E-CORRIDOR
Edge Enabled Privacy & Security Platform For Multi Modal Transport

D7.2

# Data Analytics Techniques First Maturation

## WP7 – Data Analytics techniques

**E-CORRIDOR**

*Edge enabled Privacy and Security Platform for Multi Modal Transport*

Due date of deliverable: 31/05/2022
Actual submission date: 31/05/2022

30/05/2022

Version 1.0

*Responsible partner: UTRC*
*Editor: Stefano Sebastio*
*E-mail address: stefano.sebastio@collins.com*

| **Project co-funded by the European Union within the Horizon 2020 Framework Programme** | | |
|---|---|---|
| **Dissemination Level** | | |
| **PU** | Public | **X** |
| **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) | |

*The E-Corridor Project is supported by funding under the Horizon 2020 Framework Program of the European Union DS-2018-2019-2020, GA #883135*

**Authors:** Stefano Sebastio, Riccardo Orizio, Amine Lamine (UTRC), Hoang-Gia Nguyen (CEA), Ilaria Matteucci, Gianpiero Costantino, Marco De Vincenzi, Giacomo Iadarola (CNR), James O'Rourke, Thomas Walsh (WIT), Roland Rieke, Florian Fenzl (FhG), Koussaila Moulouel, Abdelghani Chibani (PEC), Mohammed Ammara (Clem')

**Approved by:** Roland Rieke, Florian Fenzl, Christian Plappert, **(**FhG), Patrizia Ciampoli (HPE)

**Revision History**

| Version | Date | Name | Partner | Sections Affected / Comments |
|---|---|---|---|---|
| 0.01 | 11-Feb-22 | S. Sebastio | UTRC | Initial table of content |
| 0.02 | 29-Mar-22 | R. Orizio | UTRC | Sec. on passenger localization |
| 0.03 | 04-Apr-22 | K. Moulouel | PEC | Sec. on facial recognition |
| 0.04 | 20-Apr-22 | S. Sebastio | UTRC | Introduction |
| 0.05 | 25-Apr-22 | R. Rieke, F. Fenzl | FhG | Sec. on automotive IDS |
| 0.06 | 26-Apr-22 | S. Sebastio | UTRC | Contribution to the goals of the AT pilot and the project |
| 0.07 | 26-Apr-22 | K. Moulouel | PEC | Sec. on activity recognition |
| 0.08 | 27-Apr-22 | J. O'Rourke, T. Walsh | WIT | Sec. on itinerary planning and CO2 analytics |
| 0.09 | 28-Apr-22 | I. Matteucci, M. De Vincenzi | CNR | Sec. on driver DNA and (initial contribution to) private secure routine |
| 0.10 | 03-May-22 | G. Costantino, I. Matteucci | CNR | Sec. on private secure routine and automotive IPS |
| 0.11 | 09-May-22 | S. Sebastio | UTRC | Executive summary and conclusion |
| 0.12 | 12-May-22 | A. Lamine, S. Sebastio | UTRC | Sec. on video analysis |
| 0.13 | 16-May-22 | G. Iadarola | CNR | Contribution to ISAC goals and analytics |
| 0.14 | 20-May-22 | M. Ammara | CLEM | Contribution to ISAC goals |
| 0.15 | 23-May-22 | H.-G. Nguyen | CEA | Sections on FHE-based solutions |
| 0.16 | 23-May-22 | G. Iadarola | CNR | Contribution to gait analysis |
| 0.17 | 24-May-22 | R. Rieke, F. Fenzl, C. Plappert | FhG | Internal review from FhG |
| 0.18 | 25-May-22 | P. Ciampoli | HPE | Internal review from HPE |
| 1.0 | 29-May-22 | S. Sebastio | UTRC | Comments integrated and additional fixes |

# Executive Summary

This document, along with the available software components, constitutes the second deliverable related to the WP7 activities and reports on the first maturation cycle of all the data analytics techniques available in the E-CORRIDOR framework. The discussed analytics stem from technologies previously owned by the project partners. That said, the proposed solutions have been co-designed and customized thanks to a constant exchange with experts of the multi-modal transportation represented by the three pilots, namely AT (Airport-Train), S2C (Smart cities and car sharing) and MMT-ISAC (Multi-Modal Transportation Information Sharing and Analysis Center). Requirements and constraints expressed by the pilots have shaped the initial design of the analytics and the constant exchange of ideas is continuing non-stop during the ongoing customization, integration, and maturation stages.

In the following, for each of the analytics, the proposed approach is recalled and then the current status is reported. As the analytics contribute to fulfilling the requirements of the E-CORRIDOR architecture (as reported in D5.1), tables summarize the relationship between requirements and the proposed solution. The overall goal of this deliverable is to document the ongoing efforts in the maturation and integration of the technologies.

Despite the COVID-19 pandemic has forced all the partners to reorganize their work (e.g., limiting the access to their labs for testing), the WP7 activities are on track. Currently, all the analytics run as standalone components, even if not for all of these the full set of features is available yet. Two analytics have been fully integrated and tested in the E-CORRIDOR framework. All in all, the status of the analytics along with the integration and validation efforts planned in cooperation with the pilots are on target to allow the project to reach the forthcoming milestone MS5 “First Validation of the pilots and of the E-CORRIDOR platform”.

# Table of contents

# 1. Data Analytics Techniques

The overarching scenario considered in the E-CORRIDOR project foresees the presence of users willing to access to cyber-security and advanced services related to multi-modal journeys. In such a context, users act as data *prosumers* for the E-CORRIDOR framework (i.e., they both produce-share their data and consume-analyze them through the analytics). By considering the three pilots involved in the project activities, the role of the user is covered by a multitude of actors, e.g., driver, passenger, and transportation service provider (airport, train station, car sharing and city bus).

Data analytics components define the above-mentioned services. Recalling the E-CORRIDOR framework, the analytics are part of the Information Analytics Infrastructure (IAI) subsystem and are included in the *toolbox*. Figure 1 (from D5.2) recalls the E-CORRIDOR framework and draw attention to the analytics toolbox subcomponent.

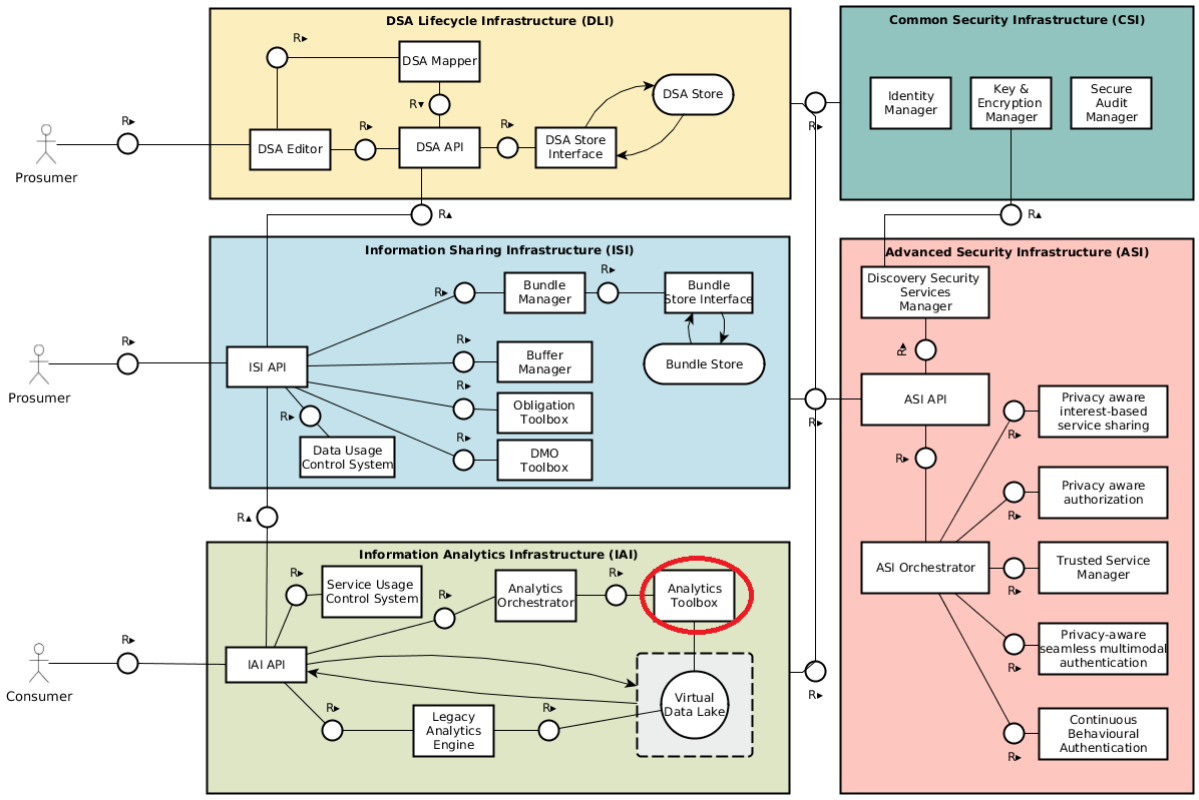


**Figure 1 The E-CORRIDOR framework - marked in red the analytics toolbox including all the data analytics discussed here**

## *1.1. Data Analytics Workflow*

Analytics in the toolbox perform classification, prediction and knowledge extraction from the data shared through the Information Sharing Infrastructure (ISI) subsystem. Assuming that data required by a given analytics were previously shared in the ISI, the workflow executes as follows:

1. The data consumer (either human or a software agent) invokes the execution of an analytics through the IAI API (Application Programming Interface).

2. The analytics orchestrator takes care of instantiating the software container and calling the analytics. In case the requested analytics is constituted by a pre-defined combination of multiple services, the orchestrator transparently processes the configuration of such a workflow (constituted by parallel and/or serial composition) and oversees its execution (please refer to D6.2 for a detailed discussion on the analytics orchestrator and its current maturation).
3. Before starting the execution of the analytics, the IAI subsystem interacts with the ISI subsystem to locate, prepare (according to the specified *Data Sharing Agreement – DSA*) and manage a virtual data lake. Such a temporary data space is used by the analytics only during its execution.
4. The analytics can then complete its execution in an asynchronous fashion. This requirement arises considering the heterogeneity of the analytics included in the toolbox.
5. Output data are stored again in the ISI according to the DSA in place and returned to the data consumer.

## *1.2. Analytics Toolbox Goals*

As already discussed in D7.1, activities performed in WP7 aim at contributing to the following project objectives (Obj.): (Obj. 2) define edge-enabled data analytics and prediction services in a collaborative, distributed and confidential way; (Obj. 3) define a secure and robust platform safe from cyber-attacks and able to ensure service continuity; (Obj. 4) improve, mature and integrate existing tools provided by the partners and tailored to the needs of platform and pilots; (Obj. 5) provide mechanisms for a seamless access to multimodal transport; (Obj. 6) deliver pilot products.

A few highlights about the contribution of the WP7 activities to the fulfillment of the project objectives are the following:

- once integrated in the toolbox, the analytics can exploit the information sharing subsystem to work on data shared in a collaborative fashion under the policy specified by the data producer (Obj. 2). According to the expressed pilots' requirements and constraints, some analytics runs on lightweight devices (like a Raspberry Pi or the car infotainment system e.g., the Driver DNA [E-CORRIDOR-IAI-SR]), on mobiles (like smartphones e.g., the Passenger location and flow optimization [E-CORRIDOR-IAI-PL]), on cloud infrastructure (less limited in terms of available resources e.g., the Passenger: Identification, Behavior, Context [E-CORRIDOR-IAI-PBI]) or in a hybrid edge-cloud fashion
- analytics presented in the following provide, to the transportation entity providers that will adopt the E-CORRIDOR framework, cyber-security solutions (Obj. 3) through a set of intrusion detection technologies – See Section 6
- and (Obj. 5) authentication mechanisms for passenger and driver – See Section 2
- all the designed analytics are the results of constant interactions between technology solution providers and pilots and, at the time of writing this deliverable, two analytics (namely, Driver DNA [E-CORRIDOR-IAI-SR] and Passenger: Identification,

Behavior, Context [E-CORRIDOR-IAI-PBI]) have been fully integrated and two more (Face recognition [E-CORRIDOR-IAI-FR] and Automotive Intrusion Detection [E-CORRIDOR-IAI-CANIDS]) are under integration in the E-CORRIDOR framework (Obj. 4)

- the adopted design approach (mentioned in the previous bullet point) aims at accelerating the maturation of the technologies and at easing a sub-sequent inclusion in products to be tested in the pilot environments (Obj. 6).

The data analytics solutions are meant to be integrated in the confidential and privacy-preserving framework for data sharing built as part of the WP5 and WP6 activities to fulfill Obj. 1. At the time of writing this deliverable a first version of such framework is available and implemented in the pilots (MS3 “Setup of the running Pilots” and MS4 “First version of the integrated platform and pilots”).

A first set of analytics have been containerized, integrated, tested and published on the E-CORRIDOR’s Nexus repository (see Figure 2).

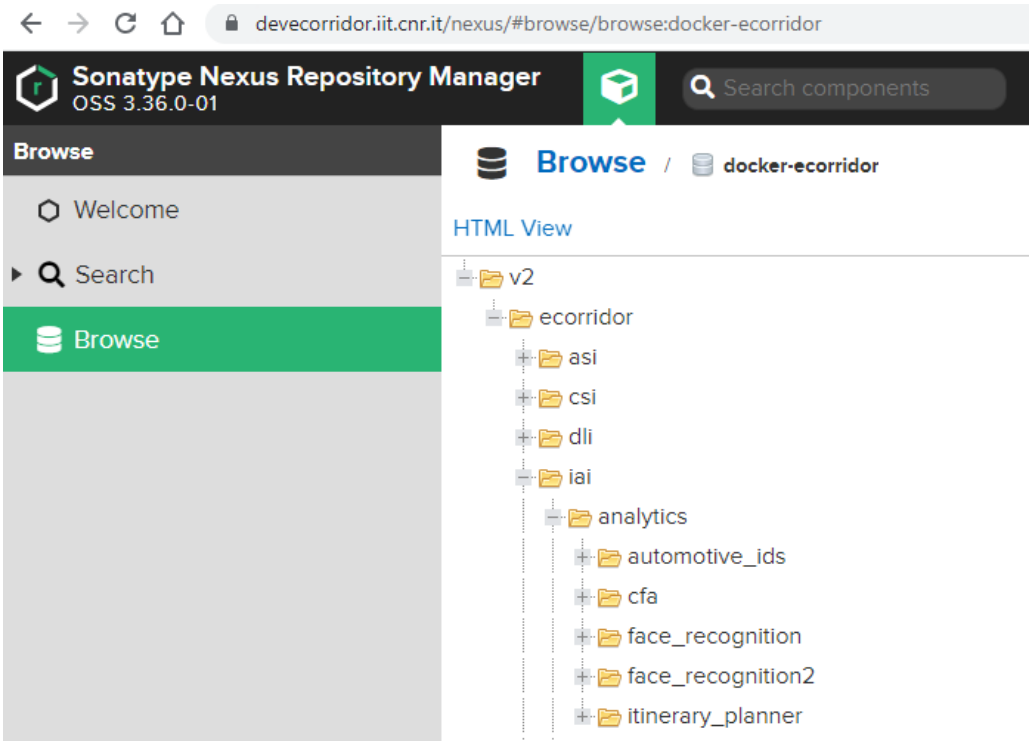


**Figure 2 The container repository for the analytics in the IAI toolbox**

## *1.3. Structure of the Deliverable*

The remaining of this deliverable is structured as follows. Sections from 2 to 6 describe the data analytics grouped by purpose, namely (Sec 2) passenger and driver identification, (Sec 3) privacy preserving itinerary planning, (Sec 4) privacy preserving security analytics, (Sec 5) carbon footprint analytics and (Sec 6) intrusion detection technologies. Figure 3 represents this logical organization of the analytics in the toolbox corresponding to the different tasks part of WP7 activities. For each analytics component its main goals, features and operating principles

are reported along with its workflow considering prosumers and pilot scenarios. Then, the current maturation status is discussed with a focus on the completed activities, the integration in the E-CORRIDOR framework and the fulfilled requirements. Each sub-section concludes with a brief plan for testing and final maturation. The progress on the pilot specific analytics is reported in Section 7. To clarify the contribution of the analytics in the pilot scenarios, the match between pilot use cases and analytics is schematically reported in Section 8. The overall contribution of the WP7 activities to the E-CORRIDOR objectives is summarized in Section 9, whereas a conclusion on the first maturation of the analytics is presented in Section 10. Bibliographic references are reported in Section 11. The Appendix contains a list of all the acronyms used in this document and a list of the included figures.

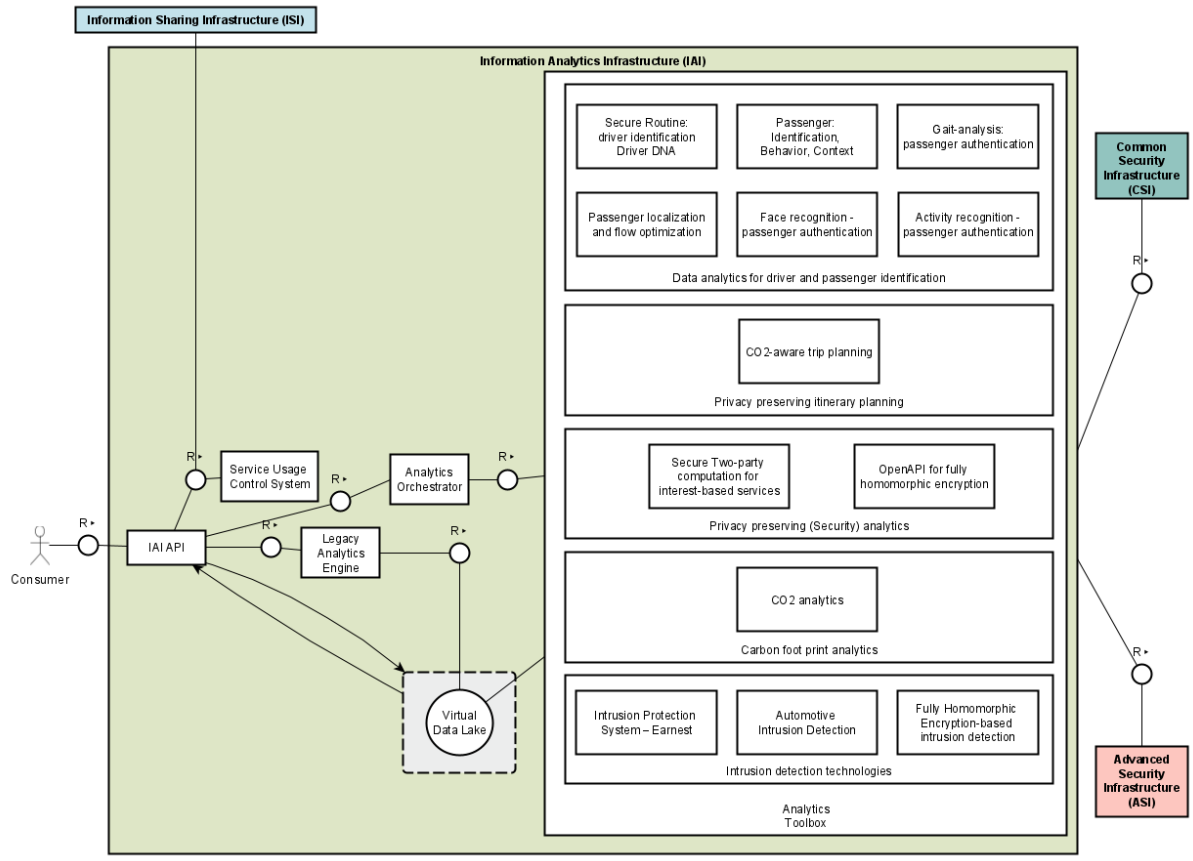


**Figure 3 A pictorial representation of the logical grouping of the data analytics available in the toolbox**

***Note to the reader***: This deliverable aims at reporting on the progress of the data analytics - previously identified in D7.1 "Data Analytics Techniques Requirements and Architecture" - with respect to design (by considering the pilot requirements), implementation, integration in the E-CORRIDOR framework and maturation. To keep the document self-contained the main goals and operating principles of each analytics are summarized in the following.

It is assumed that the reader is already familiar with the main concepts of the E-CORRIDOR framework discussed in D5.2 "First version of the E-CORRIDOR architecture" and of the sharing and analytics infrastructures discussed in D6.2 "Sharing and Analytics Infrastructures first maturation". Moreover, a general understanding of scenarios and requirements expressed by the three pilots (in D2.1, D2.2, D3.1, D3.2, D4.1 and D4.2) may be needed to firm grasp some design choices of the data analytics.

# 2. Data Analytics for Driver and Passenger Identification – Task 7.1

Project pilots (S2C and AT pilots in particular) have expressed an extensive interest in technologies for identifying and later authenticating users exploiting sensors already adopted by them. The analytics developed in this task make use of machine learning and artificial intelligence techniques to process data from a multitude of sensors (such as cameras, accelerometers and gyroscopes, Bluetooth beacons, and OBD-II diagnostic data) to build models supporting the identification of driver (in the S2C pilot) and passenger (in the AT pilot). The performed identification is used for authentication purposes or for offering more advanced and customized services.

