667	0.006933333	0.04355
01.05.2018-20:00:00	0.082466666	0.073	0.0002	0.031683333	0.00705	0.043533333
01.05.2018-21:00:00	0.082299966	0.073	3.33E-05	0.031733333	0.00715	0.043383333
01.05.2018-22:00:00	0.138633334	0.073	0	0.088066667	0.007116667	0.04345
01.05.2018-23:00:00	0.082150034	0.073	1.67E-05	0.031816667	0.007366667	0.04295

Figure 6. 1 hour interval

2.1.2.1.2. AIR QUALITY

The model tested with real work data through apis. The following figures concrete example is that testing the model on plum labs api.

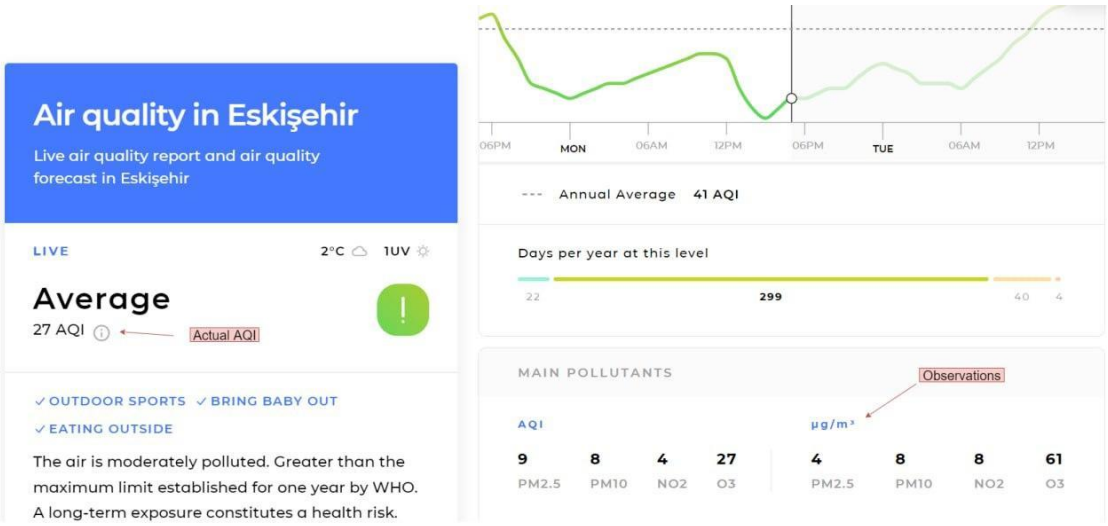


Figure 7. Plum labs API

```
# one obs test
X_test = [{
    'no2': 4,
    'o3': 61,
    'pm10': 8,
    'pm2_5': 4
}]

obs = pd.DataFrame(X_test)
y_pred = rf.predict(obs)
print(y_pred)
```

✓ 0.5s

```
[27.]
```

Figure 8. Result of air quality algorithm

2.1.2.2. RESULTS

2.1.2.2.1. ENERGY MONITORING

The input of the model is electricity cost per hour, desired device usage time interval, usage duration of the device and power consumption of each sensor. Input panel designed for the time scheduling. The output of the DQN Algorithm is that action values which takes binary values as 1 or 0 refers to switch on or off respectively.

The DQN time interval suggestion system aims to provide a user-friendly suggestion table that is the interpreted form of the optimization plots. Optimization plots are the visual output of the DQN algorithm and display the insight of the produced results. However, sometimes users may not have enough background to interpret the graphs. In order to solve this problem, we provide a solution that interprets the plots at the background and informs

the users what are the optimal ranges for usage of the devices. As an example of the recommended optimal hours of the devices presented in Figure 9.

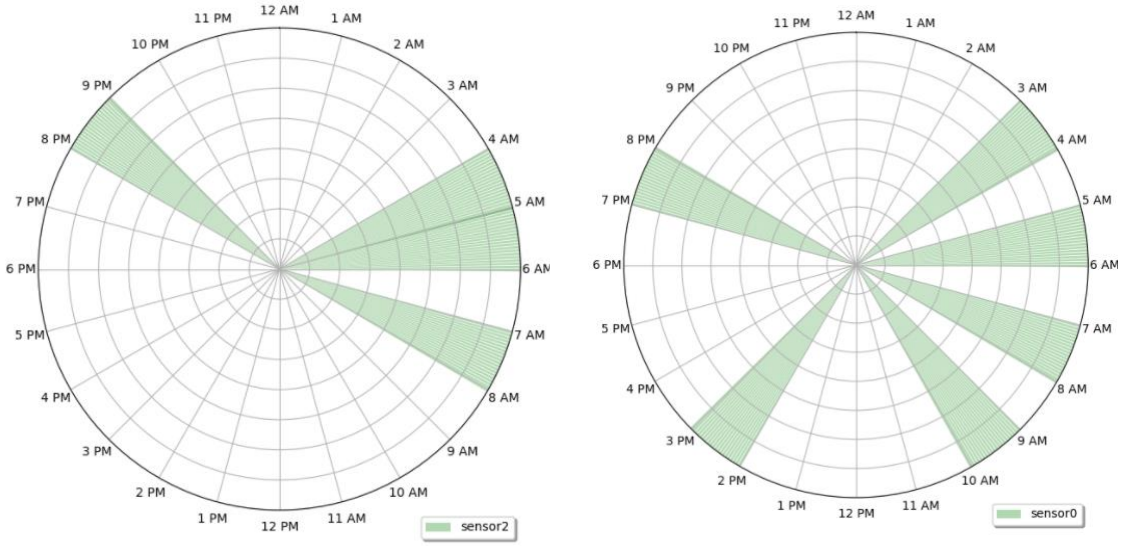


Figure 9. Recommended time intervals

In order to look at the optimal hours more precisely, the optimal hours are also presented in the tabular form. Figure 10 is an example of the tabular form of the optimal hours.

Optimal Time Interval for Use of the Devices

time_start	time_end	message	sensorid
13:00	14:00	Optimal usage interval	sensor2
19:00	20:00	Optimal usage interval	sensor2
20:00	21:00	Optimal usage interval	sensor2
21:00	22:00	Optimal usage interval	sensor2
22:00	23:00	Optimal usage interval	sensor2
23:00	00:00	Optimal usage interval	sensor2

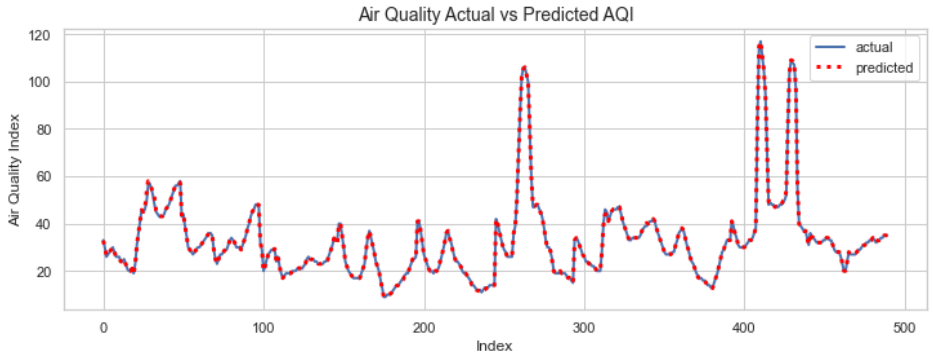
Figure 10. Optimal time intervals in tabular form

2.1.2.2.2. AIR QUALITY

Random forest algorithm is utilized for air quality index assessment. In this context NO2, O3, PM10 and PM2_5 is used for input in order to predict in air quality index. Air quality index is predicted as output for determining the quality of the air.

Output of the algorithm is understandable. Predicted air quality index values converted into categories such as very unhealthy, hazardous, unhealthy, sensitive group, moderate and good. Air quality index describes the level of air pollution that can be identified with the index value between 0 to 500. Higher air quality index value indicates that the higher pollution.

The trained random forest algorithm over the air pollution dataset which is gathered from different APIs is presented in the following figures.



	actual	preds
0	33.0	32.81
1	30.0	30.38
2	26.0	26.48
3	27.0	26.86
4	28.0	28.10
5	29.0	29.00
6	30.0	29.98
7	28.0	28.13
8	26.0	26.03
9	26.0	26.00

Figure 11 Actual vs. Predicted AQI

The accuracy of the regression task is measured by RMSE(root mean squared error). In this case achieved RMSE is 2.61.

2.1.3. BEIA

Air quality monitoring is a fundamental task in order to check the collection of the data from specific sensors from the station. It is important to ensure continuous monitoring and to ensure that the implementation is carried out correctly and works as expected. This entails that:

- The point of presence, where BEIA is collecting data, is correctly placed on the station.
- The devices that are going to process the air quality are powerful enough to perform with the expected load of pollution from users

In order to check these aspects, it is necessary to perform a systematic and continuous analysis of the data collected from the air quality station. Multiple tests in controlled scenarios allow us to determine which is the expected profile and load of pollution collected per device.

Outdoor air pollution comes mainly from the incomplete combustion of fuels used in transport, energy and industry, as well as other human activities. The urban population is exposed to some extent to concentrations of pollutants in excess of the limit values set by EU legislation and those in World Health Organization (WHO) guidelines.

2.1.4. MANTIS: SOCIAL MEDIA MONITORING AND ANALYTICS ADAPTOR

We have used Twitter API to monitor social media data including message, user names, location, user description, tweet hashtags, and reply counts. To this end, a stream based data monitoring and analytics platform were implemented as in figure below. In this architecture, Elasticsearch was used for data warehouse and analytics engine, RabbitMQ was adopted as a message broker, Logstash was employed as data processing pipeline, and Kibana was used for visualization.

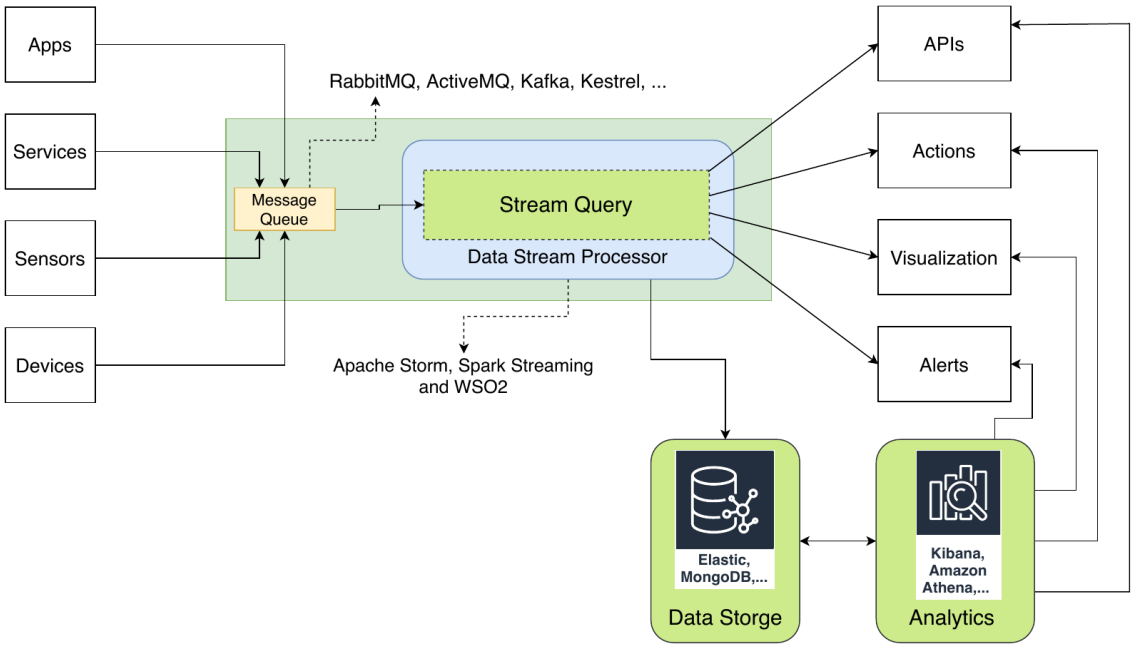


Figure 12. General architecture of data monitoring and analytics system

2.1.4.1. TESTS

To evaluate the performance of the monitoring and analytics system, we have used a collection of 1,000,000 tweets that has been obtained through two years (2020 and 2021). All these data were sent to the monitoring system for evaluating its write performance in different workloads. The experiment was conducted using Elasticsearch with 3 nodes and 6 shards (2 shards per node). In Elasticsearch, an index is split into elements known as shards that are distributed across multiple nodes. Each node (machine) possessed 16 GB of RAM and a 6-Core CPU in our experiment. The performance of the system was evaluated on the write and read workloads.

2.1.4.2. RESULTS

The monitoring capability of the system is measured in the first experiment. The following figure shows the throughput of the system for write-only workload over different number of client threads. As can be seen, increasing the client threads can boost the system's write throughput by 5500 write operations per second using 32 threads.

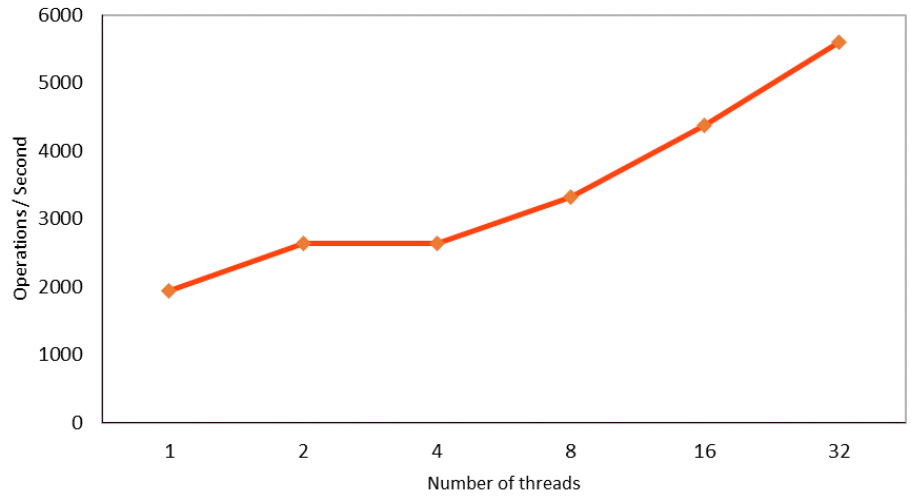


Figure 13. Write throughput the monitoring system in the write workload

To show the monitoring performance of the system, we also reported the averaged-write latency over different number of records as follows.

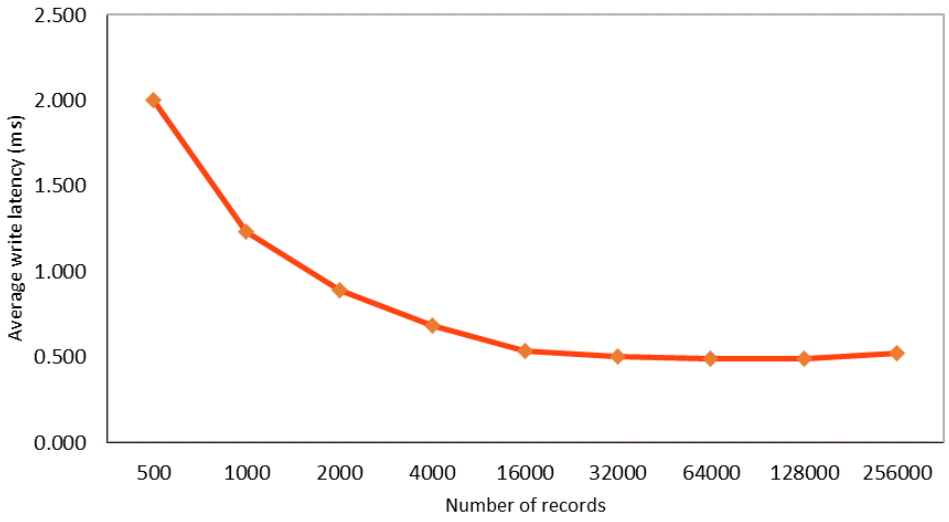


Figure 14. Averaged-write latency over different number of records

In the second experiment, the analysis capabilities of the system were taken into account in read-only and read-write workloads using an aggregation query. The following figure shows the averaged read latency in read-only and read-write workloads. As can be seen, the averaged read latency increases in both workloads as the number of records grows beyond 4000. As a result, the proposed system has a low read latency for aggregation queries with fewer than 4000 records. This performance can be enhanced by increasing system resources.

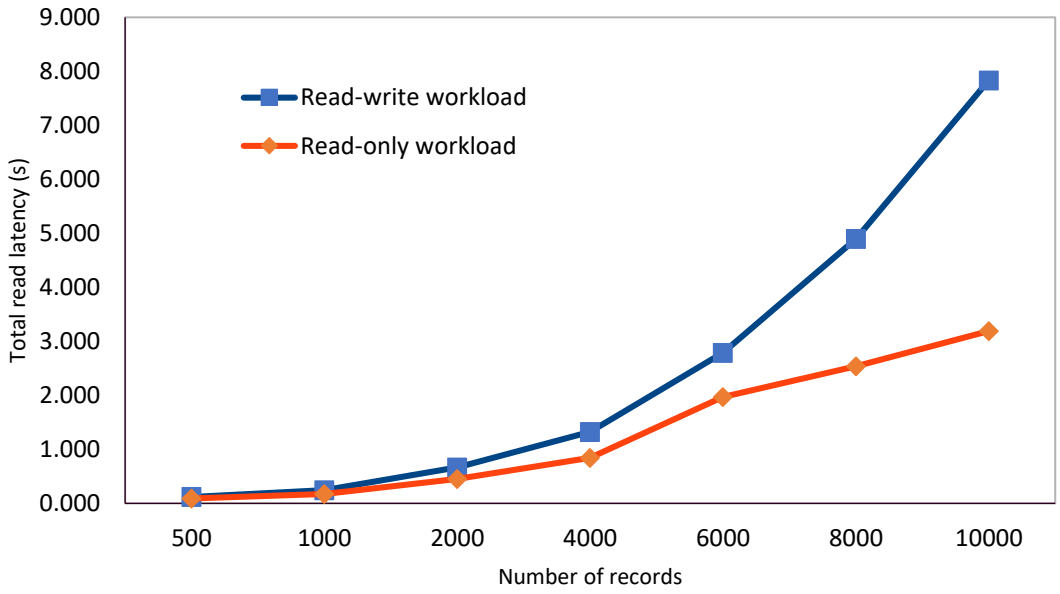


Figure 15. Total read latency for read-only and read-write workloads in the proposed analytics system

2.1.5. NOMMON

The solution developed by Nommon consists in estimation and prediction of tourist flows based on the fusion of mobile phone records and other relevant data sources.

Below we provide a list with a short description of the different data sources gathered and analysed by Nommon in the scope of the project.

Mobile phone data.

The main data source used for the extraction of tourism activity indicators is anonymised mobile phone records.

Mobile phone data includes the geolocated positions of mobile devices (at a network cell level), both for active events (calls, SMS, Internet connections) and for certain passive events (coverage area changes, network updates, etc.). Additionally, sociodemographic information is available on the operator's customers, such as their age and gender.

This information is organised in three different files:

- An events file, this is a csv file that contain the information related with the recorded event: anonymous user ID, time at which the event occurs, cell ID of the cell at which the device was connected and other information stored for billing purposes (e.g type of device, operative system, etc.) but not relevant for the current development.
- Cell file, this is a csv file that contains the information of the position for each cell identifier.

- Users profile, this is also a csv file which contains information related to the sociodemographic characteristics of the agent.

Land use data

Land use data from the Spanish Land Occupation Information System (SIOSE) is used to improve the characterization and precision of the activities identified from the mobile phone data.

SIOSE is a database that provides land occupation information for the entire national territory at a cartographic scale equivalent to 1:25,000. The territorial division does not correspond to any existing classification, but through divisions similar to those of a puzzle it describes the territory indicating the different coverages present in the SIOSE.

