MAtchUP

D4.24: Urban platform Integration and Interoperability in Antalya

Final version

WP 4, T 4.6.3

30 November, 2020 (M38)

Authors: Authors: M.Serdar Yümlü, Cenk Güreken, Serdar Serdaroğlu, Caner Tosunoğlu, Selçuk Ilıkcan(SAM)

MAtchUP - SCC-1-2016-2017

Innovation Action - GRANT AGREEMENT No. 774477

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

Technical References

Project Acronym	**MAtchUP**
Project Title	MAximizing the UPscaling and replication potential of high level urban transformation strategies - MAtchUP
Project Coordinator	Ernesto Faubel Ayuntamiento de Valencia efaubel@valencia.es
Project Duration	1 October 2017 - 30 September 2022 (60 Months)

Deliverable No.	D4.24
Dissemination Level	CO[1]
Work Package	WP 4-Demonstration in Antalya
Task	T 4.6.3- Integration and Interoperability
Lead beneficiary	16 (SAM)
Contributing beneficiary(ies)	A4 (ANT)
Due date of deliverable	30 November 2020
Actual submission date	19 November 2020

[1] PU = Public
PP = Restricted to other programme participants (including the Commission Services)
RE = Restricted to a group specified by the consortium (including the Commission Services)
CO = Confidential, only for members of the consortium (including the Commission Services)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

Version	Person	Partner	Date
1.0 Final (M24)	M. Serdar Yümlü, Çağan Sevim, Ali Demir Akgünlü Cenk Güreken	SAM	24 September 2019
1.1	Cenk Güreken, Selçuk Ilıkcan, Eser Karakaya	SAM	27 August 2020
1.2	M. Serdar Yümlü, Cenk Güreken, Serdar Serdaroğlu, İbrahim Acar Caner Tosunoğlu	SAM	29 September 2020
1.3	Cenk Güreken	SAM	15 October 2020
1.4	Pau Arce Villa	UPV	02 November 2020
1.5	M. Serdar Yümlü Cenk Güreken	SAM	10 November 2020

Disclaimer

The information reflects only the author’s view and the Commission is not responsible for any use that may be made of the information it contains.

Table of Content

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

List of Tables

This project has received funding from the European Union's Horizon 2020
research and innovation programme under Grant Agreement N°774477

List of Figures

This project has received funding from the European Union's Horizon 2020
research and innovation programme under Grant Agreement N°774477

Abbreviations and Acronyms

Acronym	Description
API	Application Programming Interface
AUP	Antalya Urban Platform
BI	Business Intelligence
CKAN	Comprehensive Knowledge Archive Network
CSW	Catalogue Services for the Web
DAQ	Data Acquisition
DCAT-AP	Data Catalogue vocabulary – Application Profile
EIP	European Innovation Partnership
EMT	Municipal Transport Company (Empresa Municipal de Transporte)
ESB	Enterprise Service Bus
EV	Electric Vehicle
GDPR	EU General Data Protection Regulation
HTTP	HyperText Transfer Protocol
ICT	Information and Communications Technology
IDABC	Interoperable Delivery of European eGovernment Services to public Administrations, Businesses and Citizens
IoT	Internet of Things
JSON	JavaScript Object Notation
KML	Keyhole Markup Language
KPI	Key Performance Indicator
LOPD	Organic Law on the Protection of Personal Data
LTS	Long Term Support
NGSI	Next Generation Service Interfaces
OMA	Open Mobile Alliance
POI	Point of Interest
REST	Representational State Transfer
SQL	Structured Query Language
STH	Short-Term Historic
VM	Virtual Machine
WMS	Warehouse Management System
XML	Extensible Markup Language
IPG	Interoperability Principles Guide
KDEP	Short Term Action Plan for Digital Transformation of Turkey

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

Abstract

This deliverable is reporting on the current state of development of the Urban Platform integration and interoperability in Antalya in the third year of MAtchUP Project (M25 – M38). This task includes the new projects and services to improve city's operations. Also another important task is to increase the connection between the City of Antalya and its citizens.

The developments during MAtchUP follow the same principles: ensuring open data, interoperability through open APIs developments and assessing the evaluation process by considering the requirements of Antalya's monitoring plan. Therefore, new operations and services must guarantee interoperability between the different components involved.

Moreover, it is needed to take into consideration the new European General Data Protection Regulation (GDPR). Security and privacy aspects should be taken into account. It is important to publish non-sensible and anonymized data for learned by citizens. Also developers want to make use when they will start creating innovative services for the city.

Integration architecture is one of the most critical aspects of the urban platform to sustain consistency and communication between several internal and 3rd party components. Integration architecture and methodologies are introduced in this version of the platform integration and interoperability.

As far as the cities in the project share a common objective, this deliverable D4.24 shares a common structure with the analogous deliverables of WP2, which is D2.24, and WP3, which is D3.24. Furthermore, these deliverables are due in M38 as a third and final version of the documents DX.10 and DX.23, respectively.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

1 Introduction

1.1 Purpose

Definition of an interoperability test plan for Antalya is to be established from a harmonised approach in the three demos, as well as testing of the new modules developed under the Urban Platform, ensuring that they interoperate, following the open API concept of the Urban platforms of Dresden and Valencia.

The target of this deliverable is to report on the current situation of development on urban platform integration and interoperability in Antalya. This deliverable includes services to improve city operation and decision making processes.

At the beginning, it is important to describe the concept of interoperability and differentiate between horizontal and vertical interoperability. Interoperability makes the communication easier between modules. Defining the data models is also important. These data models are leading to the exchange of data from one module to another module. Well defined and strong data models will increase the interoperability between components. Additionally, datasets used in the project are also defined in this deliverable.

- this deliverable provides the main components of the Antalya urban platform and presents the integration architecture among the internal modules.
- platform uses a modular architecture based on virtualization of layers, communication, integration and messaging between different components is provided in this deliverable.
- data modelling based on NGSI defines all the interventions and data models used in the implementation of specific actions in Antalya.
- open data integration is explained within the platform.
- data security, data ingestion and anonymization is also addressed.
- interoperability tests approach and the updated plan is provided.

Lastly, privacy and security issues are mentioned. European General Data Protection Regulation (GDPR) helps to protect privacy and security. This deliverable also focuses on the need to evaluate security mechanisms in the application endpoints, to control access to the city platform, and to improve security strategies, such as encrypted communications or the anonymization process, among others.

1.2 Contribution of partners

The following table depicts the main contributions from participant partners in the development of this deliverable.

Participant short name	Contributions
SAM	Integrator for the MAtchUP platform, All ICT interventions and integrations among components
ANT	Data Provider, KPIs and integration perspectives with legacy systems
ANP	Data Provider, KPIs and integration perspectives with legacy systems
AKD	Data controller, KPIs
DEM	Data controller, KPIs
TAY	Data controller, KPIs
UPV	Reviewer

Table 1 Contribution of partners

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

2 Interoperability definition

In general, interoperability is the capability of the system's components or subsystems to be able to exchange and use the information in a harmonized and homogeneous way. More specifically, in information technology, “the ability to share data between different components or machines, both via software and hardware”, or it can be defined as the exchange of information and resources between different computers through local area networks (LANs) or wide area networks (WANs) [1].

Components can be said to exhibit syntactic interoperability if they are able to communicate by using specified data formats and/or communication protocols like XML or SQL standards, or even more than that, when they manage to communicate if they are able to autonomously interpret the data exchanged relevantly and validly to conclude to meaningful results as end users of both components.

In systems which have many subsystems like Smart City’s Urban Platforms, it is necessary to obey both syntactic and semantic interoperability. Antalya project obeys both interoperability principles as well. Next sections focus on the interoperability specifications in the context of Antalya MAtchUP project.

2.1 Interoperability in the Smart City context

European Innovation Partnership (EIP)’s definition of Smart Cities is that “solutions that help cities to digitize and manage their services” for the Urban Platforms group of the integrated infrastructure action cluster. Furthermore, Urban Platforms are an essential block of all data which cities manage. With security and privacy, interoperability is one of the core expectations of Urban Platform context. According to the EIP, ensuring the vertical interoperability and the development of reusable and interoperable services is the main priority instead of the horizontal interoperability. Therefore, reusability and scalability must be at the top of the development plan. Both vertical and horizontal interoperability is the topic of MAtchUP project. With the EIP suggestions, the development processes are mainly focused on vertical interoperability instead of horizontal interoperability. The main reason about the prioritization of vertical interoperability in Smart City concept is sharing data among fragmented services for each sector of energy, mobility, ICT. For example, transportation service is in mobility sector but its data is needed in the energy sector, and vice versa, transportation service may need data of energy sector. Horizontal interoperability refers to data exchange between different Smart Cities which has similar services.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement N°774477

Figure 1 Vertical Interoperability

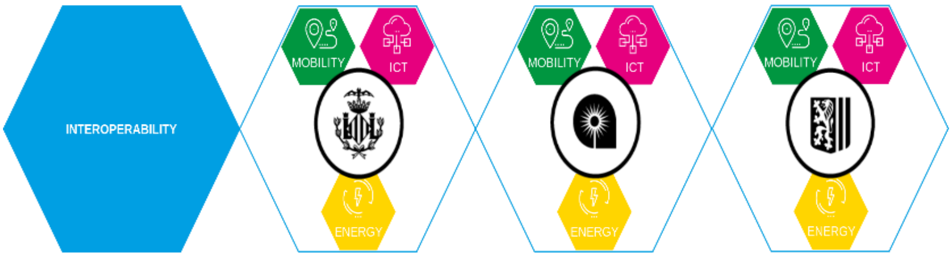

Figure 2 Horizontal Interoperability

2.2 State of the art and related projects

One of the challenges faced by the project is to process and visualize data from heterogeneous systems and environments, such as sensors or the other platforms involved. All these systems must be integrated in a common platform, and for this purpose, it becomes necessary to guarantee the interoperability of data.

This project has received funding from the European Union's Horizon 2020
research and innovation programme under Grant Agreement N°774477

In the context of Smart Cities, information is obtained from elements of the city, such as street lamps or buses, which use different protocols to send information. FIWARE[2] is in charge of translating all the information gathered from different sources into a common language, the Next Generation Services Interface (NGSI). Moreover, the use of FIWARE allows to replicate and scale this platform in other cities without major adaptation efforts, since FIWARE is a standardized solution and is mainly based on components called Generic Enablers that facilitate interoperability.

Besides the FIWARE project, recently there have been some new projects related to interoperability, such as the INTER-IoT[3] project, which is a European project framed in Horizon 2020 programme with the aim to provide an interoperable framework architecture for seamless integration of different IoT architectures present in different application domains.

2.3 Regulatory framework for interoperability

Antalya MAtchUP Project includes integration of the various systems that can cause some difficulties about interoperability. Regulatory framework aims to eliminate the problems which can rise against interoperability by standardization of principles. This section will continue with national and international standards which project is going to obey for interoperability.

2.3.1 National interoperability standards

Republic of Turkey has announced a “Short Term Action Plan for Digital Transformation of Turkey” (KDEP) under the management of State Planning Organization. KDEP aim is described as [1] “Easily serving to citizens in electronic environment, by avoiding bureaucracy” [2]. That plan obviously pointed a need for regulatory framework on interoperability standards to exchange information among institutions. Within and as an outcome of this need the Prime Ministry announced Circular 2005/20 with a guide named “Interoperability Principles Guide” (IPG)[4] regulating the National Interoperability Scheme. This circular describes the interoperability and procedures of open data sharing and exchanging of data standards for both governmental and private organizations.

IPG includes three aspects of the interoperability. The first part mostly includes context of the interoperability, descriptions and interoperability policies. The second part states technical expectations and standards for interoperability. Technical standards part includes topics like open data share and exchange, interconnections, data integration and content management, security, geographical information systems, development of

---

[2] https://www.fiware.org/

[3] https://inter-iot.eu/

[4] http://www.bilgitoplumu.gov.tr/wp-content/uploads/2014/04/Birlikte_Calisabilirlik_Esaslari_Rehberi_2.1.pdf

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement N°774477

information systems and management. The last part includes sections about institutional architecture works, data vocabulary, examples and easy access. This guide will be a reference for Antalya MAtchUP Project in related and covered parts.