## *2.1. Secure routine for driver identification - Driver DNA [E-CORRIDOR-IAI-SR]*

### 2.1.1. Component Description

Nowadays, road vehicles can be considered four-wheel smart devices connected to the Internet and the road infrastructure. This scenario allows carmakers and private providers to offer customized services, based on vehicle and user data, to influence the drivers, for example, to a more eco-friendly or safer driving style. These opportunities motivated us to develop an Android app that can retrieve data from the vehicle to provide customized services and influence the driving style.

The app is based on the concept of Driver DNA, which is a metric to evaluate some users' attitudes during the driving. As defined in [1] the metric is composed of four main parameters retrieved from the vehicle: vehicle speed, engine revolutions per minute (RPM), breaking, and turning. In our scenario, the Android app retrieves some parameters to compute the Driver DNA and should compare the result with the other drivers' values on the same predefined road segment. This operation enables the infrastructure or a provider to offer a customized service and incentive specific driving styles. The app is driver-friendly with an intuitive interface, and, during the driving, it enables users to easily ask for customized services like a price discount or a maintenance service.

### 2.1.2. Workflow in Action

The Driver DNA app structure can be divided into two main parts: the front-end, with which users interact, and the back end, invisible to users. Following, the back end could be divided into two more areas: the in-vehicle and the out-of-the-vehicle.

**Front-end:** The Driver DNA app is designed for the Google's Android operating system, and it can be installed on every Android device with at least Android 4.0.1 Level 14 version *Ice Cream Sandwich* (2011). In particular, the app is designed to be used on the infotainment system of vehicles. For this reason, it is designed to work horizontally in *Landscape* mode and with a minimalist and clean layout to allow drivers to select options also during the driving without compromising their safety. For these reasons, our app could be considered driver-friendly, but the app can be installed also on every Android smartphone satisfying the minimum required Android version.


**Figure 4 Driver DNA app's home page**

Figure 4 shows the app's home page which allows users to select the Driver DNA services. During the driving, the user must start the app to collect the vehicle's data and the app displays some useful information as shown in Figure 5. In particular, in the middle of the page, the actual speed, the RPM, and the engine starting time are displayed. At the bottom-left, it is displayed the protocol used to communicate with the vehicle, as we explain in the Back-end description.

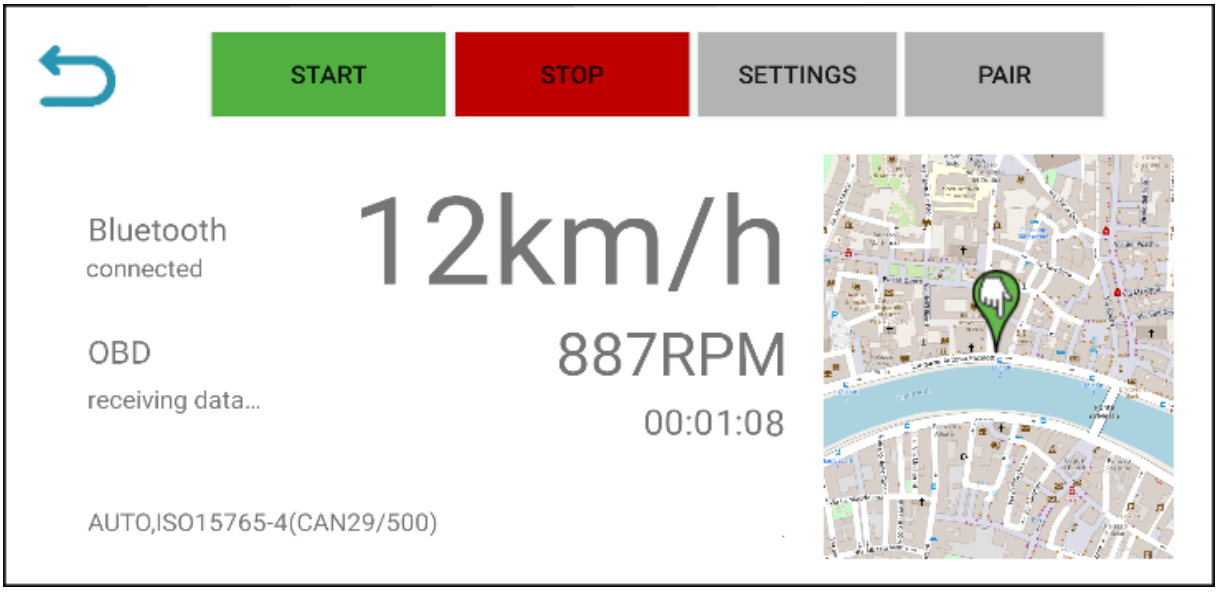


**Figure 5 App main screen which displays real-time data and an OpenStreetMap with a pin on the current position.**

If the user needs to ask for a specific service. It could go to the service request page and select the needed service. In particular, it can choose from different services like customized prices for a recharge, in a restaurant, or service maintenance as shown in Figure 6.

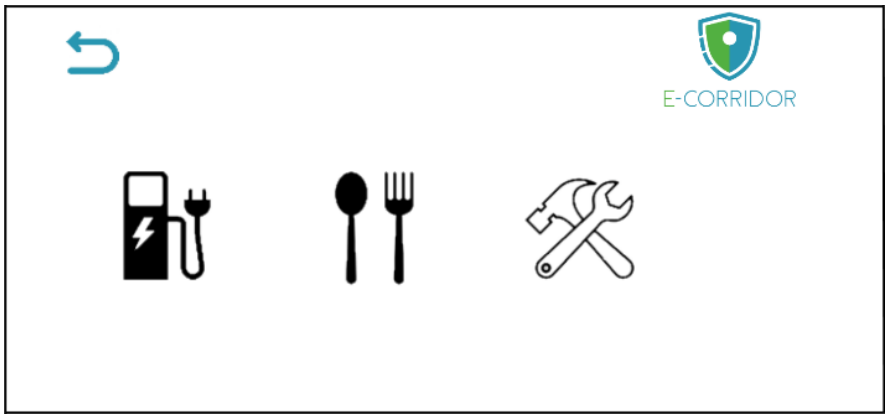


**Figure 6 Service request page.**

## Back-end:

*In-vehicle*: The Driver DNA should be installed on the user's device (infotainment system or smartphone), and it can retrieve the vehicle's data through a Bluetooth connection using an OBD-II port ELM327 dongle. The Bluetooth dongle is a small device connected to the OBD-II port, which is usually located under the dashboard, beneath the steering wheel column. Since 1996 the port is mandatory in the USA for all new vehicles and in the following years, almost all countries of the world transposed the rule. The OBD-II is used to access the vehicle's data for various purposes like emission tests and diagnostics. From this port, it is possible to retrieve different data by sending OBD-II parameter IDs (PID), standardized by SAE J1979. To summarize, the in-vehicle infrastructure is composed of the dongle that sends data via Bluetooth to the user's device, which then resends the received data to the external infrastructure using a broadband network. The dongle can collect different vehicle data like the ambient air temperature, the vehicle speed, or the throttle position. For our analytics, we retrieve some data like the vehicle speed and the engine RPM in addition to the latitude and longitude that can be obtained from the user's device. All these data are collected with a defined frequency (for example every 4 seconds) and saved in a JSON file. When the file reaches a defined dimension (less than 1 Mega Byte), it is sent out of the vehicle, the buffer of the vehicle is cleaned, and the vehicle starts to write data in a new JSON file.

*Out-of-the-Vehicle*: Figure 7 shows the external infrastructure with the two main phases of the service request. The first is the *Storing Process*, where data are sent out of the vehicle and saved in the Information Sharing Infrastructure (ISI). As described in the previous section, the vehicle's data are saved in a JSON file and sent to the ISI with a POST call to be saved. The user's device receives back the DPO Id, which is a unique identifier of the uploaded file.

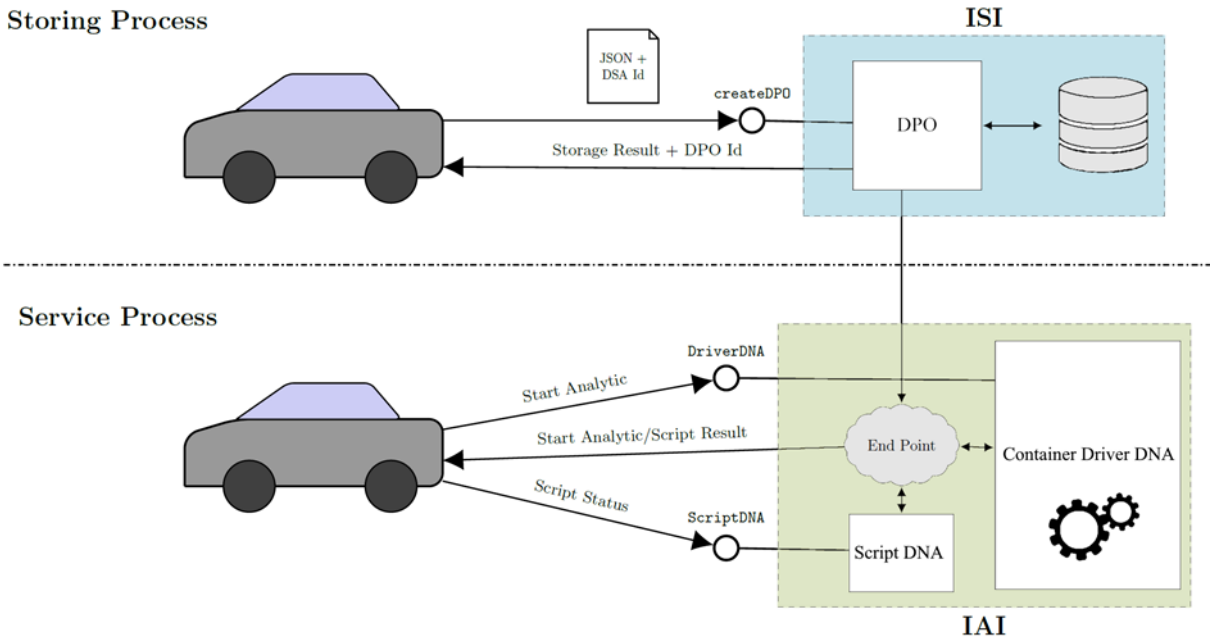


**Figure 7 Out-of-the-vehicle infrastructure**

The second phase is the *Service Process*, where the user asks for a service and the app requests the start of an analytic. In particular, as shown in Figure 7, the IAI exposes a service called *DriverDNA*, so the app can make a GET call to require a service. Immediately, the IAI, which is the analytic core of the system, starts the Driver DNA computation process, asking the End-Point to retrieve the users' stored data in the ISI. Besides, the End-Point sends a string to inform the app about the start of the analytics. During this phase, the app asks periodically with a polling operation the status of the analytics. When the analytics is finished, the app receives the result and sends it to the Service Provider to attain a customized service based on the Driver DNA value.

### 2.1.3. Integration and Maturation Status of the First Release

The previous description reports the actual status of the app and of the analytics integrated into the E–CORRIDOR framework. The app can collect data from the vehicle, store data in the ISI and request the start of the analytics from the IAI, currently using only the RPM values. The next steps to complete the app are:

1. Defining more complex and complete analytics using all four Driver DNA parameters
2. The definition of the call to the external service provider (e.g., recharge station, restaurants, or maintenance services) is not yet implemented. Two options will be evaluated: call the provider directly from the IAI or let the user requests a service to the external provider from the app.
3. Following the previous point, we have to define the database (DB) containing the drivers' data to be compared. The DB could be centralized, for example, in a unique server, or distributed, for example one in each area or road segment.
4. The vehicle's data in the JSON file are in a clear format and, consequently, stored in the ISI in cleartext. In agreement with pilot's requirements, as next step, we will implement a private version of the Driver DNA analytics, where data are encrypted inside the

vehicle and the analytics are computed over encrypted data. Thus, we collapse in this activity the concepts we described in D7.1 Section 4.2 about the usage of the secure multiparty computation algorithm to identify a driver.

### 2.1.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | **Status** | **Rationale for the status** |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-SR-01 (E-CORRIDOR-DM-01, E-CORRIDOR-Sec-RC-01** | MUST | Data used to establish the driver DNA may require to be obfuscated or anonymized, e.g., data coming from vehicles that may provide information on drivers | In progress | Actually, data saved in a JSON file are not encrypted. Studying the possibility to use encryption methods |
| **E-CORRIDOR-IAI-SR-02 (E-CORRIDOR Ope-02)** | MUST | Driver DNA analytics can be run at the edge. | Partially | Currently, Driver DNA analytics runs on the E-CORRIDOR server, but the IAI can be also located in an edge node to perform the analytics |
| **E-CORRIDOR-IAI-SR-03 (E-CORRIDOR-Tst-S2C-01, E-CORRIDOR-Tst-S2C-02)** | SHOULD | The In-Vehicle Infotainment (IVI) or Electronic Control Units (ECUs) may be used for collecting GPS and driving behavior data. | Completed | GPS data are collected using the infotainment system, connected to the web. Vehicle data are collected from the OBD II port which receives the data from the ECUs. |

### 2.1.5. Plan for Testing and Final Maturation

After the integration, the next steps are the testing with the application of the encryption methods to evaluate the computational time and costs. Another test is to use the app directly on a vehicle during the driving to simulate in a real scenario the collection of data and the request for a customized service.

## *2.2. Passenger location and flow optimization [E-CORRIDOR-IAI-PL]*

### 2.2.1. Component Description

The passenger location and flow optimization component aims at helping the travelers reach relevant indoor locations during their (waiting) time in the train station or airport. For example, at the airport, the passenger will receive information about the directions to reach all her Points of Interest (PoIs), from the baggage drop point to the boarding gate. Upon request of the passenger, information related to the journey may be retrieved in a privacy-aware manner from the E-CORRIDOR framework, through the ISI component subsystem.

Considering that the component will have to guide passengers in an indoor environment, the Bluetooth technology has been chosen over the GPS. Indeed, GPS signals are not powerful enough to penetrate buildings. A more detailed explanation of this design choice has been provided in D7.1 [2]. The passenger will be using an Android device running the E-CORRIDOR app with Bluetooth capabilities. Multiple IoT (Internet of Things) or Bluetooth beacons are deployed in the airport and train station, in particular closer to the PoIs, and are able to interact with the passengers' devices. The devices deployed in the premises of the transportation entity are referred to as beacons from now on.

The communication between the Bluetooth IoT devices, specifically the Received Signal Strength Indicator (RSSI) values, will allow the component to estimate the position of the passengers. The position information will be locally computed and not shared to provide the guidance service while preserving the privacy of the passenger, unless the passenger explicitly opts to share such information. A special case is constituted by the Passengers with Reduced Mobility (PRM) requiring the use of assistive devices, such as a wheelchair. In this case the IoT device is attached to the wheelchair property of the airport/train station to optimize localization and planning.

In the event the location information is shared in the E-CORRIDOR framework, it may be used to further strengthen the authentication capabilities offered by other analytics and security services offered by the framework. Furthermore, various businesses in the airport and train station can have their own beacons. Therefore, targeted advertisements could be provided to the passengers that will enable such a service in their preferences.

### 2.2.2. Workflow in Action

We assume that each passenger's device and each beacon has the ISI subsystem running on it. Once the passenger reaches the indoor environment, her device will start receiving messages from the nearby beacons, allowing the (approximate) localization to begin. The more beacons are located within a given area, the more accurate the localization will be. At the same time, information about the passenger and her journey (e.g., flight or train number) is gathered from the ISI installed on the personal device. This will enable the component to create the indoor path and guide the passenger to the PoI through privacy-aware ad-hoc messages sent by the beacons.

Figure 8 shows a complete scenario where all the possible interactions between devices and beacons are represented. First, all passengers will receive messages from the beacons advertising their locations. The more beacons are available nearby the passenger, the more messages will be received and with higher accuracy the location can be computed from the app running on the passenger's smartphone. According to the airport/train station policy, PRM passengers using a wheelchair may automatically have their position shared with the E-CORRIDOR framework to ease the supply of the requested special assistance service (e.g., to

help the localization of the subcontractors providing assistance to different connected means of transportation like airport and train). This choice is depicted in the figure with a circle around each passenger. If this is the case, the passengers will communicate, back to the ISI subsystem of each beacon, their locally computed approximated position.

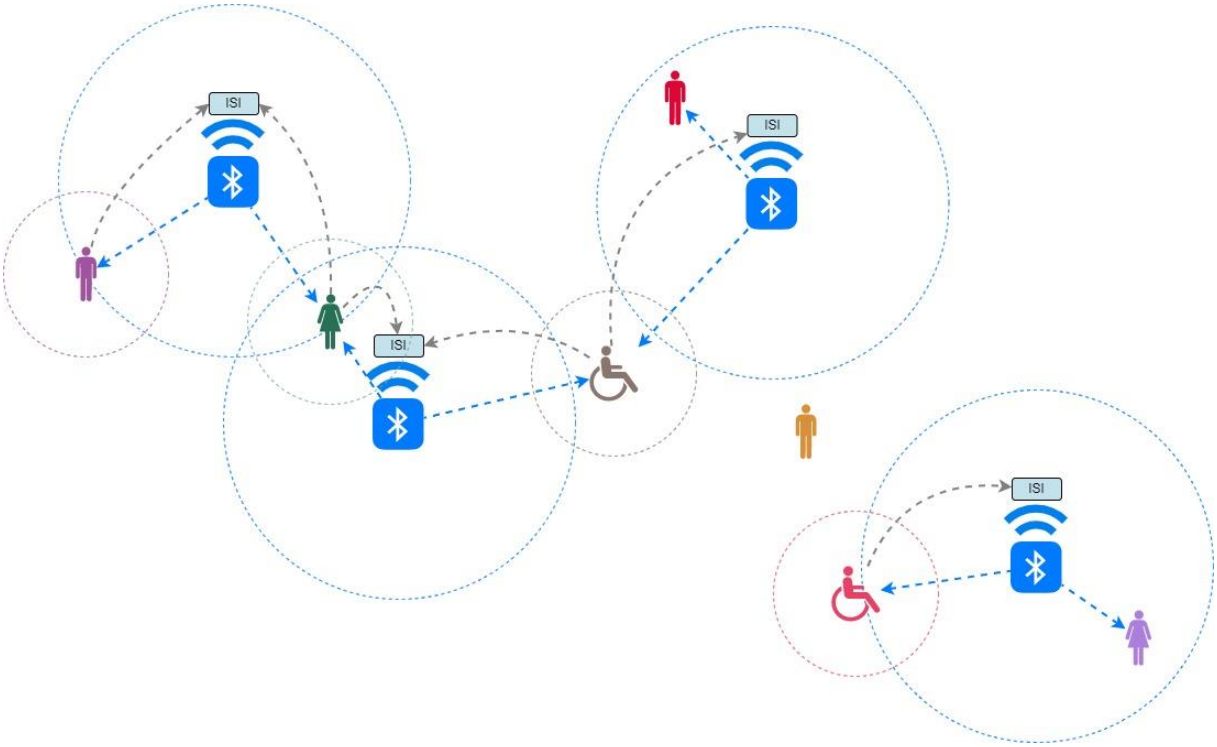


**Figure 8: Interactions between passengers and beacons running the localization app**

It is possible that the passenger could enter in a zone not covered by any beacon. In such a case, the earlier guidance messages should ensure to provide enough information to guide the passenger out of the signal blackspot and to her next PoI where the coverage, and therefore all the localization services, recommence working normally. All the information collected from the passengers will then be used to estimate the passengers' flow and to enhance the airport services by redistributing operators in the highest crowded areas.

### 2.2.3. Integration and Maturation Status of the First Release

We are currently able to generate personalized beacon messages that will help the passengers in identifying their position in the airport terminal or train station. We focused on generating beacons that will provide relevant information about the beacon itself. We decided to deploy two different technologies for the beacons, iBeacon and Eddystone, which both work well and are easily interpretable by most of the IoT devices.