Through the database you can access the download files in shp and dbf format, which allow maps to be drawn with the corresponding attributes for each zone, the most important fields being:

- ID_POLYGON: unique identifier of each SIOSE polygon
- SIOSE_CODE: code that describes each coverage and attributes by means of labels.
- SURF_HA: surface in hectares of each SIOSE polygon

Census data

The National Institute of Statistics (INE) publishes annually, on January 1 of each year, the population of all census sections disaggregated by gender and age.

The gender field can take as values "men" and "women" and the ages are grouped into the following groups, 00-04, 05-09, 10-14, 15-19, 20-24, 25-29, 30 -34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85-89, 90-94, 95-99, +100.

This information con be directly downloaded from the portal and comes in csv files.

Income distribution

The National Institute of Statistics (INE) has published the atlas of household income distribution based on household surveys of some 16,000 households spread throughout the national territory. The most up-to-date data available is from 2017, although work is being done on expanding the surveys to update the available data. This type of data will start to be annual from the next update.

The indicators generated consist of:

- Average income per person
- Average income per household
- Income sources (in %) for the census section. Five categories are considered: salary, pensions, unemployment benefits, other benefits, other income.
- Average income and distribution of income per consumption unit:
    - Percentage of population with income per consumption unit below certain fixed thresholds (5,000, 7,500, 10,000).
    - Percentage of the population with income per consumption unit below certain relative thresholds expressed as a percentage of the median (thresholds considered, obtained from the distribution of income per consumption unit: 40%, 50%, 60%, 140%,160% and 200% of the median)
- For the income distribution indicators per consumption unit, disaggregated data is also offered according to these classification variables:
    - Sex
    - Age groups (under 18 years old, from 18 to 64 years old, 65 years old and over)
    - Nationality (Spanish, foreigners)
- The variable 'sex' is crossed with 'age groups' and 'nationality'.

This data is also downloaded directly from the source and comes in csv files.

From these data, shp files are generated with the income information as an attribute of the spatial object which allows their subsequent integration with the user profiles.

Statistics of tourist movements at the border (FRONTUR)

The INE (National Institute of Statistics) performs monthly statistics on tourist movements at borders based on surveys carried out on non-resident users in Spain who enter or leave our country.

The study population is all non-resident visitors in Spain who enter the country, whether they have stayed overnight or not, in addition to visitors in transit in Spain.

From these statistics it is possible to obtain statistics on visits to Spain, such as the number of entries and exits by nationality.

The total segmentations offered by this data are:

- Type of traveler: non-resident tourist, non-resident hiker

- Way of entry: Highway, Airport, Ferry/Ship, Train
- Country of habitual residence: Belgium, France, Germany, Ireland, Italy, the Netherlands, the Nordic countries, Portugal, the Rest of Europe, the Rest of the world, Russia, the United Kingdom
- Main destination of the trip: All autonomous communities, as well as visitors in transit
- Main accommodation: Hotels and similar, Rental accommodation, Camping, Rural house, Cruise, Other market accommodation, Home ownership, Family/friends housing, Other non-market accommodation
- Reason for the trip: Leisure/holidays, Business, Studies, Personal (health, family), Other reasons
- Tourist package: Yes, No
- Duration of the trip: None night, 1 night, From 2 to 3 nights, From 4 to 7 nights, From 8 to 15 nights, more than 15 nights

This as all the data from INE can be downloaded from the institution's portal and comes in csv format.

InsideAirbnb

Airbnb is a digital platform dedicated to offering accommodation to individuals and tourists, through which hosts can advertise and rent their properties to their guests. Hosts and guests post reviews publicly for reference by future users. Currently the offer is about 2,000,000 properties in 192 countries and 33,000 cities.

InsideAirbnb is a non-commercial tool that allows access to rental data on Airbnb. Information is collected from all establishments, by city, and data is provided on the location of the accommodations offered on Airbnb, along with their identifier, as well as the owner's identifier. In addition, the information is known on how many times the accommodation is rented per month, how many times it has been rented in total and the number of days per year in which the apartment is available.

All the information can be directly downloaded from the portal in csv format.

In the same way that the characterization of land uses allows to improve the identification of activities, the location of the accommodations allows to identify the areas with the greatest capacity to receive visitors. This data will allow an improvement in the characterization of the overnight stay areas of visiting users.

Currently the data is not available for all establishments registered on Airbnb, which limits the use of this information to the cities for which this information is available. In Spain, the cities for which this information has been published are Barcelona, Madrid, Malaga, Mallorca, Seville and Valencia.

Figure 20 shows the distribution of Airbnb accommodation in Madrid, categorized by type of accommodation, with entire apartments in red, private rooms in green, and shared rooms in blue.

From this data it is possible to calculate the density of accommodations published on Airbnb by zone, dividing the number of accommodations by the area of the zone, allowing this data to be included when assigning an overnight area to visitors in the zones.

As already mentioned, all the data listed above has already been collected by different organizations and except from the mobile phone records can be downloaded directly from the web page of these organizations and hence none adapters are required for their collection.

In the case of mobile phone data due to the sensibility of the information contained in this data sources and to comply with the GDPR data cannot leave the premises of the data provider and all analysis should be performed in their platform. From the analysis only the results, aggregated data will come out.

2.1.5.1. TESTS

Although the data collection does not entail the deployment of sensors and data adapters, the gathered data needed to be pre-processed to be ingested by the solution. This work of data adaptation involved data integrity analysis and data standardization. Below we briefly describe the process followed for each data source.

Mobile phone data:

- Data standardization. The raw data provided by the network operator was process such that only relevant information is analyzed and all other fields are removed.
- Representativeness analysis. The distribution of sample (users) along the territory was analyzed to identify and correct possible biases in the data. The distribution of the different nationalities was also analyzed to ensure that all nationalities were well represented.

Land use data

This data is used to better characterize the activities performed in the different zones of the regions of study. To facilitate the use of this data the region of study is divided in smaller zones and a weight representing its capacity to accommodate activities by visitors and/or users in general is assigned to each zone according to its land attributes. based on its.

Census data

To facilitates the use of the data provided by the INE shp files are generated with the total population by age and gender as an attribute of each geographical object.

Income distribution

The same procedure applied to the census data was applied to the income information to facilitate the use of these data.

Since the files of InsideArbnb and Frontur were easily manageable in their original form no standardization process were applied to them.

2.1.5.2. RESULTS

Mobile phone data:

Standardization. After the standardization process each event was represented in the following format

{"agent_id":"XXXX","time_stamp":"2021-10 02T00:25:30","cell_id":"214030715039271"}

Representative analysis. Figure 16 the distribution of users along the Spanish territory. As it can be seen the sample is homogenously distributed with higher values in main cities.

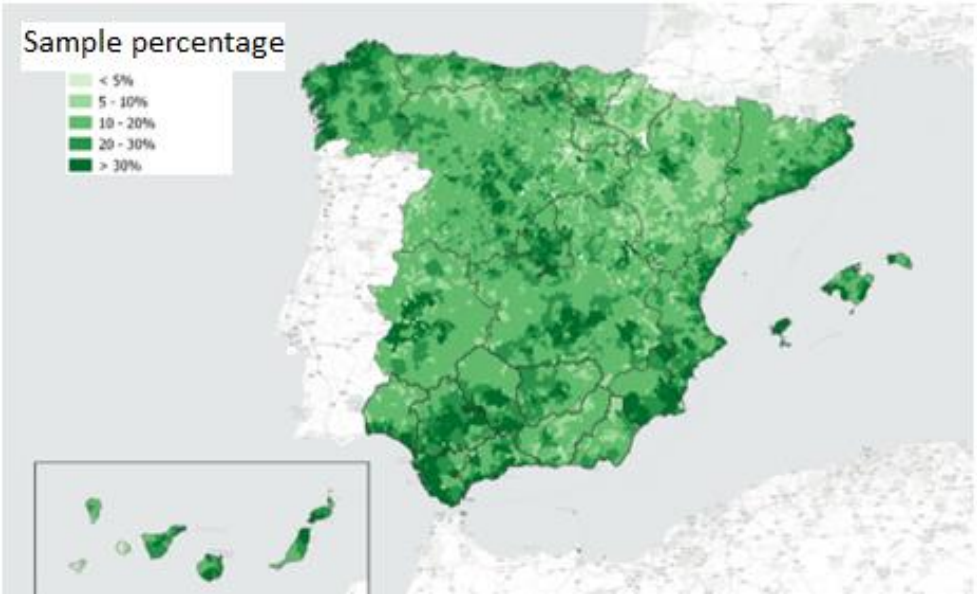


Figure 16 Distribution of users along the territory

Land use data

Bellow, in Figure 17 we can see we can see a fragment of the map of Spain with the weight of each zone according to its main land use. We can see that the zones inside the cities are well represented by commercial and residential land use while in the peripheries of the cities and farther away the predominant land use is rural.

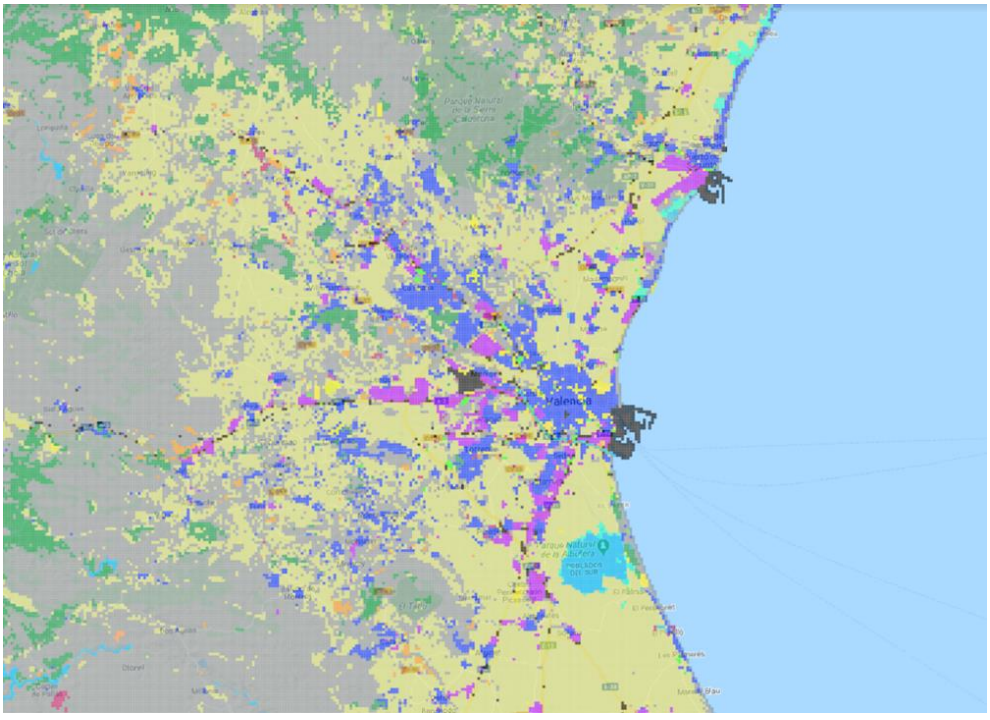


Figure 17 Fragment of the meshed map of Spain showing the main land use for each square. Purple and pink represent commercial and residential zone, while green and yellow represent more rural zones.

Census data

Figure 18 shows the representation of the population obtained from the shp generated with the census information. We can see that most of the population is concentrated in the cities while the rural areas are almost empty. This is exactly what is observed in the raw data and hence we can conclude that the generated shp file is correct.

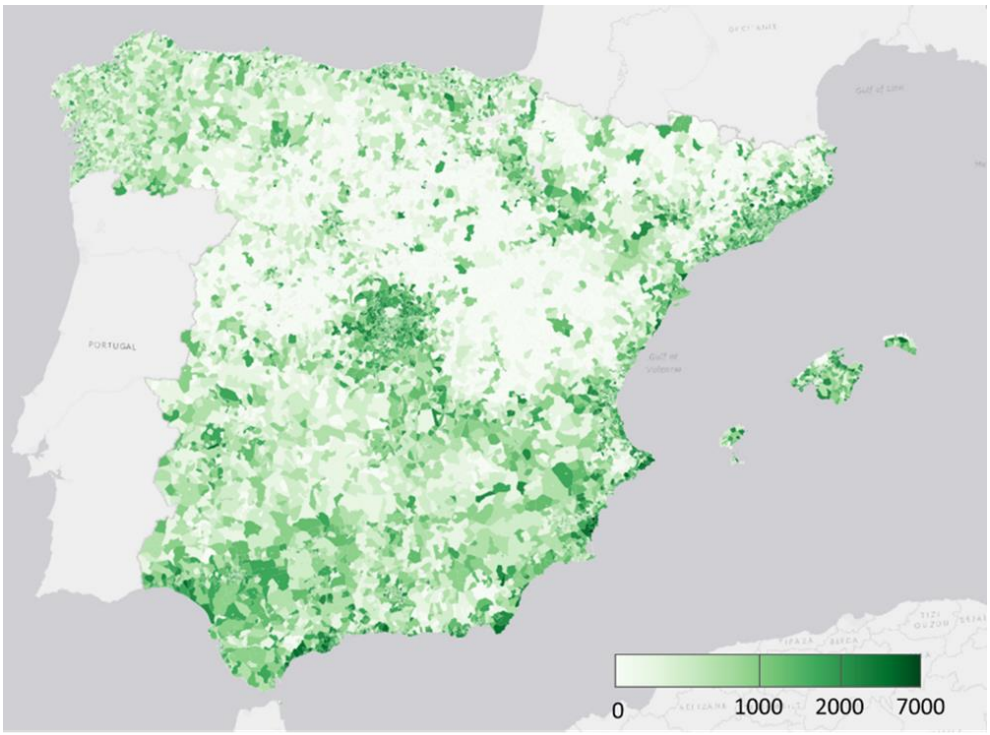


Figure 18. Map of Spain with the autonomous communities categorized by inhabitants

Income distribution

Figure 19 shows the income distribution in the Spanish territory extracted from the shp generated with the income information. The income distribution corresponds to what is known, i.e. higher incomes are concentrated in Madrid (the capital of Spain) and in the northern part of the country. that observed in the raw data and hence we can conclude that the generated shp file is correct.

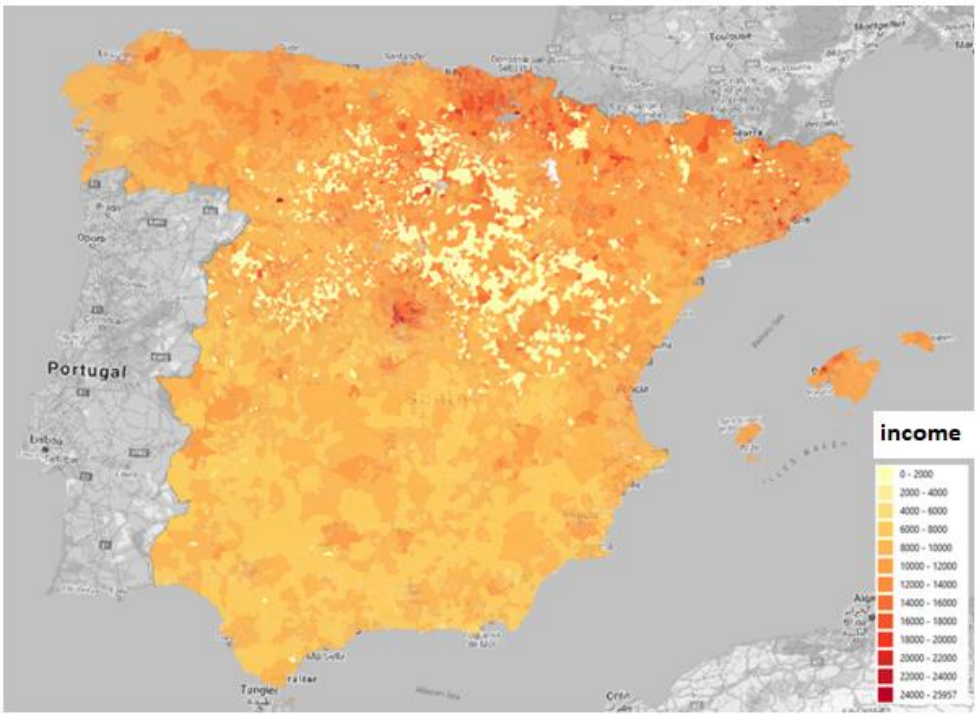


Figure 19. Income distribution in the Spanish territory.

2.1.6. QUOBIS

The voice recognition and sentiment analysis tool provides real time information about positive or negative sentiments in a conversation. As first step it identifies the language and then it analyses the voice transcription, providing anonymized information about the sentiments. The conversation or speech can come from an audio collection system or from a video stream. The audio sample is processed in real time using an algorithm of detection. Once specified the language of the speech, the audio is processed in 30 seconds long fragments, where the text is transcribed.

Those words in the transcription that have subjective value are captured, giving information about the stay, positive and negative. Then, the analyzer counts the amount of each part and makes a comparison between them. The output data is a result's sequence of the audio stream analysis giving the endpoint information about the analysis. This result is compounded by two numeric values: the polarity and the subjectivity.

- The subjectivity expresses the relevance of the information in terms of the customer's tastes, if the speech includes information about its experience this value will be closer to 1.
- The polarity is the final resolution of the analysis: if it is closer to -1, the speech tells about a negative experience, and if it's closer to 1, a positive one.

2.1.6.1. TESTS AND RESULTS

To evaluate the performance of the voice recognition and sentiment analysis system, we have used a collection of videos talking about a variety of themes from touristic issues to institutional messages, the impact of COVID pandemic, etc. with positive, negative and neutral messages.

This table shows the values obtained for each test:

Test video	Subjectivity	Polarity
25 Most beautiful destinations in Europe	640	0.15162283819000
7th Global Tourism Crisis Committee	79	0.05899925317401
Paradise lost mass tourism leads to pollution crisis	-11	0.02010968921389
The impact of COVID-19 on tourism around the globe	30	0.01512859304084
Coronavirus: How big is the damage to the tourism	53	0.02625061911837
Global Health Summit	4	0.01261829652996
20 years of ITEA	326	0.15825242718446
ITEA project reporting	53	0.03840579710144
ITEA Guidelines for preparing an FPP	143	0.03459956448100
ITEA General Information	79	0.05796038151137

Figure X. Results obtained during tests

The results are aligned with expectations providing a value according to each test. Here we found some differences between the institutional and formal messages and the dissemination or informal messages.

It is important to take into account that the sentiment analysis is not going to be accurate 100% of the time. This is because a normal conversation could have some symbolism on it, for example, the sarcasm expressed by the intonation of some words, has a different meaning for speakers, but the analysis system would take it literally.

Another result we found is that to make the sentiment analysis more trustful, the speech recorded is divided into 30 seconds long fragments, in order to produce more accurate results.

2.2. DATA ADAPTERS

The *Data Acquisition Layer* is responsible for implementing the processes necessary to incorporate information from several data sources to achieve the requirements of the target environment. This information could come from different types of sources, so that it is necessary to develop specific software modules, capable of adapting to each case.

Specific adapters should be implemented according to the information and systems to be integrated into the platform.

2.2.1. POLDER DATA ACQUISITON ADAPTERS

In order to acquire and share data through POLDER open platform, FCC, responsible of the architecture design and platform development, has developed two data acquisition mechanisms. These mechanisms have been considered by **active request** to a data provider or as a data reception **service provider**. This way, adapters that work by active request, invoke some devices or systems in order to collect data on demand according to the protocol or interface designed by the device manufacturer or the system owner. On the other hand, adapters that work as services providers allow the data ingestion by external systems that must meet the **NGSI-v2 specification**. In both cases, *Data Acquisition Layer* will send data to the *Knowledge Layer* meeting this specification.