2.3.2 International interoperability standards

European Framework for Interoperability has published “European Interoperability Framework for Pan-European e-government Services” (IDABC) for administrations, business and citizens. Furthermore, European Parliament and the Council have published a guide for public administrations in September 2009 with the decision number 922/2009 including action plans for e-government on interoperability by aiming to inform policies of open data sharing, reusing and cooperation.

There are also some open standards publicly available. These are standards that are global and regardless any individuals, legal or technical situations that limit their use.

DCAT-AP: The DCAT Application profile for data portals in Europe (DCAT-AP) is a specification based on the Data Catalogue vocabulary (DCAT) for describing public sector datasets in Europe [3].

CSW: Catalog Service for the Web (CSW), sometimes seen as Catalog Service Web, is a standard for exposing a catalogue of geospatial records in XML on the Internet (over HTTP). The catalogue is made up of records that describe geospatial data (e.g. KML), geospatial services (e.g. WMS), and related resources. [4][5][6][7][8]

INSPIRE: The INSPIRE Directive aims to create a European Union spatial data infrastructure for the purposes of EU environmental policies and policies or activities which may have an impact on the environment. [9]

REST (Representational State Transfer): Standard interface for web applications approached in 2000 and is generally used to develop APIs. For example, the common NGSI language used in FIWARE provides a REST API via HTTP for obtaining data or performing operations on data stored in FIWARE.

Schema.org: It aims to produce, maintain, and promote schemas for structured data on the Internet. The schemes have a specific standard vocabulary for different entities. For example, the Place scheme can be used in a city for fixed places such as squares or parks.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement N°774477

3 Smart City modules definition

Antalya Metropolitan Municipality provided a smart city vision to include several aspects for a safe, green and vibrant Antalya Smart City. This Smart City Framework (Figure 3) provides the main aspects including policies, citizen centric approaches, business/economy, safety, sustainability and mobility features. Several urban platform components and ICT infrastructure play a basic role in this picture. MAtchUP is an additional layer to support this city wide initiative and with the demo district all interventions related with MAtchUP are also depicted in the figure as yellow dots.

All the related information is described in deliverable D4.22 with details. This section includes more general informatics about Smart City Modules.

Figure 3 Antalya Smart City Framework

3.1 Requirements for Urban Platform Integration and Interoperability

The general structure of the Antalya urban platform is depicted in Figure 4 and the details are already explained in the deliverable D4.22. On top of the existing modules, the input modules, output modules as well as the core modules are listed in the following sections based on the general structure of the requirements.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

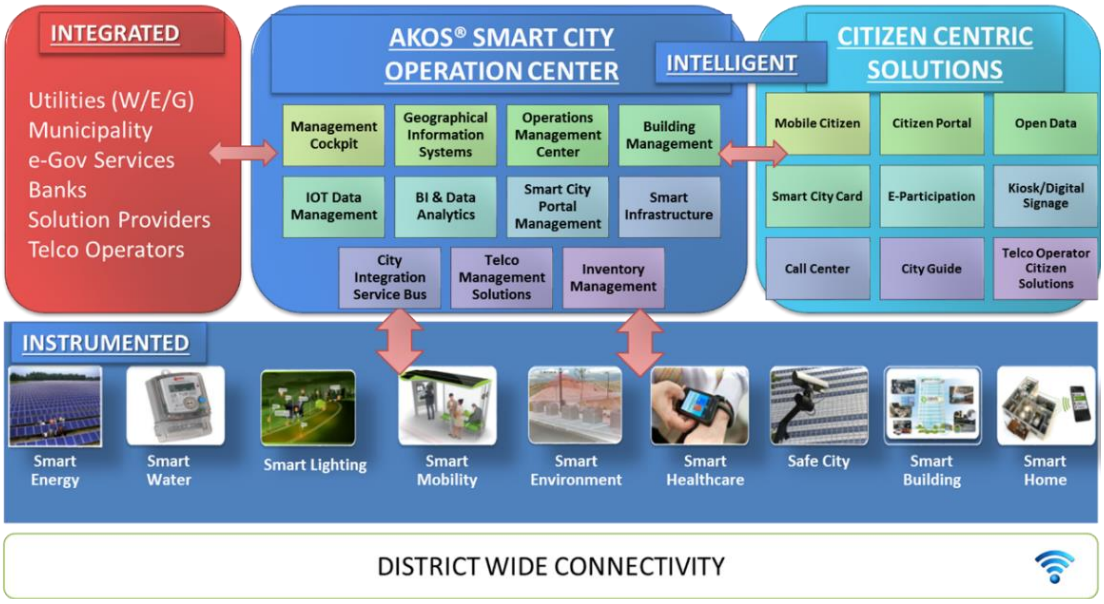

Figure 4 Antalya Urban Platform

3.2 Input modules

In the Smart City concept, the Urban Platform should offer a bunch of useful services to citizens, governmental and private institutions and entrepreneurs. To produce these services, the Urban Platform needs raw input data in order to process and deliver to Core Module to be transformed into meaningful data. The rawest kind of data is city life itself in the Smart City concept. However, this kind of data comes in many different ways and formats. Therefore, the most useful and efficient data must be targeted to collect. The Urban Platform capability of the targeted collection is limited by technology. In Antalya's MAtchUP project, a number of technologies and services are used to collect targeted data:

- **IoT Sensors, Agents:** This module handles the input from different FIWARE data models in Antalya's interventions. This part is depicted as the instrumented layer in Antalya urban platform (figure below). Currently, the IoT sensors are synthetic in the platform which are:
    - Charge Stations
    - Smart Meters
    - Public Lighting
    - Solar Panel

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

Figure 5 Instrumented Layer

- **Open Data:** This is the open data portal of Antalya that is implemented during MAtchUP which is based on CKAN data management system.
- **Legacy Systems:** These are existing systems, databases and services that Antalya is currently using to provide e-Government services. All these services are explained in detail in 4.22 Urban platform adaptation specifications in ANTALYA.
- **APIs and Services**: APIs and services that will be integrated to Antalya's Urban Platform. Existing APIs and new APIs/Services considered to be implemented in MAtchUP are given in details in 4.22 Urban Platform adaptation specifications in Antalya.

3.3 Output modules

Output modules are the main basic concept and key success factor of the Urban Platforms. Data is firstly processed in input modules. After cleaning and homogenization, the data is delivered to core modules for the last big data modifications. Homogenised data is delivered to output modules and transformed into analytical solutions, business intelligence solutions, dashboards and KPIs via:

- Websites/Web Apps: Kepez Santral websites that provide up-to-date information mainly for citizens
- Open Data Portal: Antalya's Open Data Portal that is implemented during MAtchUP.
- API Services: New API and services to be provided through the Urban Platform
- Mobile App: Mobile apps for citizen engagement and direct access to the platform
- Co-creation apps: Decidim based Antalya co-creation portal for citizen engagement activities.
- Big Data visualization: The visualization of big data and urban platform outputs. This is based on Elasticsearch and Kibana.

3.4 Core modules

Core modules are the orchestrators between input and output modules. These are the modules planned in Antalya MAtchUP:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

- **MQTT broker:** MQTT is a machine-to-machine "Internet of Things" connectivity protocol. It was designed as an extremely lightweight publish/subscribe messaging transport.
- **Hadoop:** To be used for big data functionalities and distributed processing:
- **Elasticsearch and Kibana:** Kibana will be used aligned with Elasticsearch for the visualization of the platform.
- **Kafka:** Kafka will act as one of the central messaging and integration element between other components and Elasticsearch/Hadoop as depicted in the preliminary architecture (Figure 6).
- **Open Data Portal:** The Antalya Open Data Portal that is implemented during MAtchUP.
- **Co-Creation Portal:** Decidim based Antalya co-creation portal for citizen engagement activities.

The relation and integration of the modules are designed within a visualization Architecture framework given as below:

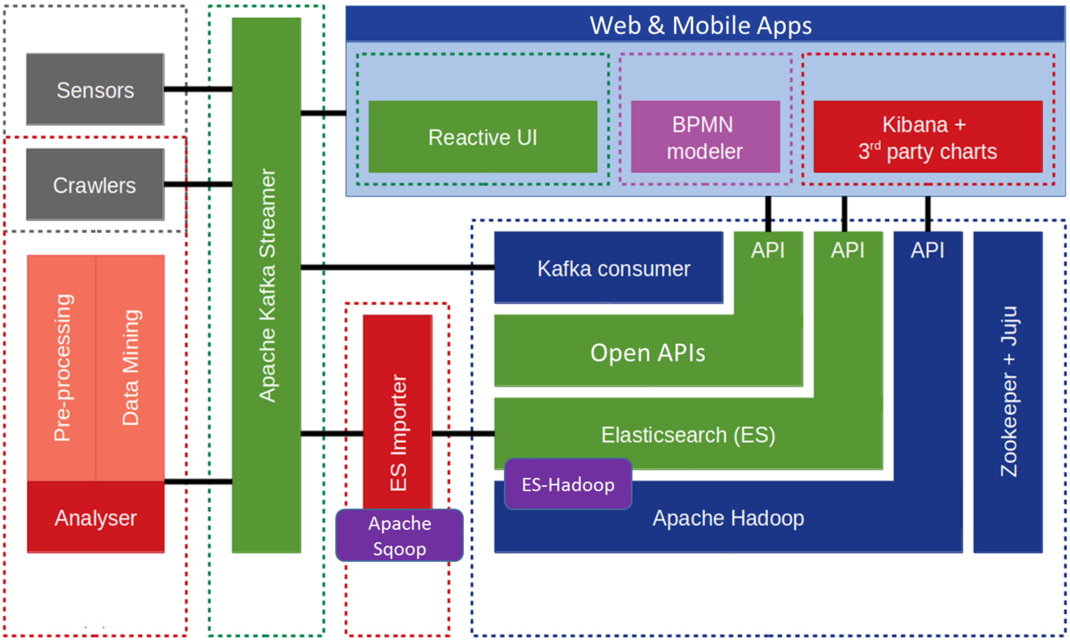

Figure 6 Conceptual Big Data Visualization Architecture for Antalya

To briefly mention the core modules of the architecture:

MQTT: MQTT is a machine-to-machine "Internet of Things" connectivity protocol. It was designed as an extremely lightweight publish/subscribe messaging transport.

Apache Hadoop: Apache Hadoop is an open source software framework for distributed storage and processing of very large data sets, based on a modular cluster distribution with several nodes. Hadoop is one of the most popular and widely used

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

platforms today. It provides a software framework for distributed storage and processing of big data using the MapReduce programming model.

Elasticsearch: Elasticsearch is a flexible scalable data analytics and text search engine. It allows to analyze, search and store big data very fast and near real time. It is generally used as the underlying engine/technology that powers applications that have complex search features and requirements.

Kibana: Kibana is a general purpose graphing and visualization tool from the Elastic Team. Some of the benefits of using Kibana to visualize your log data are: it's built to work easily with Elasticsearch, it provides many included visualization widgets like pie charts, line graphs, and more, and it provides an easy dashboard setup and sharing method.

Kafka: Apache Kafka [20] is a distributed messaging system and a robust queue that can handle a high volume of data and enables the user to pass messages from one end-point to another. Kafka is suitable for both offline and online message consumption. Kafka is built on top of the ZooKeeper synchronization service. Kafka is a unified platform for handling all the real-time data feeds.

Using these components Antalya Urban platform will support the integration architecture of the Antalya ICT interventions given in Figure 7.