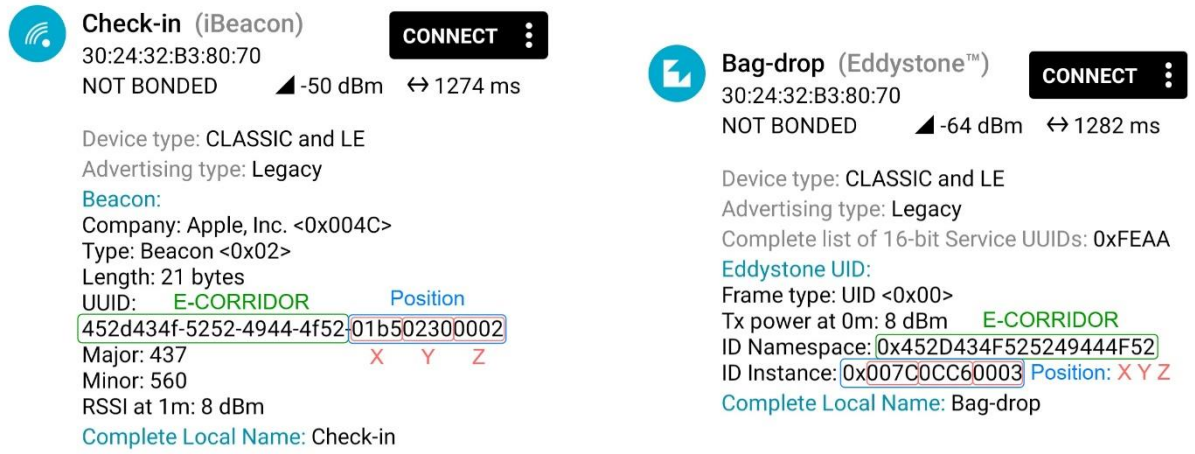


**Figure 9: Beacon messages and annotated specification (iBeacon on the left and Eddystone on the right)**

Figure 9 shows the messages related to two beacons that can represent different points of interest in an airport. Both beacons broadcast information about their position respecting the specification of the two technologies that we implemented. On the left part of the figure, we created a beacon with the iBeacon protocol related to a "Check-in" station, identified by the name of the project paired with its position (with x, y, and z coordinates). Instead, on the right-hand side of the figure, we created an Eddystone beacon related to a "Bag-drop" location identified in the same way.

Moreover, we are working on the passenger flow estimation capabilities of the component, trying to roughly estimate the number of unique devices around each area.

Among the features considered in the design phase of this analytics some have been already implemented in the version released with D7.2. More features will be added for the final maturation of the component planned for the next year. The following table summarizes features and efforts expected for this component and mark if these are already available or will be ready for D7.3.

| Feature | Current version | Final release |
|---|---|---|
| Beacon centered messages and related localization | ✓ | |
| Preliminary lab tests | ✓ | |
| Path and related messages generation for indoor guidance service | ✓ | |
| User centered messages and related ISI exchanges | | ✓ |
| Extended lab tests | | ✓ |
| On-site tests | | ✓ |
| Integration and tests with E-CORRIDOR framework | | ✓ |
| Integration with business beacons and passengers' preferences | | ✓ |
| Passengers' flow estimation | | ✓ |

### 2.2.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | **Status** | Rationale for the status |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-PL-01 (E-CORRIDOR-DS-06, E-CORRIDOR-DS-07, E-CORRIDOR-DS-16, E-CORRIDOR-DS-17, E-CORRIDOR-DS-24)** | MUST | Accuracy and effectiveness of passengers' localization and specific guidance messages depend on the DSA defined and options selected by each passenger. | To be deployed on the final release | To preserve the privacy, passenger information is stored on the ISI only if they opt to activate the service. |
| **E-CORRIDOR-IAI-PL-02 (E-CORRIDOR-DS-27, E-CORRIDOR-DA-05, E-CORRIDOR-DA-07)** | MUST | The passengers' flow estimation will rely on the information gathered by all the passengers that allowed their data to be analyzed. | To be deployed on the final release | If the passenger allows to share her position, this information will be shared with the ISI and then become accessible for further analysis. |
| **E-CORRIDOR-IAI-PL-03 (E-CORRIDOR-DA-10, E-CORRIDOR-Ope-05)** | SHOULD | The guidance messages directed to specific passengers will be generated considering the privacy of the passenger. | Completed | The guidance messages will be encrypted and accessible only to the passenger. |
| **E-CORRIDOR-IAI-PL-04 (E-CORRIDOR-Use-02)** | SHOULD | The estimated position of each passenger should be used to enhance her authentication. | Completed | If the passenger allows to share her position, this information will be stored in the ISI and then become accessible to the authentication mechanisms. |
| **E-CORRIDOR-IAI-PL-05 (E-CORRIDOR-DS-10, E-CORRIDOR-DA-11, E-CORRIDOR-Sec-IS-02)** | MUST | The inferred passenger location is transmitted and stored (only for the needed time) in a privacy-aware and secure manner. | To be deployed on the final release | The messages containing the position of each passenger (if shared) will be encrypted and an appropriate data retention policy is applied. |

| | | | | |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-PL-06 (E-CORRIDOR-Ope-01, E-CORRIDOR-Ope-02)** | SHOULD | The analytics to locate the passenger can run either on the personal device (i.e., at the edge) or on the cloud. | To be deployed on the final release | The current version will focus on the passengers' device computation. The same procedure can be performed on the cloud in an aggregated form after pseudo anonymization. |

### 2.2.5. Plan for Testing and Final Maturation

Referring to Figure 8, we envision to test the scenarios with Android and Raspberry Pi devices for passengers and beacons respectively. We will test the scenarios in a lab setting, varying the position of the beacons as well as the path walked by the passengers. We envision to have various beacons and an increasing number of passengers moving around the environment, to test various possibilities and potential complexity derived from the environment architecture and the location of the beacons. Potentially, the scenario will be recreated in the airport internal laboratory for more precise and ad-hoc tests.

Finally, a further extension will include potential businesses at the airport running the beacons. According to user preferences, tailored messages and enhanced guidance service could be provided, still retaining the focus on the privacy of all the exchanged messages. At first, this will be tested in an internal lab and then potentially be demonstrated on the test-site at the airport emulating the presence of different businesses.

## *2.3. Passenger: Identification, Behavior, Context [E-CORRIDOR-IAI-PBI]*

### 2.3.1. Component Description

Before the halt due to the COVID-19 pandemic, large airports were used to receive tens of millions of passengers per year with a growing trend. As restrictions have been lifted in many countries, the ongoing recover of the travel sector is expected to boom and projected to return to pre-pandemic levels already in 2023 (and grow even further in the next 10 years). To perform performance analysis and identify security and safety risks, transportation service providers install video cameras throughout the airport and train station.

In such a scenario, our analytics aims at performing a continuous and automatic monitoring of the transportation facilities by means of video analysis. The output measures can be used to perfect the provided services (e.g., by identifying a bottleneck in the passenger flow), analyze and mitigate issues potentially affecting the passenger experience (e.g., by summoning the operators to open an additional information desk if the queue is increasing) and extract knowledge in the way passengers approach to travel and access to the service.

Achieving the above-mentioned goals requires the analysis of the passenger behavior along with an understanding of the context in which they move. The challenge becomes more difficult as the areas are generally crowded causing frequent occlusions and interactions between passengers and groups (e.g., particularly in case of a family or people travelling together).

### 2.3.2. Workflow in Action

The developed analytics is constituted by a combination of computer vision and image processing techniques. The analytics is based on deep learning systems and combines YOLO v4 [3] and the DeepSORT [4] model for object recognition and tracking. The designed pipeline foresees object detection, classification, localization, tracking and knowledge extraction. The goals for the above steps are the following:

- *Object detection*: identification of the relevant subjects in the scene. In the considered scenario constituted by the airport terminal and train station, we assume that baggage, passengers and possibly their assistive devices (like a wheelchair) are relevant for our analysis. As the environmental camera covers a quite large visual field, our analytics has to deal with two main issues: complexity of the scenario (constituted by many people) and distorted images.
- *Object classification*: for each of the identified subject, it is important to extract features useful for their classification and therefore understand the kind of subjects composing the scene. For instance, different kind of baggage can be identified (such as back-pack, trolley, etc.). This processing is useful to later perform the knowledge extraction.
- *Object localization*: this step deals with the position of the object in the scene. It is worth to highlight that the usual cameras adopted for monitoring are installed in the cornice area of the rooms and generally don't have any depth sensor. Considering the distorted view, determining the right distance is challenging. Our analytics transforms an arbitrary perspective into a bird's eye view.
- *Object tracking*: in this step the analytics keep using the same object identifier while the subject moves in the environment. The test setting turned out to be particularly challenging due to the presence of numerous people that often overlaps from the camera's viewpoint and that may move at a brisk pace. We would like to remark that our analytics is not meant for surveillance. It is therefore intended to operate in a privacy-preserving fashion and can then be used after having performed facial redaction operations through the corresponding DMO (please see D6.2 and Sec 4.3.1 therein for a detailed description of the component).
- *Knowledge extraction*: this is the final step of the analytics. Metrics useful to expand the knowledge about the use of the transportation services are collected and presented in a systematic way. Further automatic or manual processing by the service providers can be performed, and the appropriate mitigation and enhancement can be put in place.

The expected workflow is reported in Figure 10. In details the steps are as follows:

1. Site managers at the airport and train station define DSA to regulate the access to and the use of the video recordings, including an appropriate data retention policy. For instance, read restrictions can be imposed according to location (e.g., only from the control room), specific group of employees (e.g., only security officers) or system state (e.g., if there is an ongoing emergency).
2. Environmental cameras are connected via IP (or else) to edge devices. These devices are resource-constrained and run only the needed subsystems of the framework (i.e., only the ISI). Videos are preprocessed and shared following the DSA specified in the first step. To preserve the passenger's privacy, the DSA may require the execution of

the face anonymization DMO before allowing the transmission of the data to the master node in the terminal.

3. The analytics is executed, and the output metrics are shared according to the DSA.

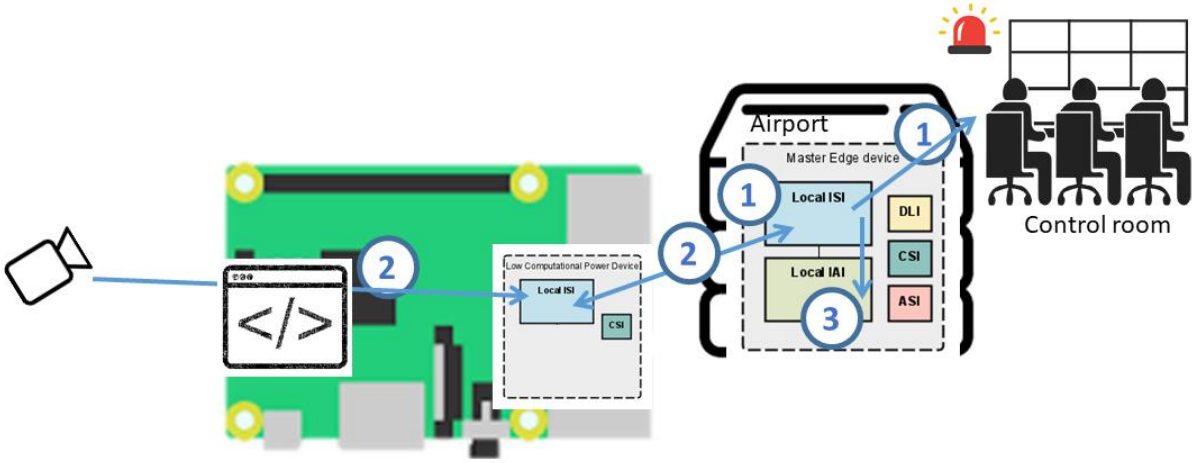


**Figure 10 The workflow of the camera feed analytics as intended to be used in the AT pilot**

Through a simple GUI, operators in the control room may instruct the analytics about their area of interest. We are currently providing in output statistics about number of passengers in an area of interest (e.g., a kiosk) and the respect of the COVID-19 safe distance.

### 2.3.3. Integration and Maturation Status of the First Release

The above-described workflow constitutes the final goal for validating the capability of the analytics. At the time of writing this deliverable our component has almost all of its features functioning, and it has already been fully integrated and tested in the E-CORRIDOR framework.

A screenshot extracted from the running analysis is represented in Figure 11. In the example, domain experts have identified three different areas. The analytics can identify the distance among passengers (e.g., to recall to the passengers to respect the COVID-19 safe distance rule through a loudhailer) and count the average number of people in each area (e.g., near the ticket vending machine on the left-hand side of the image).


**Figure 11 The camera feed analytics measuring the performance in three areas**

In the following table features and efforts planned for the components are summarized, remarking whether these have been already delivered or are expected for the next year.

| Feature | Current version | Final release |
|---|---|---|
| Runtime detection and identification | ✓ | |
| Classification of relevant subjects (passengers and baggage) | ✓ | |
| Localization and distance estimation | ✓ | |
| Metrics in output | ✓ | |
| Integration and tests with E-CORRIDOR framework | ✓ | |
| Extended metrics and output in standard JSON format | | ✓ |
| Optimization of the performance of the whole pipeline | | ✓ |
| Extended test with more videos | | ✓ |
| Connection with lab/pilot cameras | | ✓ |

### 2.3.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | **Status** | **Rationale for the status)** |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-PBI-001 (E-CORRIDOR-Use-02)** | SHOULD | The collected features of every passenger should be used to enhance the authentication process. | In progress | Some of the metrics provided in output can be used as attributes by the context analysis (please see Sec 2 in |

| | | | | |
|---|---|---|---|---|
| | | | | D8.2). E.g., information on the environment, and the presence of baggage or wheelchair. |
| **E-CORRIDOR-IAI-PBI-002 (E-CORRIDOR-DS-10)** | MUST | The cloud service, after finishing the travel journey must delete the subject information. | Completed | We successfully designed and tested a DSA specifying a data retention policy for the video |
| **E-CORRIDOR-IAI-PBI-003 (E-CORRIDOR-Sec-IS-02)** | MUST | The collected passenger features are transmitted and stored in a privacy-aware and secure manner. | Completed | This component can be used along with the face anonymization DMO to ensure the passengers' privacy. DSA regulates who can access the video |

### 2.3.5. Plan for Testing and Final Maturation

As described above the analytics is fully functional and integrated in the E-CORRIDOR framework. Currently, tests have been performed with a limited set of pre-recorded and publicly available videos. For the next maturation cycle we plan to extend this set of experiments and possibly connect with cameras installed in a laboratory environment provided by the AT pilot.

Additional efforts will be devoted to the performance optimization and to the extension of the metrics extracted by the analysis following the discussion with the pilot. Finally, we will present the extracted knowledge in a consistent format (e.g., JSON) to allow further processing by other analytics internal or external to the E-CORRIDOR framework.

## *2.4. Gait analysis – passenger authentication [E-CORRIDOR-IAI-GA]*

### 2.4.1. Component Description

Biometric authentication solutions that can leverage personal devices are always demanded as can provide options to passengers that would like to avoid the touch of public surfaces. This analytics presents a solution for the continuous authentication of the users based on a gait analysis performed through the sensors deployed in the mobile devices. It exploits inertial sensors and Recurrent Neural Network (RNN) models for a deep learning-based classification. The component handles all the stages needed for a continuous authentication: starting from data collection to data preprocessing, classification, and policy enforcement. The designed tool along with some experimental results have been also presented in [5].

### 2.4.2. Workflow in Action

Every day, smartphones, tablets, smartwatches, and other wearable devices collect, produce and store a plethora of users' sensitive data, and at the same time, they give access to operations that are critical for users' privacy, and device integrity [5]. Continuous authentication is considered an enabling technology of utmost importance, allowing verifying a user's identity continuously and in real-time, making device access seamless to authorized users, while still ensuring protection from unauthorized access attempts.

However, continuous user authentication is a challenging task that requires continuous monitoring of the user's unique parameters, such as behavior and physical features. To be performed in a seamless way, this control must be non-intrusive, i.e., requiring minimal to no active interaction. One promising direction is the usage of gait analysis, i.e., monitoring through inertial sensors the features of the user's walking style and feeding them to classifiers trained to authenticate specific users. Our tool is based on the usage of deep learning classifiers, specifically Recurrent Neural Networks, to analyze different gait types of smartphone users aiming at uniquely authenticating a specific user by classifying her walking collected through smartphone inertial sensors. In detail, the tool is formed by different modules that handle the collection (sensing), pre-processing, gait type (action) recognition, classification, and on-device enforcement. The enforcement is based on the Usage Control paradigm, particularly fitting for continuous enforcement of policies.

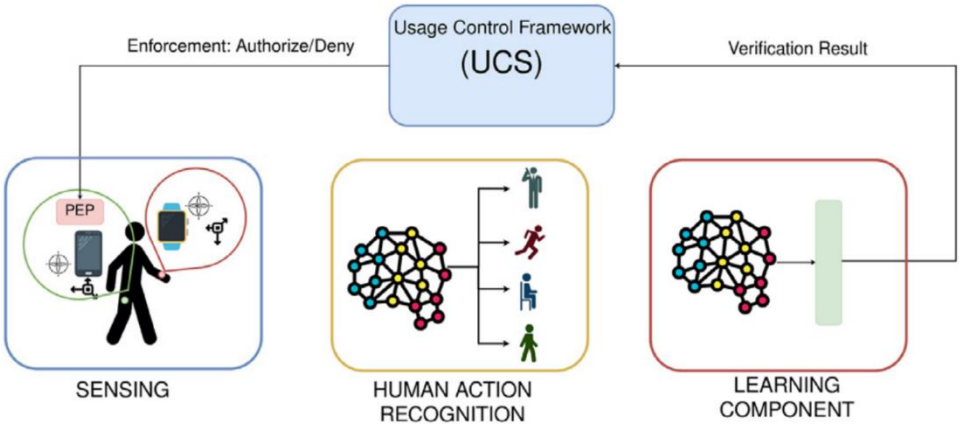


**Figure 12 The pipeline of the gait analysis component**

The UCS model (Usage Control System) and the User Verification component handle the authorizations based on the input from the Human activity recognition component (HAR). The latter is a deep learning model based on Recurrent Neural Network (RNN) that analyzes the inertial signal provided by the preprocessing component and infers the specific movement or action that a person is performing. The verification process and the attributes in the policy enforce the authorization.

### 2.4.3. Integration and Maturation Status of the First Release

The tool has been developed and tested locally. Experiments were performed to test the accuracy of the methodology, and we are working to integrate the tool computations into the cloud and the E-CORRIDOR framework. The component is available for testing and research purposes on GitHub [6].

### 2.4.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | **Status** | **Rationale for the status** |
|---|---|---|---|---|
| **E-CORRIDOR-GA-DM-01 (E-CORRIDOR Ope-01)** | MUST | The gait analysis system provides interaction through the edge device (smartphone) and the cloud service for the walking path analysis. | Completed | We implemented and successfully tested the APIs and an Android application for performing the gait analysis. The tool distinguishes between 4 different activities and enforce user authorization. |
| **E-CORRIDOR-GA-DM-02 (E-CORRIDOR-DS-10)** | MUST | The cloud service, after the authentication process must delete the subject information. | In progress | We are integrating the analytics in the E-CORRIDOR framework finalizing the integration in the IAI toolbox. Data in the ISI will be stored and deleted following the defined DSA policy. |
| **E-CORRIDOR-GA-DM-03 (E-CORRIDOR-DS-11)** | MUST | The gait analysis system allows to control the access to every transportation hub through the walking path user verification. | To be deployed on the final release | We are working to extend the training dataset and the model on the use cases identified by the pilots (e.g., S2C or AT). |

### 2.4.5. Plan for Testing and Final Maturation

After the full integration in the E-CORRIDOR framework the component will be tested in the pilot environment. In particular, the dataset used in the learning stage will be extended considering users operating in the pilot setting. E.g., a smart tunnel deployed in the airport can be used to match the extracted features with the ones collected at home during the enrollment process. Moreover, the gait-based authentication can be a building block of more complex multi-factor authentication through the ASI (see Sec 2 in D8.2)

## *2.5. Face recognition - passenger authentication [E-CORRIDOR-IAI-FR]*

Research on face recognition has become a popular topic in recent decades. Numerous studies exploring its applications on video surveillance, criminal identification, and building access

control, etc. have made significant progress. Face recognition can be explained as the process of recognizing a person's face using biometric features.

### 2.5.1. Component Description

The face recognition process is often described as a task that can be divided into four steps: (i) face detection, (ii) face alignment, (iii) face extraction, and (iv) face recognition. The first step is essential in charge of locating human faces in images and video frames and detecting if these faces are real. In the second step, the unstructured information (a face) is normalized into a set of digital information (data) to be consistent with the database. The third step consists of extracting face features from the real ones. The fourth step verifies if two faces belong to the same person [7], [8].

Face recognition analytics is based on a ResNet deep learning model trained on a dataset of about 2.7 million faces; this dataset is composed by the union of the FaceScrub and the VGG datasets [9], [10]. The former contains 106,863 face images of 530 people, whereas the latter contains 2.6 million images of 2,622 people. To illustrate the accuracy of the model, the Labeled Faces in the Wild (LFW) benchmark, which is designed to study the problem of unconstrained face recognition, is exploited. More than 13,000 images of faces are collected from the web in this benchmark. The name of the person pictured is associated to every face. Moreover, at least two photos of 1680 people are included in the dataset. The performance of the model in terms of accuracy is of 99.38%, meaning that the model correctly predicts if the images are of the same person. The model based on a ResNet network is composed of 29 convolutional layers corresponding to the ResNet-34 variant presented in [11].