The architecture proposed for **POLDER's Open Urban platform** is based on a scheme that is aligned with the layer model defined in official regulations and widely extended interoperability recommendations such as **ITU-T Y.4200 (2018)** and **ITU-T Y.4201 (2018).**

Information collected by the *Data Acquisition Layer* is provided to the *Knowledge Layer* through the **Orion Context Broker**, which provides a FIWARE **NGSI v2** Restful API, enabling to perform updates, queries or subscribe to changes on context information.

The adapters implemented so far are working as **service provider**, so that they publish several entry points that can be invoked by external systems to supply their data to the platform, using the following syntax:

http://<IP-Address>:<Port>/<methodName>

<*methodName*> is one or more names, strings, acronyms, etc. that identifies the action to be done. If the action is to supply information, it must be provided by means of a JSON file according to FIWARE-NGSI v2 format.

A preliminary common structure has been agreed with the partners to include several mandatory attributes into de JSON file.

2.2.1.1. TESTS

The tests performed verify that both the loading and the flow control of the data is correct.

For the data loading tests and by means of the adapters, it is verified that the data is correctly ingested into the platform, specifically, it is verified that the data is stored in the OCB.

For the data flow tests and using the platform's REST APIs, it is verified that the data once are in the OCB, they can be read, modified, updated and that, through the subscriptions mechanism, it is historicized in an external database.

As general procedure, a set of unit tests are performed in a local environment. If any deviations or bugs appear, the software is updated. Once the tests have been successfully passed in the local environment, the version of the platform hosted on the project server is updated.

Unit testing steps to follow for each adapter:

1. The services of the corresponding adapter to be checked, are published
2. A subscription is created (if applicable)
3. The corresponding service is called for data insertion.
4. Through the platform's REST API calls, it is verified that the data reading works and therefore they exist in the OCB.
5. Using a database tool, check that the subscriptions have worked by verifying that the corresponding data have been historicized.
6. Data are updated/modified
7. For each action of the previous point, the process from point 1 to 5 is repeated.

The tests described above are first performed in a local environment in order to verify the suitability of the software and then, each partner will perform them on the project server, once the platform version is updated.

More information about the tests performed by partners using these adapters can be found in Deliverable 6.3: *Implementation and integration of adapters*.

2.2.1.2. RESULTS

To verify the test results, curl commands or REST client tools, such as Insomnia, can be used.

In FCC's tests the Insomnia tool has been used.

To perform the tests, dummy entities are created (with the appropriate format) and it is verified that these entities, previously sent by means of the corresponding adapter, are correctly

displayed in the OCB. If this is fulfilled, the loading process has been successful. The same process is repeated for any data modification, update or delete action.

To check the subscriptions have worked, it must be verified the historization has been completed, using a database management tool compatible with the database being used.

Using Insomnia, FCC checks if the REST methods of the platform API are working correctly. For this action, the read methods are called, and the OCB results (entities and their attributes) are checked against the originals sent via adapters.

2.2.2. FORTEARGE GENERIC DATA ADAPTER

On the other hand, FORTEARGE has developed their own generic data adapter component which is based on Apache NiFi platform. This section describes the tests performed with this component.

On the Apache NiFi platform, where data flow processes within the scope of the project are managed, purpose-designed data adapters consist of a combination of NiFi processors that perform different functions.

2.2.2.1. TESTS

The first processor that creates the adapter is the processor that enables the data in CSV format to be received from the server to be included in the system flow. Here, a file can be retrieved from a local directory or more functionally, the SFTP protocol can be used. SFTP is a protocol for transferring files using SSH. Unlike FTP, it provides a secure transfer and works over TCP. In order to increase the scope of use of the adapter, it is more convenient to use the SFTP protocol. Processor-specific tests are performed on LogMessage, which is connected to the processor and keeps logs.

After the data is received, the data file in CSV format needs to be parsed in rows. Each row represents a data entry/record in the system. This process is completed with the SplitText processor on the NiFi. Processor-specific tests are performed on LogMessage, which is connected to the processor and keeps logs. Each parsed row can be observed in these logs.

Each data record parsed as a row must be assigned to attributes by separating their values. Thus, the data parsed in lines from the main file will be parsed into columns, and the entire parsing process will be completed. Parsing the data is done by regex extraction. As a result of this process, where the Extract Text processor is used, each value is assigned to the attribute variables defined internally in the processor. Processor-specific tests are performed on

LogMessage, which is connected to the processor and keeps logs. Each variable assignment and assigned values can be monitored in these logs.

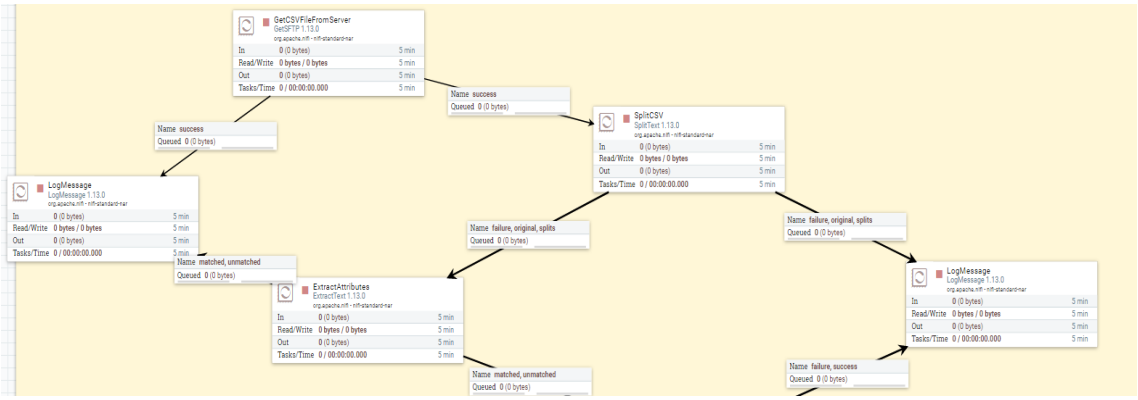


Figure 20. A data adapter instance for CSV file

After the parse process is completed, the values assigned to the attribute variables are sent to a new processor to be converted to the standard message format used in the system. In this process, where the ReplaceText processor is used, the data obtained as a result of processing the CSV file is converted to NGSI Context format to be sent to Orion Context Broker. Using this structure, parsed messages can be converted to messages in any standardization structure. Processor-specific tests are performed on LogMessage, which is connected to the processor and keeps logs. Each NGSI Context message created in these logs can be observed.

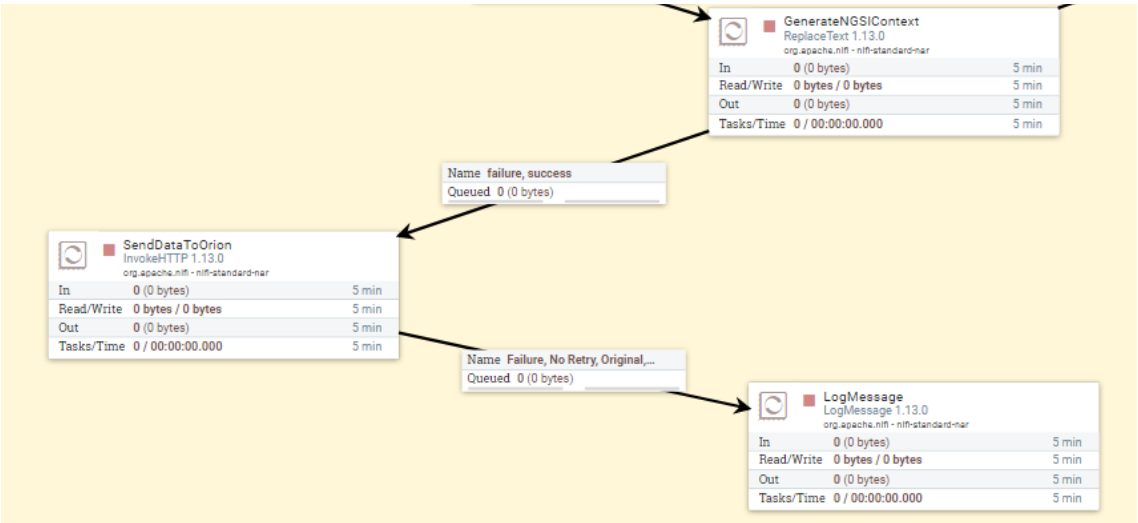


Figure 21. Generic data adapter - Sending to Orion

2.2.2.2. RESULTS

The prepared data can be sent to the desired component or database as an HTTP post request. It has been verified that applications on Apache NiFi can send data to MongoDB, PostgreSQL, and Spark.

Testing the adapter as a whole is to observe whether the desired results are obtained by repeatedly starting the dataflow processes. During this observation, the logs in the LogMessage processors connected to each processor and the records expected to be formed in the database are examined.

	Date/Time	Type	FlowFile Uuid	Size	Component Name
	04/13/2022 15:08:03.763 GMT	DROP	defbad55-f37e-44d8-9f55-4c849932eead	359 bytes	LogMessage
	04/13/2022 15:01:28.907 GMT	DROP	3653bafd-e2a4-4484-9d9f-b69b983f33eb	357 bytes	LogMessage
	04/13/2022 14:56:28.543 GMT	DROP	d203c7ad-218e-42c9-905d-33c808dd4224	357 bytes	LogMessage
	04/13/2022 14:49:42.296 GMT	DROP	0a575a93-e292-4de1-b51b-efb1347c866a	357 bytes	LogMessage
	04/13/2022 14:41:51.588 GMT	DROP	7db32f12-f46f-4f12-8a06-908b55457cc4	359 bytes	LogMessage
	04/13/2022 14:41:51.447 GMT	DROP	2974fde3-3209-46fe-88d4-cb7583475159	358 bytes	LogMessage
	04/13/2022 14:41:51.396 GMT	DROP	9214e449-4d7d-44e7-8134-f518f0a6a338	359 bytes	LogMessage
	04/13/2022 14:41:51.345 GMT	DROP	5634bc6d-79b5-4762-9eda-4d6b017eaf1e	357 bytes	LogMessage
	04/13/2022 14:41:51.234 GMT	DROP	d649766c-65fd-452e-8552-bfaef47eaab1	359 bytes	LogMessage
	04/11/2022 14:40:44.483 GMT	DROP	7c21bbd3-3bf5-4fd6-874c-7ae4158859b1	256 bytes	LogMessage
	04/11/2022 14:40:36.151 GMT	DROP	db69c253-8c74-4714-b64e-eac3b30e4f1f	337 bytes	LogMessage
	04/11/2022 14:19:30.292 GMT	DROP	54ca02d9-f169-4f23-a3c3-c7bac2648f11	338 bytes	LogMessage
	04/11/2022 14:19:26.215 GMT	DROP	64a780f5-d0c2-4065-8a2c-8ef5ac5c34a9	337 bytes	LogMessage
	04/11/2022 14:17:49.071 GMT	DROP	021dbb58-1a18-49ab-8dbc-09f56725f506	337 bytes	LogMessage
	04/11/2022 14:16:09.499 GMT	DROP	cd0de77c-46f2-43ab-8311-d41beff15615	338 bytes	LogMessage
	04/11/2022 14:14:41.067 GMT	DROP	08799dad-54c8-4c02-985b-ccf3c53e1396	338 bytes	LogMessage
	04/11/2022 14:13:16.526 GMT	DROP	724a6dce-d772-4527-94dc-81fd202f6949	338 bytes	LogMessage
	04/11/2022 14:12:17.508 GMT	DROP	0fa10b7b-bbb3-4088-bac8-41106f010c54	337 bytes	LogMessage

Figure 22. Log records for testing

3. KNOWLEDGE LAYER TESTS

The information collected in the data acquisition layer is sent to the knowledge layer, where an intelligence and learning process is carried out inside the advanced analytics phase. Then, the processing results are stored in the context broker.

3.1. ORION CONTEXT BROKER

A Context Broker Generic Enabler is the core and the only mandatory component of any platform or solution "Powered by FIWARE".

POLDER uses FIWARE's Orion Context Broker (OCB) to orchestrate all context information.

Orion Context Broker allows to manage context information in a highly decentralized and large-scale way. It also allows context producers and context consumers to interact in a decoupled way using different synchronous or asynchronous communication paradigms.

3.1.1. TESTS

The Polder Platform software implementation is based on Docker container technology. Therefore, OCB enabler is a Docker container. Additionally, the OCB need to have an associated database to work, in our case is Mongo.

A simple way to check if the OCB and his associated database is working properly, is to verify if both containers are running and healthy using a simple Docker command line.

`docker ps` or `docker ps -a`

3.1.2. RESULTS

Example:

Checking the OCB container, we obtain the following result:

```
fiware/orion:3.1.0    Up 2 months (healthy)    0.0.0.0:1026->1026/tcp          fiware-orion
```

Checking the Mongo database container, we obtain the following result:

```
mongo:4.4             Up 2 months (healthy)    0.0.0.0:27017->27017/tcp        db-mongo
```

3.2. ADVANCED ANALYTICS

Performing advanced analysis on the data collected allows you to extract the most relevant and interesting information from them, so it is advisable to perform tests to ensure that the information extracted is reliable and meets certain requirements in terms of accuracy, for example.

Among these analyses can be found:

- Real-time analysis tools to manage data streams.
- Event flows management thanks to a CEP (Complex Event Processing) component to extract higher level information and derive conclusions from it.
- Specialized tools for prescriptive and predictive analysis.

3.2.1. ACCURO

In one frame per second, all the objects are detected, and counters are implemented in the algorithm for each object in order to send the number of objects per each classification.

The output of the data is a json format that contains the number of objects detected, for example, number of people, number of suitcases, etc. Also, other information is included in the list such as location of the detection, date and hour.

This information can be used for other partners because they information provides the location of the detection, the date, the hour, and the number of objects. Thus, a partner can apply another algorithm for knowing in the same hour and location other information from the same instant of time.

3.2.1.1. TESTS

When the algorithm is detecting the objects, can be visually represented in the image using boxing boxes. During this test, it has to be verified that the level of confidence (from 0 to 1) is appropriate in order to avoid false positives. For example, if we see that some objects are detected well with a confidence of 0.6 then, in the algorithm this value can be implanted and

only objects detected with minimum 0.6 can be passed and send it to the database. Otherwise, if we want a high accuracy of the object detection, the level of confidence should be more than 0.8.

3.2.1.2. RESULTS

Using the validation of the test through the adjustment of the confidence, good results can be obtained but there are some circumstances that need to be taken into account. For example, Figure 4 shows a total number of 8 people. If we analyze the image, we can see a group of people in the center with 5 people, but in reality, there are 3 people behind of 2 people, a situation that can be noticed if we take into account people's legs. In this case, we can see that the algorithm detects 2 bodies and 1 face, it cannot detect the 5 people, but it is because in this frame we cannot see well the characteristics of the other 2 people.

Therefore, in order to detect all the objects, the camera has to have a good angle to capture all the characteristics of the object. Also, it is important to avoid visual illusions for avoiding false positives such as mannequins or sculptures with human figures that can be in a street.

Moreover, the distance is also important to cover only distances up to 30 m; otherwise, if the distance is too high, the algorithm cannot identify the objects' details and, thus, they cannot be detected.

Finally, even in difficult situations like Figure 4, we can see that the algorithm can detect objects with an accuracy of 80 % considering a maximum distance of 30 m. In ideal situations the algorithm can detect 100% of the objects.

3.2.2. BEIA

BEIA used the air quality monitoring station for pollutants frequently exceeded in cities (PM1, PM2.5, PM10), combining data from sensors with algorithms. These technologies have been implemented to provide air quality information with a resolution that allows accurate identification of areas where pollutants exceed limit values.

3.2.3. NOMMON

The data described in section 2.1.5 was use to generate the following descriptive (observed) and predicted indicators:

Overnights: Number of tourist or residents sleeping in each zone, of a previously defined zoning od the region of study, segmented by nationality, age, gender and income (for national visitors), type of visitor (resident, national, international), trip length (number of days in the region of study).

Daily presence: number of unique visitors to each zone during the day segmented by nationality, age, gender, income and type of activity (for national visitors), type of visitor (resident, national, international), trip length (number of days in the region of study) and visit duration (hours spent on the given zone).

Hourly presence: number of unique visitors to each zone during the hour of the day segmented by nationality, age, gender, income and type of activity (for national visitors), type of visitor (resident, national, international), trip length (number of days in the region of study) and visit duration (hours spent on the given zone).

In addition to the extraction of indicators the solution developed includes the characterisation of the different zones according with the type of visitors it attracts and the activities (time of the visit are performed in them).

3.2.3.1. TESTS

Different experiments were performed to test, either directly or in an indict way, the accuracy and plausibility of the results obtained with the developed methodology.

The descriptive solution, i.e. the estimation of the defined indicators, was tested in an indirect way. There is not ground truth to be compared with the results obtained for the estimation of indicators, this is why the developed methodology is so relevant. Hence the estimation was tested in an indirect way and aided by the judgement of experts. The disaggregated indicators obtained with the developed solution were aggregated and compared with available aggregated official statistics. Also the visits patterns obtained for different nationalities were analyzed and contrasted with contextual information (e.g national holidays, etc.) to assess the plausibility of the observed trends.

The methodologies developed for the zones characterization and prediction of indicators were tested in Madrid for the month of April 2019. In order to test perform this test the city of Madrid was divided in zones of 1 km^2 and the defined indicators were calculated for each zone for the month. There indicators were used to characterize the zones according with the observed behavior of tourist for the defined indicators. A series of relevant touristic points of interest were identified by experts and the resulting classification was analyzed to assess the plausibility of results.

The last week of April was used to test the predictive model. Predictions of the three defined indicators were made for each zone and compared with the actual estimated values from this comparison both the absolute and the relative error were computed.

3.2.3.2. RESULTS

Descriptive solution:

Figure 23 and Figure 24 exemplify the comparative analysis performed between the total number of UK and France, respectively, residents' entries (to the country) segmented by length of stay and (overnights in the country) and main final destination in the country obtained from official statistics and estimated with mobile phone data. Each point of the data represents a segment and the value is the share of this segment over the total. It can be seen that there is a high correlation, around 0.96 for UK and 0.95 for France, between both data sources. This analysis was performed for all identified nationalities and the results are equally good.