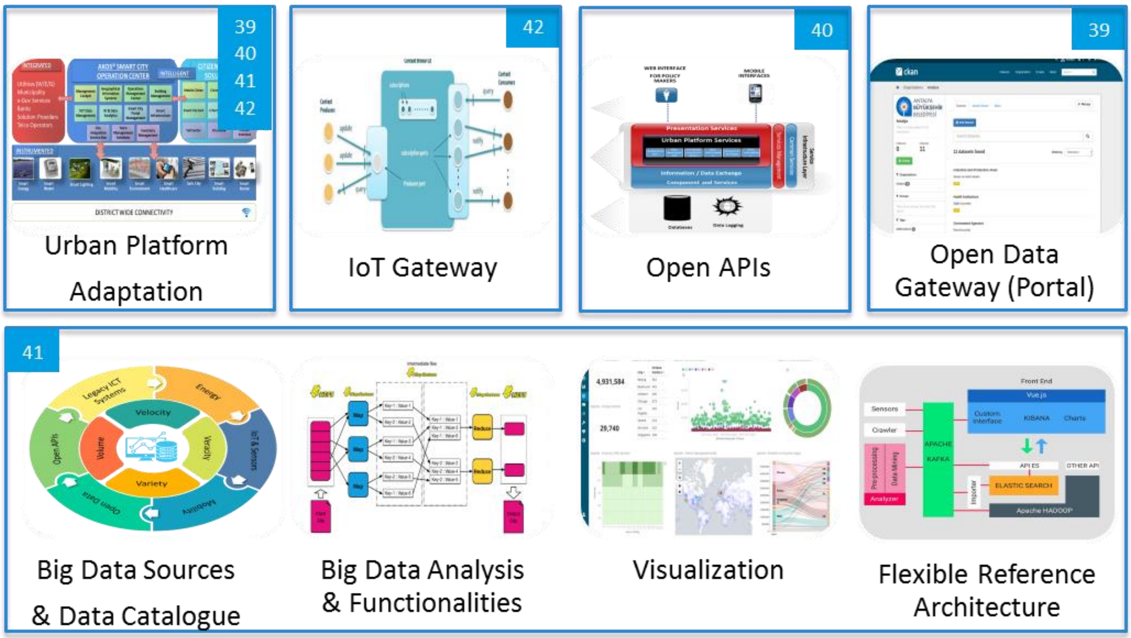

Figure 7 AUP Components and Actions

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

3.5 Legacy modules

Legacy modules are the modules that are co-created with third party institutions and embedded into the system for both getting input and giving output. Legacy modules in Antalya are: (see D4.22: Urban platform adaptation specifications in Antalya for a detailed version of Antalya services)

- Municipality: Municipality legacy systems and services
- Antalya e-Government Services (www.antalya.bel.tr)
    - ATASEM Application and Other Services
    - ATABEM Application and Other Services
    - Wholesale Market Hall Application
    - MUBİM Application, Requests and Complaints
    - Fast Payment Application
    - Municipal Debt Inquiry
    - Purchase Requests Login Application
    - Those Who Deceased Today
    - Duty Pharmacies
    - Application for Information
    - Creating Wish, Request, Complaint
    - Theatre E-Ticket Application
    - E-signature Query and Verification
    - Assembly and Council Decisions
    - Wedding Reservation and Status Inquiry
    - E-Declaration
- E-government system: Central e-Government services and integration points
- Banks: Services for payment purposes
- Solution Providers: 3rd party solution providers in Antalya Urban Platform
- Telco Operators: Telco operators and their services such as SMS and mobile APIs

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

4 Integration Architecture

This section presents the integration approach and architecture used in Antalya urban platform. Section starts by explaining the virtualization infrastructure and gives details regarding the integration architecture.

4.1 Integration of the Urban Platform

The Antalya Urban Platform uses VMs to provide the virtualisation of resources. Ubuntu 18.04.1 LTS was used as the base image for the VMs. The core set of services used to build the service platform are Apache Hadoop, Apache Kafka, Apache Spark, Apache Flume, Apache Zookeeper, Elasticsearch, and Kibana. Each of these will be described in the following subsections. Some of these services provide web UIs for developers and end users to interact with the platform. The currently exposed ports are also listed in Figure 8.

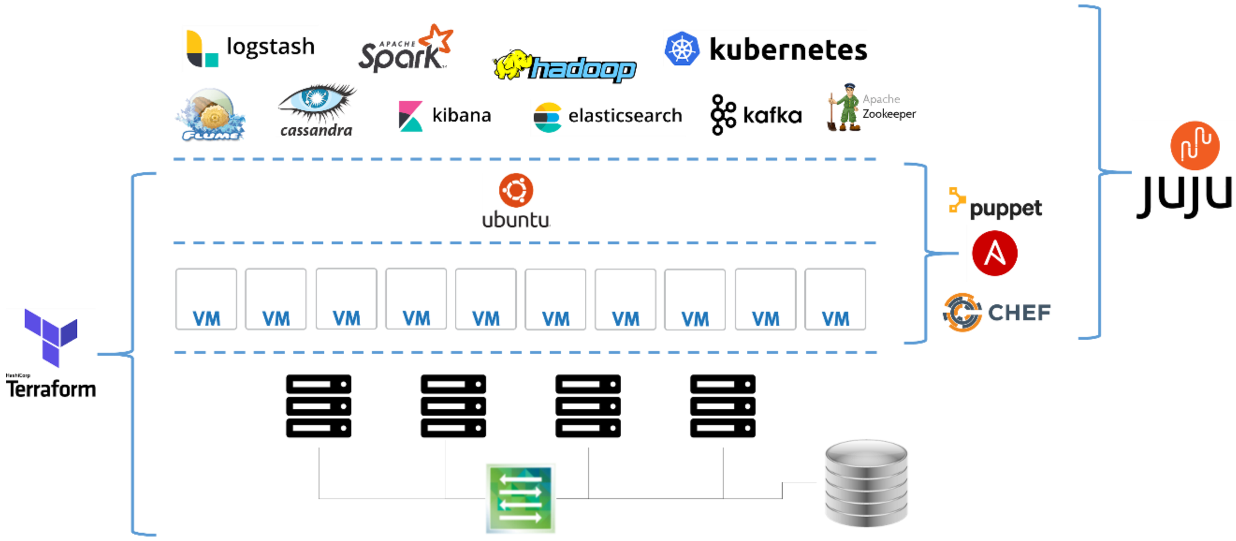

Figure 8 General System Architecture for the Urban Platform

4.1.1 Communication among big data components

The ElasticSearch Logstash Kibana (ELK) stack in Antalya Urban platform is implemented over a cluster, including a master node, data nodes and a coordinator node, that also hosts the Kibana service. This allows us to scale if required for high load of data in the future deployments of interventions. Clusters are encryption enabled. The X-pack feature provided by Elasticsearch enables encrypting traffic to, from and within the cluster.

This project has received funding from the European Union's Horizon 2020
research and innovation programme under Grant Agreement N°774477

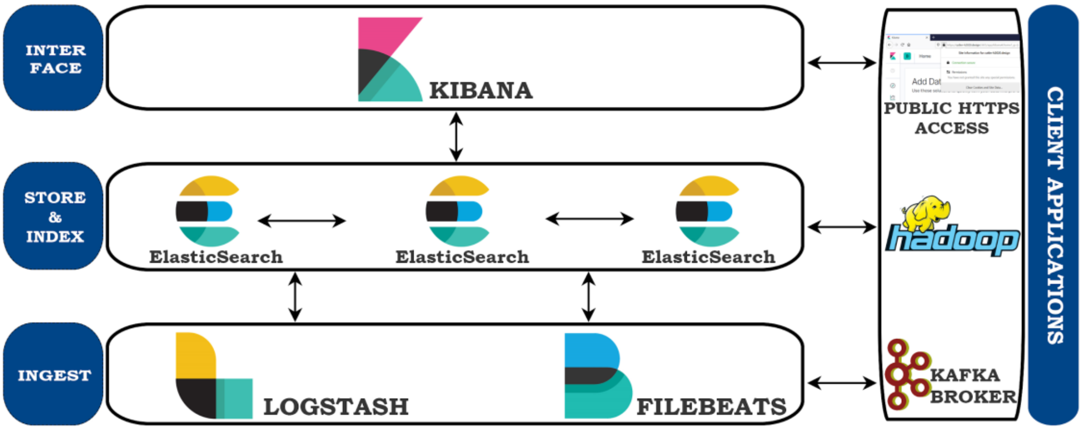

Figure 9 - All communication links in ELK cluster are encrypted

The first step to perform encryption and authentication in ELK stack is to generate certificates and certificate authorities (CA). This is done using the built-in utility provided by Elasticsearch, elasticsearch-certutil. This utility generates X.509 certificates and simplifies the process of generating certificates for the ELK stack. It takes care of generating a CA and signing certificates with the CA. This utility is used to first generate a CA for Elasticsearch and Kibana services and then output a keystore for each service that contains a node certificate, a node key and the CA. For Elasticsearch, as the certificates are only used for communication within the cluster, they are generated with no hostnames or subject alternative names (SAN). This means that the same certificate and keystore can be used for each of the nodes in the cluster. For Kibana, the certificate is generated with the hostname and SAN fields set, as it is used for communication with end-users and their browsers, which requires hostname verification for certificate validation.

Integration and Messaging

As described in “D4.22: Urban platform adaptation specifications in Antalya”, Antalya urban platform uses Apache Kafka for messaging and integrating components via these messages. Kafka is a distributed, scalable and robust publish-subscribe messaging system for transferring large amounts of data. Data is transferred as **messages**. These messages are assigned to categories called **topics**. The elements that are used during Kafka distributed messaging are as follows[5]: Kafka is capable of sharing Avro and JSON messages.

Avro is an open source data serialization system that helps with data exchange between systems, programming languages, and processing frameworks. Avro helps define a binary format for your data, as well as map it to the programming language of

[5] https://kafka.apache.org/intro

your choice. This makes Avro is a better way of streaming data. Avro has a JSON like data model, but can be represented as either JSON or in a compact binary form. It comes with a very sophisticated schema description language that describes data.

Avro is reported to be a better choice for a number of reasons[6]:

- It has a direct mapping to and from JSON
- It has a very compact format. The bulk of JSON, repeating every field name with every single record, is what makes JSON inefficient for high-volume usage.
- It is very fast.
- It has great bindings for a wide variety of programming languages so you can generate Java objects that make working with event data easier, but it does not require code generation so tools can be written generically for any data stream.
- It has a rich, extensible schema language defined in pure JSON
- It has the best notion of compatibility for evolving your data over time.

Kafka supports Avro[7] messages and allows us to work with Schemas and Schema Registry as well as given in Figure 10.

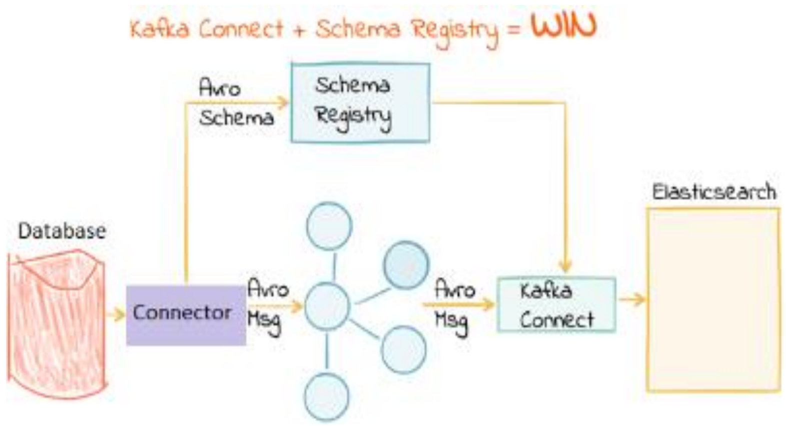

Figure 10 Kafka Connectors with Elasticsearch

The Elasticsearch connector allows moving data from Kafka to Elasticsearch. It writes data from a topic in Kafka to an index in Elasticsearch and all data for a topic have the same type.

Avro allows us to define schema sets for managing the data model and the data structures that will be defined in the next section. For example, the example given below provides the data model for a charging station.