To distinguish between a real and a fake face image, face recognition analytics allows to monitor the left and right eye, nose, and the left and right ear landmark positions. Using a stereo camera like the Intel RealSense D435 depth camera, we can create a point-cloud and the 3D model of the face. Therefore, we can distinguish between real and fake face images with high probability. Additionally, face recognition analytics also allows to monitor the eye landmark positions to distinguish between live and static face image, e.g., a recorded video or an image. An Eye Aspect Ratio (EAR), characterizing the eye opening in each video frame, is exploited for detecting fake face images. EAR is computed using the height and width of the eye based on six 2D landmark locations representing eye contours. An EAR differs from real to fake face image where this ratio remains constant all over the fake frames. A Support Vector Machine (SVM) classifier is exploited to detect EAR values pattern in a short temporal window. We adopted a cross-dataset validation on two public datasets: it means that the classifier is trained on the Eyeblink8 dataset and tested on the ZJU dataset and vice versa [12], [13]. ZJU dataset consists of 80 short videos of 20 subjects while Eyeblink8 dataset consists of 8 long videos of 4 subjects; these datasets contain ground-truth annotations of blinks. The performance of the classifier in terms of precision and recall for the two datasets are of approximately 97% and 94%, respectively.

Face recognition analytics exposes an HTTPS API web service to comply with the IAI toolbox specifications and it has been packaged as a docker container.

### 2.5.2. Workflow in Action

Face recognition analytics takes two inputs, namely a ground-truth face image and a video stream recorded in the airport/train station, for recognizing the passenger. The ground-truth face image is collected from the passenger’s identity document (either passport or identity card) e.g., by means of the E-CORRIDOR Android application, through the support of an Identity

Management System. A video stream is recorded in the train station kiosk and processed by the local instance of the E-CORRIDOR framework. Face recognition analytics first exploits the video stream to select from it the frames where the passenger face appears. Moreover, for privacy purposes, a saliency detection is employed to select only the section of the image containing the face of the passenger in front of the train/airport station kiosk, in case multiple persons behind him/her are present.

To summarize the face recognition analytics workflow, Figure 13 shows the different steps involved in the authentication of the passenger.

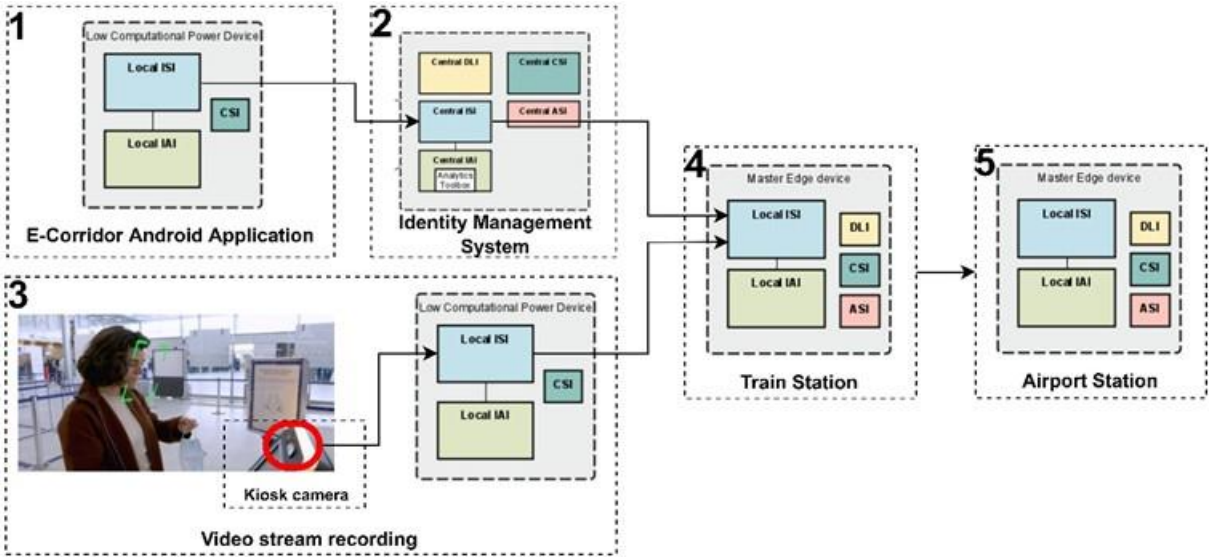


**Figure 13. The proposed face recognition analytics workflow**

In details, the steps of the face recognition analytics workflow are as follows:

- Step 1: E-CORRIDOR Android application aims, mainly, to collect the passenger personal information as well as the choice of his/her ticket journey by following different guiding instructions.
- Step 2: the passenger performs the sharing of his/her travel documents information before starting the journey. The Android application is developed in the context of AT pilot. To be enrolled in the face authentication mechanisms, the passenger must explicitly opt-in and agree to share his/her personal information with the transportation service provider (e.g., the train station). Such an agreement is represented by a Data Sharing Agreement (DSA) rule. The information is stored in the framework in an encrypted form and made accessible only to the kiosk devices deployed in the airport and train-station.
- Step 3: it consists of the recording of a video stream required for the face recognition analytics; this video stream is stored in an encrypted format in the local ISI.
- Step 4: face recognition analytics is invoked according to the above-mentioned DSA rule in the train station.
- Step 5: similarly to step 4, this step consists in performing face recognition using the passenger personal information collected in step 1 and a new video stream recorded in the airport station kiosk. The transmission of the tokenized information from the local

ISI of the train station to the one located in the airport is performed according to a DSA rule. Finally, after performing face recognition in the airport station, the same DSA rule allows to delete all the information for privacy purposes.

### 2.5.3. Integration and Maturation Status of the First Release

The containerization of the face recognition analytics is achieved, and the resulting Docker is available on the E-CORRIDOR's Nexus repository. Several tests were conducted to check the connection between the local ISI and the IAI in a local setting. Ongoing efforts are oriented at testing the integration with a DSA policy defined according to the workflow which specifies authorizations and obligations rules.

### 2.5.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | **Status** | **Rationale for the status** |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-FR-01 (E-CORRIDOR-DS-05, E-CORRIDOR-DS-06, E-CORRIDOR-DS-07, E-CORRIDOR-DS-17, E-CORRIDOR-DS-23, E-CORRIDOR-DS-24)** | MUST | The accuracy and effectiveness of passengers' face recognition is dependent on the DSA specified by each passenger, passengers' activity, and on contextual/environmental properties. | To be deployed on the final release | The DSA defined by the passenger DSA regulates the applicability of the data sharing needed to perform the face recognition in subsequent touchpoints. |
| **E-CORRIDOR-IAI-FR-02 (E-CORRIDOR Ope-02)** | MUST | Face recognition can be performed at the edge. | Completed | If the passenger allows to share his/her personal information using the developed E-CORRIDOR Android app. |
| **E-CORRIDOR-IAI- FR-03 (E-CORRIDOR-Tst-AT-02, E-CORRIDOR-Tst-AT-03)** | MUST | Intel RealSense stereo camera, Light Detection and Range camera are used for face recognition to identify passengers | To be deployed on the final release | Future tests will exploit the same Intel RealSense camera in use for the activity recognition analytics. |

| | | | | |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-FR-04 (E-CORRIDOR-Use-02)** | SHOULD | The face recognition of each passenger should be used to enhance the seamless authentication. | Completed | The current version allows authenticating the passenger using his/her personal information. |
| **E-CORRIDOR-IAI-FR-05 (E-CORRIDOR-DS-10)** | MUST | The inferred passenger face information is transmitted and stored (only for the needed time) in a privacy-aware and secure manner. | Completed | The message containing the authenticity of the passenger is encrypted and the DSA regulates the data retention period. |

### 2.5.5. Plan for Testing and Final Maturation

Further efforts for the face recognition analytics will be devoted to the design of the Android application for sharing passenger information. Moreover, we are evaluating the possibility to adopt a depth camera to detect with more accuracy the presence of a real person in front of the camera.

We are currently performing integration tests to evaluate the whole workflow. Finally, more experiments in challenging/real environments will be performed to further maturate the analytics.

## *2.6. Activity recognition - passenger authentication [E-CORRIDOR-IAI-AR]*

Human activity recognition refers to the classification of the person's activities using data collected by wearable, environmental, or vision sensors. An activity is defined as the movement of one or more parts of the person's body. Human activity recognition (HAR) technology has been used to develop context-aware applications in a variety of fields, including Internet of Things (IoT), Ambient Assisted Living (AAL), and healthcare. As part of the E-CORRIDOR project, the HAR analytics is used in order to obtain contextual information for enhancing authentication obtained by other components, such as face recognition and gait analysis.

### 2.6.1. Component Description

In HAR, activities are generally recorded with specific motion systems, such as optical motion capture systems. However, in an airport setting, we cannot generally track a traveler with wearable sensors (with perhaps the sole exception of passengers that install an application on the smartphone and are willing to share sensor data). Therefore, we focus on image-based activities recognition analysis.

Our analytics works in two essential steps. The first step is related to the data collection. This is the most critical and challenging step of the entire human activity recognition task. Through the proposed methodology, we must ensure that the reached accuracy and representativeness allow us to correctly perform HAR in the test environment. At this stage, we used multiple Intel

RealSense D435 series depth cameras, each able to capture color, depth, and two stereo infrared streams. We first segment the image to extract different persons in the frame, then use a framework called MediaPipe [14] to identify the human body's critical data points. More specifically, *MediaPipe's Pose* calculates and estimates the center point between the person's hips, the size of the circle circumscribing the entire person, and the incline based on the Vitruvian man. This framework produces in output 33 points, but we are interested just in the following 25 landmarks: the nose and the 24 Landmarks (right and left) of the eyes, shoulders, elbows, wrists, fingers, thumbs, indexes, hips, knees, ankles, heels, and hallux. We process with this framework the output streams generated by the Intel RealSense cameras to extract the users' activity sequences and critical data points.

After having generated the labeled datasets, the second step is the learning phase. We feed the data to machine learning algorithms to classify the activities based on extracted normalized coordinates of each landmark (25 landmarks), macro features which describe the general characteristics of an activity, such as speed, cadence, and step length and micro features that measures the small variations in gait (like shivering). Support-vector machine (SVM), Random Forest (RF), Naïve Bayes (NB) and Long Short-Term Memory (LSTM) algorithms are explored for preliminary results. Once the best parameters of each model are found, we move on evaluating their performance and choosing the best one in terms of two measures: Precision and Area Under the Curve (AUC).

### 2.6.2. Workflow in Action

Activity recognition analytics takes as input video streams recorded in the airport/train station, for recognizing the activity of the passengers. These videos are collected by means of the D435 cameras connected to Raspberry Pi devices running an edge instance of the E-CORRIDOR framework. The videos are then sent to the Information Sharing Infrastructure (ISI) subsystem, encrypted, and made available for processing by the activity recognition analytics. Following the same approach adopted by the face recognition analytics, to enroll the passenger in the activity recognition mechanisms, he/she must explicitly opt-in by accepting a DSA with the transportation service provider (e.g., the train station). By following the DSA, from this point on only the activity recognition analytics will be able to perform analysis on those videos.

To summarize the activity recognition analytic workflow, Figure 14 shows the different steps performed for authenticating the passenger.

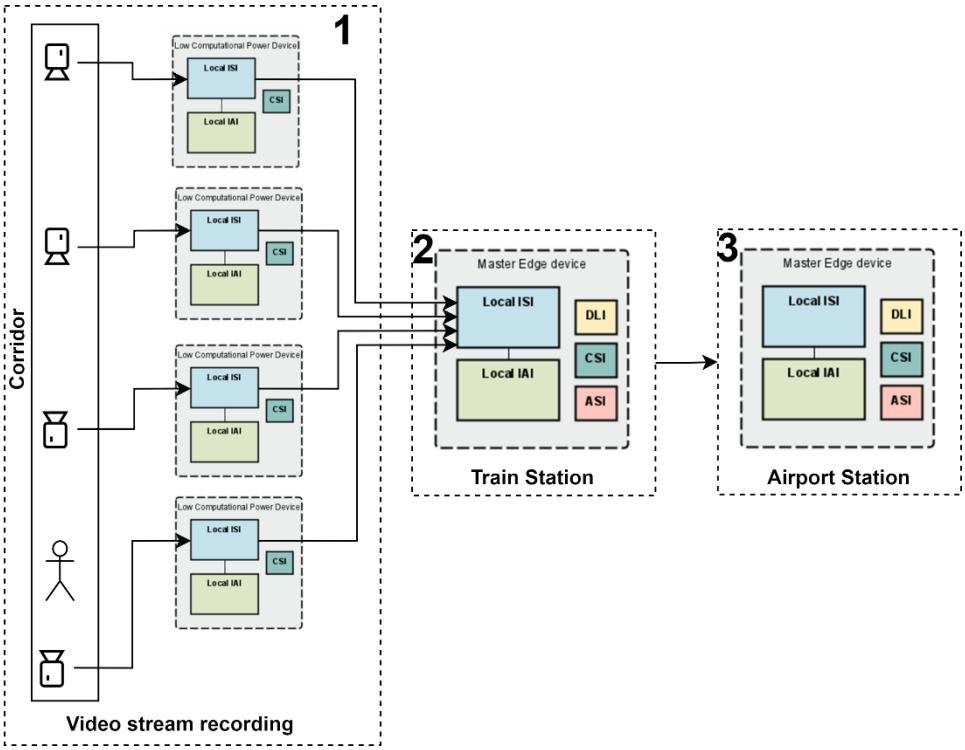


**Figure 14 The proposed activity recognition analytics workflow**

The steps of activity recognition analytics workflow are as follows:

- Step 1: it consists of the recording of a video stream required for the activity recognition analytics for each of the depth camera; these video streams are stored in their local ISI in an encrypted format.
- Step 2: activity recognition analytics is invoked according to a DSA rule in the train station. In addition to allowing the data processing only to such analytics, the DSA will pose additional constraints such as on location and organization for which the analysis can be performed.
- Step 3: similarly to step 2, this step consists of performing activity recognition using the videos collected in step 1 and new video streams recorded in the airport station. Finally, after performing activity recognition in the airport, a DSA rule enforces the deletion of all these data for privacy reasons.

### 2.6.3. Integration and Maturation Status of the First Release

We performed some initial tests in our labs whose setting was as follows:

- Five healthy adults (our research colleagues working in the project) participated in the data collection with different morphological characteristics and various clothes.
- The dataset targets actions commonly observed in an airport environment. Therefore, each participant completed five activities: walking, walking with a bag, walking with a suitcase, sitting, and standing

- The objects utilized in the activities have been placed in such a way that they stimulate the adoption of various postures. However, no instructions were provided to the participants on how to complete each activity.

Each trial consisted of performing one of the above mentioned five activities. We recorded streams (color, depth, stereo infrared) from multiple Intel RealSense D435 cameras. Each participant performed all five activities, with two consecutive trials per activity, i.e., ten trials per participant. The time required to each participant to complete the recording was of approximately 1 hour, including recording time, break time, and repetition time due to problems in the recording. Therefore, the dataset includes a diversity of participants and several trials of activities performed in different ways to consider inter-and intra-individual movement variability.

The data for this preliminary test were collected in a laboratory environment (laboratory corridor and university hall). In the laboratory corridor, each of the four cameras were connected to a computer to perform the data collection. The cameras were positioned at a height of 2 meters: at the entrance, in the middle, and at the end of the corridor. In the university hall, we adopted three cameras positioned in a well-studied corner to have a good viewing angle close to the conditions in the airport. Figure 15 shows how the hall was equipped with these cameras positioned at different angles of the room, to capture information of the participant acting at about 3 meters from different points of view.


**Figure 15 Camera installation in the university hall**

Among the above-mentioned classification algorithms, results from our experiments show that the SVM provides better results in the activity classification with an accuracy of 91%. Whereas, for identifying the person performing the activity, results show that the Random Forest generated the best model, with an accuracy of 96%.

Currently some features are available and preliminary tests have been performed. During the next year additional features will be added, and tests performed. A summary of the performed efforts and a plan for the next year are reported in the following table.

| Feature | Current version | Final release |
|---|---|---|
| Primary laboratory tests | ✓ | |
| Skeleton data extraction | ✓ | |
| Primary classification using deep learning algorithms | ✓ | |
| Extended lab tests | | ✓ |
| On-site tests | | ✓ |
| Improving the accuracy of the deep learning methods in online tests | | ✓ |
| Integration of new skeleton data arriving on-the-fly | | ✓ |
| Integration and tests with E-CORRIDOR framework | | ✓ |

### 2.6.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | Status | Rationale for the status |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-AR-01 (E-CORRIDOR-DS-05, E-CORRIDOR-DS-06, E-CORRIDOR-DS-07, E-CORRIDOR-DS-17, E-CORRIDOR-DS-23, E-CORRIDOR-DS-24)** | MUST | The accuracy and effectiveness of passengers' activity recognition is dependent on the DSA specified by each passenger, and on contextual properties. | To be deployed on the final release | Macro features (e.g., color of the clothes, and skeleton data obtained by the depth camera) will be shared on the ISI if passengers accept the service beforehand. |
| **E-CORRIDOR-IAI-AR-02 (E-CORRIDOR Ope-02)** | MUST | Activity recognition can be performed at the edge. | To be deployed on the final release | If the passenger allows to use activity recognition through the E-CORRIDOR Android app. |
| **E-CORRIDOR-IAI- AR-03 (E-CORRIDOR-Tst-AT-02, E-CORRIDOR-Tst-AT-03)** | MUST | Intel RealSense stereo cameras and Light Detection and Range camera are used for activity recognition | To be deployed on the final release | A setting with few cameras has been tested in the lab |

| | | | | |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-AR-04 (E-CORRIDOR-Use-02)** | SHOULD | The activity recognition of each passenger should be used to enhance the seamless authentication. | To be deployed on the final release | The current version allows authenticating the passenger based on his/her skeleton data. |
| **E-CORRIDOR-IAI-AR-05 (E-CORRIDOR-DS-10)** | MUST | The inferred passenger activity information is transmitted and stored (only for the needed time) in a privacy-aware and secure manner. | To be deployed on the final release | The message containing the authentication of the passenger is encrypted and associated with a data retention policy. |

### 2.6.5. Plan for Testing and Final Maturation

We are now focusing on enhancing the accuracy of our activity classification and person identification based on the developed end-to-end deep learning algorithm. Firstly, extended tests will be performed in the laboratory for on-the-fly skeleton data extraction as well as online activity recognition including a fourth depth camera located in the hall. Finally, further testing will be possibly conducted in the airport and train station.

# 3. Privacy Preserving Itinerary Planning – Task 7.2

Nowadays, people often choose multi-modal travels. This task aims at providing analytics to support the trip planning by inferring the best itineraries according to the traveler's interests and preferences such as CO2 footprint, price, time, and number of connections. While planning the best trips, the analytics exploits public and anonymized data and self-adapt the plans to context changes or encountered critical situations.

## *3.1. CO2-aware Trip Planning [E-CORRIDOR-IAI-MMIP]*

### 3.1.1. Component Description

The trip planning can exploit multiple mobility solutions such as electric vehicle, car sharing and on-demand bus services along with public transit information. To improve the user experience, users can customize their trip with their own preferences.

The analytics uses the mature A* algorithm [15] to route the passengers to the desired location. This is aided by the preferences supplied by the user but, as some journeys require walking further, or due to scheduling the arrival time might not be reached, these parameters are modeled as soft constraints.

The main use case for this analytic is the S2C-US-03 "Trip planning and Carbon footprint" described in D3.1. To fulfill the pilot requirement, the Multi-Modal Trip Planner interacts with the CO2 analytics [E-CORRIDOR-IAI-CFA] (please see Sec. 5.1) to provide the end user with CO2 calculations on their journey. Furthermore, following recent trends on the public transport, to achieve a more equitable and efficient mobility in European cities, the trip planner foresees the integration with a micro subsidies calculation engine. More precisely, the Rideal's micro subsidy engine [16] will be integrated to manage, monitor, and control rider incentive programs. By calling the external micro subsidy engine with limited information, the analytics will then be able to inform the user about the availability of a micro subsidy for the desired trip. In such a way the users, from a variety of available routes, can better choose also considering discounts/offers that may be available.

Using a RESTful API, the S2C pilot environment can access the service and get the required data back in a simple JSON format. To customize their trips, users may include their preferences such as transit mode or distance walked, and the analytics will consider these in the computation of the best fitting options. In any case, these options have default values. Figure 16 shows the exposed API.