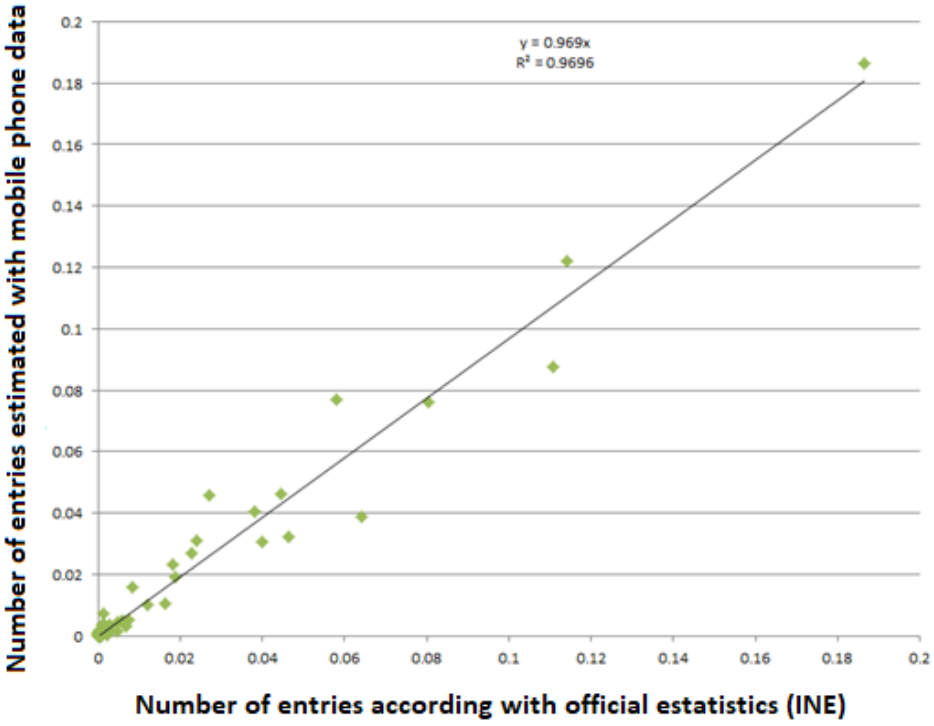


Figure 23. Comparison of arrivals and departures for users residing in the United Kingdom segmented according to nationality, length of stay and main destination Autonomous Community

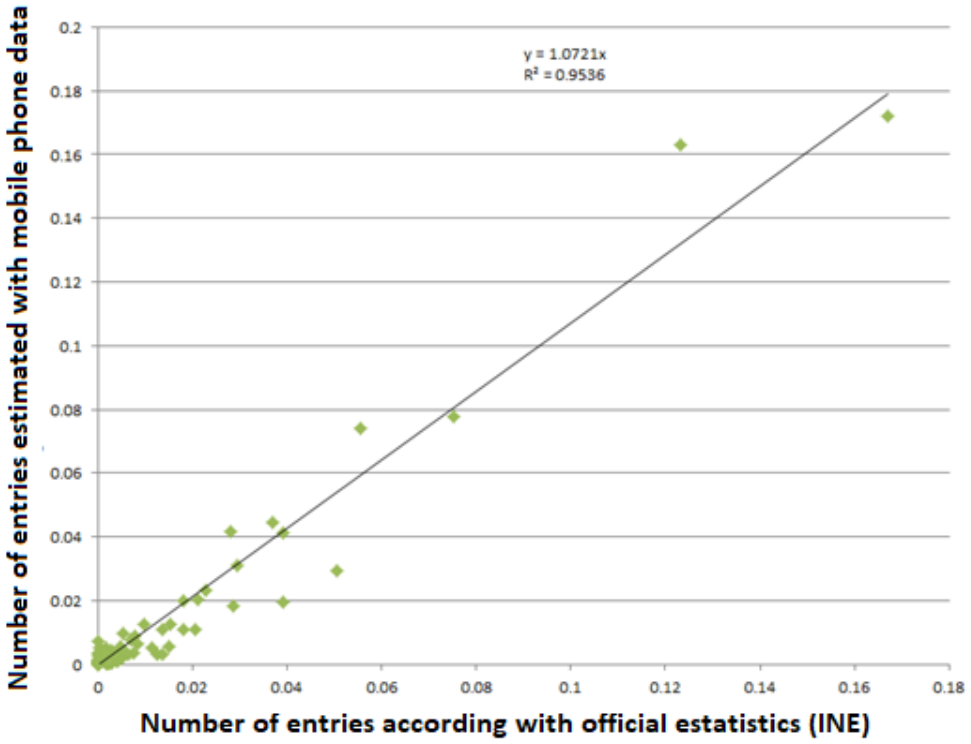


Figure 24. Comparison of arrivals and departures for users residing in the France segmented according to nationality, length of stay and main destination Autonomous Community

Figure 25 shows the profile of overnight stays of Portuguese users in Spain during the month of April. It is interesting to analyze two peaks. The first one that occurs between April 6 and 13, 2019, which means a significant volume of overnight stays corresponding to stays of between 4 and 7 overnight stays and between 8 and 15 overnight stays. As it has been verified, on April 8, Monday, a one week holidays period begins in Portuguese schools, this being the cause of the high volume of stays of Portuguese users that begins on April 6, Saturday. The duration of the stays in Spain fits with the duration of this school's holiday period.

On the other hand, a high peak of stays of between 2 and 3 days can be seen, beginning on April 25, Thursday. In Portugal, April 25 is a holiday, coinciding with the celebration of the carnation revolution. Because the holiday starts on Thursday, the high peak corresponds to people visiting Spain between Thursday and Sunday.

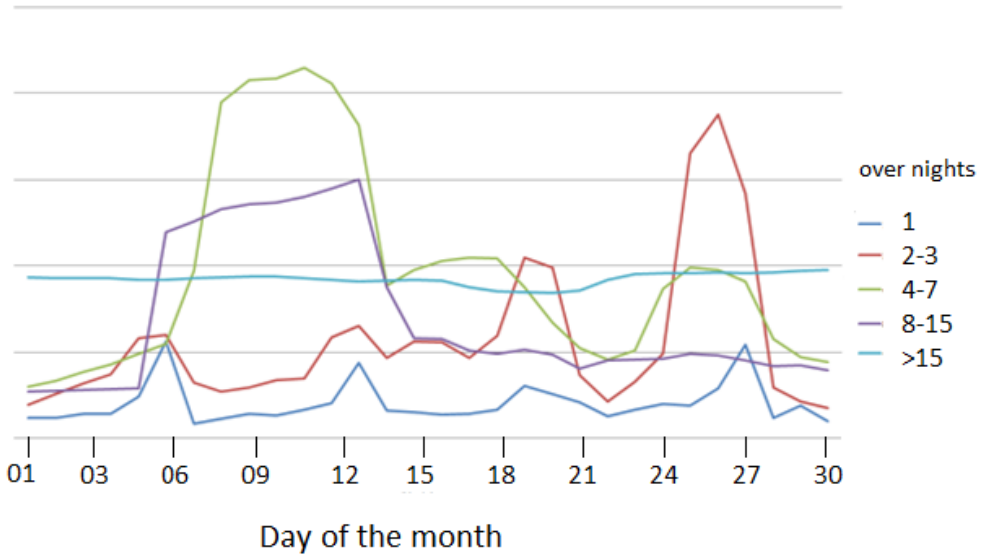


Figure 25. Number of Portuguese visitors sleeping in Madrid during the month of April, segmented by the number of overnights.

This analysis was performed with for all identified nationalities and serve to test and validate the solution.

Zones classification:

For each of the defined indicators different classes were obtained for the test case of Madrid. For overnights 7 classes, also called clusters, where identified

- University and logistics cluster (cluster 0): overnight stays of mainly national residents and visitors. Cluster with little tourist interest.
- Tourism cluster (cluster 1): overnight stays by foreign visitors mainly.
- Residential cluster (cluster 2): majority overnight stays by residents.
- Peripheral cluster with few overnight stays (cluster 3): peripheral areas with overnight stays mostly by residents. Cluster without tourist interest.
- Cluster without overnight stays by visitors (cluster 4): only overnight stays by residents.
- Cluster without overnight stays (Cluster 5).
- Mixed cluster (cluster 6): balanced overnight stays of residents, national and foreign visitors.

The classification for daily presence consists of 5 clusters

- Mixed cluster (cluster 0): the presence of residents, national visitors and international visitors is balanced. Labor, residential and a little tourist cluster.
- Residential cluster (cluster 1): presence of national residents and visitors.
- Tourism cluster (cluster 2): greater presence of visitors, mainly foreigners.

- Logistics cluster (cluster 3): Mercamadrid and other relevant logistics areas, with mainly national residents and visitors. Cluster without tourist interest.
- Residential cluster (cluster 4): majority presence of residents.
- Cluster of foreign visitors (cluster 5): satellite building of Terminal 4 of the Barajas airport, with a predominant presence of foreign visitors.

Finally, the classification for hourly presence consists of 9 clusters:

- Cluster of industrial estates and isolated logistics, commercial and sports areas (cluster 0).
- University cluster (cluster 1): contains the public universities of Madrid and IFEMA.
- Mercamadrid cluster (cluster 2).
- Residential cluster (cluster 3): residential areas in the south and north of Madrid.
- Cluster of the Wanda Metropolitano Stadium (cluster 4): contains only the stadium area.
- Airport logistics cluster (cluster 5): contains only one airport logistics area.
- Airport logistics cluster (cluster 6): contains only one airport logistics area.
- Cluster with restaurant and isolated recreational area (cluster 7): contains only one area near El Pardo.
- Labor and tourism cluster (cluster 8): downtown areas with offices and shopping areas.

To assess the plausibility of the characterisation of each zone accordingly with the previously mentioned classes contextual data was used (number of restaurants, hotels, universities, etc.). Table 1 present a list of PoI, their main touristic characteristics and the class they belong to. We can see that those areas with tourist attraction belong to the touristic classes for each of the indicators. This serves as a validation the obtained results.

Table 1. Cluster location of the selected POIs for the three zone classifications

Type attribute	Name	Presence cluster	Overnights cluster	Hourly presence cluster
Museums	Prado National Museum	2	1	8
	Thyssen-Bornemisza Museum of Art	2	1	1
	Reina Sofia Museum	2	1	8
	Museum Sorolla	2	1	8
	National Archaeological Museum	2	1	1

Type attribute	Name	Presence cluster	Overnights cluster	Hourly presence cluster
Tourist attractions	Retiro Park	0 and 2	1 and 6	1 and 8
	Royal Palace of Madrid	2	1	8
	Santiago Bernabéu Stadium	2	6	8
	Gran Via	2	1	8
	Plaza Mayor	2	1	8
	Plaza de Cibeles	2	1	1
	Neighbourhood Salamanca	0 and 2	1	1
	Debod Temple	0	1	8
	El Capricho Park	2	1	3
	San Miguel Market	2	1	8
	Cibeles Palace	2	1	1
	Plaza Santa Ana	2	1	8
	Puerta del Sol	2	1	8
	Crystal Palace	0	6	8
	Roof of Círculo de Bellas Artes	2	1	8
	Puerta Alcalá	2	1	1
	Almudena Cathedral	2	1	8
	Atocha Station	0	1	8
	San Antón Market	2	1	8
	Platform of Chamberí Station	2	1	8
	Plaza de España	0 and 2	1	8
	Neighbourhood La Latina	2	6	3

Type attribute	Name	Presence cluster	Overnights cluster	Hourly presence cluster
	Botanic Garden	2	1	8
	Plaza Callao	2	1	8
	Plaza de Oriente	2	1	8
	Fuencarral Street	2	1	8
	Congress Madrid	2	1	8
	Plaza de Colón	2	1	1
	Calle de Alcalá/Goya	0 and 2	1 and 6	1 and 8
	Real Theatre/Opera	2	1	8
Universities	Universidad Autónoma de Madrid (UAM)	1	0 and 2	1
	Universidad Complutense de Madrid (UCM)	0	0	1
	Universidad Politécnica de Madrid (UPM)	0	0	1
	Universidad Rey Juan Carlos - Vicálvaro Campus (URJC)	1	2	3
Logistic areas	Mercamadrid	3	0	2
	Valdemingómez recycling plant	3	0 and 1	3 and 8
Airport	Airport	2	1	0,1,5,6,8
	Airport T4S (satellite building of Terminal 4)	5	1	8
Other points of interest	Madrid Río Park	1	0	3
	Plaza de Toros de las Ventas	0	6	3
	El Rastro	2	1 and 6	3
	Cable Car/Teleférico Madrid	0	6	8

Type attribute	Name	Presence cluster	Overnights cluster	Hourly presence cluster
	Trade Fair Madrid IFEMA	3	0	1
	Attraction Park Madrid	1	2	3
	Parque del Oeste/ East Park	0 and 3	6	8
	Planetarium Madrid	1	0	3
	Caixa Forum	2	1	8
	Zoo Aquarium of Madrid	0	6	8
	Matadero Madrid	1	0	3
	Casa de Campo Lake	0	0	8

Show the distribution of official accommodation among all the clusters belonging to a given class. We can see that the zones belonging to the touristic class (cluster 1) for the overnights indicator cover 71% of the offer of hotels, 95% of guest houses, 21% of pubs and 45% of the available Airbnb accommodations.

Overnights cluster	Hotel	Guest house	Pub	Airbnb
0	0.06	0.02	0.13	0.08
1	0.71	0.95	0,21	0.45
2	0.02	0.01	0.36	0.19
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0.21	0.02	0.30	0.28
Total number	224	254	657	17236

Predictive solution:

Figure 26, Figure 27 and Figure 28 present prediction of the three defined indicators: overnight, daily presence and hourly presence, respectively, and the relative error and the absolute error obtained by comparing the predictions obtained with the actual values. It can be seen that those areas with a higher absolute error correspond to areas with a higher number of visitors. Hence although high in relative terms this error represents a low percentage of the total volume. Each dot represents a zone, the horizontal bars represent the absolute error and the value in the horizontal axis represents the relative error (the percentage of the total). It can be seen that those areas with a higher absolute error correspond to areas with a higher number of visitors. Hence although high in relative terms this error represents a low percentage of the total volume.

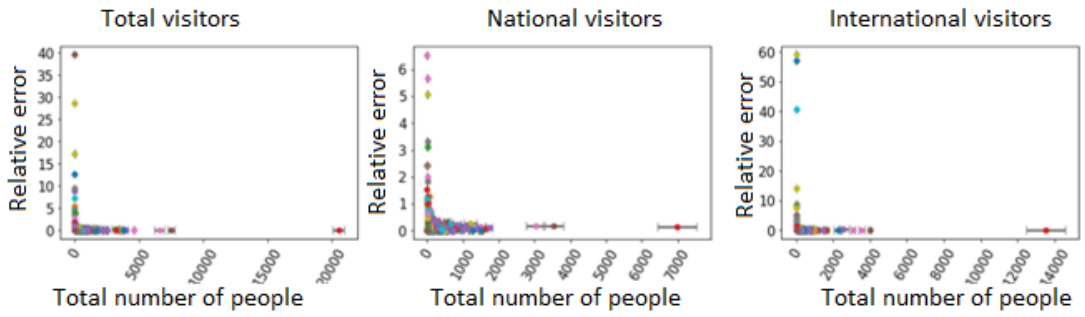


Figure 26. Relative error of the predicted number of visitors sleeping in each zone of the city as a function of the predicted volume. The bars indicate the absolute error.

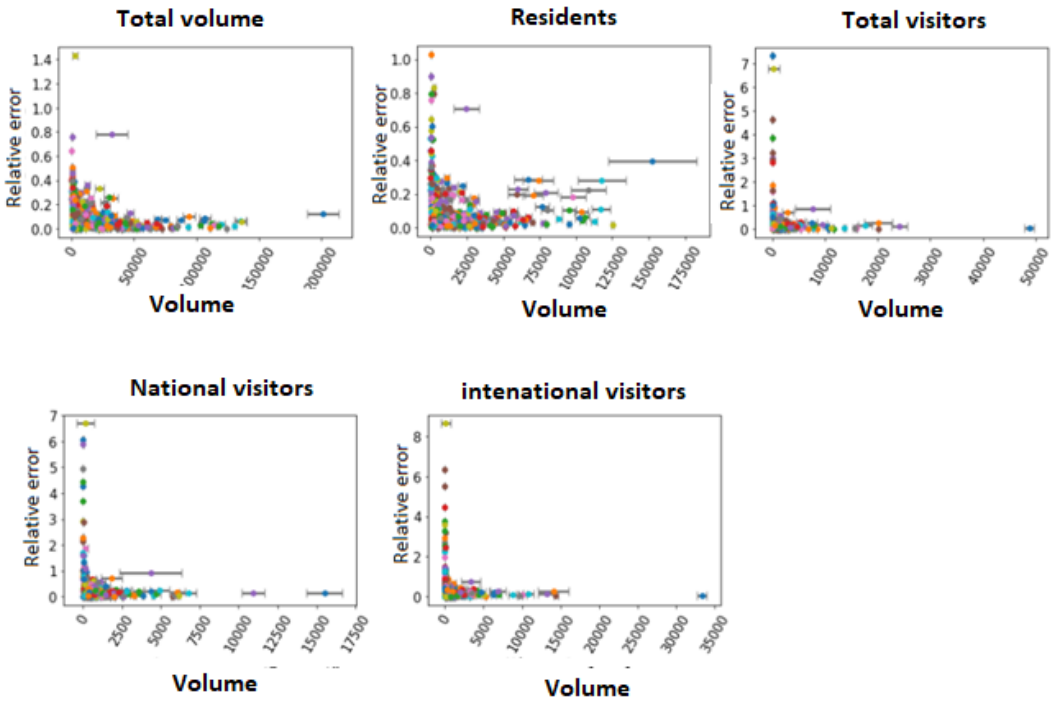


Figure 27. Predictive errors for estimating presence at the zone level for April 24, 2019.

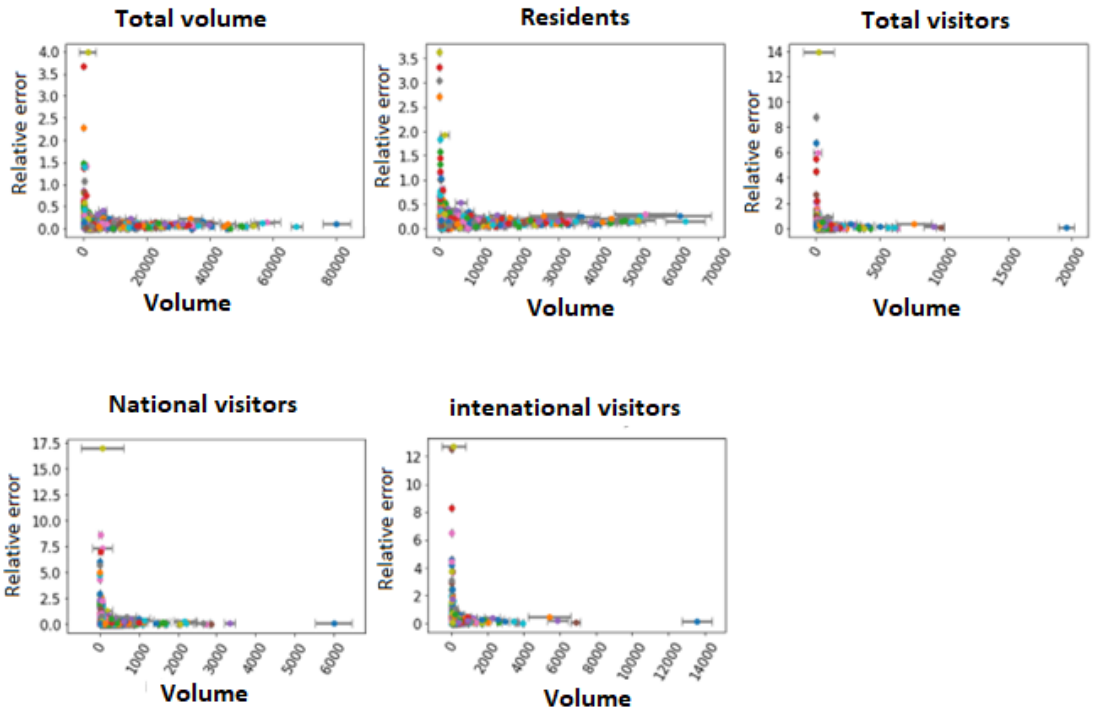


Figure 28 Predictive errors for estimating presence at the zone level for April 24, 2019 at 12:00.

4. INTEROPERABILITY LAYER TESTS

The interoperability layer provides a standard interface to allow third parties to interact with the platform so that they can access data, consume information or make use of the facilities provided by the platform.

Authentication and authorization aspects are going to control access according to the established terms of use and security.

One of the functional units included in this layer is a visualization module that provides some monitoring and reporting mechanisms.

4.1. POLDER PLATFORM DASHBOARD

Within this layer is the platform with which the end users will interact and with which they will be able to access the information obtained from the analysis and contained in the context broker, in the knowledge layer.

This platform is a dashboard that gives access to the services offered by the data providers, as well as to the visualization tools and control panels developed by them.

4.1.1. TESTS

<List and describe the tests performed>

4.1.2. RESULTS

<Following the same schema as for the tests, provide a description and analysis of the tests' results>

4.2. DASHBOARDS AND PLATFORMS FOR DATA VISUALIZATION AND ANALYSIS

This section contains the tests and validation of the visualization tools developed by the data providers, which are integrated into the platform dashboard.

4.2.1. ACCURO

ACCURO has worked in a total amount of four use cases framed in the Smart Tourism domain where each one is prepared for a situation. The specific objective of each of the use cases are described below along with the main graphics that compose it.