[6] https://www.confluent.io/blog/avro-kafka-data/

[7] http://avro.apache.org/docs/current/

```
{
  "type": "record",
  "doc":"This event records the active data stream from a charging station",
  "name": "ChargingStationEvent",
  "fields" : [
●       Charge Model: int
●       Voltage: float
●       Current: float
●       Battery: float
●       Speed: float
●       Location: geopoint
●       Date: datetime
●       Id: int
    {"name":"charge_model", "type":"int", "doc":"The charging station model"},
    {"name":"voltage", "type":"float", "doc":"The current voltage parameter"},
    {"name":"battery", "type":"float", "doc":"The current battery status"},
    {"name":"speed", "type":"float"},
    {"name":"location", "type":"geo_point"},
    {"name":"date", "type":"datetime"},
    {"name":"id", "type":"int", "doc":Identifier of the charging station"},
    {"name":"payment",
     "type":{"type":"enum",
             "name":"payment_types",
            "symbols":["cash","mastercard","visa"]},
     "doc":"The method of payment"}
  ]
}
```


Schemas and Schema Registry let the producers or consumers of data streams know the right fields are need in an event and what type each field is. They document the usage of the event and the meaning of each field in the “doc” fields and protect

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement N°774477

downstream data consumers from malformed data, as only valid data will be permitted in the topic. Schemas and Schema Registry provides robustness, clarity and semantics, compatibility, conversion method and allows us to reduce errors by standardizing messaging. Figure 11 provides an example of how Schema registry fits into Kafka messaging layer.

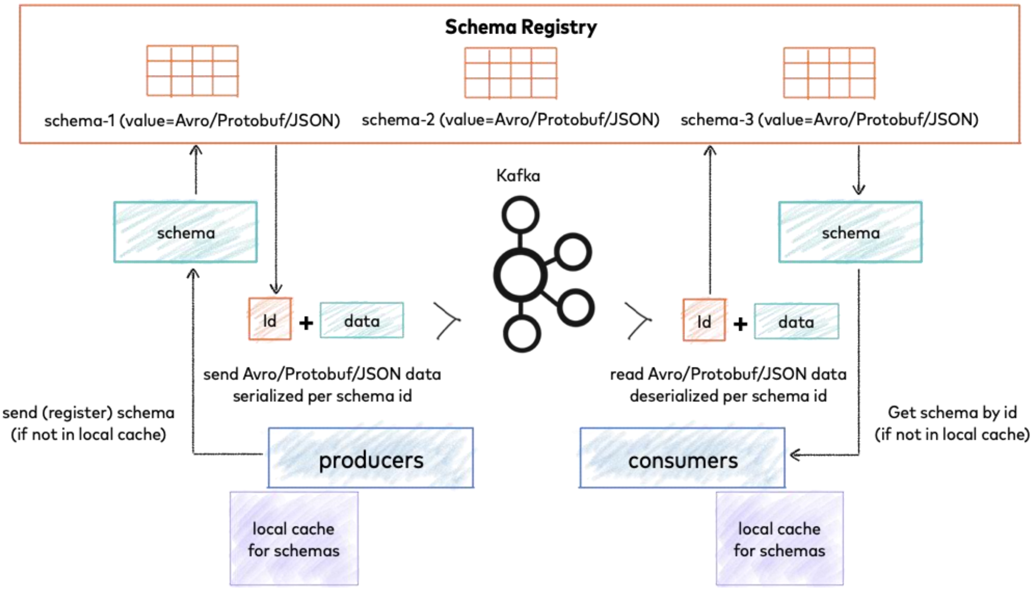

Figure 11 Confluent Schema Registry for storing and retrieving schemas

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

5 Data modelling

Data is the raw material of projects that need micro or macro decision-making. The data used for the projects can be stored in databases. However, in the projects including multiple sources and huge amounts of data like Smart City Urban Platforms, data is diverse and needs to be modelled. Data structure can be designed and modelled so that data can be used by different users from stored databases. Databases can be structured according to different criteria or goals of projects. Data models evolve with time to meet with new prospects of data usage and manage. Hierarchical data modelling and the network modelling were the first to be established and followed by the relational data modelling. After these, a model evolved named entity-relationship model which compensated real world's data expectations. Nowadays, the most common model is the non-relational data modelling, it goes over the limitations of relational data models when data used in projects is getting bigger and bigger.

Hierarchical modelling aims to model the data in a tree-like structure. The data are connected to each other through links and these links are connecting at the top of the tree. In network modelling, there is a flexible approach to represent objects and their relationships, which is that the schema is viewed as a graph where relationships are represented as arcs and objects as nodes. Networking model is different than others because it is not restricted to be a hierarchy or a framework. Moreover, in relational modelling there is the conceptual basis of relational databases. This is a process of designing data using relations, which are grid-like mathematical structures consisting of rows and columns. Lastly, in non-relational modelling the data are designed in key-value pairs. NoSQL databases are generally used in real time web applications.

In the Matchup project non-relational model is employed because of the large amount of data and its unstructured nature. Furthermore, in this model, the system can be measurable and observable. All data in the project collected from the city are going to be both real time or in batch mode also, some part of the data will be used for analyzing in real time and the other parts will be for historical data and open data catalogue.

FIWARE [10] data models are going to be reference models in Antalya because of interoperable structures for storing. Moreover, they allow easy data exchanges. These data models have been harmonized to enable data portability for different applications including, but not limited, to Smart Cities.

5.1 Data model standards

Smart City Standards for APIs & Data Models

This section introduces the NGSI-9 and NGSI-10 interfaces. Another relevant aspect is the usage of common data models in order to increase interoperability and reusability of applications. Within SynchroniCity, the FIWARE Data Models project that provides a wide set of data models based on NGSI specification was chosen.

5.1.1 NGSI data models and interfaces

NGSI is a protocol developed by OMA to manage Context Information, which provides the following functionalities:

- Manage the Context Information about Context Entities.
- Access (query, subscribe/notify) to the available Context Information about Context Entities.

Context Entities are entities that are described by Context Information and also described by the Context Information Model.

The Context Information Model details how Context Information is structured and associated to Context Entities in order to describe their situation. In this model, Context Information is organized as Context Elements, which contain set of Context Attributes and associated metadata. Figure 12 gives also some examples on entities that can be used as Context Entities.

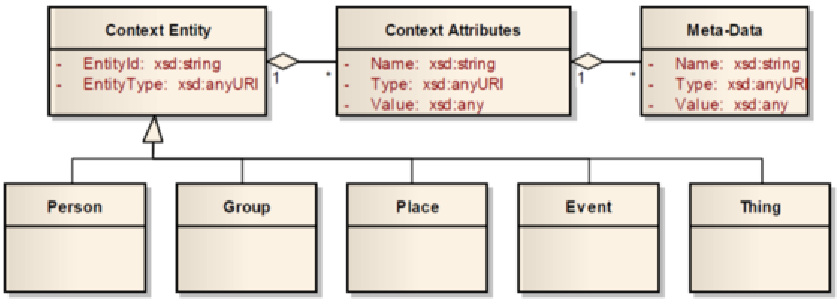

Figure 12 NGSI Context Information Model.

NGSI defines the following two interfaces for managing information based on this Context Information model:

- NGSI-9: provides operations to obtain the availability information about context entities and their attributes; this interface contains operations to register context entities and to discover context information providers.
- NGSI-10: provides operations for exchanging information about entities and their attributes; this interface contains operations to perform queries, update or activate subscription on context entities.

Besides the adoption of common APIs, the adoption of common models is a fundamental step to allow for a scalable data ecosystem that supports the interoperability and re-usage of data and applications working on top of data. The FIWARE Community is promoting the development of re-usable and harmonized Data Models under the umbrella of the FIWARE Data Models initiative. These data models are reusing and extending the work performed under the GSMA IoT Big Data initiative. The work conducted under FIWARE is evolving on a daily basis by taking into consideration requirements from Data Models' adopters.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

FIWARE has already harmonised multiple sets of data models within many domains such as Environment, Street Lighting, Civic Issue tracking, Device etc. A short description of the current data models can be found below:

- **Environment**. A model to enable the monitoring of air quality and other environmental conditions for a healthier living. In particular, covered entity types include:
  - *AirQualityObserved*: an observation of air quality conditions at a certain place and time.
  - *AeroAllergenObserved.* It describes aero allergens observed at a given

  location and related overall allergen risk.

- **Transportation**. Transportation data models for smart mobility and efficient management of municipal services. The covered entity types include:
  - *TrafficFlowObserved*: a recorded observation of traffic flow.
  - *Road*: a geographic and contextual description of a Road.
  - *Vehicle*: a specific vehicle instance.
  - *VehicleModel*: a model of vehicle, capturing its static properties such as dimensions, materials or features.
  - *EVChargingStation.* a public charging station supplying energy to electrical vehicles.

- **Street Lighting Data Models**. Streetlights, commonly known as 'lamp-posts', are designed to make the streets safer for pedestrians and drivers. It encompasses the following entity types:
  - *Streetlight.* a particular instance of a streetlight. A streetlight is composed by a lantern and a lamp. Such elements are mounted on a column (pole), wall or other structure.
  - *StreetlightGroup.* a group of streetlights being part of the same circuit and controlled together by an automated system.
  - *StreetlightModel.* a model of streetlight composed by a specific supporting structure model, a lantern model and a lamp model. A streetlight instance will be based on a certain streetlight model.
  - *StreetlightControlCabinet.* an automated equipment, usually on street, typically used to control a group(s) of streetlights, i.e. one or more circuits.

- **Weather**. Weather observed, weather forecasted or warnings about potential extreme weather.

5.1.2 ETSI-CIM

In 2018, the ETSI Industry Specification Group for cross-cutting Context Information Management (ISG CIM) released its first specification GS CIM 004 [21]. This specification defines a simple way to send or request data and its context such as the meaning, related information, source or licensing of that data. Smart cities will be the

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

first ones to benefit from this specification. Group Specification CIM 004 defines a standard Application Programming Interface (API) for Context Information Management enabling close to real-time access to information coming from many different sources. Furthermore, the ETSI ISG CIM also released the NGSI-LD API [22] as the Context Information Management API. The rationale is to reinforce the fact that the API leverages on the former OMA NGSI 9 and 10 interfaces and FIWARE NGSIv2 to incorporate the latest advances from Linked Data. Also in SynchroniCity, this specification is perceived as the next step after NGSIv2 adoption.

FIWARE data models consists of twelve data models to answer different expectations of Smart City concept varying by the city elements where each one has a set of particular attributes and relationships:

Alarms: Events related to risk or warning conditions which require taking action.

Parks & Gardens: Data models intended to make an efficient, effective and sustainable management of green areas.

Environment: Enable to monitor air quality and other environmental conditions for a healthier living.

Point of Interest: Specific point locations that someone may find useful or interesting. For instance, weather stations, tourist landmarks, etc.

Civic Issue tracking: Data models for civic issue tracking interoperable with the de-facto standard Open311.

Street Lighting: Modelling street lights and all their controlling equipment towards energy-efficient and effective urban luminosity.

Device: IoT devices (sensors, actuators, wearables, etc.) with their characteristics and dynamic status.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

Transportation: Transportation data models for smart mobility and efficient management of municipal services.

Indicators: Key performance indicators intended to measure the success of an organization or of a particular activity in which it engages.

Weather: Weather observed, weather forecasted or warnings about potential extreme weather conditions.