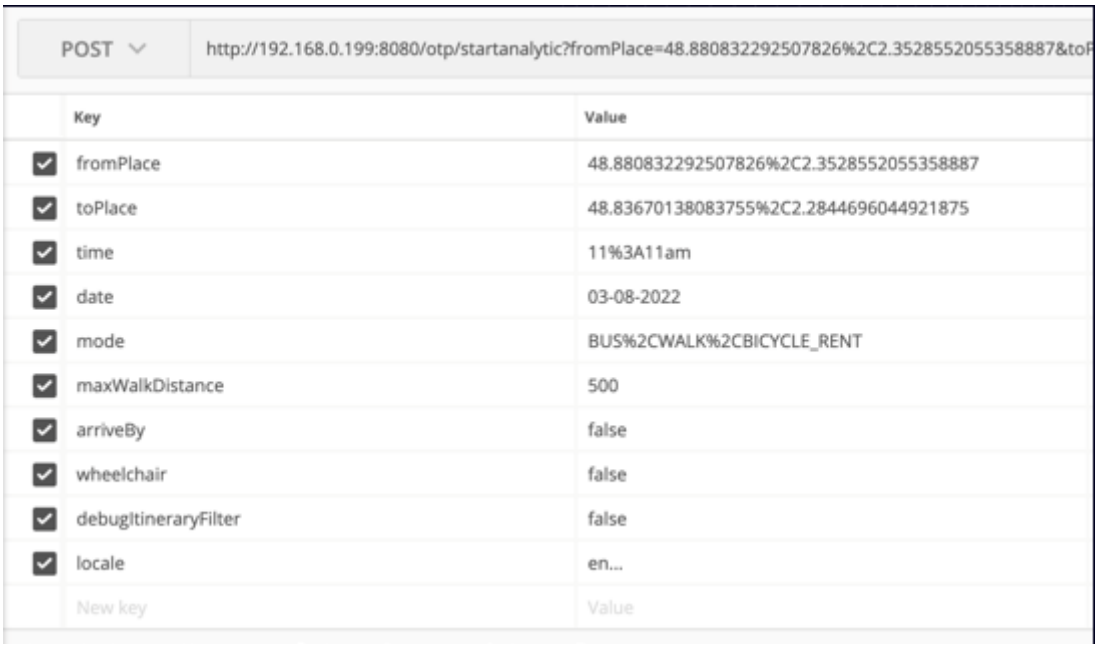


**Figure 16 The *startanalytics* endpoint with the Itinerary Trip Planning parameters**

### 3.1.2. Workflow in Action

When coupled with the S2C pilot application, the Multi-Modal Trip Planner will respond with the data required to show the path on the map. The path is represented using polylines, and the info box will return the itinerary of the planned trip including the CO2 calculation, and the likelihood of receiving a micro-subsidy.


**Figure 17 The S2C pilot application for the Multi-Modal Itinerary Planning**

Having developed further upon the service, the Multi-Modal Trip Planner has integrations with the IAI interface and can be accessed through the */startanalytic* interface with its relevant parameters.

As the integration with the CO2 calculation and the micro subsidies engines progress the trip planner analytics will be able to expand the options available to the user as well as return more detailed results to the end user.

### 3.1.3. Integration and Maturation Status of the First Release

Implementation and integration efforts are summarized in the following table as well as if these are ready or will be part of the final maturation.

| Feature | Current version | Final release |
|---|---|---|
| Primary integration into IAI | ✓ | |
| Primary Integration of Micro subsidies | ✓ | |
| Primary Integration of CO2 Calculation Analytic – task 7.4 | ✓ | |
| Full integration in the E-CORRIDOR framework (ISI interaction) | | ✓ |
| Access data for GTFS/GBFS feeds externally | ✓ | |
| Micro subsidies integration | | ✓ |
| Integration of the CO2 calculation analytics | | ✓ |

### 3.1.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | **Status** | Rationale for the status |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-MMIP-01 (E-CORRIDOR-DS-20, E-CORRIDOR-Tst-Int-S2C-02)** | SHOULD | The multi-modal trip planning tool should be able to pull data (such as public transit feeds) from external sources, with specified polling intervals. | Completed | Once supplied with a feed address the multi-modal trip planning tool pulls for the requested data at a specified interval |
| **E-CORRIDOR-IAI-MMIP-02 (E-CORRIDOR-DM-01, E-CORRIDOR-DM-02, E-CORRIDOR-Sec-RC-01, E-CORRIDOR-DS-10)** | MUST | Data used to automate trip planning needs to be obfuscated or anonymized to enhance the privacy preservation. It should be also deleted after a certain amount of time. | Completed | The multi-modal trip planning tool uses only data seen to be public (location) and does not save user data or associate location data with user data. |

### 3.1.5. Plan for Testing and Final Maturation

Moving forward during the final tests the Multi-Modal Trip planner will get incremental updates and complete the integration with IAI API. Moreover, the integration plan with the E-CORRIDOR framework foresees that the data to and from the multi-modal trip planner will pass through the ISI subsystem. The end goal is to have seamless access to trip planning services with CO2 calculations and micro subsidy information embedded in the E-CORRIDOR framework.

As the planning aspect of the multi-modal trip planner already uses mature algorithms, future iterations of this component will mainly involve the exploitation of the ISI and any small changes to the CO2 integration.

Using the S2C pilot environment, the API responses will be tested to ensure that the user ends up with the best options for their given route. This feature testing will involve checking different routes across the day. Such a testing will involve the verification and validation of the:

- Integration with S2C pilot front-end
- Micro subsidies integration results
- CO2 integration results

Although these final tests cannot be carried out through the test suite, the current test suite for the routing does ensure that the routing planner works as expected.

# 4. Privacy Preserving (Security) Analytics – Task 7.3

Analytics based on privacy-preserving techniques such as Fully Homomorphic Encryption (FHE) and secure two-party computation (2PC) can ensure privacy over the user data even at execution time. By considering the pilots' scenarios and the related security and privacy constraints, the two techniques have been used to customize analytics for the S2C and the AT pilots. In particular, the FHE has been used to check the validity of the driving license against a repository of valid numbers, whereas the 2PC has been exploited to match the passenger preferences with the services offered in the airport terminal.

## *4.1. OpenAPI for Fully Homomorphic Encryption [E-CORRIDOR-IAI-FHEC]*

Fully Homomorphic Encryption (FHE) is a form of encryption that allows one to perform computations directly on the encrypted data. Such a technique is finding application in cloud computing services to increase security and respect the privacy of the users. However, FHE processing is generally orders of magnitude slower than the corresponding plaintext operations. Such an issue can limit its applicability in real-time systems or more generally in all the application where time is critical.

### 4.1.1. Component Description

To mitigate the above-mentioned performance limitation, we propose an efficient *pattern search* algorithm deployed on a computer cluster. Such an algorithm can check if a given encrypted information is present in an encrypted dataset without revealing any additional information on the plaintext data. Application domains span from secure analysis and monitoring to private search in a database. In accordance with the S2C use cases, in this section we applied our solutions to verify the validity of a driver license number against a database of expired, not valid, or suspended cards.

The component is composed by two main subsystems:

- The homomorphic encryption engine: in charge of evaluating the function e.g., to check the validity of the driving license.
- The supporting functions (exposed as Open API) to handle requests to the database i.e., for storing, deleting, modifying a ciphertext.

In general, the component requires the generation of the secret and the public keys (plus additional special keys useful for optimization purposes). As first step, the database constituting the ground truth is encrypted. Later, whenever there is a need to check the validity of a driving license, the latter is encrypted before being analyzed. The evaluation engine takes in input the ciphertexts of the driving license(s) under analysis and the one of database, and returns the presence of a match (i.e., an intersection among the two sets). Thanks to the adoption of the FHE technique, no other information is reveled, and the data are encrypted even during computation.

### 4.1.2. Workflow in Action

In the E-CORRIDOR framework, the needed keys are generated by the CSI (Common Security Infrastructure) subsystem. As discussed above, the prosumer uses these keys to create a new reference database with the list of valid driving licenses (the keys are later used to manage such a database). The API exposed by the analytics is presented in Figure 18.

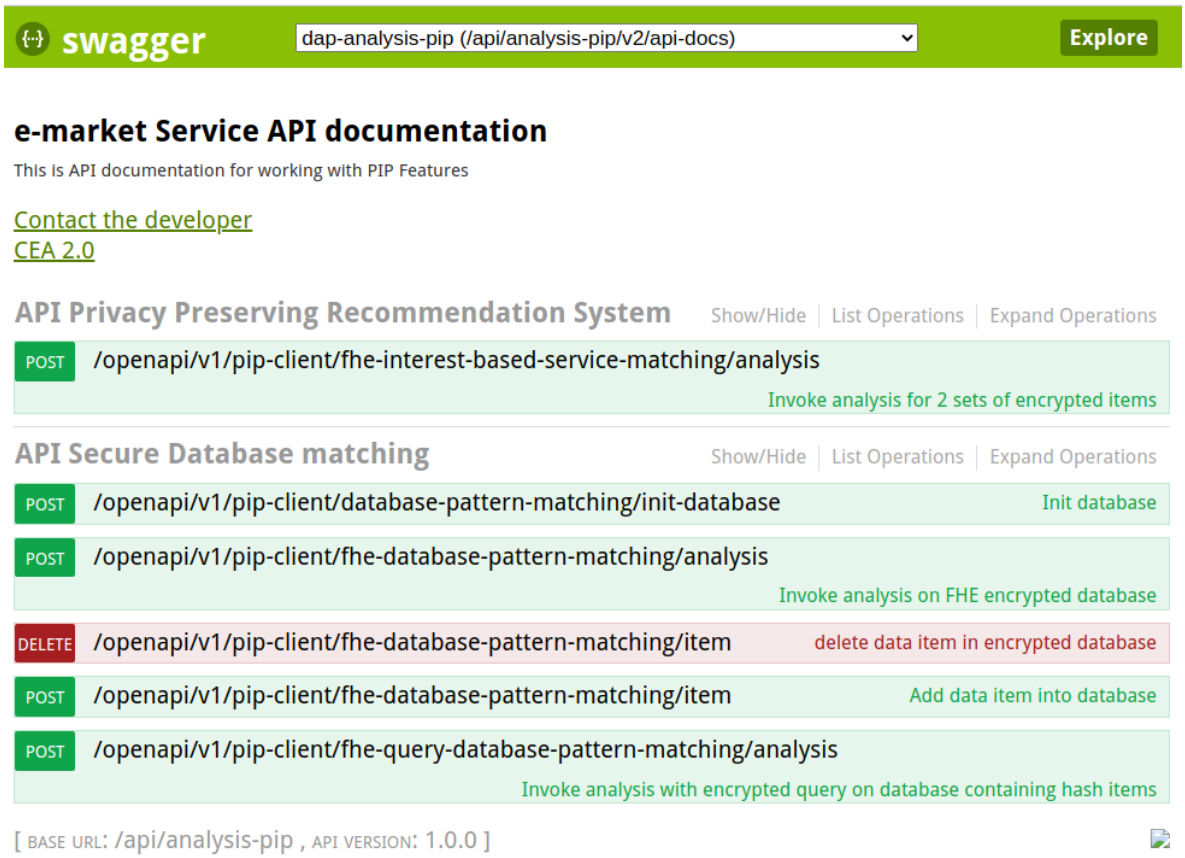


**Figure 18 Pattern Search Analysis API**

To exemplify the process, the database with the ground truth is first created by calling the *init-database* API. Then, in order to check if a driving licence is valid (i.e., if it is part of the database created in previous step), the user invokes the *analysis* API providing in input the driving license number that he/she wants to verify (with the *privacy-pattern* attributes highlighted in Figure 19). Such an input is encrypted before being sent to the E-CORRIDOR server for analysis.

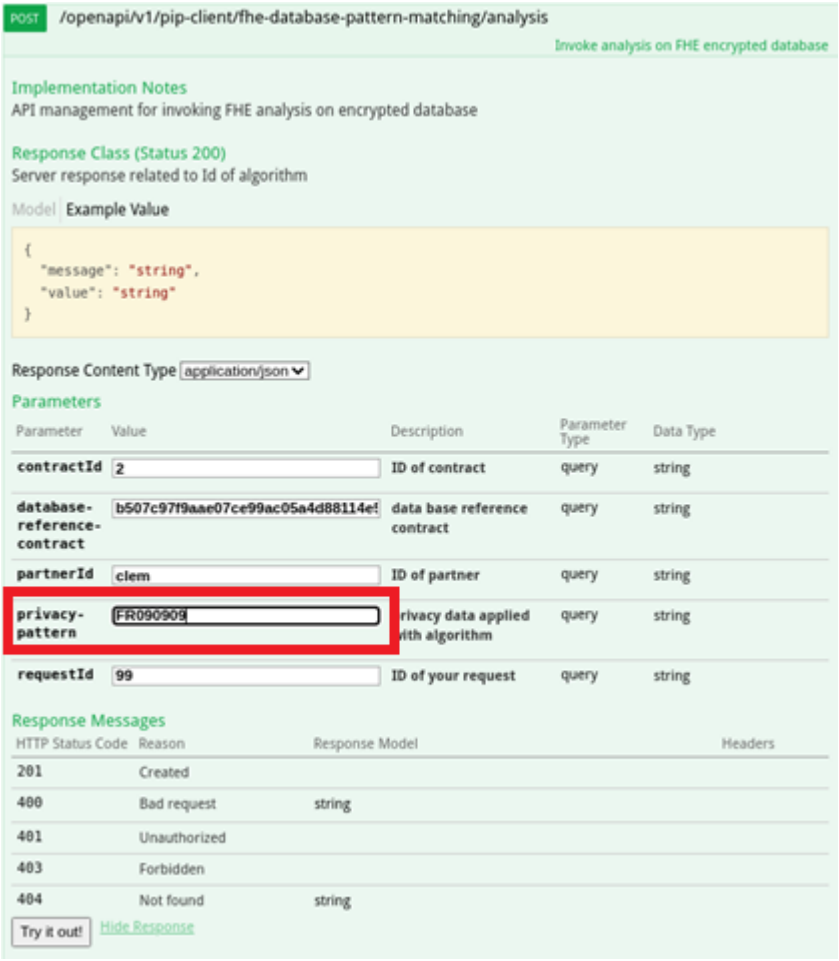


**Figure 19 Driving license checker API**

As in case of large databases the analysis can require several seconds, an asynchronous approach has been adopted. Therefore, upon submitting the request, the user receives a reference identifier later used in the result manager API (see Figure 20). By invoking the *analysisReferenceId* and providing the reference identifier, the analysis results can be retrieved (i.e., a Boolean value representing the presence/lack of a match in the database).

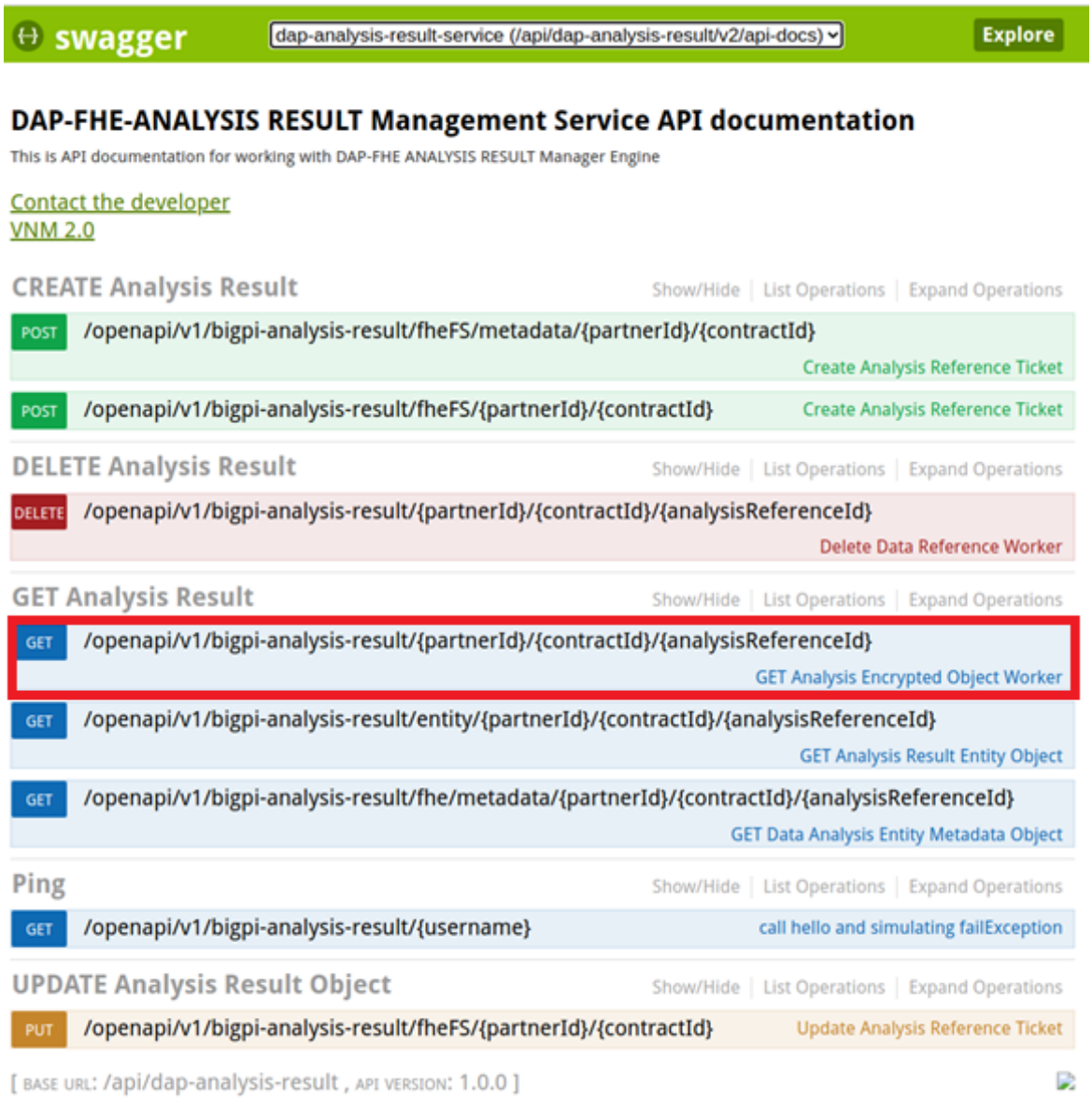


**Figure 20 Pattern Search - Result API**

### 4.1.3. Integration and Maturation Status of the First Release

At the time of writing this document, the algorithm and APIs have been implemented and their functioning tested on the CEA server. The integration in the E-CORRIDOR framework is in progress as well as the containerization of the component.

The experiments of our pattern search algorithm have been conducted on an Intel Core i7-7600U with 16GB of memory and Ubuntu OS installed. Results show that a set of 102 ciphertexts can be checked in less than 0,5 seconds. However, some security parameters of the algorithm can be tuned to work with a larger set of ciphertexts or to improve the security level. These changes will affect the performance of the algorithm and further optimization tasks would be needed.

### 4.1.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | **Status** | **Rationale for the status** |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-FHEC-01 (E-CORRIDOR-DA-03, E-CORRIDOR-DA-05)** | MUST | The mapper functionality ensures that the data sharing constraints expressed in the DSA are enforced at any time during the analysis. | Completed | FHE processes data on encrypted format. Therefore, data stored in the cloud are always encrypted. |
| **E-CORRIDOR-IAI-FHEC-02 (E-CORRIDOR-DA-02)** | MUST | Data used to identify drivers may require to be transformed into a common data format to work with the analytics. | Completed | The FHE engine can work with any basic data format such as integer, string, etc. Driving license are usually a string of alphanumeric characters thus the translation is straightforward. |
| **E-CORRIDOR-IAI-FHEC-03 (E-CORRIDOR-Tst-S2C-01, E-CORRIDOR-Tst-S2C-02)** | MUST | The In-Vehicle Infotainment (IVI) or Electronic Control Units (ECUs) may be used for collecting GPS and connection behavior data. | Under evaluation with the S2C pilot | If the driving license should be checked along additional attributed (such as the car position), the analytics will be extended to accept multiple parameters. This option/need is still under discussion with the S2C pilot. |

### 4.1.5. Plan for Testing and Final Maturation

To deal with large datasets, a distributed version of the analysis is under development. The designed approach will exploit multiple virtual machines for the execution of the evaluation function. Moreover, the component will be integrated in the S2C pilot, and its customization will be finalized according to the received feedback.

## *4.2. Secure Two-party-computation for interest-based services [E-CORRIDOR-IAI-MPCSR]*

### 4.2.1. Component Description

In this section, we describe an interest-based analytics based on the usage of the secure Two-Party Computation (2PC) technique. In particular, the analytics uses the 2PC-based service sharing part of the Advanced Security Subsystem (ASI) subsystem and described in T8.3 of D8.2. The goal of Interest-Based 2PC is to provide a way to run analytics that will use passengers' private information, however without disclosing those private details to the other party.

Note that with respect to the D7.1, we refined this analytics in agreement with the pilot's requirements. Initially, we were considering applying the Secure Multiparty computation to the driver identification in the S2C pilot scenario. Within the second year of the project, we noted that part of the concepts introduced in corresponding section of D7.1 (please see Section 4.2 there) can be used to make the Driver DNA analytic private, as described above in Section 2.1. Hence, in this section we describe the analytics we are developing for the AT pilot that has shown a huge interest in the usage of the S2PC protocol to provide customized services to passengers according to their interests.

### 4.2.2. Workflow in Action

The analytic works considering two main actors involved, they are:

- A Service Provider that is represented by the airport and/or train station.
- A Prosumer that is represented by a passenger during his/her multimodal travel.

The Prosumer aims to run a particular service offered by the Service Provider. When running a service both parties involved wish to keep their data private. In particular, the offered service will make use of the prosumers' interests.