It is important to note that all the dashboards – which were already presented in Deliverable 7.2: *Software design and integration document* – are accompanied by a summary table of the main goal by hours and a series of filters through which the results are analyzed according to needs.

4.2.1.1. CASE 1

In this case, the objective is to detect the number of people, backpacks, bags and suitcases in order to estimate the number of tourists walking through the area.

In this case, people who have backpacks and suitcases can give more information to estimate the number of tourists, since it is not common for a native person to walk with a suitcase in a tourist area. Therefore, these objects have more weight in the KPI.

The visualizations that make up the first use case are focused on searching for occupancy levels. These visualizations are comprised of:

- A line diagram showing the different items that can be detected, bags, suitcases, backpacks, and people.
- A pie chart that shows the distribution of objects over certain pre-established periods of hours. Distributing the density of objects by hours.
- A line diagram of the same elements as the previous one but showing the values per minute, achieving a more precise follow-up of the evolution of the values.

CASE 1 TESTS

To corroborate the use of these graphs, a series of indicators have been generated: the first one shows the time of maximum occupancy; the second indicator shows the best time to make a visit, that is, the moment in which the occupation has the lowest level. These indicators collect the values shown in the different graphs in a concise manner and with the main objective of this use case.

4.2.1.2. CASE 2

In this case, a method to be able to calculate the safety distance in an area of the desired image has been developed. In this use case, new objects such as bicycles, cars, buses, trucks, motorcycles, horses, etc. have been detected.

This type of detection can be used to find out whether said vehicles can travel or park in certain areas. For instance, if a camera is installed in a pedestrian zone and suddenly one of these vehicles appears where the passage is prohibited, an alarm can be sent indicating that these vehicles cannot travel or park in prohibited areas.

The main objective of this case is to detect unauthorized people and vehicles, together with the breach of an established safety distance between people, following COVID-19 regulation. For this, a total of four graphs have been used:

- A line diagram that detects the density of people per minute.
- A line diagram that detects tourists per hour.
- A diagram that combines both lines and bars to perform an analysis of authorized and unauthorized vehicles.
- A line diagram for pet detection.

CASE 2 TEST

The conclusion of these graphs has been made with the representation of several indicators, summarizing the main objectives of the visualization:

- A first indicator that shows the time at which there is a greater number of individuals not complying with the established safety distance.
- A second indicator that shows the number of unauthorized vehicles in the selected period.

4.2.1.3. CASE 3

Due to the safety measures adopted for the health emergency of the last two years, caused by COVID-19, this use case focuses on the detection of facemasks in people and the analysis of periods of maximum occupation to avoid times with high density of unmasked population.

This use case's visualization panel consists of four graphs:

- The number of people per hour who access a place without a mask is represented through a line diagram.
- A pie chart that shows the visits without a mask distributed by certain previously generated schedules.
- A line diagram showing people with and without masks. It is different from the first graph since the first only shows people who are not wearing a facemask, while in this case it detects all the people who appear on camera.
- A table with the hours established in the pie chart and shows the visits made throughout the week without a mask. The table runs from Monday to Sunday and distributes the visits along the generated schedules. With this information, a more detailed view of the visits per hour is obtained.

CASE 3 TEST

For this use case, only one indicator has been used, it shows the maximum occupancy time of people without masks. This result combines both the population density in the analyzed area and the moment of greatest probability of contagion.

4.2.1.4. CASE 4

For this last use case, two independent algorithms have been carried: one to detect people in a pedestrian area and another to detect swimmers.

Two visualization panels have been created with the aim of establishing the maximum density of people in a determined area, differentiating between bathing areas (swimming pools and beaches) and pedestrian areas. This time clustering has been made by groups of people.

The first visualization, corresponding to pedestrian areas, consists of three graphs, being these the following:

- A bar diagram that represents the density of people and vehicles per hour.
- A pie diagram that represents the average number of groups per hour, different distinctions have been generated according to the size of the group.
- A bar chart that represents people per minute.

For the second visualization (bathing areas), an analysis of bathers in a certain area has been carried out, mainly linked to density. Secondly, an analysis of unauthorized vehicles by time slot has been carried out. Three graphs have been used:

- A line diagram that shows the bathers and unauthorized vehicles per hour.
- A table that distributes by hours and days of the week the average number of vehicles there are in the different established time slots.
- A plot of lines representing the average number of swimmers per hour.

CASE 4 TEST

For the first visualization, a single indicator has been used, which shows the time with the greatest grouping of groups of more than six people.

For the second visualization, two indicators have been used, the first of which is focused on bathers and shows the time of maximum bather occupancy; the second indicator shows the time with the highest occupancy of unauthorized vehicles.

4.2.2. ACD

The reason why we prefer Grafana is that it is able to extract data from PostgreSQL and able to draw time series graphs. PostgreSQL is a free and open-source relational database management system emphasizing extensibility and SQL compliance.

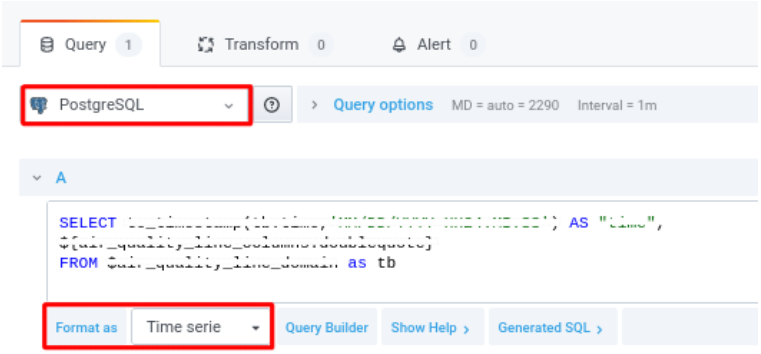


Figure 29. Grafana

Grafana is found to be compatible with PostgreSQL and fiware components.

4.2.2.1. TESTS

4.2.2.1.1. ENERGY

These graph represent the total cost by the devices and the cost separately for each device. The time data of this drawing are samples taken at intervals of 5 minutes in a day. Results can be displayed from the platform.

When area graph and energy after cost optimization are selected:

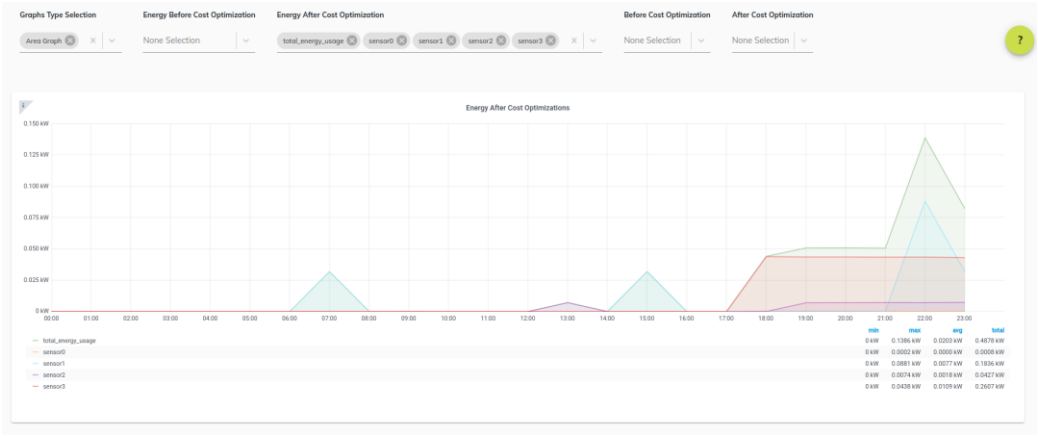


Figure 30. Energy line plot

When table (raw data) and energy after cost optimization are selected:

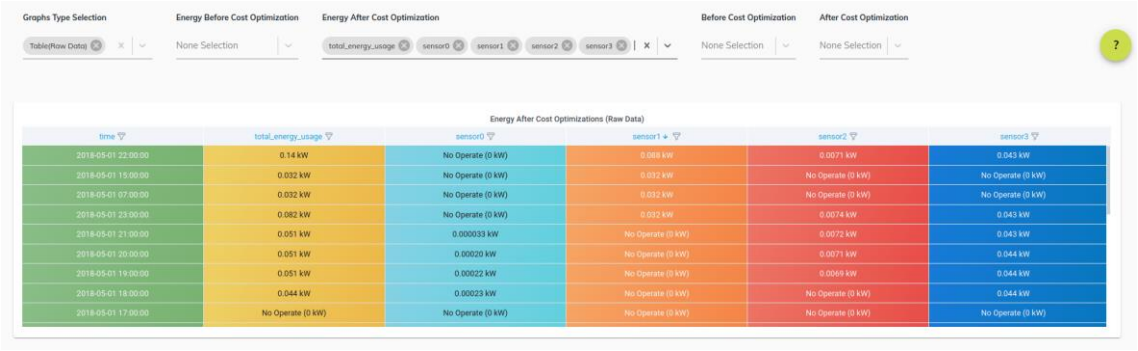


Figure 31. Energy table

When chart and energy after cost optimization are selected:

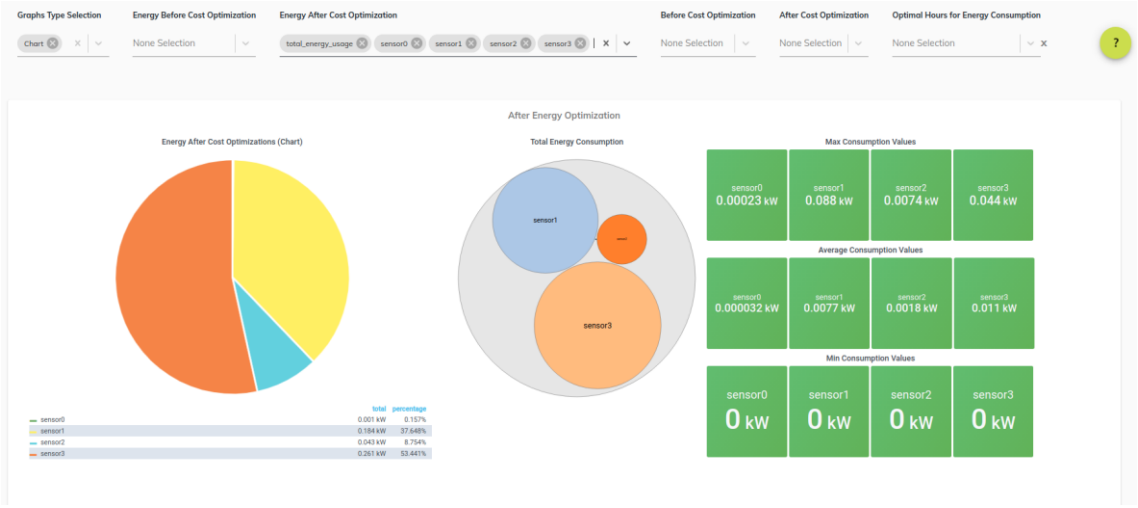


Figure 32. Energy pie chart

4.2.2.1.2. AIR QUALITY

Analyzing and detecting abnormal values of the sensors for air quality is a crucial policy for health protection. Detection of anomalies, sending warning messages and visualizing the sensor values helps to suppliers for management of the air quality. In addition, anomalies in the data coming from the sensors are detected and kept in table form in the "Logs" section. Thus, feedback is provided to the user if an anomaly occurs.

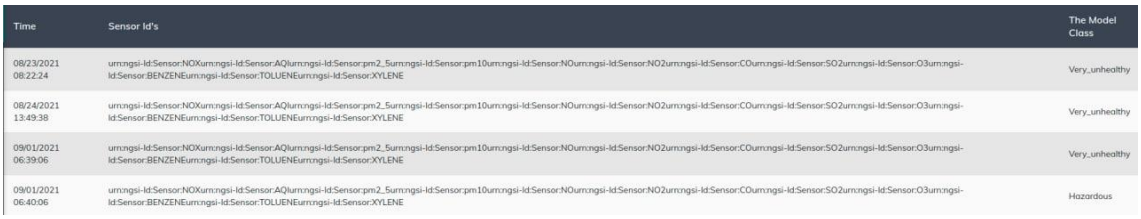


Figure 33. Log file

If selected Notification System, there are three graphic types. These; Line, Distribution and Histogram. In this context Benzene,CO, NO, NO2, NOX, O3, PM10, PM2_5, SO2, Toluene and Xylene are Notification System Attributes.

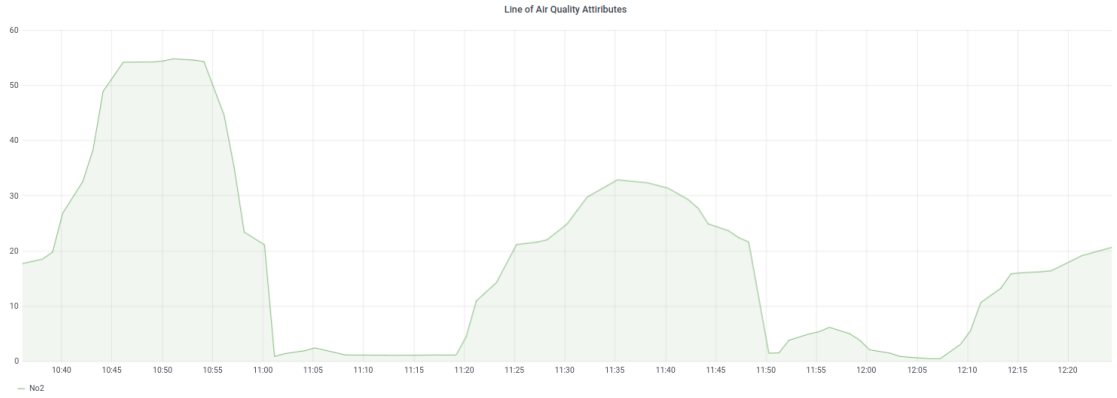


Figure 34. Grafana line chart

Due to DQN algorithm is a reward based approach, it can not be measured performance by using accuracy metric. It can be evaluated intuitively. We used DQN algorithm is a recomender system to provide a optimal usuage interval. Looking at the plots shown below, the proposed algorithm recomends cheaper time interval among the user specified desired interval.

4.2.2.2. RESULTS

4.2.2.2.1. ENERGY

In order to observe energy consumption and cost values before and after DQN algorithm more clearly, the graphs in which these values are plot separately. Looking at the plots shown, the proposed algorithm recommends cheaper time interval among the user specified desired interval.

4.2.2.2.2. AIR QUALITY

Outputs of the algorithms are understandable. The platform visualizations provide information about air quality to users. When an abnormal condition occurs, the Notification system provides warnings and provides convenience to the user.

4.2.3. ARD

4.2.3.1. TESTS

Project Name:	POLDER	**Web Site Link:**	http://smartraffic.ardsistem.com.tr/#/login

Table 3. Test scenarios for ARD's data analytics tool

User Story / Requirement ID	Test Scenario ID	Test Scenario Description	Page	Priority
US0001	TS0001	Verify if a user will be able to login with a **valid username** and **valid password**.	Login	High
US0002	TS0002	Verify if a user cannot login with a **valid username** and an **invalid password**.	Login	High
US0003	TS0003	Verify the login page for both, when the field is **blank** and Submit button is **clicked**.	Login	Medium
US0004	TS0004	Verify the messages for invalid **login**.	Login	Medium
US0005	TS0005	Verify the 'Remember Me' functionality.	Login	Medium
US0006	TS0006	Verify if the data in password field is either visible as asterisk or bullet signs.	Login	Medium
US0007	TS0007	Verify if a user is able to login with a new password only after he/she has changed the password.	Login	Medium
US0008	TS0008	Verify if SQL Injection attacks work on the login page. The application should not be vulnerable to SQL injection attacks	Login	High
US0009	TS0009	Verify the search functionality by selecting the valid data from combobox(es).	traffic-estimation	Medium
US0010	TS0010	Verify the message when system not able to search any result.	traffic-estimation	Medium
US0011	TS0011	Verify that count of searched result showing or not.	traffic-estimation	Medium
US0012	TS0012	Verify the search functionality by entering the valid searched keyword in search box.	traffic-density	Medium
US0013	TS0013	verify that when user click on search button without entering any parameter.	traffic-density	Medium
US0014	TS0014	Check the time taken by system to display the result in grid.	traffic-density	Medium
US0015	TS0015	Verify the message when system not able to search any result.	traffic-density	Medium
US0016	TS0016	Verify the no of result in each page	traffic-density	Medium

User Story / Requirement ID	Test Scenario ID	Test Scenario Description	Page	Priority
US0017	TS0017	Verify that count of searched result showing or not.	traffic-density	Medium
US0018	TS0018	Verify the loading symbol when page take more time to display the result in grid.	traffic-density	Medium
US0019	TS0019	Verify the order of result are showing as per requirement or not.	traffic-density	Medium
US0020	TS0020	Check that user able to edit, delete, view …the search result or not.	traffic-density	Medium
US0021	TS0021	Verify the pagination, when there is more result in grid than default.	traffic-density	Medium
US0022	TS0022	Verify the Next and previous functionality working as per requirement or not.	traffic-density	Medium
US0023	TS0023	Verify the result when user enter partial searched keyword and click on search.	traffic-density	Medium
US0024	TS0024	Verify the clear search box when user click on clear.	traffic-density	Medium
US0025	TS0025	Verify the search functionality by selecting the valid data from Combobox(es).	analysis	Medium
US0026	TS0026	Verify the search functionality by entering the valid searched numeric up/down.	analysis	Medium
US0027	TS0027	Verify the search functionality by entering the invalid searched numeric up/down.	analysis	Medium
US0028	TS0028	Verify the message when system not able to search any result.	analysis	Medium
US0029	TS0029	Verify that count of searched result showing or not.	analysis	Medium
US0030	TS0030	Verify the search functionality by entering the valid searched keyword in search box.	notification	Medium
US0031	TS0031	verify that when user click on search button without entering any parameter.	notification	Medium
US0032	TS0032	Check the time taken by system to display the result in grid.	notification	Medium
US0033	TS0033	Verify the message when system not able to search any result.	notification	Medium
US0034	TS0034	Verify the no of result in each page	notification	Medium
US0035	TS0035	Verify that count of searched result showing or not.	notification	Medium
US0036	TS0036	Verify the loading symbol when page take more time to display the result in grid.	notification	Medium

User Story / Requirement ID	Test Scenario ID	Test Scenario Description	Page	Priority
US0037	TS0037	Verify the order of result are showing as per requirement or not.	notification	Medium
US0038	TS0038	Check that user able to edit, active and passive, the search result or not.	notification	Medium
US0039	TS0039	Verify the pagination, when there is more result in grid than default.	notification	Medium
US0040	TS0040	Verify the Next and previous functionality working as per requirement or not.	notification	Medium
US0041	TS0041	Verify the result when user enter partial searched keyword and click on search.	notification	Medium
US0042	TS0042	Verify for uploaded excel path	file-upload	Medium
US0043	TS0043	Verify excel upload functionality with excel files of different extensions (e.g. xls, xlsx)	file-upload	Medium
US0044	TS0044	Verify excel upload functionality with excels having space or any other allowed special character in file name	file-upload	Medium
US0045	TS0045	Verify duplicate name excel upload	file-upload	Medium
US0046	TS0046	Verify excel upload with excel size greater than the max allowed size. Proper error message should be displayed.	file-upload	Medium
US0047	TS0047	Verify if file selection dialog shows only supported files listed	file-upload	Medium
US0048	TS0048	Verify if user is able to use the uploaded excels	file-upload	Medium
US0049	TS0049	Verify text box required	camera	Medium
US0050	TS0050	Verify combo box required	camera	Medium
US0051	TS0051	Verify the required fields by not filling any data	camera	Medium
US0052	TS0052	Verify the optional field when do not fill data	camera	Medium
US0053	TS0053	Verify if blank spaces are passed in required fields.	camera	Medium
US0054	TS0054	Verify user can verify its Camera Code unique	camera	Medium
US0055	TS0055	Verify camera is displayed in maps.	map	Medium

User Story / Requirement ID	Test Scenario ID	Test Scenario Description	Page	Priority
US0056	TS0056	Verify for person view click on 'Street View' button and view should be displayed.	map	Medium

4.2.3.2. RESULTS

Table 4. ARD's data analytics tool test results

Test Case ID	Test Case Description	Test Steps	Test Data	Pass/Fail
TC0001	Check User Login with **valid** Data	1 Go to site http://smartraffic.ardsistem.com.tr/#/login	Email Adress = polder@ardgrup.com Password = 11!22?44.Polder	Pass
		2. Enter UserId		Pass
		3. Enter Password		Pass
		4. Click Login		Pass
TC0002	Check User Login with **invalid** Data	1 Go to site http://smartraffic.ardsistem.com.tr/#/login	Email Adress = polder1@ardgrup.com Password = 11!22?44.Polde1r	Pass
		2. Enter UserId		
		3. Enter Password		
		4. Click Login		
TC0003	Check User Login with **blank**	1 Go to site http://smartraffic.ardsistem.com.tr/#/login	Email Adress = Password =	Pass
		2. Enter UserId		
		3. Enter Password		
		4. Click Login		
TC0004	Verify that count of searched	Verify Camera combobox is loaed as expected	Camera : Kızılay Select Time : 20 Min	PASS
		Verify Select Time combobox is loaed as expected		
		Verify that the **search** button is displayed as expected		

Test Case ID	Test Case Description	Test Steps	Test Data	Pass/Fail
	result showing or not.	Click **search** button when no search criteria is entered		
		Verify that when a user types a query, the system generates auto suggestions as you type.		
TC0005		Check for the Window title of Import file/"Save as" window.	Camera : Kızılay Date: 01.06.2021	PASS
		Verify by importing the file with valid file extension. (xls, xlsx)		PASS
		Verify by clicking on the Browse button to select a valid path to import the file.		PASS
		Verify by importing Invalid file.		PASS
		Verify the message displayed when user imports the file successfully.		PASS
		Verify the display of the imported file records in the window.		PASS
		Verify the imported records.		PASS
		Verify by importing the file after entering max length characters in all the fields.		PASS
		Verify by entering invalid characters in the file and import.		PASS
		Verify by selecting invalid path to import the file.		PASS

4.2.4. NOMMON

The developed algorithms for indicators estimation and prediction as well as the zones classifications were integrated in an interactive visual analytic tool for the facilitation of tourism monitoring and scenarios analysis for tourism planning and management. To make more agile the results consultation from the visualization dashboard the integration was done via a data based. The results obtained from the described solutions are pre-computed and downloaded to a data base that is integrated with the visual interface.