Antalya will follow the data model developed according to a reference FIWARE data model scheme as given below:

Data Model	Parameters
Charge Stations	• Charge Model: **int** • Voltage: **float** • Current: **float** • Battery: **float** • Speed**: float** • Location: **geopoint** • Date: **datetime** • Id: **int**
Public Lighting	• Decibel: **float** • Heat: **float** • Working hours of light: **int** • Spending of lamb: **int** • Open/close: **boolean** • Consumed electricity: **int** • Carbon dioxide measure: **int** • Location: **geopoint** • Date: **datetime** • Id: **int**

Table 2 Data Models Developed

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

Table 3 ANTALYA Interventions and FIWARE Data Model Dependency

Action Pillar	Action Context	MAtchUP Action	Data Source Type	Collected Data / Parameter
District/ Buildings	**Domotics & Smart Controls**	**A6** Smart controls & domotics (SAM)	• Alarms • Environment • Point of Interest • Device	• Location: geopoint • Date: datetime • ID: int • Open/close: boolean • Alert: (text/URL): string • Indoor Air Quality: string • VOC (mg/m3): float • CO2: float • Temperature: float • Moisture: float • Pressure: float • Outdoor Temperature: float
	Building Integrated RES	**A3** PV installation for public buildings (ANT) **A4** Solar thermal collectors' installation for residential buildings (ANP)	• Alarms • Environment • Point of Interest • Device	• Location: geopoint • Date: datetime • ID: int • Open/close: boolean • Energy generation (kW): float • Consumption value (kW): float • Solar PV production (Daily, Monthly and Annual - kWh): float • Solar PV performance index: string • Radiation and module temperature: float

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

	Storage	**A5** Electrical storage for buildings and charging stations (ANT)	• Alarms • Environment • Point of Interest • Device	• Location: geopoint • Date: datetime • ID: int • Instant battery charge / discharge power (kW): float • Battery SoC value (kWh or%): float • Battery cycle count: int • Battery efficiency rate: float
City Infrastructures	**Smart grids**	**A7** Smart meters (SAM)	• Alarms • Environment • Point of Interest • Device	• Location: geopoint • Date: datetime • ID: int • Open/close: boolean • Alert: text/URL: string • Electricity Consumption: float • Hot Water Consumption: float
	Public lighting	**A8** Led Public Lighting (ANP) **A9** Smart control of public lighting (ANP)	• Alarms • Environment • Parks & Gardens • Point of Interest • Street Lighting • Device	• Location: geopoint • Date: datetime • ID: int • Open/close: boolean • Heat: float • Working hours of light: int • Spending of lamb: int • Consumed electricity: int • Carbon dioxide measure: int
	Urban RES	**A10** PV systems with a total capacity 5MWp (ANT)	• Alarms • Environment • Point of Interest • Device	• Location: geopoint • Date: datetime • ID: int • Stored Energy: float • Spent Energy: float

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

				• energy generation (kW): int • consumption value (kW): int • Solar PV production (Daily, Monthly and Annual - kWh): int • Solar PV performance index: text • Radiation and module temperature: int
		A11 LFG and electricity generation (ANT)		• Location: geopoint • Date: datetime • ID: int • Alert: text/URL: string • Energy generation value (kW): float • Energy consumption value (kW): float • Incoming Solid Waste (tonnes): float • Stored Solid Waste (tonnes) : float
Mobility	**Electric vehicles**	**A13** 4 Hybrid buses (ANT)	• Alarms • Point of Interest • Device • Transportation • Parking	• Location: geopoint • Date: datetime • ID: int • Battery Status: float • Damage and Overturning: float • Instant Speed: float • Average Speed: float • Total Distance: float • Error Conditions: float • Time of Use: float
		A14 20 e-vehicles for municipality fleet (ANT)		• Date: datetime • ID: int • Battery Status: float • Damage and Overturning: float • Instant Speed: float

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

				- Average Speed: float - Total Distance, : float - Estimated Remaining Distance: float - Instant Operation Mode: int - Error Condition: float - Time of Use: float
		A15 30 e-Scooters (ANT)		- Location: geopoint - Date: datetime - ID: int - Battery Status: float - Damage and Overturning: float - Number of Users: fint - Instant Speed: float - Average Speed: float - Total Distance: float - Internal Battery Temperature: float - Estimated Remaining Distance: float - Instant Operation Mode: int - Error Condition: float - Time of Use : float
	Charging stations	**A16** 5 e-vehicle charging points (ANT) **A17** 5 e-scooter charging stations (ANT)	- Alarms - Environment - Point of Interest - Device - Transportation - Parking	- Location: geopoint - Date: datetime - ID: int - Charge Model: int - Voltage: float - Current: float - Battery: float - Speed: float - Location: geopoint

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

				• Date: datetime
	Multi-modality	**A19** 2 multimodal hubs (ANT)	• Alarms • Environment • Point of Interest • Device • Transportation • Parking	• Awareness of location of multimodal hubs: string • Awareness of services at the mobility hub: string • Number of modal shift user: float • Frequency of modal hub utilization: float • Effect of multimodal hub on mobility behaviour of users: float • Reason to journey areas (multimodal hub stations): string • Distance between transportation modes (m) : float • Travel distance between transportation modes (min.) : float • All demographic properties of the users: string
		A20 Integrating existing light rail with eScooter station and bus transport (ANT)		• Awareness of location of integration among light rail, e-Scooter and bus transport: string • Awareness of services among integration of light rail, e-Scooter and bus transport: string • Number of modal shift user among light rail, e-Scooter and bus transport: int

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

				• Frequency of modal shift utilization among light rail, e-Scooter and bus transport: float • Effect of integration of light rail, e-Scooter and bus transport on mobility behaviour of users: string • Reason to journey areas (to integration of light rail, e-Scooter and bus transport): string • Distance among all these three modes (m): float • Travel distance among all these three modes (min.): float • All demographic properties of the users: string
		A21 Applying last mile mobility for citizens via integration of e-Scooter station with the light rail station (ANT)		• Awareness of location of integration between e-Scooter station with the light rail station: string • Awareness of services for the integration: string • Number of user: int • Frequency of utilization: float • Effect of utilization of this integration on behaviour of users: string • Reason for the utilization: string • Distance between e-Scooter station with the light rail station (m): float • Travel distance between e-Scooter station with the light rail station (min.): float • All demographic properties of the users:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

				string
		A23 2 Intelligent Intersections (ANT)		• Number of vehicles in each intersection leg: int • Decrease in vehicle delays: float • Reduction in CO2, float • Reduction NOx: float • Reduction HC: float • Reduction PM e: float • Traffic volume for each leg (veh/hr): float • Fossil fuel save (per each vehicle): float • Vehicle types : string
Non-technical actions	**Citizens' engagement**	**A35** Citizens' feedback channel (SAM)	• Point of Interest • Civic Issue tracking • Street Lighting • Device • Transportation • Parking • Questionnaires • Surveys	• Location: geopoint • Date: datetime • ID: int • Age: int • Income: float • Household income spent on energy: float • Number of Family Members: int • Commuting Method: string • Environmental awareness level: string • Number of Citizens involved in policies formulation: int • Local community participation level: string

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

5.2 Data model interoperability

FIWARE data models are flexible to be coherent with European data standards vocabulary DCAT [6, 7] and metadata standards DCAT-AP, related outcomes of Turkish Interoperability Principles [17] and International standards like oneM2M [18].

5.3 Open data catalogues

Antalya Open Data Portal is developed with CKAN infrastructure. Its user interface is customized and the platform translated into the Turkish language. The platform is ready to be opened in which the beta test is carried out. The users will be able to select or search through the given pillars or search bar as depicted below:

Figure 13 Antalya Open Data Portal

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

Open data portal search box allows users to search for specific datasets in the platform by typing relevant keywords. If the search result is too large, users can filter the result by tags, formats, organizations. Antalya open data portal allows to categorize different data sets and provides these with explainable symbols for end-users to directly go for a specific category (e.g. agriculture) based on its symbol. Free text search also allows to find the relevant tagged data sets easily.

The following section describes the user and organization structure followed by Antalya.

Users:

Each user belongs to an organization and has a different role like member, editor or admin.

- Members of the organization can see the private datasets of the organization.
- Editors can add, edit or delete datasets of the organization.
- Admins can manage anything like user management, configuration options, dataset management.

Organizations:

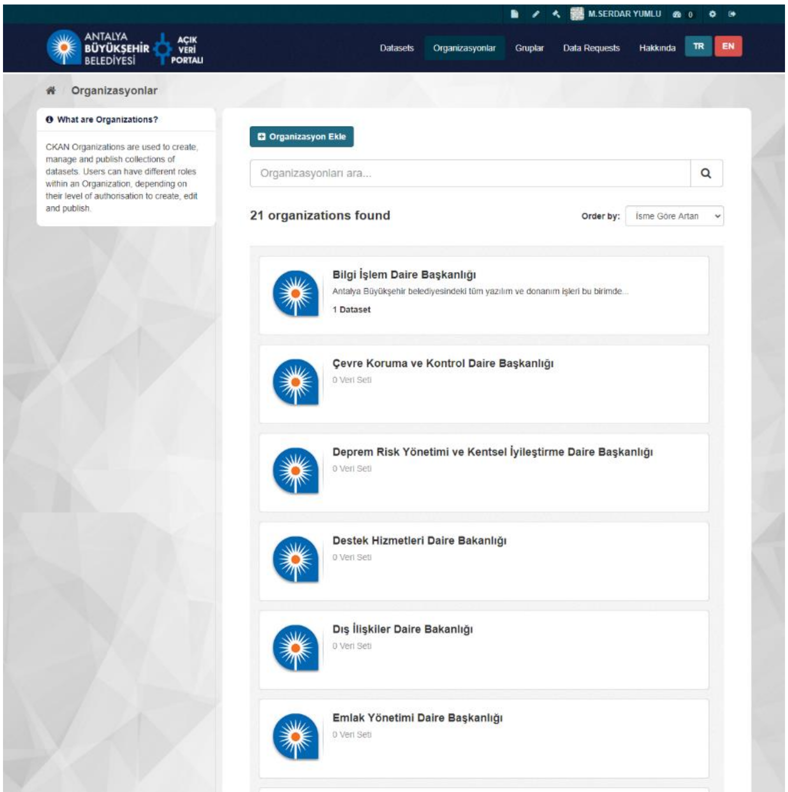

Figure 14 Antalya Open Data Portal-Organization List

Organizations are units of the Antalya's Metropolitan Municipality. Each organization has their users and datasets. Organizations can make their datasets public or private.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

Public datasets can be seen by anyone visiting the portal. Private datasets are only for the members of the organization.

Datasets:

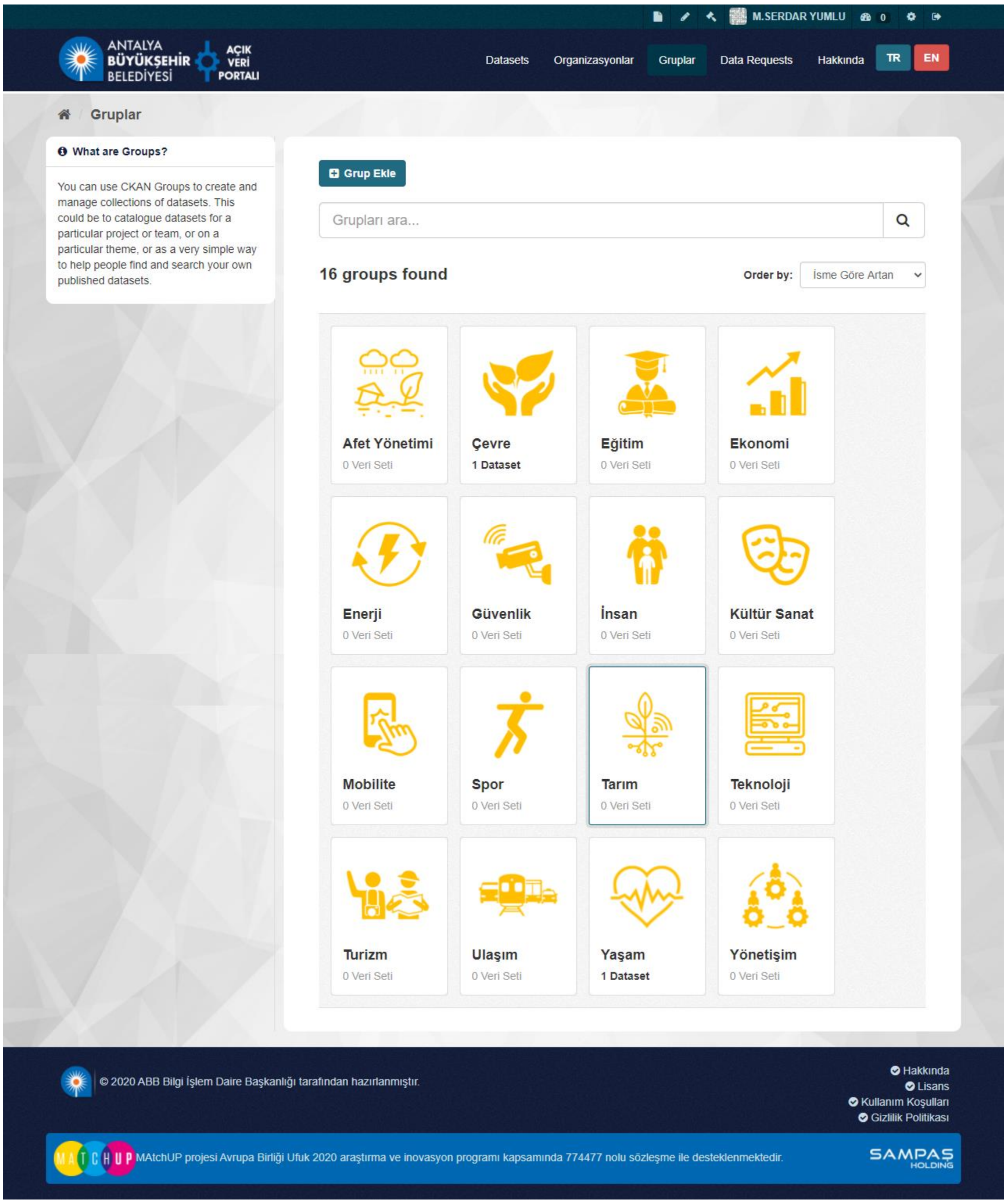

Figure 15 Antalya Open Data Portal-Datasets

Users can upload many types of datasets like XLS, PDF, PNG, JSON, GeoJSON, CSV, ODP, etc. All datasets can be visualised together in a datasets page or each dataset can be visualised in its organization page.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement N°774477

Users can categorize the datasets according to organizations, groups, tags, formats, licenses.

After selection of a dataset, users can see the information about the selected dataset like created date, last update, licence, format, id, size, etc.

Datasets can be viewed in different options (graph or maps). If a dataset has geojson information, it can be visualised as a map. As shown in the figure below, the map view shows the points of the data, if we click on each point, a modal appears and we can see other information about that point. Other view formats are table format or raw data format.

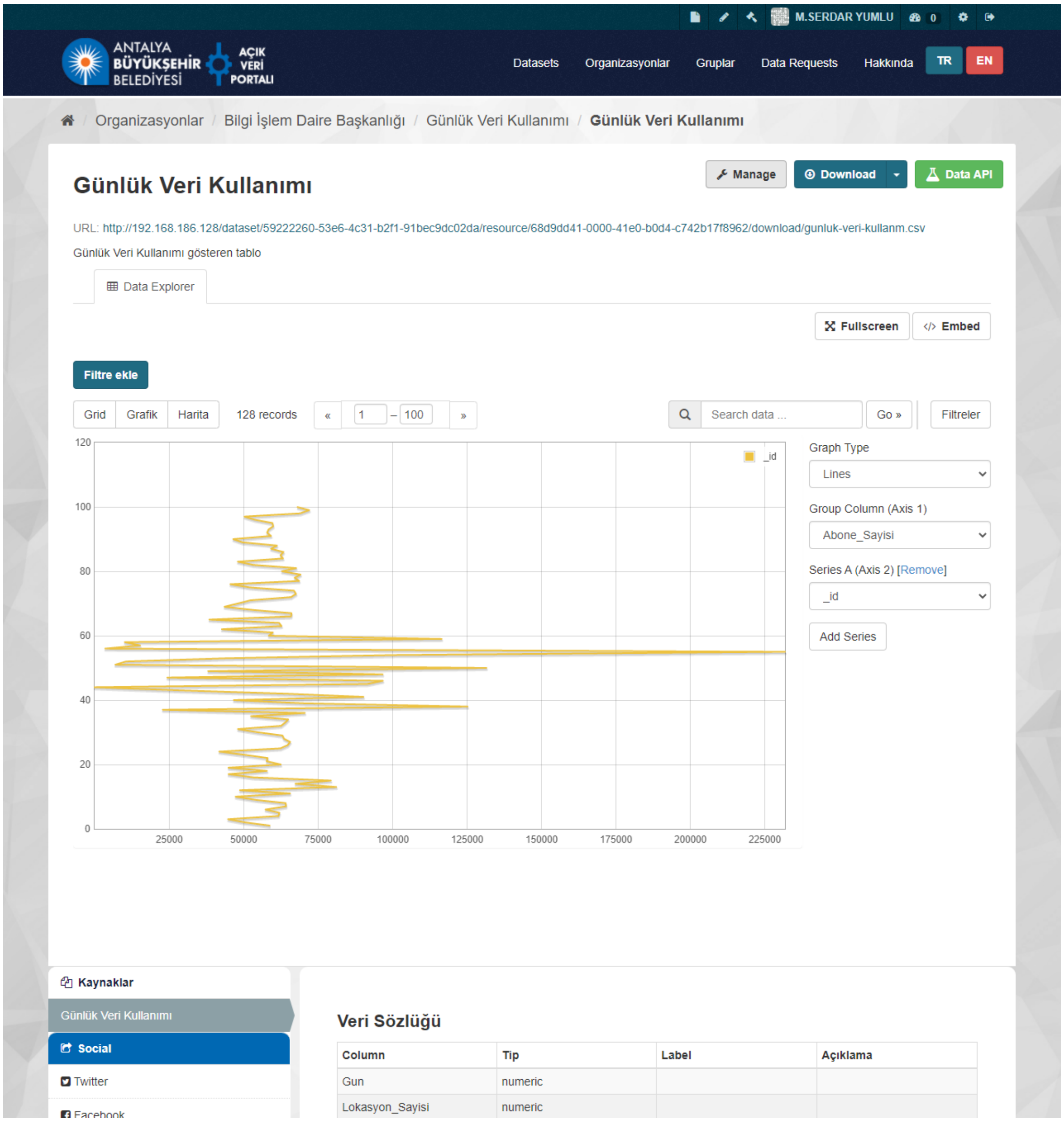

Figure 16 Open Data Portal Graph View

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

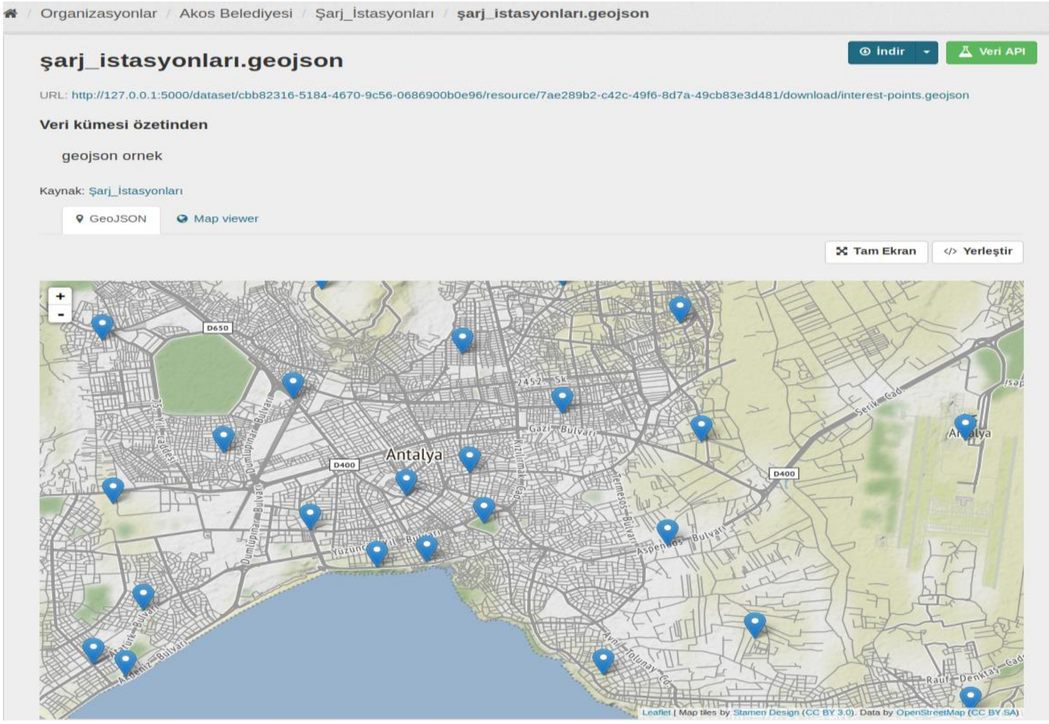

Figure 17 Antalya Open Data Portal-Map View

Users can download datasets to their devices and there is an API link for developers to use the datasets in their applications.

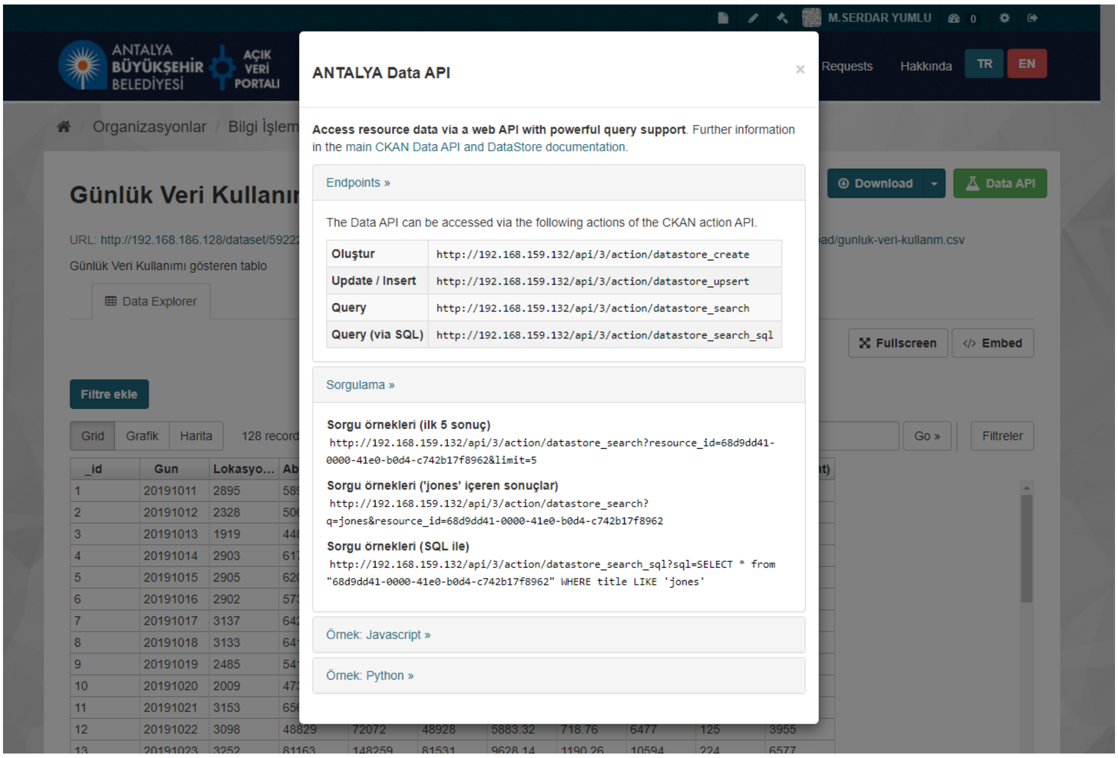

Figure 18 Antalya Open Data Portal-API

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

6 Security aspects

In smart city initiatives and the evolving world of social media and Internet of Things in our daily lives, privacy is becoming one of the most important aspects. Privacy deals with the protection of personal data. In Europe, all smart city initiatives and R&D projects under Horizon 2020 shall comply with ethical principles and relevant national, Union and international legislation, including the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights and its Supplementary Protocols. According to Data Protection Directive (95/46/EC) of the EC, personal data “means any information relating to an identified or identifiable natural person; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one of more factors specific to his physical, physiological, mental, economic, cultural or social identity”[8].