The scenario that is under consideration for the AT pilot foresees passengers involved in a multimodal travel that move from, for instance, a first part of the travel done by train to the second part, which involves a connection flight. The passenger has some free time to spent to go to a restaurant before the departure. When the passenger is approaching the airport, she will be able to use her smartphone to run the interest-based analytic that uses the 2PC technique to keep private the involved data. The service provider hosting at the airport side will allow the passenger to know all the restaurants that match the passengers' interests.

**Table 1 Set of interests that a passenger can consider with examples**

| Parameter | Description | Value types |
|---|---|---|
| Menu | It considers the type of restaurant | Italian, American, Indian, Japanese |
| Location | How far from the passenger position | In a range of 200 meters |
| Cost | It expresses the type of restaurant in terms of costs. For instance, a fast-food or an expensive restaurant | Cheap, expensive |
| Time to wait | It approximately indicates the maximum amount of time to wait until the passenger is served | Less than 10 minutes |

Table 1 shows the list of interests currently considered to be part of the interest-based analytics. A passenger has to set up those parameters before running the analytics. For each interest, a passenger can provide her degree of preferences, and these will be matched in a private manner with the information available by the service provider.

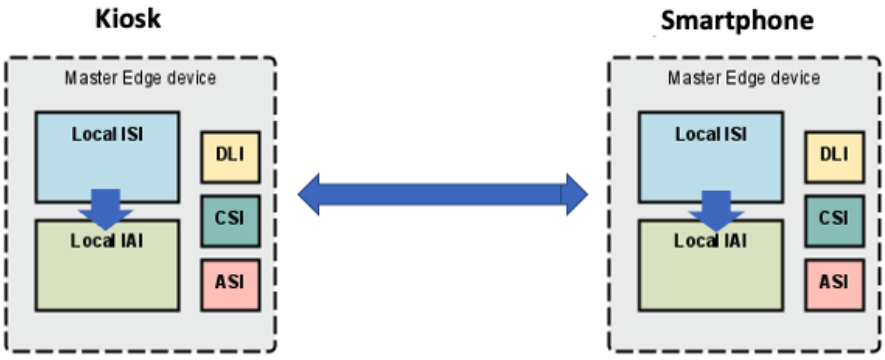


**Figure 21 Interest-based 2PC service between a Kiosk and Smartphone**

In Figure 21, we show an interaction between a self-service kiosk, which resides at the airport side and acts as Service Provider party, and a passenger's smartphone. All data related to the passenger's interests and the available restaurants are first stored in the local ISI. Then, when the analytics that involves the 2PC–based service sharing component is run, the 2PC module available in the ASI subsystem is triggered.

### 4.2.3. Integration and Maturation Status of the First Release

At the time of writing this document, the analytics works as standalone component and its integration with the AT Pilot and the E-CORRIDOR framework is in progress. The current implementation is written in Java, and it is based on CBMC- GC tool [17]. It is composed of two main parts: the compiler that translates functions written in "C" into garbled circuits, and the interpreter that can execute compiled functions [18]. Thus, CBMC-GC offers a very flexible high-level language that allows developers to express a wider range of functions compared to simpler techniques, which for instance only focus on simple private matching operations.

To work with passengers' smartphones, we have extended and adapted CBMC- GC to work on the Android operating system (OS).


**Figure 22 2PC – based service main window**


**Figure 23 2PC– based service price customization**

In Figure 22 and Figure 23, we show the 2PC-based service app customized to work in a standalone fashion for the AT pilot. With the app, a passenger will be able to get a list of restaurants based on his/her preferences customized through the app itself. Once the passenger has indicated the preferences, he/she will get the list of restaurants that will be selected according to the preferences in a complete privacy-preserving manner, i.e., the passengers' preferences will not be disclosed with the service provider and vice versa.

Our version of CBMC- GC extended to work as 2PC – based service sharing is able to work on a scenario that foresees the usage of two devices based on the Android OS, or in a hybrid case in which an Android OS and an edge node or cloud node that supports the Java virtual machine are involved.

### 4.2.4. Requirements Traceability Matrix

As remarked at the beginning of this section, the privacy-preserving computation capabilities originally proposed for this analytics have been merged in the Driver DNA analytics (in Section 2.1) considering the privacy requirements of the S2C pilot. On the other hand, there was a surging need and opportunity in the AT pilot to match the passenger preferences with the services offered in the terminal. In the following table, the contribution of the analytics to the platform requirements has therefore been adapted.

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | **Status** | **Rationale for the status** |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-MPCSR-01 (E-CORRIDOR-DS-09)** | MUST | Stakeholders may require running analytics expressing conditions to preserve confidentiality over the shared data. | Completed | The analytics adopts the 2PC-based service sharing provided by the ASI subsystem to assure confidentiality on the user's and service provider's data |
| **E-CORRIDOR-IAI-MPCSR-02 (E-CORRIDOR-DM-01, E-CORRIDOR-Sec-RC-01)** | MUST | Data used to identify the user may require to be obfuscated, anonymized or other privacy-preserving technologies must be adopted. | Completed | The 2PC technique assure privacy on the data even during computation |
| **E-CORRIDOR-IAI-MPCSR-03 (E-CORRIDOR-DA-02)** | SHOULD | Data used to identify users may require to be transformed into a common data format to work with the analytics. | Next release | We are working with the AT pilot to understand the best format to express the services offered in the airport |
| **E-CORRIDOR-IAI-SR-02 (E-CORRIDOR Ope-02)** | MUST | Analytics can be run at the edge. | In progress | The analytics may run on two Android smartphone or with devices running the Java virtual machine. |
| **E-CORRIDOR-IAI-SR-03 (E-CORRIDOR-Tst-S2C-01, E-CORRIDOR-Tst-S2C-02)** | SHOULD | The In-Vehicle Infotainment (IVI) or Electronic Control Units (ECUs) may be used for collecting GPS and driving behavior data. | Not applicable: by considering the requirements of the S2C and AT pilots, the privacy part has been included in the Driver DNA analytics (please see Section 2.1) and here we focus on the airport information kiosk | |

### 4.2.5. Plan for Testing and Final Maturation

Our current effort is oriented to the improvement of the interest-based service sharing analytics and to its development and customization for the AT pilot scenario. Moreover, we are working on the integration of the underlying 2PC-based component as service in the ASI subsystem to make the service reachable by the ASI gateway.

Then, we will progress on the maturation of the 2PC–based service sharing component to properly work as an ndroid application with the support of the CBMC-GC framework. In this way, this analytics will be able to work both at the user side and at the edge of the E-CORRIDOR framework.

# 5. Carbon Footprint Analytics – Task 7.4

In a world where ethical choices are on the rise, people are more aware of their impact relating to their choices, be it in travel or home life. To aid with this trend this CO2 calculator uses data collected by the EU and a blend of technologies designed in E-CORRIDOR to provide the information needed to make informed choices on carbon footprint in relation to travel plans.

## *5.1. CO2 analytics [E-CORRIDOR-IAI-CFA]*

### 5.1.1. Component Description

As user plans his/her trip, the main goal of this analytics is to provide information on CO2 usage on that trip so that he/she can make an informed decisions on the trip choices. Thanks to this information users can choose a route having a lower carbon footprint and help keeping CO2 emissions down. Presented with this information hopefully users will become more aware of their carbon footprint over their journeys and will end up choosing the greener option.

The CO2 calculator analytics is accessed using an API endpoint that provides the following information: vehicle type (bus, car, truck, etc.), vehicle maker, model, and year. Based on these values the component currently has multiple ways of calculating CO2 usage (g/km):

- CO2 Approximator: the CO2 calculator uses an approximate evaluation for CO2 footprint per passenger on a bus over a distance travelled.
- Database look-up: the CO2 usage is calculated using the EU provided dataset for passenger vehicles (available at: https://www.eea.europa.eu/data-and-maps/data/co2-cars-emission-20).

After the above calculation/lookup has been carried out the result is returned to the caller.

### 5.1.2. Workflow in Action

The CO2 calculator communicates over the IAI API using the /startanalytic endpoint. The list of the available parameters along with their description are reported in Figure 24.

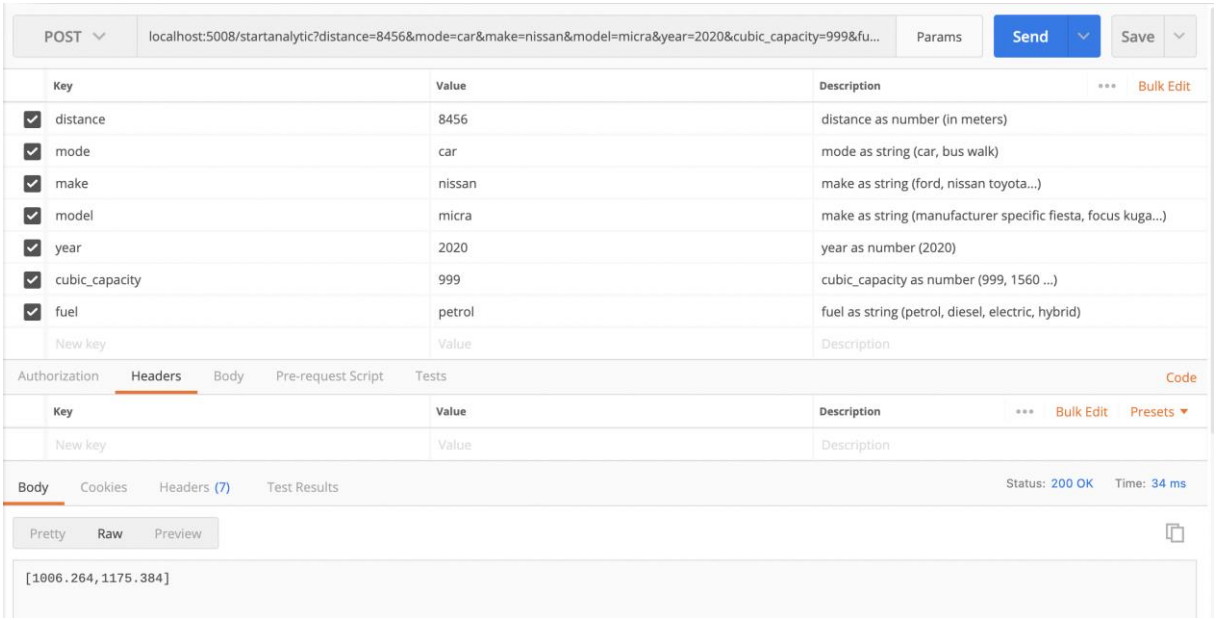


**Figure 24 The CO2 calculator API**

When the data are sent via the API, the parameters are parsed and checked for correctness, the calls are made to the database if needed and the response from the database is checked to ensure it validates the desired output. At this point the response is formed for the trip planner and it is then sent back. At present, the response is returned to the sender. In a future iteration of the component the CO2 calculator could store its response in the ISI, where the data could be used by multiple analytics e.g., for providing the end user with common CO2 usage for a planned route. Figure 25 represents the described workflow.

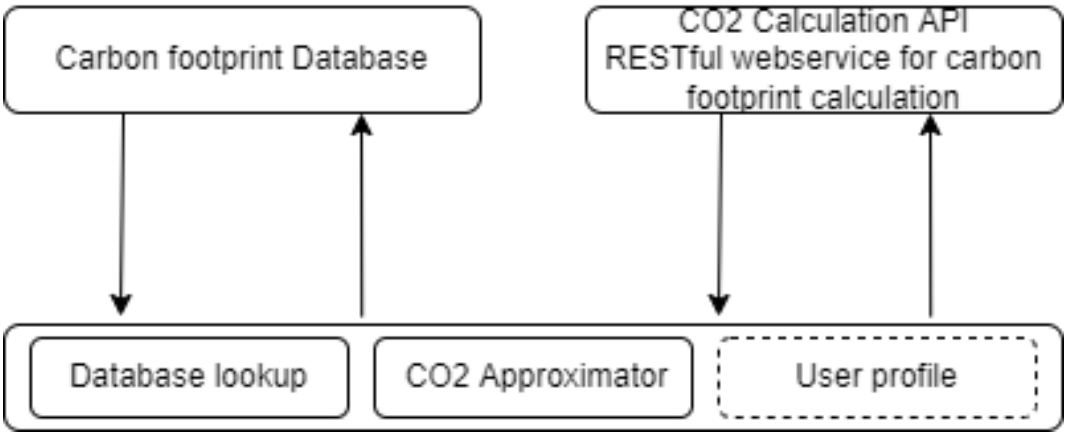


**Figure 25 CO2 analytics - API Flow Diagram**

When the end user plans a trip, the current version takes distance for a bus journey and calculates the CO2 usage over that journey. As a bus can carry many passengers, the calculations are shown on a per passenger basis, based on the most common bus types used for public transit. The calculator then returns this value to the multi-modal trip planner, where the value is shown on the itinerary. This allows the end user to choose a route that best suits their needs but also their ethical view towards their carbon footprint. Figure 26 shows how the data are presented to the end user.

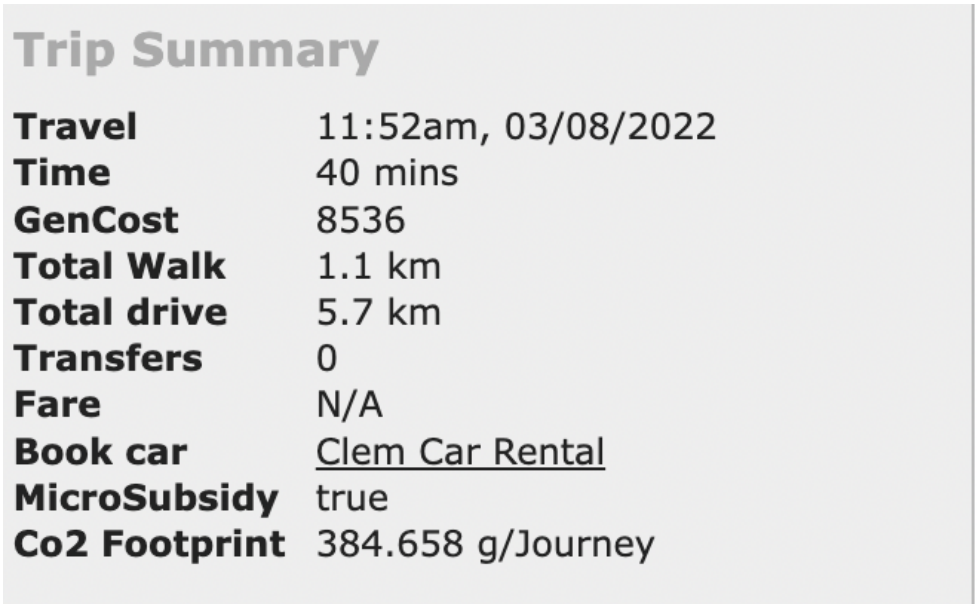


**Figure 26 CO2 calculation shown to end user**

### 5.1.3. Integration and Maturation Status of the First Release

Currently, we are adopting the EU environment agency CO2 emissions from new passenger cars database to compute the CO2 footprint. During the verification process we noticed that, on the supplied information, there were some ambiguities for some vehicles. Indeed, multiple entries in the database matching a given vehicle could correspond to a wide range of CO2 values. E.g., the query:

*distance=8456&mode=car&make=ford&model=focus&year=2020&cubic_capacity=999&fuel=petrol*

returns the results reported in Figure 28. Currently, our analytics return a range of CO2 values as further research will be needed to identify how to handle this correctly.

| Make | Commercial name | Category of the ... | Category of the ... | Total new registrations | Mass in running orde... | WLTP test... | Specific $CO_2$ E... | Specific $CO_2$ E... | Wheel base... | Axle width... | Electric range... | Axle width... | Fuel type | Fuel mode | Engine capacity... | Engine power... |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FORD | FOCUS C - MAX | | M1 | 1 | 1392 | | 129 | | 2640 | | | | DIESEL | M | 1560 | 80 |
| FORD | FOCUS C - MAX | | M1 | 1 | 1423 | | 154 | | 2640 | | | | DIESEL | M | 1560 | 80 |
| FORD | FOCUS C-MAX | | M1 | 1 | 1391 | | 127 | | 2640 | | | | DIESEL | M | 1560 | 66 |
| FORD | Focus | M1 | M1 | 1 | 1518 | | 113 | 133 | 2700 | 1581 | | 1576 | DIESEL | M | 1995 | 110 |
| FORD | Focus | M1 | M1 | 1 | 1383 | | 101 | 131 | 2700 | 1581 | | 1576 | PETROL | M | 999 | 92 |
| FORD | Focus | M1 | M1 | 1 | 1431 | | 135 | 166 | 2700 | 1581 | | 1576 | PETROL | M | 999 | 92 |
| FORD | Focus | M1 | M1 | 1 | 1395 | | 117 | 141 | 2700 | 1581 | | 1576 | PETROL | M | 999 | 92 |
| FORD | Focus | M1 | M1 | 1 | 1559 | | 111 | 136 | 2700 | 1581 | | 1576 | DIESEL | M | 1995 | 110 |
| FORD | Focus | M1 | M1 | 1 | 1395 | | 117 | 141 | 2700 | 1581 | | 1576 | PETROL | M | 999 | 92 |
| FORD | Focus | M1 | M1 | 1 | 1395 | | 117 | 142 | 2700 | 1581 | | 1576 | PETROL | M | 999 | 92 |

**Figure 28 An example where for the same vehicle multiple CO2 emissions are returned**

Moreover, the database has a rich set of parameters to identify a vehicle correctly (e.g., weight, fuel type, or presence of innovative technologies). As all these values might not be available to the calling services, in case of similar vehicles matching the request a range of CO2 consumption is returned.

The current version of the component runs standalone. During the next year the component will be fully integrated in the E-CORRIDOR framework and improve the API exposed for the end user. The table below summarizes status and plan.

| Feature | Current version | Final release |
|---|---|---|
| Standalone | ✓ | |
| Internal API | ✓ | |
| Primary integration into IAI | | ✓ |
| Data polling with intervals | | ✓ |
| End user API access | | ✓ |

### 5.1.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | **Status** | Rationale for the status |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-CFA-001 (E-CORRIDOR-DS-20, E-CORRIDOR-Tst-Int-S2C-02)** | SHOULD | The carbon footprint analytics tool should be able to pull data (such as carbon profiles of different vehicles) from external sources, with specified polling intervals. | In progress | For final integration there will be CO2 consumption data of different vehicles and then the analytics will calculate CO2 values for a route based on the vehicle selected |

### 5.1.5. Plan for Testing and Final Maturation

We are designing a set of test cases to validate that the data coming to the CO2 calculator and from the different external calculators are correct. We will then focus our efforts on the integration with the trip planner and the analytics toolbox.

The driving style data could be considered as another factor affecting the CO2 results. We will then evaluate the chance of exploiting the driving style built by the Driver DNA [E-CORRIDOR-IAI-SR] (that considers RPM, breaking, speed/speeding, etc.) to create a CO2 usage profile customized for each user.

# 6. Intrusion Detection Technologies – Task 7.5

To strengthen the cyber-security of the transportation service providers in ever more connected ecosystems, the E-CORRIDOR framework provides a set of solutions to prevent and detect real attacks. In such a way, mitigation strategies and solutions can be adopted, as well as targeted notifications can be fostered exploiting the collaborative analysis capabilities of the E-CORRIDOR framework.

This task proposes analytics to perform intrusion detection and prevention in network connected systems with a particular focus on the automotive domain.

## *6.1. Automotive Intrusion Detection [E-CORRIDOR-IAI-CANIDS]*

### 6.1.1. Component Description

The Automotive Intrusion Detection component aims at providing interfaces to both create and utilize classification models for automotive network data, such as recorded CAN bus communication. For example, we consider a car sharing provider with a large fleet of similar vehicles as a possible user of this service. The provided interfaces enable such a client to create a set of models based on the normal behavior of his individual vehicles and evaluate newly recorded driving sessions in a simple performant manner, to recognize anomalous behavior or contaminated systems early.

Currently, the intrusion detection service provides a manner to create and refine Neural Network-based models built on the Python 3 *Tensorflow* framework, from recorded CAN bus traffic. In its final iteration, we will provide a set of different classification mechanisms for several data types that are typically used in the automotive context, such as CAN bus or Automotive Ethernet.

Then, the system can use these trained models to evaluate recordings of arbitrary length on possible anomalies. All evaluation results will then be made available to the operator that initiated the analytics step in the form of a copy of the recording with each message flagged according to the evaluation result. In the next iteration this step will be refined to provide intrusion reports in different configurable formats.

The intrusion detection approach implemented by us has been developed over the course of our previous publications [19, 20, 21] on automotive intrusion detection methods. We have discussed the possible threats and attack surfaces of the automotive domain [22, 23] and evaluated the in-vehicle network regarding privacy-awareness [24] and trust [25]. Regarding the next iterations, we work towards integrating new communication protocols to be compatible with the Automotive Intrusion Detection analytics service and we have discussed SOME/IP regarding security [26]. In order to integrate the different approaches into a generic automotive architecture and to evaluate how they could successfully interact and cooperate with one another, we have designed several security patterns as a collaboration with different E-CORRIDOR partners [27].