4.2.4.1. TESTS

To carry out the integration tests of the developed tool, the repository data was be compared with the data shown by the visualization. The repository data is disaggregated, so to facilitate comparison. The presence and overnight stay data have been loaded into a dynamic table in Excel to perform aggregation for the testing exercise. The use of a dynamic table allows us filtering according to the different segmentations of the data. In addition, it will allow us to view the time profiles of hourly presence.

In carrying out the integration tests, the following variables have been taken into account to generate the widest possible test battery. The variables considered are: the indicator, daily presence, hourly presence and overnight stays; study population, residents, national visitors, foreigners or a combination of these; and segmentations of the indicators. The defined tests attend to a battery of possible combinations of these variables. The validation tests must cover each of the possible values of these variables.

The following tests have been defined:

- Validate the daily presence values for the entire population in one or several specific areas.
- Validate the daily presence values for the national population in an area for a given segmentation.
- Validate the hourly presence values for the foreign population based on a specific segmentation.
- Validate the hourly presence values for the national population according to a certain segmentation.
- Validate the values of overnight stays for the entire population in one or several specific areas.
- Validate the values of overnight stays for the resident population in an area for a given segmentation.

4.2.4.2. RESUTLS

Aggregated results from the use of Excel pivot tables are presented along with results presented in visualization using screenshots.

Daily presence for the entire population in one or several specific areas:Table 5 and Figure 35 show the results obtained:

Table 5. Daily presence for the entire population in different areas obtained from the Excel dynamic table.

Day	Zone id	Volume
2019-04-06	31	93.910
2019-04-06	32	77.515
2019-04-06	33	99.742

Figure 35. Daily presence for the entire population in various areas. As visualised in the platform

- Daily presence for the national population in an area for a given segmentation: Table 6 and Figure 36 show the results obtained.

Table 6. Daily presence for national population segmented by gender in different areas obtained from the Excel dynamic table

Day	Zone id	Gender	Volume
2019-04-06	31	Femenino	28.803
2019-04-06	32	Masculino	31.535

Figure 36. Daily presence for national population segmented by gender as presented by the platform

- Hourly presence for the foreign population according to a specific segmentation: Figure 37Figure 38show the hourly profiles obtained for the activity duration segmentation.

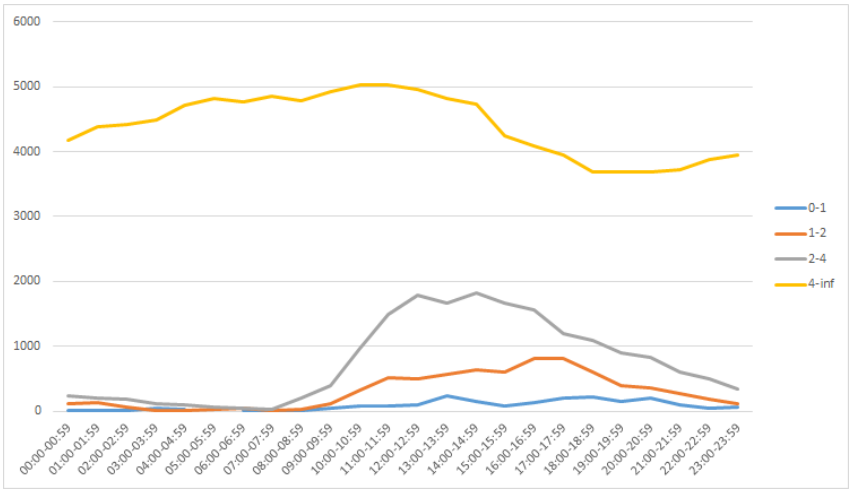


Figure 37. Hourly presence for the foreign population according to the duration of the activity obtained in Excel

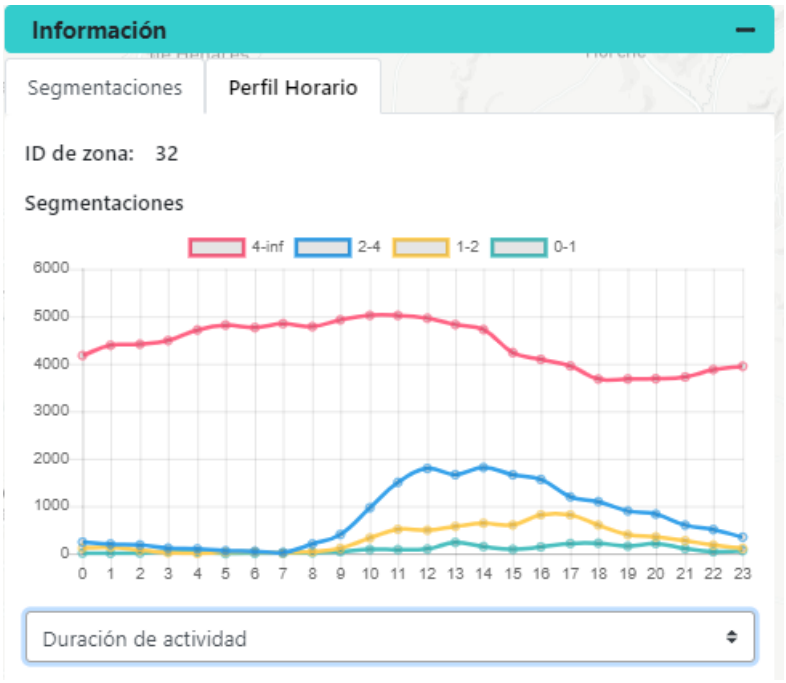


Figure 38. Hourly presence for the foreign population according to the duration of the activity in the visualization tool

- Hourly presence for the national population according to a certain segmentation. Figure 39 and Figure 40 show the hourly profiles obtained for the type of activity segmentation.

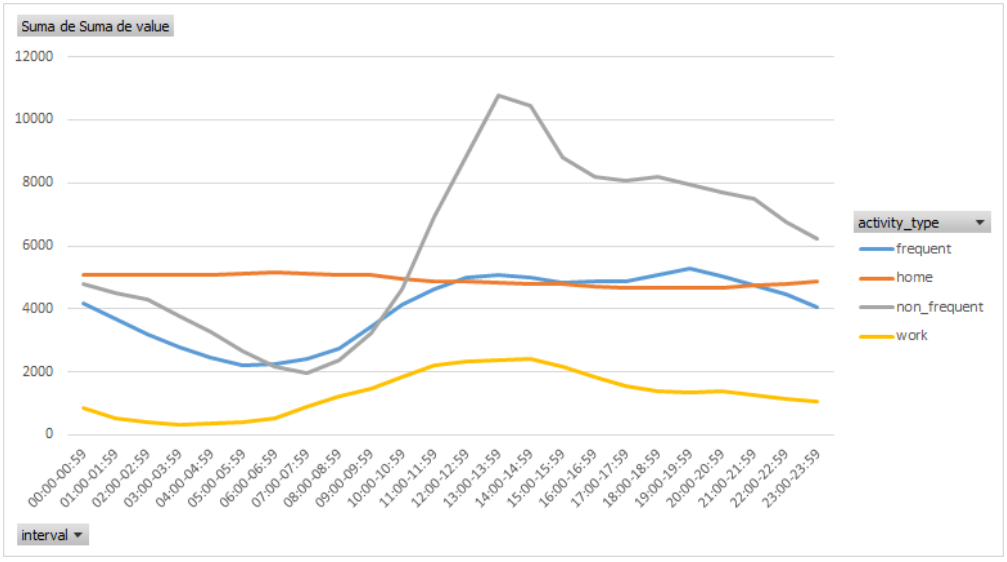


Figure 39. Hourly presence for the national population according to the type of activity from Excel

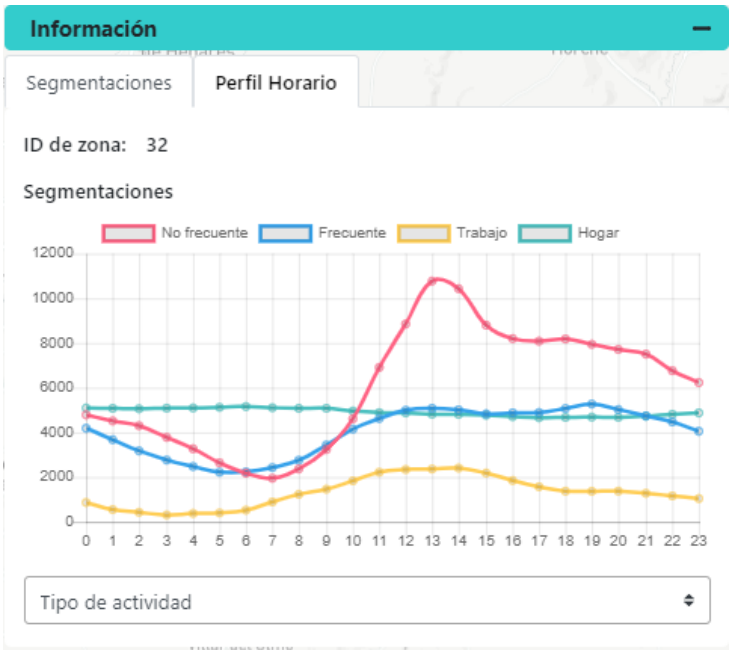


Figure 40. Hourly presence for the national population according to the type of activity as visualised in the platform.

- Overnight stays for the entire population in one or more specific areas. In Table 7 and Figure 41 you can see the results obtained from the comparison:

Table 7. Daily overnight stays for the entire population in different areas.

Day	Zone id	Volume
2019-04-06	31	21.627
2019-04-06	32	13.896
2019-04-06	33	23.911

Figure 41 Visualisation of the daily overnight stays for the entire population in different areas.

- Overnight stays for the resident population in an area for a given segmentation. In Table 8 and Figure 42 you can see the results obtained:

Table 8. Number of overnights for Spanish residents

Day	Zone id	Age	Volume
2019-04-06	32	A0	3290.522824
2019-04-06	32	A1	2824.984831
2019-04-06	32	A2	2653.159724
2019-04-06	32	A3	2412.473561

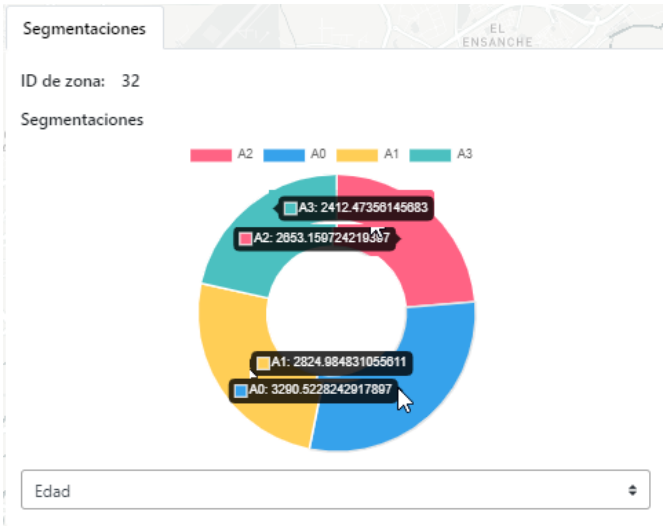


Figure 42. Number of overnight for Spanish residents represented in the visualusation tool.

As can be seen in the tables and figures above, in all cases the values obtained from the visualization and the existing data in the repository coincide.

5. INTELLIGENT SERVCES LAYER

The intelligent services layer provides services to end users or stakeholders interested in the management or optimization of some city/community resources handled through the platform. Smart mobility, air quality, water, energy, smart tourism, security, urban planning, etc. are some of the issues that could be running over POLDER Platform or they can be external services that publish or consume information. Services like web applications where resources can be monitored and controlled, alarm notification tools, administration services tools, training suites, etc. are some examples of services that can be developed to be provided by this layer.

Since it was not possible to test the platform in a real scenario such as the one in Lloret de Mar at the beginning of the project, different scenarios have been developed in which the partners

can demonstrate the capabilities of POLDER. The following sections show the results of these scenarios and the tests carried out in them.

5.1. SCENARIO 01: TRAFFIC AND PEDESTRIAN FLOW IN FRONT OF A BAR

VEXIZA and ACCURO have collaborated on the design and implementation in a laboratory environment of a common use case to demonstrate the potential of reusing and cross-referencing data from different domains.

As partners using object recognition in video images, have used images from the same source, but each extracting different features.

5.1.1. OBJECTIVES

This use case combines data from both the general use cases of Smart Tourism and Smart Monitoring, in particular population dynamics monitoring (ACCURO) and traffic monitoring (VEXIZA).

The partners have used as a data source a camera located at the entrance of the Tini Martini bar in Florida, which broadcasts its recording in real time on YouTube, freely accessible to everybody[1]. Five fragments of this broadcast of between 30 seconds and 4 minutes in length have been recorded, showing people walking down the street, people entering and leaving the bar, vehicles driving on the road and vehicles parking in front of the bar.

Data cross-referencing aims to analyze the density of people on the street and the density of cars, cross-reference this information and predict possible situations that can help restaurant operators analyze possible demand over a period of time, provide estimated waiting times for queuing customers or offer potential customers information on parking availability or traffic to get there.

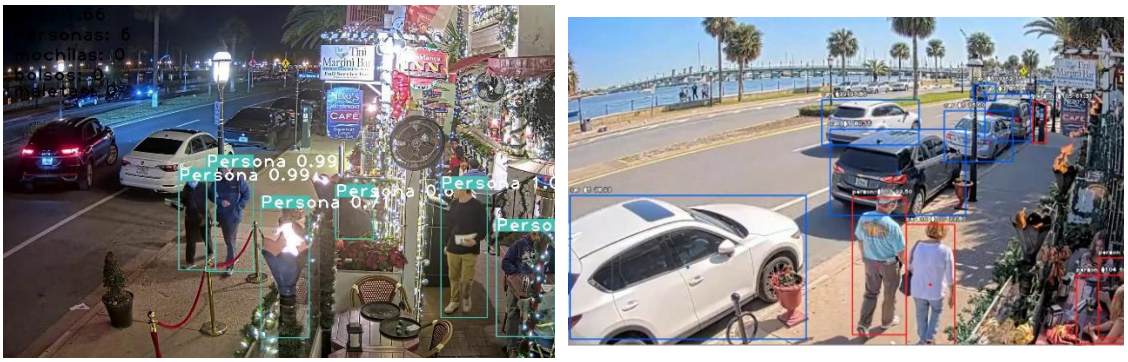


Figure 43. Detection of people (left) and vehicles (right) in the Tini Martini Bar

[1]The Tini Martini bar camera can be accessed via the following link: Tini Martini Bar St. Augustine FL - YouTube

Figure 44. Dashboard with crossed data from VEXIZA and ACCURO

The data obtained has been represented in a dashboard (Figure 44) in which the number of people detected per minute, the total number of vehicles identified per minute, and the type of vehicles detected at each moment can be observed. In addition, a table is included with the occupation of the street every hour and two indicators: one indicating the hour with the highest number of pedestrians and the other with the highest number of vehicles (maximum influx of visitors and time with the most traffic).

5.1.2. TESTS

The main tests that have been carried out have been with the intention of verifying that both VEXIZA and ACCURO had access to each other's data through the POLDER platform, whose ultimate goal is to guarantee access to the data provided by different partners in order to be able to cross-reference them.

The exchange of data could be carried out without any problem through this platform, with the exception that, as the analysis of the images was carried out at a different time, the data of both partners were not synchronized.

To solve this problem, ACCURO downloaded the data from VEXIZA and modified locally (not on the platform) the date of the records, taking into account the start and end time of the analysis of each video, and its own duration.

5.1.3. RESULTS

The tests were successful, allowing both partners to access the shared data without any problems.

It should be noted that the problem of data synchronicity is due to having worked in a laboratory environment independently; this problem would not exist if both had access to the same information at the same time in a real environment, so it cannot be considered a problem of the Platform, but of implementation in a real environment.

Finally, it should be noted that the main objective of the use case has been met: to demonstrate the possibility of cross-referencing data from different domains to provide stakeholders with a more complete view of what is happening in cities.

5.2. SCENARIO 02: AIR QUALITY AND ENERGY

For energy monitoring, real smart-city datasets were obtained from https://www.kaggle.com/. These datasets were converted to NGSIv2 format and sent to the platform.

With the scope of the air quality index prediction system, real sensor values are gathered through different APIs such as PlumeLabs. In this context, the sensor data and air quality index of 113 countries are gathered to train a random forest algorithm. The range of the AQI values is considered in order not to bias toward a narrow range. Therefore, the values are collected to allow the model to learn the pattern for all real-world scenarios.

5.2.1. OBJECTIVES

The objective of this scenario is making sure the accuracy rate is robust for the algorithms by working with real datasets.

5.2.2. TESTS

The tests performed are shown in Section 2.1.2.1.

5.2.3. RESULTS

The results performed are shown in Section 2.1.2.2.

5.3. SCENARIO 03: CITY MONITORING

5.3.1. OBJECTIVE

The goal of this scenario is to text the behavior of city monitoring tool developed by FORTEARGE under changing user loads.

5.3.2. TESTS

For testing the performance and behavior of city monitoring application, JMeter application is used. JMeter is an open-source test tool from Apache used to analyze and measure the performance of applications, different software services and products.

First, a typical application page is selected. This page is a dynamic resource which shows some data from database.