6.1 Regulations and Security

There are some regulations to protect data collected through a variety of channels which are the data sources described in D4.22. Besides the technical issues such as architecture, interoperability, and approachability cornering open data of the Urban Platform, regulations of security play an important role when collecting data from the city, personal or private data of citizens.

The General Data Protection Regulation has been identified as the most relevant to smart city projects from a Personal Data Protection and Privacy perspective, even though reference should also be made on Directive 2016/680/EC to the extent that personal data are processed in the context of criminal investigations and prosecutions by competent public bodies. In addition, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector, as modified by Directive 2009/136/EC (e-Privacy Directive) should be followed. Procedures and appropriate technologies, appropriate European and Internet security standards from ISO, ITU, W3C, IETF and ETSI have to be used to ensure the protection of personal data.

6.1.1 EU General Data Protection Regulation (GDPR)

GDPR is a regulation which protects data and privacy for all individuals within the European Union (EU) and the European Economic Area (EEA). The GDPR was adopted on 14 April 2016 by the EU Parliament. Subsequently, it was enforced on 25 May 2018. GDPR protects primary rights and freedom of natural persons and in private their right to the protection of personal data.

[8][1] STEP H2020, D2.3 Guidelines for Ethical legal issues and data protection

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement N°774477

Since Antalya is a non-EU lighthouse cluster (Turkey) in MAtchUP, data transfer to non-EU countries will require agreements and specific authorizations from the national data protection authority.

6.1.2 Data Security

Data should be secured from viruses, hacker attacks, forgery, etc. Security means protection of information and information systems by ensuring confidentiality, availability, integrity, authentication, and non-repudiation [19].

- Confidentiality: Information is not made available or disclosed to unauthorized individuals and entities.
- Availability: Data/information has to be available, only authorized persons can remove it, in accordance to law.
- Integrity: only authorized persons can modify the data/information, in accordance with law.
- Authentication must be preserved (data/information must be authentic).
- Non-repudiation – participants will not be able to successfully challenge the authorship of the data provided.

The following table presents the importance of each EU level legislation for MAtchUP Antalya lighthouse project and how it will be considered during the implementation.

Legislation/Act/Article	MAtchUP Impact
Directive 95/46/EC on protection of personal data (Data Protection Directive)	Processing of personal data and collection during Antalya implementation
Directive 2002/58/EC on privacy and electronic communications (e-Privacy Directive)	Protection of privacy in web and mobile based communication channels in Antalya Confidentiality of information
Directive 2009/136/EC (Cookie Directive)	Management of hidden information Exchange between the platform and users
Accessibility (Article 21)	Increasing accessibility issues in Antalya
Data security (2005/222/JHA)	Confidentiality, integrity, authentication and non-repudiation features to be provided during the implementation in Antalya

Table 4 EU level legislation for MAtchUP Antalya lighthouse project

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

6.1.3 Turkey KVKK 6698 Law

In non-EU countries, instead of GDPR, national personal data protection laws are valid. In Turkey, Act. 6698 (Turkish Personal Data Protection Law no. 6698) provides regulations for all personal data related initiatives including smart city projects. Act. 6698 aims to protect the fundamental rights and freedom of persons, privacy of personal life in particular, while personal data are processed, and to set forth obligations of natural and legal persons who process personal data and procedures and principles to comply with for the same.

KVKK was published on 7 April 2016 and became effective on 7 October 2016. ‘The purpose of this Law is to protect the fundamental rights and freedom of persons, privacy of personal life in particular, while personal data are processed, and to set forth obligations of natural and legal persons who process personal data and procedures and principles to comply with for the same.’

Within the context of compliance with Law No. 6698, institutions must analyse the life cycle of the domain very well and establish and operate technological infrastructures that can manage it in the process. In practice of this law, the following terms are being defined within the scope of data protection:

- **Personal Data:** Any information relating to an identified or identifiable natural person;
- **Explicit Consent**: Freely given specific and informed consent;
- **Anonymization**: Rendering personal data by no means identified or identifiable with a natural person even by linking with other data;
- **Data Anonymisation** aspect is to preserve sensitive information by first identifying the sensitive data and their attributes, and then analyzing the data/attributes and applying pre-defined privacy preserving mechanisms (pseudonymization, generalization, suppressing, k-anonymity, differential privacy) accordingly. This also takes into account the trade-off between privacy and utility;
- **Data subject:** Natural person whose personal data are processed;
- **Processing of personal data:** Any operation which is performed upon personal data such as collection, recording, storage, preservation, alteration, adaptation, disclosure, transfer, retrieval, making available for collection, categorization or blocking its use by wholly or partly automatic means or otherwise than by automatic means which form part of a filing system;
- **Data processor:** Natural or legal person who processes personal data based on the authority granted by and on behalf of the data controller;
- **Data controller:** Natural or legal person who determines the purposes and means of the processing of personal data, and who is responsible for establishment and management of the filing system.

Data Ingestion and Anonymization of Big data flow and ingestion is provided in Figure 19.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement N°774477

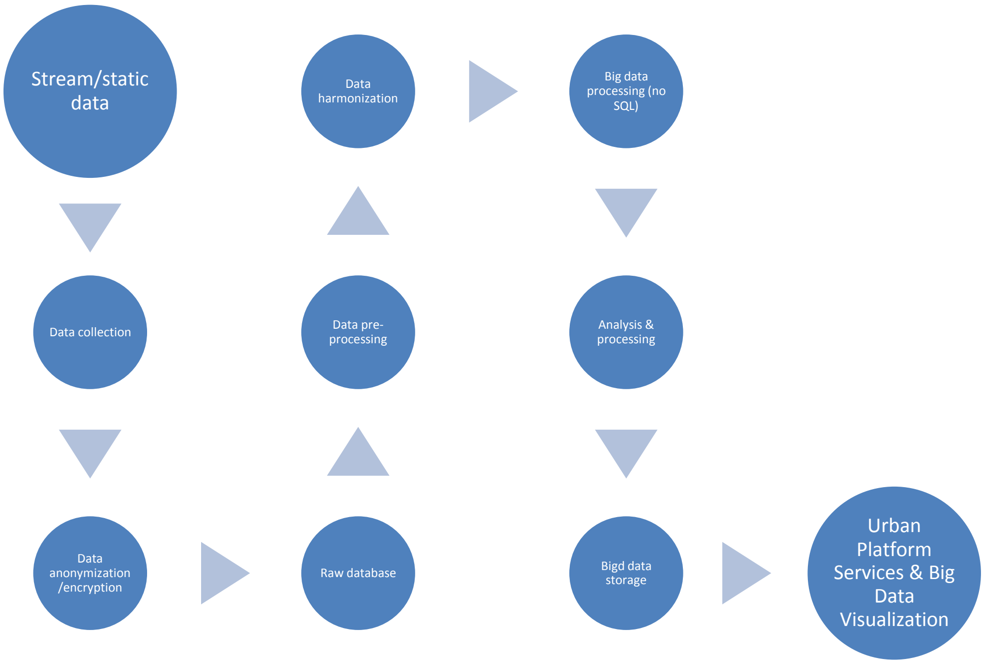

Figure 19 Big Data Ingestion & Data flow

6.1.2 Data Ingestion

For data ingestion and integration, Antalya Urban Platform uses Apache Kafka as a messaging gateway as explained in the general architecture in Section 3.4.

Apache Kafka was introduced, in order to provide real-time streaming data pipelines that reliably get data between applications. The Kafka cluster stores streams of records in categories called topics which is a category or feed name to which records are published. Each record consists of a key, a value, and a timestamp. Each crawler consists of a producer and a consumer. The producer API allows an application to publish a stream of records to one or more Kafka topics and the consumer API allows an application to subscribe to one or more topics and process the stream of records produced to them. Once the data is ready it is then pushed to Elasticsearch through the use of the Elasticsearch API[9]. A typical structure of a crawler and the process executed in the crawler is:

1. Start the producer of the crawler
2. Gather the data from the source or open file
   2.1. Data Quality control (Are the data complete? Are the data consistent? Are there outliers? Are there duplicates? etc.)
3. Harmonization (pattern handling)

[9] https://www.elastic.co/guide/en/elasticsearch/reference/current/docs.html

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

3.1. Enrich the dataset with information needed in Elastic & Kibana (e.g. station geolocation)

3.2. Rename parameters to match the environmental data model protocol established in Section 5.1

3.3. Metadata enrichment to include other needed fields such as station geolocations and also some values (e.g. station names) needed to be represented with a Latin character set

3.4. Transform data to perform data aggregations

4. Broadcast the data in the Kafka cluster
5. The consumer of the crawler is triggered and consumes the data from the Kafka topic

Injection of the data into Elasticsearch in the form of JSON documents.

This is a general process that is used in all cases. However, there are instances where the data are accessed through requests to certain APIs and they refer to near real-time data that need to be collected daily. In the production servers of the pilots, cron jobs have been set up in order to run the producers on a daily basis.

The planned data anonymization and pseudonymization example are as follows:

With the pseudonymization, the personal data is transformed in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information. This means that the personal identifiers are removed from the data and stored in a separate database:

ID	First Name	Last Name	Date of Birth		ID	First Name	Last Name	Date of Birth
1	Kimberly	White	09-25-1987	→	1	Eujixlyx	Ciysn	04-26-1993
2	Anthony	Redding	02-15-1974	→	2	Hxhzhlt	Zwoqlju	01-04-1986
3	Patricia	Smith	04-19-1982	→	3	Zzrdmgqh	Twigh	12-15-1976
4	Craig	Farmer	06-07-1991	→	4	Hfmoz	Dlqxan	07-29-1995
5	Edward	Palmer	05-17-1985	→	5	Vixzgi	Mthdyg	04-17-1979
6	Christine	WIlliams	11-13-1992	→	6	Hloiihkkg	Gzuzpkzr	09-25-1991
7	Ruth	Brown	02-01-1977	→	7	Cjnl	Bqaql	07-17-1982
8	Wanda	Wilson	05-18-1993	→	8	Mnywn	Jauebo	09-14-1990
9	Juan	Miller	03-22-1986	→	9	Ubtc	Ockwdt	06-24-1988
10	Christopher	Johnson	10-22-1988	→	10	Nbybjhrwkck	Cfpjqkm	09-02-1978

Figure 20 Data Pseudonymization

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

Regarding the anonymization, the end result can be depicted as:

Name	Token/Pseudonym	Anonymized
Clyde	qOerd	xxxxx
Marco	Loqfh	xxxxx
Les	Mcv	xxxxx
Les	Mcv	xxxxx
Marco	Loqfh	xxxxx
Raul	BhQl	xxxxx
Clyde	qOerd	xxxxx

Figure 21 Data anonymization

This project has received funding from the European Union's Horizon 2020
research and innovation programme under Grant Agreement N°774477

7 Interoperability tests for interventions

This chapter describes the interoperability testing methodology and approach that is used in order to perform the technical testing and validation of the MAtchUP implementation.

The objective of Interoperability tests for interventions is to setup a strong evaluation framework (see also WP5) to be deployed in the three lighthouse cities with the aim to assess the effectiveness of the proposed intervention, deployed in the associated individual actions. This framework addresses this performance not only from a pure technical perspective but also considering an economic and social assessment as well. CITYKEYS and SCIS initiatives are the main references for the identification of KPIs and procedures. It supports the definition of the monitoring programs within MAtchUP mainly focusing to the selection of DAQ and data sets. The purpose of the technical testing was to ensure the correct and reliable operation and the interoperability of components. The testing procedure includes technical testing processes including unit testing and integration testing and load testing for the implementation of interventions and for the validation and interoperability among MAtchUP implementation in Antalya aligned with Valencia and Dresden.