### 6.1.2. Workflow in Action

As a prerequisite for the usage of the Automotive Intrusion Detection component we assume that clients are constantly recording and uploading CAN bus communications from one or many vehicles of their fleets to the E-CORRIDOR framework through the ISI subsystem.

Initially there will be no detection model for the respective vehicle present. Thus, after providing enough initial data from a small number of driving sessions, an initial model should be trained on the uploaded data. This requires either that previously anomalies have been observed and manually identified or synthetic anomalies are introduced to the clean recordings. Though the initial training on synthetic intrusions may impede the performance of the model at the beginning, we assume that this will be the common approach for initial training, because correctly labelled intrusion data for a specific vehicle are hard to create manually. For the initial training, the client then chooses a set of recordings including both normal driving behavior messages as well as (synthetic or real) anomalies or recorded intrusions and invokes an analytics process using the *train* parameter. The respective anonymized files will then be provided to the IDS detection system by the ISI subsystem through the Virtual Data Lake. Such data are preprocessed and used to train a first iteration of the Neural Network-based classification model specifically for the given vehicle type. The trained model will then be stored back in the ISI subsystem, and according to the specified DSA, made available to the clients for its use on other cars of their fleets.

For the basic use case of the Automotive Intrusion Detection System, we assume that the client has observed anomalous behaviors from one of his vehicles and has already performed the initial training for the respective vehicle. Therefore, an initial model capable of detecting basic anomalies is already available and stored within the E-CORRIDOR framework.

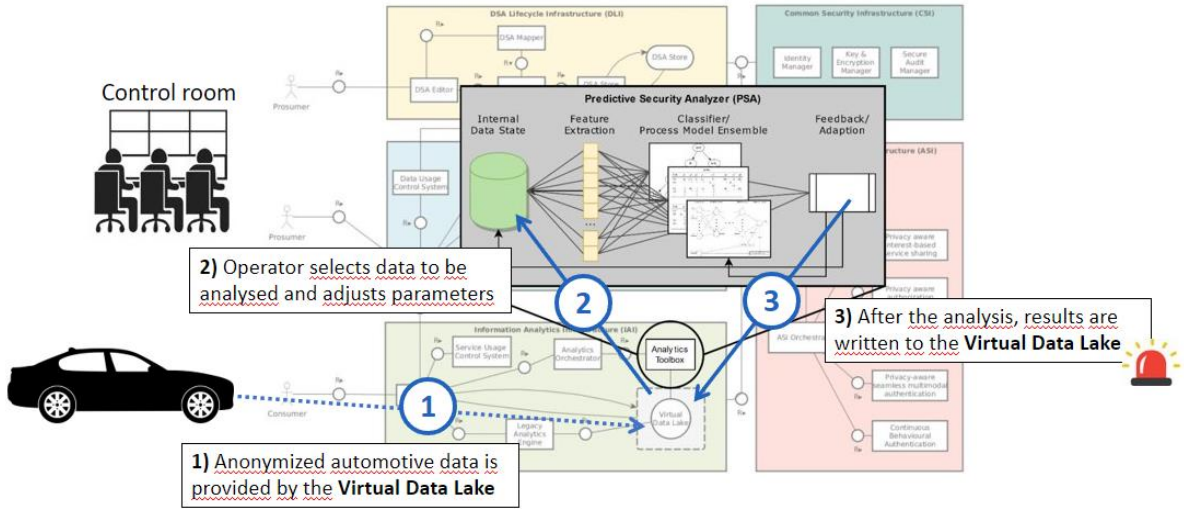


**Figure 29 Pictorial representation of the Automotive IDS scenario**

The client may then upload the potentially anomalous automotive data to the Virtual Data Lake for further processing. Next, to extract knowledge from the collected data, the client invokes the analytics process using the *classify* parameter, as well as the existing model file and file names for the recordings to classify. The Automotive IDS analytics service will then generate a labelled recording constituted by all the messages along with their classification results. In case the presence of any actual anomalous or malicious messages should be reported by the analytics, the client can then initialize further steps necessary to assure security and safety in accordance with the vehicle characteristics. Figure 29 represents the steps involved in the Automotive IDS scenario.

A further potential scenario foresees that the client is made aware of a new type of intrusion and receives an anonymized recording for such an intrusion through the ISI subsystem. The

client may then use such data to update the existing classification model for the respective vehicle by invoking the analytics using the *train* parameter, existing model, and recording of the newly observed intrusion.

### 6.1.3. Integration and Maturation Status of the First Release

The analytics service currently provides the operators with a message log containing messages with their respective classification results. An example excerpted from a recorded denial of service log file is shown in Figure 30.

```
time    ,ID    ,len ,p1   ,p2   ,p3   ,p4   ,p5   ,p6   ,p7   ,p8   ,type
...
0.7979 ,1087 ,   8 ,   0 , 64 , 96 ,255 ,126 ,203 ,   8 ,   0 , 1
0.7982 ,   0 ,   8 ,   0 ,   0 ,   0 ,   0 ,   0 ,   0 ,   0 ,   0 ,-1
0.7984 ,1088 ,   8 ,255 ,   0 ,   0 ,   0 ,255 ,203 ,   8 ,   0 , 1
0.7987 ,   0 ,   8 ,   0 ,   0 ,   0 ,   0 ,   0 ,   0 ,   0 ,   0 ,-1
0.799  ,   0 ,   8 ,   0 ,   0 ,   0 ,   0 ,   0 ,   0 ,   0 ,   0 ,-1
```

**Figure 30 Excerpt from a recorded denial of service log file example**

In this example the type of the message is determined through the classification of the intrusion detection service.

In the following iterations of our component, this log containing classification results will be replaced by alerts in the *Structured Threat Information Expression (STIX™)*. This format of alerts is more detailed containing more information on the intrusion. The following visualization in Figure 31 shows how such an intrusion report is structured.

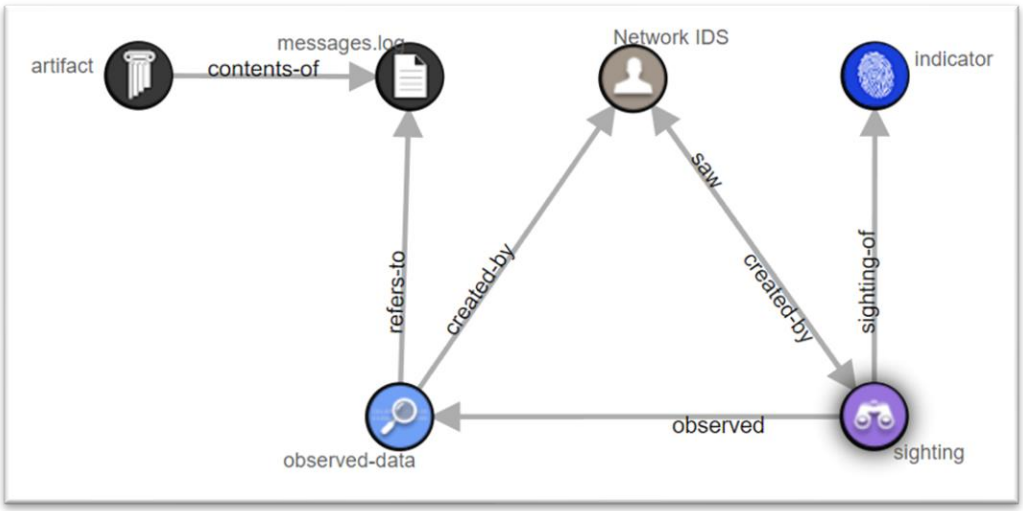


**Figure 31 Example of a STIX alert report**

Available features, integration and testing efforts as well as the ones planned for the final maturation of the component are summarized in the following table.

| Feature | Current version | Final release |
|---|---|---|
| Primary integration of intrusion detection system into IAI | ✓ | |
| Training of intrusion detection models using CAN bus communication recordings on request | ✓ | |
| Classification of recorded CAN bus traffic using Neural Network-based models | ✓ | |
| Adapt models to new situations by continuous training on new CAN bus recordings | ✓ | |
| Anonymize CAN bus recordings | | ✓ |
| Extended lab tests | | ✓ |
| Reporting of sightings using different intrusion reporting formats | | ✓ |
| Integration of other model types than Neural Networks | | ✓ |
| Integration of other types of data recordings than CAN bus | | ✓ |
| Model deployment to edge devices | | ✓ |

### 6.1.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | **Status** | Rationale for the status |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-CANIDS-01 (E-CORRIDOR-Tst-ISAC-01, E-CORRIDOR Ope-05)** | MUST | Emulate a working ECU inside the in-vehicle network to generate, collect, share, and analyze CAN bus data for S2C-UC-06, ISAC-UC-02, ISAC-UC-07. | To be deployed on the final release | The emulation of the ECU is working and being tested in our research environment and will be integrated in the following iterations of the analytics service. |
| **E-CORRIDOR-IAI-CANIDS-02 (E-CORRIDOR-Tst-S2C-02)** | SHOULD | Support device that is compatible with OBD II (or CAN BUS) for monitoring and sending GPS and driving behavior data. | To be deployed on the final release | Means of gathering GPS and driving data are being explored and currently tested. |
| **E-CORRIDOR-IAI-CANIDS-03 (E-CORRIDOR-Tst-ISAC-01, E-CORRIDOR-Tst-ISAC-02, E-** | SHOULD | Support Pilot ISAC Test Bed & Production Requirements | Complete | Initial version of the service can be used to create detection models and classify recorded data logs. |

| | | | | |
|---|---|---|---|---|
| **CORRIDOR-Tst-ISAC-03,**<br>**E-CORRIDOR-Tst-ISAC-04)** | | | | |
| **E-CORRIDOR-IAI-CANIDS-04 (E-CORRIDOR-Tst-Int-ISAC-02)** | SHOULD | Support an intrusion protection system able to authenticate the ECU in an intra-vehicle network when it aims at sending cross partition CAN frame | Complete | The service provides classification results that can be used to verify the integrity of messages. |
| **E-CORRIDOR-IAI-CANIDS-05 (E-CORRIDOR-Ope-02 edge)** | MUST | CAN IDS must work at the edge | To be deployed on the final release | Currently the service is working within the analytics cloud, but models trained there may be deployed in edge components. Means of deployment are currently being developed. |
| **E-CORRIDOR-IAI-CANIDS-06 (E-CORRIDOR-Ope-01 (both), E-CORRIDOR-Ope_03 (collaboratively in the cloud))** | COULD | CAN IDS could support deployment in cloud and edge or collaboratively in the cloud | Potentially deployed on the final release | Possible means of deployment and collaborative interaction are currently being explored. |
| **E-CORRIDOR-IAI-CANIDS-07 (E-CORRIDOR-DA-06,**<br>**E-CORRIDOR-DS-19 (push))** | SHOULD | CAN IDS should support intrusion detection reporting to the E-CORRIDOR cloud instance by means required by respective ISAC use cases. | Partially complete | Initial reporting formats are implemented, others will be added in later iterations of the service. |

### 6.1.5. Plan for Testing and Final Maturation

As shown on the Integration and Maturation Status overview (Sec. 6.1.3), we aim to make our intrusion detection analytics service capable of detecting anomalies not only using the currently

available Neural Networks-based classifier, but also several different approaches developed by us in several publications over the course of the E-CORRIDOR project.

Therefore, thanks to such novel approaches, in future iterations of the service we aim to implement also more lightweight classification methods that are designed to work at the performance level available in edge components for in-vehicle networks. This will also enable us to deploy the generated classification models on actual in-vehicle networks and edge components. To achieve this, we are also working on a workflow to train models within the cloud and then deploy them to the vehicles in the context of an E-CORRIDOR framework instance.

We have already designed several means for clients to anonymously store CAN bus data in a way that the origin of the recording cannot be traced, but the quality of the information stored within the log files is not compromised. As the next step we will refine these anonymization steps and work on integrating them into the E-CORRIDOR framework.

A further extension will include different formats of intrusion reporting in addition to the current proof-of-concept implementation, that creates a log file containing all classification results. For this we have already shown in the previously deliverable D7.1 the intrusion detection reporting format *Structured Threat Information Expression (STIX™)* which will be integrated in one of the next iterations of the service. Additionally, we are currently evaluating the AUTOSAR Intrusion Reporting System, for possible integration opportunities of our service.

Finally, we are currently working on enabling our Automotive IDS service to work not only with CAN bus data, but also with other data formats, such as Automotive Ethernet using SOME/IP or MQTT GPS and vehicle data.

## *6.2. Fully Homomorphic Encryption-based intrusion detection [E-CORRIDOR-IAI-FHEIDS]*

A network Intrusion Detection Systems (IDS) monitors and analyzes all the traffic from and to a given host with the purpose of identifying malicious packets. One of its components is constituted by the IP blacklist checker. The latter is used to detect and filter out connections to malicious or illegitimate IP address from a real-time database of IP addresses or domains known for being used to send spam, malicious or illegal content.

### 6.2.1. Component Description

To preserve the privacy of the user, this analytics adopt the Fully Homomorphic Encryption (FHE) technique to check the presence of any blacklisted IP address among the ones a host is trying to connect. It takes in input a Domain Name System blacklist (DNSBL) and a set of addresses. An alarm is raised if the IP addresses to which a host is trying to connect (or has connected in the past) is blacklisted.

The analytics is based on the same core components of the FHE-based driving license checker described in Section 4.1. The difference is constituted by the format adopted for the entries. Indeed, the IP blacklist checker works on IPv4 addresses stored as 4 bytes.

### 6.2.2. Workflow in Action

As the exposed API are the same discussed in Section 4.1.2, for the sake of brevines here we discuss only some additional functions we deem more relevant for this scenario.

The list of blacklisted IP addresses is inherently more dynamic. It means that new addresses may need to be included in the database as malicious actors move their attack resources over the web. On the other hand, this dynamism may bring the zealous inclusion of addresses that are instead genuine (e.g., because they have been victim of an attack). In such a case, owners will require an amend to the maintainers of these blacklists.

In the two above scenarios the *post-item* and *delete-item* can be used to keep the list updated without incurring in the burden or recreating the dataset. Figure 32 highlights with red boxes the above-mentioned functions.

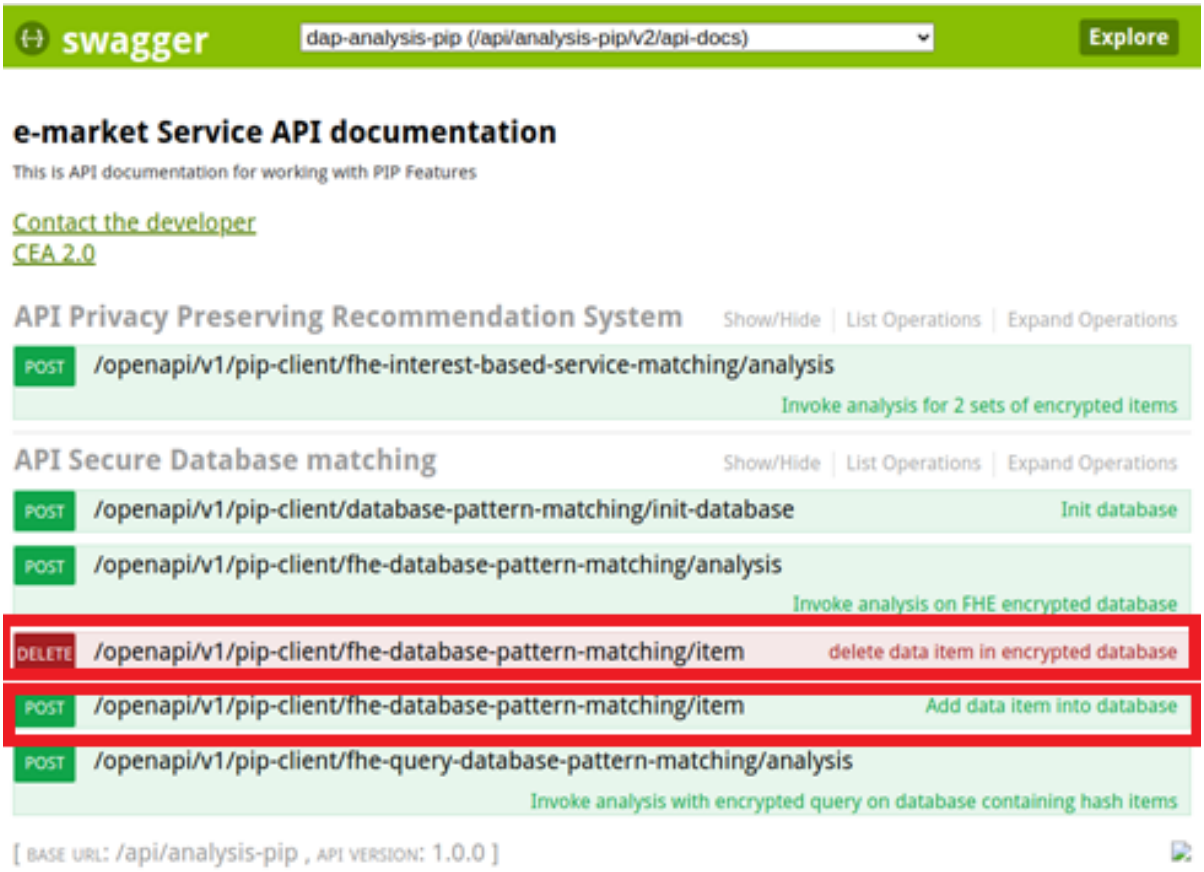


**Figure 32 Pattern Search Analysis API - add and remove from a database**

### 6.2.3. Integration and Maturation Status of the First Release

The analytics is fully functional, and its APIs are available for testing. Currently the component is not yet containerized and integrated in the E-CORRIDOR framework.

### 6.2.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | **Requirement description** | **Status** | **Rationale for the status** |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-FHEIDS-01 (E-CORRIDOR-Tst-ISAC-01, E-CORRIDOR Ope-05)** | MUST | Emulate a working ECU inside the in-vehicle network to generate, collect, share, and analyze CTI (Cyber threat intelligence), connection logs data for S2C-UC-06, ISAC-UC-02, ISAC-UC-07. | In progress | The component will be integrated in the multi-modal ISAC pilot to analyze network logs and protect against IP attacks. |
| **E-CORRIDOR-IAI-FHEIDS-02 (E-CORRIDOR-Tst-S2C-02)** | SHOULD | Support formats that are compatible with car data e.g., Automotive Ethernet using SOME/IP or MQTT GPS data. | Under evaluation | The component can be customized to work with entries expressed in different formats. |
| **E-CORRIDOR-IAI-FHEIDS-03 (E-CORRIDOR-Tst-ISAC-01, E-CORRIDOR-Tst-ISAC-02, E-CORRIDOR-Tst-ISAC-03, E-CORRIDOR-Tst-ISAC-04)** | SHOULD | Support Pilot ISAC Test Bed & Production Requirements | To be deployed on the final release | The FHE-based IP blacklist checker will be integrated as a tool in the multi-modal ISAC pilot |
| **E-CORRIDOR-IAI- FHEIDS - 04 (E-CORRIDOR-Tst-Int-ISAC-02)** | SHOULD | Support an intrusion protection system able to identify blacklisted IP addresses or spam content | Partially completed | The component already works with IP addresses and could be extended to check spam string in text messages. |
| **E-CORRIDOR-IAI- FHEIDS - 05 (E-CORRIDOR-Ope-02 edge)** | MUST | FHE IPS must work at the edge | Completed | The analytics is developed in Java and can running on any device supporting the Java Virtual Machine. |

| | | | | |
|---|---|---|---|---|
| **E-CORRIDOR-IAI- FHEIDS -06 (E-CORRIDOR-Ope-01 (both), E-CORRIDOR-Ope_03 (collaboratively in the cloud))** | COULD | FHE IPS analysis could support deployment in cloud and edge or collaboratively in the cloud | To be deployed on the final release | We will explore the use of a distributed infrastructure (see Section 4.1.5) where local and remote nodes work together to perform the analysis. |
| **E-CORRIDOR-IAI- FHEIDS -07 (E-CORRIDOR-DA-06, E-CORRIDOR-DS-19 (push))** | SHOULD | FHE IP checker should support intrusion detection reporting to the E-CORRIDOR cloud by means required by respective use cases. | To be deployed on the final release | Once integrated in the ISAC pilot, reports will be shared with the Cyber data analysis tools [E-CORRIDOR-IAI-CDA] (see Section 7.3). |

### 6.2.5. Plan for Testing and Final Maturation

The maturation plan follows the one of the main FHE pattern search component discussed in Sec 4.1.5 and not reported here for the sake of conciseness.

The tool will be integrated in the ISAC pilot as a privacy-preserving tool offered to any transportation service operators for protecting their Internet connected devices.

## *6.3. Intrusion Protection System – Earnest [E-CORRIDOR-IAI-CANIPS]*

### 6.3.1. Component Description

One of the requirements of upcoming regulation and standards on automotive security is the introduction of an Intrusion Protection System (IPS) on board of modern vehicle [28].