Then different test loads are created in the tool, which are:

- 1 user
- 5 users
- 10 users
- 20 users

When the tests run, the following response time graphs are generated by the tool for each user load:

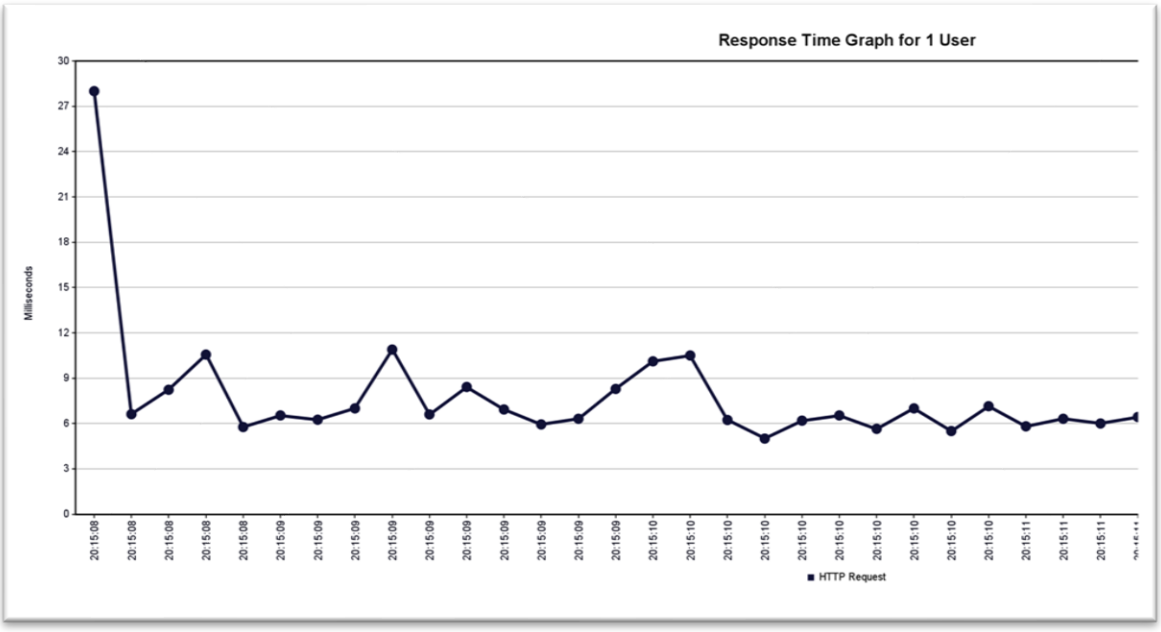


Figure 45. Response time graph for 1 user

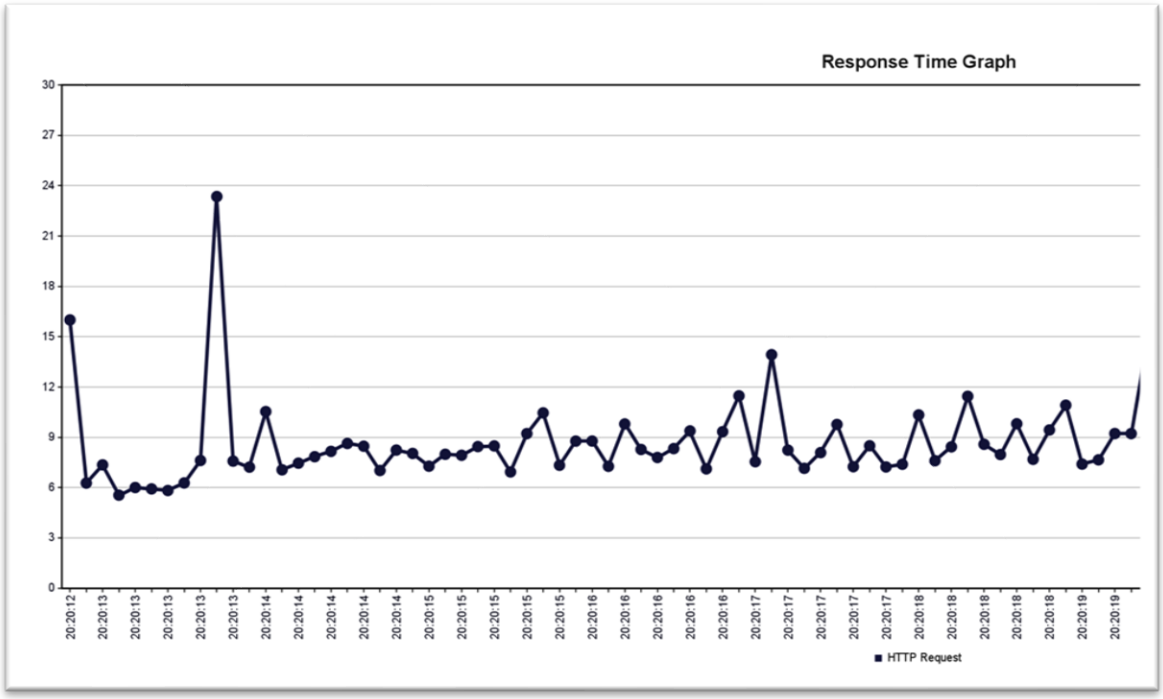


Figure 46. Response time graph for 5 users

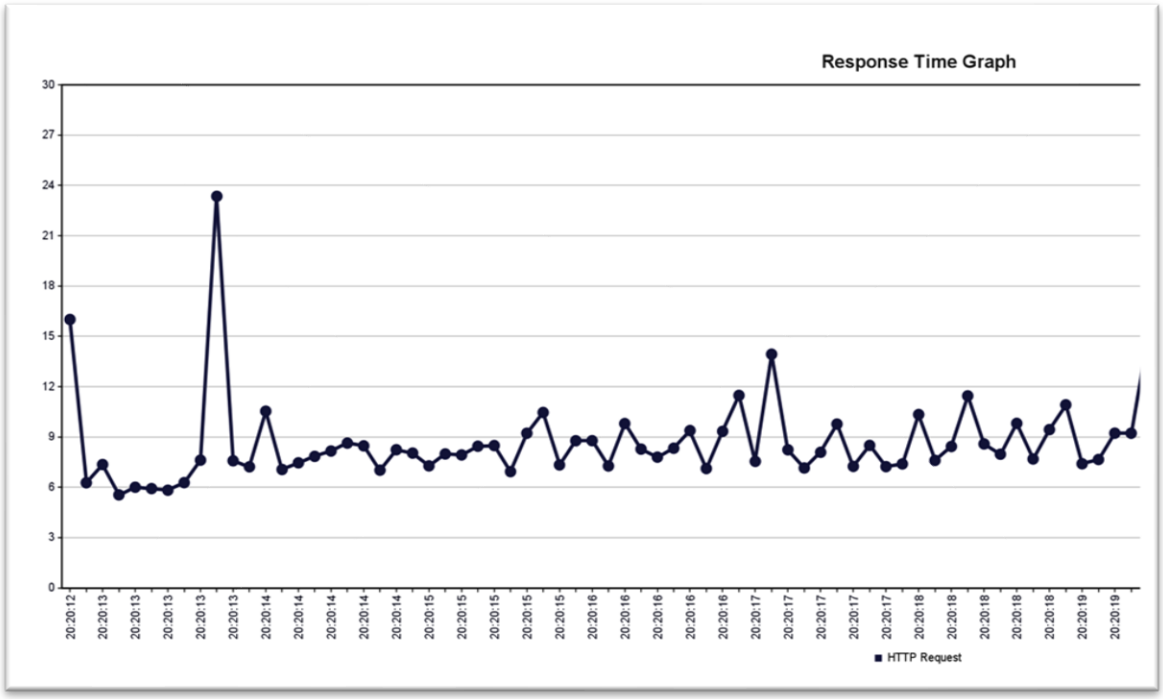


Figure 47. Response time graph for 10 users

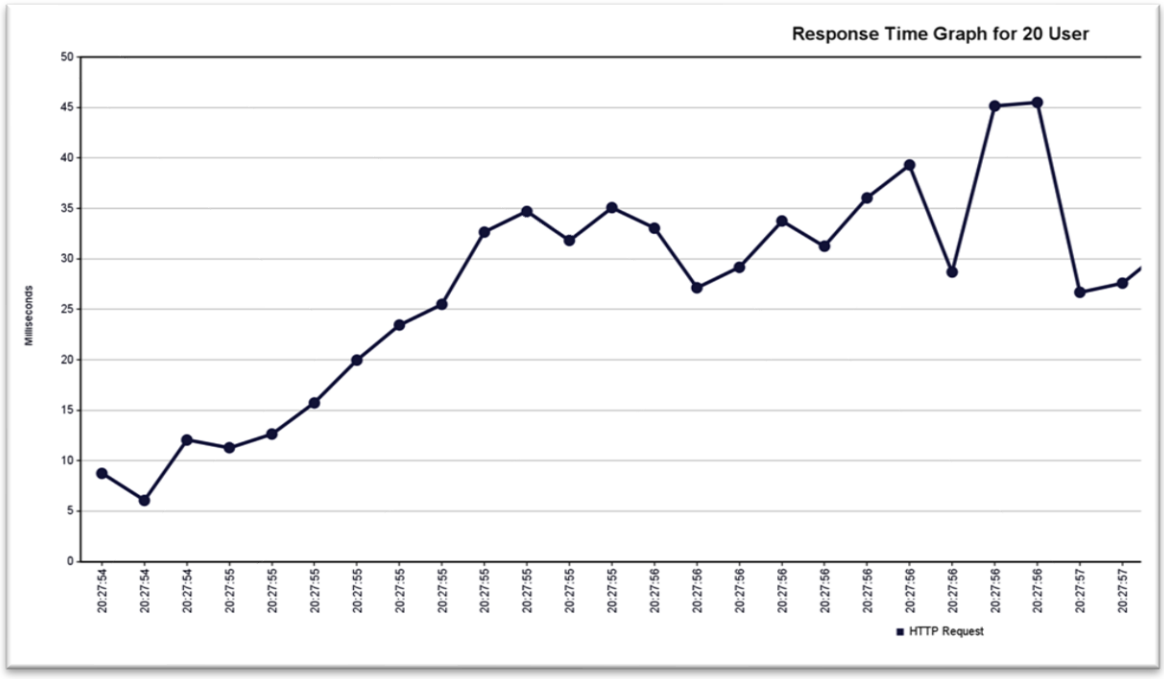


Figure 48. Response time graph for 20 users

5.3.3. RESULTS

The tests have shown that the behavior of the server application is as expected under changing user loads. The response times increase under user load in normal fashion, keeping the similar increase ratio.

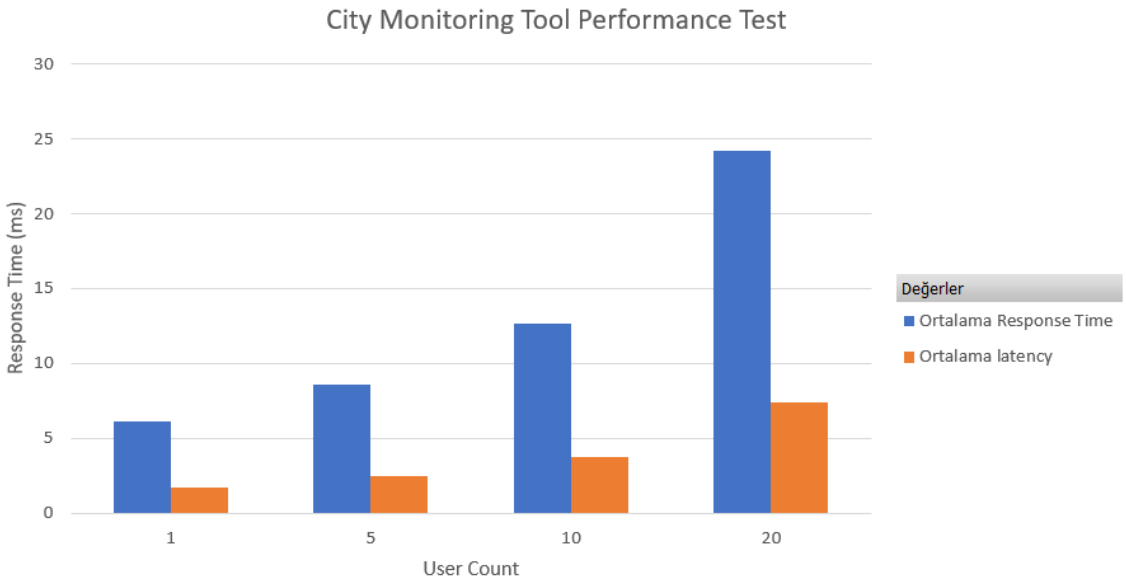


Figure 49. Average response times for different user loads

5.4. SCENARIO 04: CITY MONITORING SURVEY MODULE

FORTEARGE's City Monitoring Survey Module is an adaptive Survey Module for different use cases, used for high-level KPI monitoring. The Survey Module, on the other hand, has analysis capability; the cities can be filtered according to the indicator scores calculated from the survey, compared on the basis of indicator values, and indicator values are displayed on the map on the basis of cities.

The functional capability of the Survey Module is tested and evaluated based on scenarios.

5.4.1. OBJECTIVES

The goal of the scenario is to functionally test the components developed within the scope of the Survey Module listed below.

- Survey management and data entry screens
- High-level KPI monitoring and comparison interfaces

5.4.2. TESTS

Test definitions are determined with test scenarios to cover all functional requirements and to find out whether the requirements it validates are met by the system.

Each requirement is validated by at least one test definition.

Test steps, if any, test inputs, preconditions, outputs and success criteria were determined and traceability was ensured between requirements and test definitions. The abbreviations used in the requirement number are as follows:

- LOGIN: System Log in
- DATA: Data entering
- MGMT: System Management Operations
- MNTG: City Monitoring

Test scenarios were created to include the use cases of the Survey Module and according to the functions of the system. It has been prepared to describe the features of the system, its interaction with the users of the system and how the system will be used in general. Test scenarios define the functions to be performed on the system on a scenario-based basis.

Each test scenario defined on the system is explained in the following section. A test scenario; it is prepared on the basis of the works that an actor can perform using the system.

The system is used by two different end users:

- Admin- authorized user

- Standard user

The admin user is responsible for the management of the system, the authorization of the users and the management of the questions to be included in the evaluation forms.

Standard users are responsible for filling in the content of evaluation forms.

5.4.3. RESULTS

1. <u>Testing Survey Management and data entry screens:</u>

Test Scenario 01:	Login the system
Actor(s):	Admin – authorized user and Standard user
Descriptions:	There is local user management in POLDER City Monitoring Application.
Preconditions:	The users should be defined from within the Application before testing.

#	Test Steps of Test Scenario 01	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
1	Admin_User logs in with username and password.	Welcome screen appears and the following menus are active. • Creating an evaluation form • Management operations o Evaluation form definition o Defining User o Defining role o Dictionaries	LOGIN-01	P
2	Standard_User logs in with username and password.	Welcome screen appears and the following menus are active: • Home page • My assessment forms • Compare • Analysis	LOGIN-02	P
3	Admin_User tries to login with wrong name or password.	Admin_User cannot login.	LOGIN-01	p
4	Standard_User tries to login with wrong name or password.	Standard_User cannot login.	LOGIN-02	P

Test Scenario 02:	Evaluation Form Definitions
Actor(s):	Admin- authorized
Descriptions:	Evaluation form definitions include templates to be created for users to fill out. There are sections in the form definitions and questions under the sections
Preconditions:	The user should login to the system as admin.

#	Test Steps of Test Scenario 02	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
1	To manage the “Evaluation Forms”, click the “Evaluation Form Definition” submenu under the “Management Operation”.	The “List of Evaluations” appears.	MGMT-01	P
2	To create a new evaluation, click the "Add Evaluation” at the top right of the list.	The new evaluation form definition appears, and the following data input fields are active: • Name • Explanation • Evaluated Type Authorized Role	MGMT-02	P
3	Data input fields are filled and saved.	“Evaluation Form Description Edit” appears.	MGMT-03	P
4	To define the indicator set, click on the add on the bottom left of the page.	“Edit Section” for defining high level KPI sets appears.	MGMT-04	p
5	Data input fields are filled and saved.	Defined indicator set appears.	MGMT-04	P
6	Click on the “Questions” in the operation section of the indicator set.	“Questions area” appears.	MGMT-05	P
7	To add a question to the indicator set, click on the add under the questions area.	“Edit question” appears and the following data input fields are active: • Question text: • Question type: • Point • Explanation • Document status	MGMT-05	P

#	Test Steps of Test Scenario 02	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
		• Selection label • Choice2 labels • Important • External scoring • Read-only • Target display command • Account command		
8	Data input fields are filled and saved.	Defined question for the indicator set appears.	MGMT-06	P
9	To add a new question to the indicator set, click on again the add under the questions area.	“Edit question” appears again for new question.	MGMT-06	P
10	Data input fields are filled and saved.	Defined questions for the indicator set appears as a list.	MGMT-07	P
11	To edit any question from the indicator set, click on the edit button on the question.	“Edit Question” appears again for the question to be edited.	MGMT-07	P
12	Any data input fields are filled and saved.	Edited question appears in the question list of the indicator set.	MGMT-07	P
13	To edit any indicator set, click on the edit button on the indicator set.	“Edit Section” appears again for the indicator set to be edited.	MGMT-04	P
14	Any data input fields are filled and saved.	Edited indicator set appears in the indicator list of the “Evaluation Form Description Edit”.	MGMT-04	P

Test Scenario 03: Defining User

Actor(s): Admin- authorized user

Descriptions: Users defined in the system can be managed under the User submenu. A new user can be created here if desired.

Preconditions: The user should login to the system as admin.

#	Test Steps of Test Scenario 03	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
1	To manage the users, click the "User" submenu under "Management Operation".	"User list" appears.	MGMT-10	P
2	To create a new user, click the "Add User" at the top right of the list.	The "Definition of User" appears, and the following data input fields are active: • Name, • Surname, • Username, • Password, • Email, • Evaluated place, • Roles, • Active/Passive	MGMT-10	P
3	Data input fields are filled and saved.	The new defined user appears in the "User List".	MGMT-10	P
4	To edit any user from the "User List", click on the "Update" button on the user's line.	The "Update User" appears, and the following data input fields are active: • Name, • Surname, • Username, • Password, • Email, • Evaluated place, • Roles, • Active/Passive	MGMT-11	P

#	Test Steps of Test Scenario 03	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
5	Any data input fields are updated and saved.	Updated user appears in the “User List”.	MGMT-11	P
6	To deactivate the user's state, click on the “Update” button on the user's line.	The “Update User” appears, and the following data input fields are active: • Name, • Surname, • Username, • Password, • Email, • Evaluated place, • Roles, • Active/Passive	MGMT-12	P
7	Click on the “Active” check box to clean tick and saved.	The updated user appears in the “User List” as “passive”.	MGMT-12	P
8	To remove the any user from the system, click on the “Delate” button on the user's line.	The removed user does not appear in the “User List”.	MGMT-12	P

Test Scenario 04: Defining Role

Actor(s): Admin- authorized user

Descriptions: The roles defined in the system are managed under the role submenu.

Preconditions: The user should login to the system as admin.

#	Test Steps of Test Scenario 04	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
1	To manage the roles, click the "Role” submenu under “Management Operation”.	“List of Roles” appears.	MGMT-13	P
2	To create a new role, click the "Add Role" at the top right of the list.	The “Definition of Role” appears, and the	MGMT-13	P

#	Test Steps of Test Scenario 04	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
		following data input fields are active: • Definition, • Description, • System		
3	Data input fields are filled and saved.	The new defined role appears in the “List of Roles”.	MGMT-13	P
4	To edit any role from the “List of Roles”, click on the “Update” button on the role’s line.	The “Update Role” appears, and the following data input fields are active: • Definition, • Description, • System	MGMT-14	P
5	Any data input fields are updated and saved.	The updated role appears in the “List of Roles”.	MGMT-14	P
6	To remove the any role from the system, click on the “Delate” button on the role's line.	The removed role does not appear in the “List of Roles”.	MGMT-15	P

Test Scenario 05: Dictionaries

Actor(s): Admin- authorized user

Descriptions: The management of the dictionaries defined in the system is done from this section. The dictionaries that can be managed are as follows:

- Period
- Evaluated
- Evaluated Type
- Category

Preconditions: The user should login to the system as admin.