Testing methodology considers specific testing requirements and considerations for each technical component. The following testing items have been developed during the implementation phase and updated accordingly.

Unit testing

Unit testing refers to tests that verify the functionality of a specific section of code, usually at the function level. In an object-oriented environment, this is usually at the class level, and the minimal unit tests include the constructors and destructors. Unit testing alone cannot verify the functionality of a piece of software, but rather is used to assure that the building blocks the software uses work independently of each other. Each component will apply its own unit testing processes and will report the outcome of the unit testing of specific functionalities and specific REST APIs to the project members using MAtchUP repository. Each integration component will define its input / output definitions and will provide relevant unit testing of specific APIs.

Integration Testing

The integration testing's purpose is to verify the interfaces between components against a software design. Software components may be integrated in an iterative way or altogether. For the integration of the MAtchUP components the former method was considered a better practice since it allowed interface issues to be discovered and fixed more quickly. Integration testing will be done to check the validity of each API component. Each API will be tested internally to check whether it provides the defined interfaces and also will be tested through the integration endpoints in specific parts of the platform. Several different components are implemented in Antalya Urban Platform such as:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

- New open data gateway
- New Open API developments
- Big data Functionalities
- IoT adaptors

As aforementioned, one of the main components for the integration among the platform will be Apache Kafka. Apache Kafka manages data, getting large amounts of it from one place to another, rapidly, scalable, and reliably. In computing, a common term for transferring data is messaging, and that's how Apache Kafka would describe itself, as a messaging system, but unlike other messaging systems, Apache Kafka is tailored for high throughput use cases, where vast amounts of data need to be moved in a scalable, fault tolerant way.

Messaging has lots of sense because it establishes a fairly simple paradigm for moving data between applications and data stores; however, when it comes to a large scale implementation, traditional message systems can struggle, especially with scalability. The means to collect and distribute data as messages relies on the role of a messaging broker. There are a lot of variables that determine the reliability and performance of a messaging system, a big one being message or data packet size. Larger messages can put severe strain on message brokers, and this is a challenge because you may not be able to control messages coming from some systems. Furthermore, a messaging environment is dependent on the ability for message consumers to actually consume at a reasonable rate. There is also the challenge of fault-tolerance.

Under this perspective Kafka will used to integrate different data sources. Elasticsearch will be working on top of Kafka communication bus. Data will be fed to Elasticsearch by using Kafka micro services.

In each of these components, different APIs and integration points are identified. As MAtchUP will include and integrate large data sets the following aspects of each component will be included in the interoperability test plans:

- data collection from sensors/devices through IoT agents
- data extraction, transformation and loading processes (mainly through Orion Context Broker)
- service and API integrations through ESB and/or direct APIs
- data integration through legacy services
- integration through Kafka for producers and consumers

Load Testing (Stress Tests) & Performance Testing

The aim of this procedure is to determine how the platform performs in terms of responsiveness and stability under a particular workload, and also to validate other quality attributes of the system, such as scalability, reliability and resource usage. For stress testing of several web components, web stress testing tools like Apache JMeter [21] will be used. For the performance testing of REST APIs SOAPUI [22] based tooling will be used to check the throughput of specific components.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under Grant Agreement N°774477

The definition of an open specifications framework, ensuring data reusability and interoperability through an open data and open APIs concept will complement the ICT Urban Platform developments, assessing the openness of all data generated by MAtchUP. UPV, TUD and SAM will define the integration of these plans within the open specifications concept of the three urban platforms. MAtchUP will feed the SCIS and CITYKEYS databases, sharing all the calculated KPIs from the information gathered.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

8 Action plan

The interventions and actions of Antalya Demo site are described in their detailed status in D4.14. Among Antalya's actions, the specific action that has a technological aspect has to be studied within the ICT domain. All these actions are specified and tested from the interoperability and integration perspectives within this deliverable.

In regards of the integration and interoperability of the further developed Urban Platform, several actions and interventions of Antalya need to be interlinked not only within an intellectual concept, moreover, the technical developments are to ensure the data exchange (input/output modules) and Open API implementation defined in this deliverable and D4.22. The interdependencies of these components are and have to be further investigated and developed according to the action plan below.

Also interoperability among lighthouse cities (VAL, DRE and ANT) has to be studied. These will be linked with city and demo site level indicators in WP1 and WP5. Use cases and interventions will be described from the data exchange perspectives and these indicators will provide the guidelines for interoperability among cities. Specific actions under energy, mobility and ICT are expected to provide joint measures to follow and to be shared between VAL, DRE and ANT. In Antalya, we will be following up on the action plan below in the upcoming periods to complete the integration and interoperability topics.

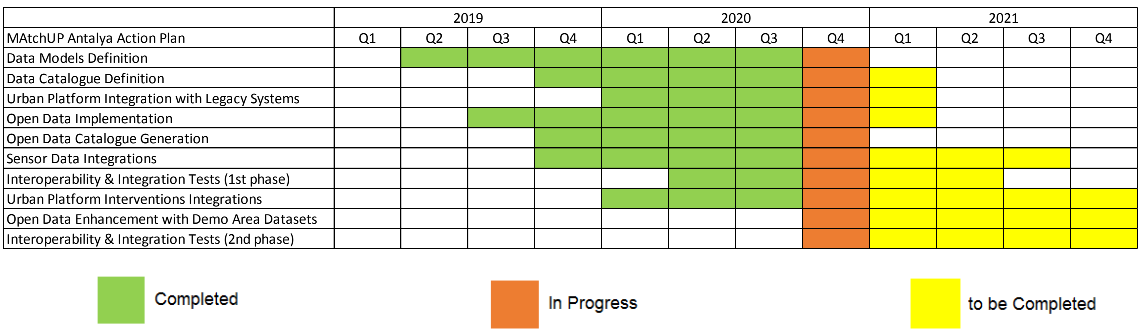

Figure 22 Antalya Interoperability & Integration Action Plans for M38-M48

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

9 Conclusions

This deliverable provides the current state concept and development process of the urban platform integration and interoperability in Antalya demo site.

First, the concept of interoperability and the regulatory framework in European Union and specifically Turkey has been defined and presented as an essential factor that facilitates communication among Antalya's urban platform modules.

Additionally, smart city modules and the requirements for Antalya's Urban Platform integration and interoperability have been studied to according to data input, output, core, integration and legacy modules. In this sense, new data models are defined for these new services. FIWARE data models are selected as the base model for Antalya's Urban Platform services implementations and specific data model dependencies are analysed for each intervention in Energy, Mobility, ICT and social pillar of the MAtchUP project in Antalya. Accordingly, the data catalogue consisting of the source, parameter and units have been defined. This deliverable also provides guidelines for the security and regulatory aspects of data giving GDPR, data security and Turkish GDPR (KVKK 6698) perspectives. In addition, it tries to suggest data protection methods through anonymization and pseudonymization.

Furthermore, first interoperability tests including unit testing, integration testing (for Kafka communication and integration gateway) and load (stress) testing have been defined in order to check the interoperability of the new services within Antalya urban platform. Integration architecture also dictates some standards on testing the compatibility using Avro type of messages and Antalya urban platform will follow the guidelines of data ingestion to ensure the data consistency and service compatibility.

Finally, the action plan is presented, with information about next steps and developments that will be carried out in order to implement the interoperability and integration tests. As seen on the action plan, the project timeline has slight shifts. Due to the unprecedented virus outbreak and pandemic, the planned work in Antalya has been affected negatively. The IoT instalments in residential units, municipality building, energy infrastructure and mobility actions are limited for quite some time and the platform integration practices in Antalya are facing delays due to the travelling restrictions and several curfews. Because of the circumstances; the developments are facing approximately 5 months of delay where the monitoring period was planned to be started at M36. Therefore, the anticipated Action plan is expected to string out to 4th year of the project 2021 through M42.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

10 References

[1] "What is Interoperability? - Definition from Techopedia." [Online]. Available: https://www.techopedia.com/definition/631/interoperability. [Accessed: 18-Sep-2018].

[2] "Kısa Dönem Eylem Planı (KDEP) (2003-2004) | e-Devlet Portali | T.C. Ulaştırma ve Altyapı Bakanlığı." [Online]. Available: http://www.edevlet.gov.tr/2015/10/13/kisa-donem-eylem-plani-kdep-2003-2004/. [Accessed: 19-Sep-2018].

[3] "DCAT application profile for data portals in Europe | Joinup." [Online]. Available: https://joinup.ec.europa.eu/solution/dcat-application-profile-data-portals-europe. [Accessed: 18-Sep-2018].

[4] "ESRI Geoportal." [Online]. Available: http://webhelp.esri.com/geoportal_extension/9.3.1/index.htm#hrvst_use_tool.htm. [Accessed: 18-Sep-2018].

[5] "GeoNetwork's cross-platform interoperability." [Online]. Available: http://trac.osgeo.org/geonetwork/wiki/interop_matrix. [Accessed: 18-Sep-2018].

[6] "GeoServer | Catalog Services for the Web (CSW) features." [Online]. Available: http://docs.geoserver.org/latest/en/user/extensions/csw/features.html. [Accessed: 18-Sep-2018].

[7] "Catalog Services for the Web (CSW) tutorial," *GeoServer 2.8.x User Manual*. [Online]. Available: http://docs.geoserver.org/latest/en/user/extensions/csw/tutorial.html. [Accessed: 18-Sep-2018].

[8] "Webkatalog." [Online]. Available: http://www.webabc.info/. [Accessed: 18-Sep-2018].

[9] "About INSPIRE | INSPIRE." [Online]. Available: https://inspire.ec.europa.eu/about-inspire/563. [Accessed: 18-Sep-2018].

[10] "Data Models - FIWARE." [Online]. Available: https://www.fiware.org/developers/data-models/. [Accessed: 17-Sep-2018].

[11] "FIWARE NGSI version 2 Release Candidate - FIWARE." [Online]. Available: https://www.fiware.org/2016/06/08/fiware-ngsi-version-2-release-candidate/. [Accessed: 18-Sep-2018].

[12] "Home - Open & Agile Smart Cities." [Online]. Available: http://oascities.org/. [Accessed: 18-Sep-2018].

[13] "IoT Big Data | Internet of Things." [Online]. Available: https://www.gsma.com/iot/iot-big-data/. [Accessed: 18-Sep-2018].

[14] "GSM Association Non-Confidential Official Document CLP.26-IoT Big Data Harmonised Data Model IoT Big Data Harmonised Data Model Copyright Notice Antitrust Notice," 2016.

[15] "DCAT: The Premier Business Development Association." [Online]. Available: https://www.dcat.org/. [Accessed: 18-Sep-2018].

[16] "Data Catalog Vocabulary (DCAT)." [Online]. Available:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477

https://www.w3.org/TR/vocab-dcat/. [Accessed: 18-Sep-2018].

[17] "Turkish Interoperability Principles." [Online]. Available: http://www.resmigazete.gov.tr/eskiler/2005/08/20050805-11.htm. [Accessed: 18-Sep-2018].

[18] "oneM2M - Home." [Online]. Available: http://www.onem2m.org/. [Accessed: 18-Sep-2018].

[19] "U.S. Department of Labor E-Government Strategic Plan: Security Privacy | U.S. Department of Labor." [Online]. Available: https://www.dol.gov/agencies/osec/e_government_plan/p23_security_privacy. [Accessed: 24-Sep-2018].

[20] "Apache Kafka." [Online]. Available: https://kafka.apache.org/. [Accessed: 24-Sep-2018].

[21] "Apache JMeter - Apache JMeter™." [Online]. Available: http://jmeter.apache.org/. [Accessed: 24-Sep-2018].

[22] "The World's Most Popular Automated API Testing Tool | SoapUI." [Online]. Available: https://www.soapui.org/. [Accessed: 24-Sep-2018].

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement N°774477