#	Test Steps of Test Scenario 05	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
1	To manage the dictionaries, click the "Dictionaries" submenu under "Management Operation".	"Definition of Dictionary" appears, and the following data input fields are active: • Period • Evaluated • Evaluated Type • Category	MGMT-20	P
2	To update the "Period", click on the "Update" button on the period's line.	The periods appear in the "List of Dictionary"	MGMT-22	P
3	Click on the "Update" button on the Period's line in the "List of Dictionary".	The "Update Dictionary" appears, and the following data input fields are active: • Definition • Start date • End date	MGMT-22	P
4	Any data input fields are updated and saved.	The updated period appears in the "List of Dictionary".	MGMT-22	P
5	To create a new "Period", click on the "Update" button on the period's line.	The periods appear in the "List of Dictionary"	MGMT-21	P
6	Click on the ""Add" at the top right of the list.	The "Update Dictionary" appears, and the following data input fields are active: • Definition • Start date	MGMT-21	P

#	Test Steps of Test Scenario 05	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
		• End date		
7	Any data input fields are updated and saved.	The new Period appears in the “List of Dictionary”.	MGMT-21	P
8	To update the “evaluated”, click on the “Update” button on the Evaluated’s line.	The Evaluated’s appear in the “List of Dictionary”	MGMT-24	P
9	Click on the “Update” button on the period’s line in the “List of Dictionary”.	The “Update Dictionary” appears, and the following data input fields are active: • Definition • Start date • End date	MGMT-24	P
10	Any data input fields are updated and saved.	The updated “evaluated” appears in the “List of Dictionary”.	MGMT-24	P
11	To create a new “Evaluated”, click on the “Update” button on the period’s line.	The Evaluated’s appear in the “List of Dictionary”	MGMT-23	P
12	Click on the “"Add" at the top right of the list.	The “Update Dictionary” appears, and the following data input fields are active: • Definition • Start date • End date	MGMT-23	P
13	Any data input fields are updated and saved.	The new “Evaluated” appears in the “List of Dictionary”.	MGMT-23	P

#	Test Steps of Test Scenario 05	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
14	To update the “Evaluated Type”, click on the “Update” button on the Evaluated Type’s line.	The Evaluated Types appear in the “List of Dictionary”	MGMT-26	P
15	Click on the “Update” button on the period’s line in the “List of Dictionary”.	The “Update Dictionary” appears, and the following data input fields are active: • Definition	MGMT-26	P
16	Data input field is updated and saved.	The updated “Evaluated Type” appears in the “List of Dictionary”.	MGMT-26	P
17	To create a new “Evaluated Type”, click on the “Update” button on the period’s line.	The Evaluated Types appear in the “List of Dictionary”	MGMT-25	P
18	Click on the “"Add" at the top right of the list.	The “Update Dictionary” appears, and the following data input fields are active: • Definition	MGMT-25	P
19	Any data input fields are updated and saved.	The new “Evaluated Type” appears in the “List of Dictionary”.	MGMT-25	P
20	To update the “Category”, click on the “Update” button on the Category’s line.	The categories appear in the “List of Dictionary”	MGMT-28	P
21	Click on the “Update” button on the Category’s line in the “List of Dictionary”.	The “Update Dictionary” appears, and the following data input fields are active: • Definition	MGMT-28	P
22	Data input field is updated and saved.	The updated “Category” appears in the “List of Dictionary”.	MGMT-28	P

#	Test Steps of Test Scenario 05	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
23	To create a new “Category”, click on the “Update” button on the period’s line.	The categories appear in the “List of Dictionary”	MGMT-27	P
24	Click on the “"Add" at the top right of the list.	The “Update Dictionary” appears, and the following data input fields are active: • Definition	MGMT-27	P
25	Any data input fields are updated and saved.	The new “Category” appears in the “List of Dictionary”.	MGMT-27	P

Test Scenario 06: Data entering

Actor(s): Standard User

Descriptions:

Evaluation forms are created by the system administrator (Admin User) and filled by the standard user. When filling out the new form:

- Evaluation form: It comes from the form definitions provided by the administrator in the system.
- Date: The date the form was completed.
- Period: The period for which the form is valid.
- Evaluated: One of the cities previously registered in the system can be selected. If it is not registered and a new record will be created, this field should be left blank.
- New evaluated: If a new evaluated record is to be created, this field must be entered.
- Project code: The project code associated with the form.
- Category.

Preconditions: The user should login to the system as standard user.

#	Test Steps of Test Scenario 06	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
1	To enter the data, click the "Evaluation Form" menu.	"Evaluation Form List" appears, and the forms filled by the user are seen as a list.	DATA-01	P
2	To fill a new evaluation form, click the "Add New Evaluation Form" at the top right of the list.	The "New Evaluation Form" appears, and the following data input fields are active: • Evaluation form • Date • Period, • Evaluated, • New evaluated	DATA-01	P
3	The form to be filled is selected from the input data field of "Evaluation Form", then the Date, Period, Evaluated, New evaluated input fields are filled and saved.	The "New Evaluation Form" as a draft and its indicator sets appear. And it is ready to fill.	DATA-02	P

#	Test Steps of Test Scenario 06	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
4	To enter the data, click on any indicator set.	The questions including input data boxes of the indicator set appear.	DATA-03	P
5	Any question input boxes are filled manually and saved.	The questions answered by the user will be scored and displayed by the system after they are saved.	DATA-04	P
6	To again access to the form, click on the “Continue” button on the Evaluation Form's line that has been started to be filled.	It is shown how many questions in total and how many were answered by the user.	DATA-01	P
7	To submit the form, click on the “submit form” after completing the data entry to the form.	An acceptance warning appears that the form will be submitted.	DATA-05	
8	Warning to be displayed accepted.	The form will be forwarded to the system administrator for approval.	DATA-06	
9	To delate form, click on the “delate form” after completing the data entry to the form.	An acceptance warning appears that the form will be deleted.	DATA-07	
10	Warning to be displayed accepted.	The form is deleted from the system and is not seen in the “Evaluation Form List”.	DATA-08	

Test Scenario 07: Evaluation Form Approval Procedures

Actor(s): Admin- authorized user

Descriptions: Evaluation forms are approved or returned to standard user who sent the form for approval.

Preconditions: The user should login to the system as Admin- authorized user.

#	Test Steps of Test Scenario 07	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
1	To approve the evaluation form, click the "Evaluation Forms" menu.	"Evaluation Form List" appears, and the forms sending by the standard user for approval are seen as a list as draft.	MGMT-30	P
2	To approve an evaluation form, click on the "Detail" button on the Evaluation Form's line in the "Evaluation Form List".	The "Evaluation Form Details" appears, and the following buttons are active: • Approval Form • Return Form	MGMT-30	P
3	Click on "Approval Form" button on the bottom of the page.	The Evaluation Form appears as "Approved" in the "Evaluation Form List".	MGMT-31	P
4	To return an evaluation form, click on the "Return Form" button on the bottom of the page.	The Evaluation Form returns to the standard user and appears as "Returned" in the "Evaluation Form List".	MGMT-32	P

2. Testing High-level KPI monitoring and comparison interfaces

Test Scenario 08:	High-level KPI monitoring of the city
Actor(s):	Standard user
Descriptions:	The questions answered in the evaluation form are scored by the system and high-level key indicators are created. Cities are monitoring according to these indicators.
Preconditions:	The evaluation form should be approved by the system administrator and the user should login to the system as standard user.

#	Test Steps of Test Scenario 08	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
1	To monitor the city, click the "Compare" menu.	"Please select the evaluation forms you want to compare" worn and "Evaluation Form List" including all approved and draft forms appears.	MNTG-01	P
2	Click on the "Choose" button on the any approved Evaluation Form's line.	The "Compare" button are active on the bottom of the page.	MNTG-02	P
3	Click on "Compare" button on the bottom of the page.	The "Indicator Sets Evaluation" of the city appears as bar graphic and list.	MNTG-03	P

Test Scenario 09: Comparing city or cities based on High-level KPI

Actor(s): Standard user

Descriptions: The city is compared with itself or with other cities on a yearly basis.

Preconditions: The evaluation forms should be approved by the system administrator and the user should login to the system as standard user.

#	Test Steps of Test Scenario 08	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
1	To compare the city, click the "Compare" menu.	"Please select the evaluation forms you want to compare" worn and "Evaluation Form List" including all approved and draft forms appears.	MNTG-01	P
2	The desired number of forms is selected by clicking the "Choose" button next to the	The selected forms appers and the "Compare" button are active on the bottom of the page.	MNTG-04	P

#	Test Steps of Test Scenario 08	Expected Result	Requirement ID	Passed, Failed or Block P/F/B
	forms to be compared.			
3	Click on "Compare" button on the bottom of the page.	The "Indicator Sets Evaluation" of the city or cities appears as bar graphic and list.	MNTG-03	P

5.5. SCENARIO 05: ANALYSIS AND PREDICTION OF POURISR FLOWS FROM MOBILE PHONE RECORDS

The developed visualization dashboard for the analysis of tourist flows enables the representation of different scenarios and analysis to support policy makers, tourist planners and service operators design different services and manage their resources according to the tourist activity of the zone.

The platform allows the filtering of different segments and aggregations of the calculated scenarios. Figure 1present a screen shot of the dashboard for its implementation in the city of Madrid. In particular we see in this screen shot the descriptive view for the visualisation of presence indicators. The screen shot show all the possible analytic and visualisation options of the dashboard.

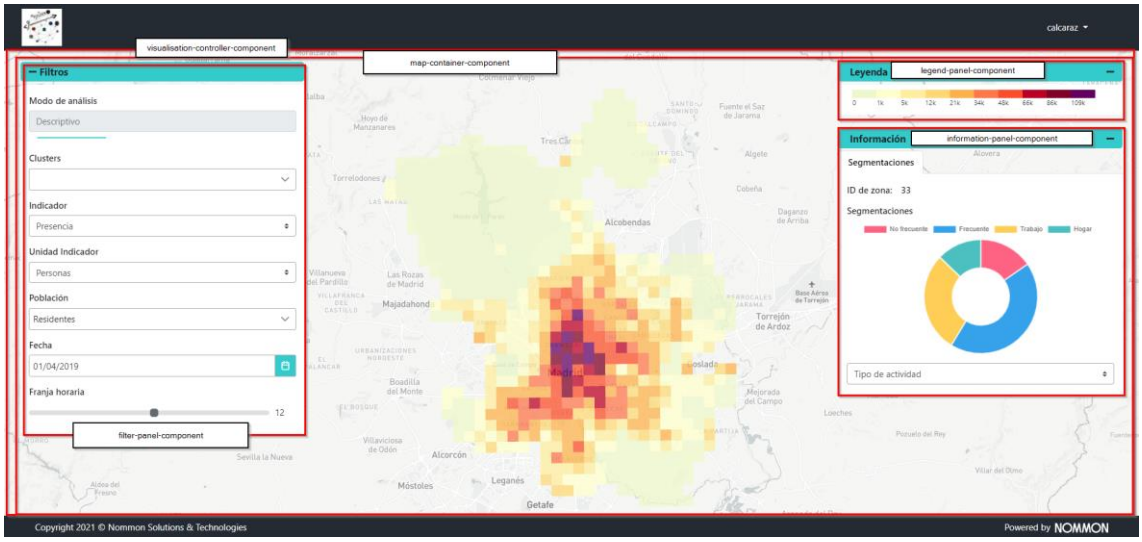


Figure 50. Screen shot of the visualisation tool presenting the view of the visualisation of the presence indicator.

5.5.1. OBJECTIVES

In order to accomplish with its main objective, supporting relevant actors in monitoring, planning and managing tourism the dashboard should accomplish with a minimum of functional requirements.

The validation of the dashboard performance from an operational point of view considers the following dimensions:

- Functionality: The indicator prediction system must function correctly, being able to collect user input and produce the relevant results without errors, warnings or unexpected behavior.
- Usability: The developed platform must provide an intuitive interface that allows fluid interaction with the system and without the need for more help than the tool's user manual.
- Agility: Changes in the application should be fluid, without sudden stops, so as to allow continuous interaction that facilitates the use of the system.
- Consistency of indicators: The results of indicators provided by the system must be consistent.

5.5.2. TESTS

In order to validate the operational performance of the dashboard the following test were defined:

Functionality

Table 2 summarizes the test to check the dashboard functionality

Table 2. Summary of functionality test

Functionality		
Test	**Description**	**Expected result**
F1	When accessing the platform website for the first time, the Login screen is displayed	The user does not access the platform directly, but must be authenticated using a username and password.
F2	After authentication, the main system window loads.	The main window of the system shows the new analysis form and a drop-down menu that allows you to “Logout”
F3	After filling in the data of the new analysis form, the analysis visualization is loaded.	The main system window displays the graphics view. Shows the filter panel, choropleth map and legend.

Functionality		
Test	**Description**	**Expected result**
F4	The user interacts with the available filters and the visualization refreshes the data presented in the rest of the elements	The choropleth map, the legend values, and the graph pane, if available, are updated with the filtered data.
F5	When the user clicks on a zone, the zone information is refreshed.	The graph panel is updated when you click on the area.
F6	The user hides any of the display elements.	Clicking the hide button on any of the panels hides it.
F7	After hiding an element from the display, the user shows it again.	When you click on the show button of any of the hidden panels, it is shown.
F8	As you vary the values of the "Indicator" selector, the options for the other selectors vary accordingly.	When changing the indicator selector, the values of the rest of the selectors are changed correctly.

Usability.

Table 3 summarizes the test defined for usability testing.

Table 3. Usability test

Usability		
Test	**Description**	**Expected result**
U1	The possible actions of the system are correctly described in the user manual.	The main functions of the system are well documented in the user manual and describe the procedure in an understandable way.
U2	The procedure to execute any action is simple. There are no actions that are difficult to access from the start menu.	The number of iterations to perform to execute any action is reasonable based on your input requirements.
U3	Interface buttons and links are easy to locate and distinguish.	Any element that can be clicked by the user is fully visible. Items whose position is most confusing are highlighted when you hover over them.
U4	The system illustrates and facilitates data entry. The system supports user data entry errors and facilitates their correction.	Discrete data entry is done through category entries (radio buttons, pull-down menus). When the user has to enter data by keyboard, the expected format is specified.
U5	The system supports user data entry errors and facilitates their correction. The system provides explanatory error messages when the entries are not correct	Basic validation of input data is handled through the use of scripts.

Agility:

Table 4 summarizes the test to assess the usability of the platform

Table 4. Agility test

Agility		
Test	**Description**	**Expected result**
Fl1	The response time during the action selection and execution process is minimal.	The delay between clicking on a button and the development of its associated action is zero, less than one second.
Fl2	The response time for changing displays is reduced.	The change in filters or zones in any of the analysis views is small, always less than 2 seconds.

Once all the validation tests summarized in this section were defined, a validation exercise protocol was developed that defines the steps that were going to be taken to carry out the development of the validation process. These steps are as follows

1. A group of validation users who had not been directly involved in the development of the tool was selected.
2. The use cases to be validated (defined in the previous sections) were established and a brief presentation was prepared to show the users before starting the tests.
3. Validation templates were created that contain the tables defined above along with an additional field to collect suggestions and other comments from users.
4. A date was selected to hold a face-to-face workshop in which the validation process was developed.

For the validation process, each user has been provided with an account that allows them to access the platform and perform the previously defined tasks. Each user was given a copy of the user manual for reference. In addition to completing the validation template, users were asked to add all the comments and suggestions they saw fit at the end.

5.5.3. RESULTS

Once the users and validation tests were selected, the opinions, errors and other comments made by the users after the completion of the process were collected. All those problems found in the functioning of the platform have been solved and some of the users have been asked to carry out the tests again to verify the effectiveness of the actions taken.

In general, the validation results have been satisfactory, giving rise to positive comments in all the validation dimensions. Some users have encountered minor problems that have been resolved by taking specific corrective actions in some aspects of the platform.

Functionality:

The functionality of the platform has been correctly validated in all its aspects. The functional requirements of the front-end of the platform have been implemented and all users have confirmed that the general design of the tool is correctly implemented in the tool. In detail, Table 9 provides the summary of each test along with the verification result and actions taken. It should be noted that two bugs related to the presentation of information have been detected and have been resolved.

Table 5. Results from the functionality test

Functionals		
Test	**Description**	**Corrective actions**
F1	When accessing the platform website for the first time, the Login screen is displayed.	
F2	After authentication, the main system window loads	
F3	After filling in the data of the new analysis form, the analysis visualization is loaded.	
F4	The user interacts with the available filters and the visualization refreshes the data presented in the rest of the elements.	
F5	When the user clicks on a zone, the zone information is refreshed.	Revenue targeting does not display the name properly. It was corrected
F6	The user hides some of the display elements	
F7	After hiding an element from the display, the user shows it again.	Minimizing the screen slides the time zone to the next line. It is corrected.
F8	As you vary the values of the "Indicator" selector, the options for the other selectors vary accordingly.	

Usability:

With regard to usability, users have validated the main features specified and have also provided additional comments to improve the tool, which are detailed below. Specifically, the users have verified the correspondence between the process of executing each action and its explanation in the user manual, suggesting small changes that contribute to improve the correspondence and facilitate understanding. of the same. Users have validated the process of locating elements and entering data, identifying small errors about the help provided by the tool for some formats that have been corrected. *Table 6*lists the results of each specific validation test and the actions taken if necessary.

Table 6. Results of the usability test

Usability		
Test	**Description**	**Corrective actions**
U1	The possible actions of the system are correctly described in the user manual.	
U2	The procedure to execute any action is simple.	
U3	Interface buttons and links are easy to locate and distinguish.	
U4	The system illustrates and facilitates data entry.	
U5	The system supports user data entry errors and facilitates their correction.	

Agility

Regarding the agility of the system, the validation process has made it possible to verify that the response times to the interactions initiated by the user are reduced, especially in the analysis data filtering processes. When viewing the data for a zone, positive feedback has been received about the provision of “spinenrs” showing that the data is loading. Table 13 summarizes the results obtained after carrying out the validation process.

Agility		
Test	**Description**	**Corrective actions**
Fl1	The response time during the action selection and execution process is minimal	Segmentations take time to load the first time a zone is clicked. 1-2s delay.
Fl2	The response time for changing displays is reduced.	

It should be noted that it has been detected that the time of the zones with the highest values take a longer time. However, the use of elements that indicate that the area is loading, make the user not lose attention from the screen. Also, this only happens on the first interactions with the zones.

6. SECURITY LAYER TESTS

This layer appears in a transversal way affecting all the previous layers, thus complying with the model defined in the UIT-T Y-4201 standard.

It supports services such as auditing, monitoring and security. Every layer of the platform needs to interact with the Support & Security Layer for each activity carried out.

The management of the permissions of the different profiles and the cryptography among others, are features that will be developed in this module. Monitoring tools for administrators will be enabled that may facilitate audits.

Developing strong encryption mechanisms are really important to ensure confidentiality of vulnerable information handled by the system.

As security layer is not completely finalized, the tests and results performed in this layer will be added at a later stage, when Wealize, the organization in charge of security, finishes its project.

7. CONCLUSIONS

As has been presented in the previous sections, the fact of not having been able to conduct a common pilot or test developments in a real environment has not been an impediment for partners to validate and demonstrate the functionality of their developments; on the contrary, they have been able to demonstrate that the obtained results are satisfactory and the versatility of their applications.

With this, the great potential of POLDER for a Smart City is demonstrated, as it provides multiple services under the same concept and the same architecture.