Aligned with such initiatives, we propose *Earnest*. It uses a challenge-based approach to filter all messages sent over the bus. Each time a given ECU (Electronic Control Unit) sends a CAN message, it is filtered by Earnest and a challenge is sent to the sending ECU. Both ECU and Earnest will perform the challenge and, in case the two challenges give the same result, Earnest will correctly forward the original message, otherwise the message will be discarded. Thus, Earnest has been designed and developed to detect and eventually protect onboard communications among ECU from Fuzzing and Replay attack that can be perpetrated by either a remote attacker or an altered ECU.

### 6.3.2. Workflow in Action

Earnest works according to the steps represented in Figure 8. When it receives a message from a non-enabled ECU, it creates a payload using the encode function. This payload contains the information, randomly generated by Earnest, necessary for the creation of a new payload on which to apply the challenge.

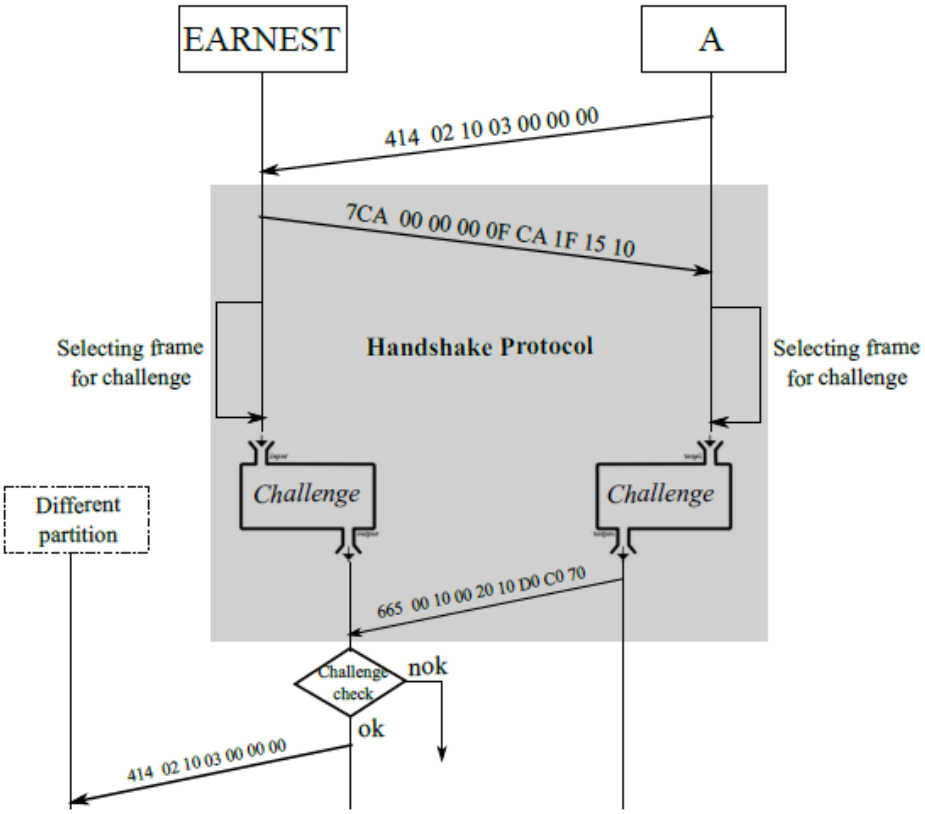


**Figure 33 Earnest Workflow**

Currently, EARNEST foresees ten challenges:

1. **Sum-To-Fifth** increases by one unit, in bytes, the element in position five.
2. **Reverse-It** reverses the order of the payload elements.
3. **Switch** exchanges the first element located in position 0 (zero) with the one in position ‘n-1’, where ‘n’ represents the maximum length of the payload.
4. **r-Rotate** rotates of three positions, intended as scrolling to the right, of all the elements of the payload.
5. **zero-Even** sets to 0 (zero) all bytes in even position.
6. **zero-Odd** sets to 0 (zero) all bytes in odd position.
7. **full-F** sets to ‘F’ bytes within position two and ‘n-1’, where ‘n’ represents the maximum length of the payload.
8. **full-Z** sets to 0 (zero) all bytes except the first and last.
9. **middle-F** sets to ‘F’ all bytes up to half the length of the payload.
10. **Odd-Even-Switch** switches bytes placed in even positions with the ones in odd positions.

The basic idea is to use the DBC, that is the database of CAN messages, as a shared secret between the in-vehicle ECU of a target vehicle and its carmaker.

The database is composed of the list of all messages regulating the vehicle's functionalities. Each message in the DBC has a pre-defined syntax: it is identified with a *BO_id* (data object identifier) and composed of several signals, one for each sub-functionalities managed by the message.

Based on these notions, Earnest prepares the challenge payload as follows:

1. randomly selects one of the 10 challenges
2. randomly selects a BO_id belonging to the DBC
3. selects a signal within the BO_id
4. assigns a random value (between the minimum and maximum value) to the selected signal
5. assigns the value (maximum + minimum)/2 to all the other signals in the BO_id

The resulting message has the structure reported in Figure 34.

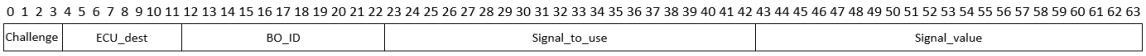


**Figure 34 Challenge Payload**

The first 4 bits identify the challenge randomly chosen by Earnest, then 8 bits identify the challenged ECU, next 11 bits the BO_id of the message, 20 bits for the signal to use and the rest of the 64bit is used for the value of the signal.

### 6.3.3. Integration and Maturation Status of the First Release

The IPS is currently at the level of design and prototypal implementation on Raspberry Pi4. It is not yet integrated into the E-CORRIDOR framework.

We implemented the workflow described above and tested our implementation to evaluate its performance in our development environments (see Figure 35).

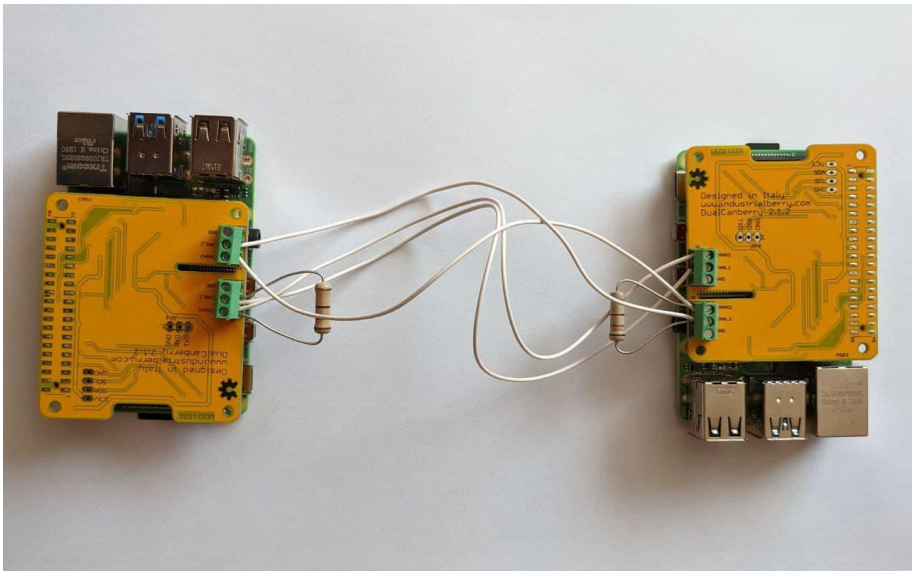

**Figure 35 Development environment**

The performance test with two ECUs, named Alice and Bob, consists of having both Alice and Bob perform 20 challenges at the same time. However, in a real setting the two ECUs, will not perform 20 challenges because after having passed a challenge they are "enabled" for "0.003" seconds (i.e., their messages will be automatically forwarded by Earnest without the need to perform any further challenge). The choice of the "enabling" time is justified by the need of simulating a sort of "dynamism": with higher "enabling" times, the ECUs would have sustained only one challenge out of twenty and consequently this test would have been not very useful. The idea is to reduce the number of challenges without causing any loss in terms of security allowing unauthorized ECU to transmit cross-partition frames.

To evaluate the impact of the introduction of Earnest into an in-vehicle network we performed some tests considering different bitrate and different number of exchanged messages. The average (AVG) we obtain is reported as last column of Figure 36.

| Bitrate | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | AVG |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1000 kb/s | 0,0483s | 0,0618s | 0,0597s | 0,0378s | 0,0614s | 0,0297s | 0,0639s | 0,0358s | 0,0439s | 0,0281s | 0,0470s |
| 500 kb/s | 0,0592s | 0,0717s | 0,0544s | 0,0615s | 0,056s | 0,0665s | 0,045s | 0,0723s | 0,053s | 0,0671s | 0,0608s |
| 100 kb/s | 0,2327s | 0,2192s | 0,2221s | 0,2183s | 0,2187s | 0,2372s | 0,2150s | 0,2235s | 0,2246s | 0,2227s | 0,2234s |

**Figure 36 Performance evaluation of the Earnest IPS**

### 6.3.4. Requirements Traceability Matrix

| ID in D7.1 (& ref to D5.1) | Priority | Requirement description | **Status** | **Rationale for the status** |
|---|---|---|---|---|
| **E-CORRIDOR-IAI-CANIPS-01 (E-CORRIDOR-Tst-ISAC-01, E-CORRIDOR Ope-05)** | MUST | Emulate a working ECU inside the in-vehicle network to generate, collect, share, and analyze CAN bus data for S2C-UC-06, ISAC-UC-02, ISAC-UC-07. | Partially completed | We implemented a protype of Earnest on two Raspberry Pi4 enhanced with CAN shields to simulate the CAN communication |
| **E-CORRIDOR-IAI-CANIPS-02 (E-CORRIDOR-Tst-S2C-02)** | SHOULD | Support device that is compatible with OBD-II (or CAN BUS) for monitoring and sending GPS and driving behavior data. | Not fulfilled yet | We do not deploy our prototype on a real ECU |
| **E-CORRIDOR-IAI-CANIPS-03 (E-CORRIDOR-Tst-ISAC-01, E-CORRIDOR-Tst-ISAC-02,** | SHOULD | Support Pilot ISAC Test Bed & Production Requirements | Not fulfilled yet | It will be evaluated for the final maturation to share vulnerabilities with the ISAC. |

| | | | | |
|---|---|---|---|---|
| **E-CORRIDOR-Tst-ISAC-03, E-CORRIDOR-Tst-ISAC-04)** | | | | |
| **E-CORRIDOR-IAI-CANIPS-04 (E-CORRIDOR-Tst-Int-ISAC-02)** | SHOULD | Support an intrusion protection system able to authenticate the ECU in an intra-vehicle network when it aims at sending cross partition CAN frame | Partially completed | We simulate a real intra-vehicle with two Raspberry Pi4 enhanced with CAN shields to simulate the CAN communication |
| **E-CORRIDOR-IAI-CANIPS-05 (E-CORRIDOR-Ope-02 edge)** | MUST | CAN IPS must work at the edge | Expected to be deployed on the final release | The collected data will be shared via the edge layer of the infrastructure |
| **E-CORRIDOR-IAI-CANIPS-06 (E-CORRIDOR-Ope-01 (both),** **E-CORRIDOR-Ope_03 (collaboratively in the cloud))** | COULD | CAN IPS could support deployment in cloud and edge or collaboratively in the cloud | Expected to be deployed on the final release | Earnest will be deployed into the vehicle |
| **E-CORRIDOR-IAI-CANIPS-07 (E-CORRIDOR-DA-06, E-CORRIDOR-DS-19 (push))** | SHOULD | CAN IPS should support intrusion detection reporting E-CORRIDOR cloud by means required by respective use cases. | Expected to be deployed on the final release | It will be integrated with the ISAC pilot in order to share a database of vulnerabilities. |

### 6.3.5. Plan for Testing and Final Maturation

As future step, in agreement with the evolution of intra-vehicle communications that are migrating to Automotive Ethernet (AE), we aim to extend and customize Earnest to deal also with AE. In addition, all the collected information about the detected and prevented intrusion will be shared with the ISAC Pilot to build a database of vulnerabilities related to the automotive domain.

# 7. Pilot specific analytics

The ISAC offers pilot-specific analytics based on the data gathered from public sources (e.g., security databases and online information) and the results produced by the security analytics provided by the other pilots. The analytics toolbox contained in the IAI subsystem instantiated in the ISAC offers three types of cyber-data analytics, that can be grouped into the same number of categories: (i) label assignment, (ii) visualization, and (iii) analysis. The data gathered from public sources are labelled and part of this information is further exploited for intuitive visualization and aggregation. Moreover, the ISAC offers interactive tools for uploading and analyzing data to get personalized results on the user's data. The data is securely uploaded on the ISAC portal through the ISI subsystem.

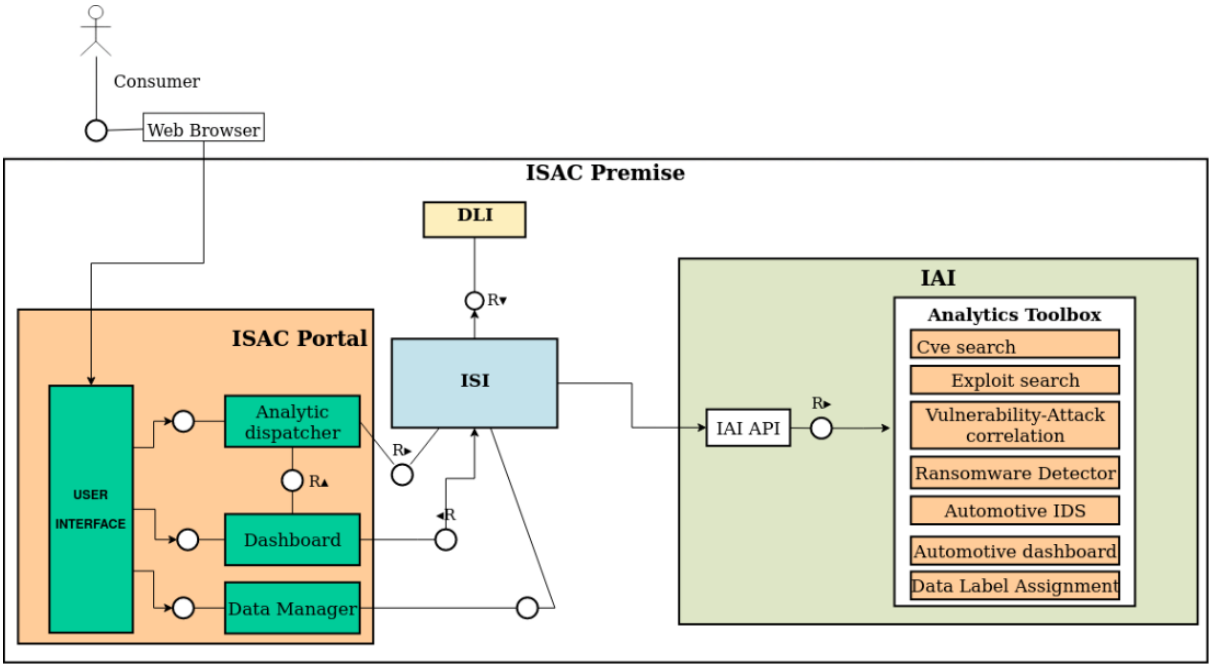


**Figure 26 Pilot-specific analytics of the ISAC pilot**

## *7.1. Cyber data label assignment [E-CORRIDOR-IAI-CDLA]*

This analytics is exploited as the main element for the ISAC cyber-threat notification workflow. The ISAC is bound to collect a huge amount of information from various providers, and at the same time, it must redistribute the collected data timely to interested stakeholders. Selecting the right information to be sent to the right partner is of high importance. In fact, failing to provide information on a relevant vulnerability might increase the exposure time of the stakeholder to cyber-attacks. On the other hand, flooding the stakeholder with non-pertinent information increases the processing times and might result in a lowered attention to actual security threats. This analytics exploits text analysis using Natural Language Processing techniques and clustering to separate the received and computed information into separate topics to which the stakeholders can subscribe. In such a way, they will be automatically notified with information that is relevant to them.

## *7.2. Cyber data visualization tools [E-CORRIDOR-IAI-CDV]*

The cyber data visualization tools are a set of macro analytics used to present and speed up the information retrieval process. It is composed of several visualization analytics tools, described in the following.

- CVE search: This service offers the possibility of searching public domain information related to known security hardware and software vulnerabilities. This service provides a general description of the vulnerabilities reporting the publishing date, a short description, the CVE score, and the impact on integrity, confidentiality, and availability properties. The consumer can research the vulnerabilities specifying an interval time or a specific keyword. Data exploited: Public vulnerability database collected from NVD (https://nvd.nist.gov).
- Exploit search: This analytics offers the possibility of searching information about the exploits performing research by date or keyword. This analytics shows date, description, and the specific platform on which the exploit is effective. Data exploited: Public exploits information collected from *exploit-db* (https://www.exploit-db.com/).
- Automotive intrusion detection visualization: This analytics is based on the automotive intrusion detection system analysis results explained in Section 6.1. (E-CORRIDOR-IAI-CANIDS). The produced result is the classification of the CAN bus messages and the potential detection of an anomaly within the vehicle's data. Such information shared by multiple vehicles with the ISAC is exploited to create a graphical representation of the intrusion information, incident, related types, and incident classes of a single vehicle or create a visualization of the correlation of the incident detected in different vehicles.
- *Driver risk behavior map*: The analytics exploits the data provided by the results of the Driver DNA analytics (E-CORRIDOR-IAI-SR). As explained in Section 2.1, the analytics provides the driver risk profile, e.g., more aggressive, speeding more frequently. Correlating these results with the geo-localization of the vehicles, the ISAC analytics can produce a map of the city reporting the average driver behavior in each road section. In such a way, it is possible to highlight the most dangerous section of the city and increase the drivers' awareness.
- Carmaker data usage: This dashboard visualizes information on which data categories each carmaker asks and declares to collect, and also information on the privacy policies enforced by the carmakers.

## *7.3. Cyber data analysis tools [E-CORRIDOR-IAI-CDA]*

While the tools above provide visualization on public/aggregated data, more tools are provided by the ISAC that perform analytics on specific data uploaded by the user/company that have access to the platform. These tools exploit the ISI and IAI subsystems to ensure privacy and confidentiality on the data and the obtained results.

The tools are described as follows:

- Vulnerability attack correlation: This analysis allows the consumer to explore the interconnection between vulnerabilities and attack patterns of specific known software or hardware. Additionally, the analytics provides recommendations for attacks and vulnerability mitigation.

- Ransomware Detector: This analysis identifies typical ransomware behaviors (such as file ciphering) into uploaded data. Differently from signature-based anti-malware, this tool is able to identify also malware whose signature is not available yet; it performs a static analysis exploiting information on the frequency of OpCode extracted from the code of the analyzed application.
- Malware Detector: The analysis detects malicious signatures in the analyzed files by scanning it with many different commercial anti-malware. The tool returns a risk index, presented as the probability that the analyzed file is malicious.
- Vulnerability Scanner: This analysis, available only for registered and certified users able to prove ownership of a virtual resource, offers a platform that allows to scan and check for any performance and security problem present on the user's systems/services. In detail, the platform offers check such as Real time controls, Periodic checks, Generation of reports and Analysis on IPv4 and IPv6.

While planning an itinerary, users are prompted with a set of options matching their travel preferences. To provide a richer set of information to the users, the S2C pilot offers the interaction with the micro-subsidies analytics.

## *7.4. Socio-geographic micro-subsidies analytics [E-CORRIDOR-IAI-MSA]*

Subsidies can play a key role in fostering the adoption of public transportation services and thus contributing to the CO2 reduction. Usually, subsidies are received by the mass transit or collective transportation operators from the public transport administration to lower the price paid by the users to access their services (e.g., according to category or geography). Such a kind of uniform concession is however not efficient as it does not target the population and does not directly consider the actual use of the services.

Micro-subsidies are a novel approach in managing the subsidies. It foresees narrowly defined categories of users (modelled according to characteristics such as age, income, disability, employment category), and features of the journey (e.g., mode of transport, type of motorization, time, location). In addition to reaching the same goals of the traditional subsidies, the micro-subsidies can also provide opportunities to better control the demand to public transportation services according to the time of the day or the geographical location.

This analytics receives in input the characteristics of the trip options matching the user preferences as entered in the trip planner analytics E-CORRIDOR-IAI-MMIP, and the user information contained into the eWallet profile (saved in the ISI and referenced through a universally unique identifier, UUID). Then, the micro-subsidies engine computes the eligibility to incentives and the corresponding amount. The information about the micro-subsidies is then displayed next to each of the available trip options. In such a manner the user can perform a more informed decision on his/her travel choice.

# 8. Contribution of the Data Analytics Techniques to the Pilot Requirements

The data analytics available in the toolbox have been designed and customized according to the pilots' requirements and constraints (ref. Obj. 4). This approach aims at fostering the prototype demonstration of the analytics in pre-operational/relevant environments and a potential generation of products for the same pilots (ref. Obj. 6).

The following Table 2 summarizes the different relevant scenarios presented by the pilots through the block diagrams and the role covered by the analytics in achieving the proposed goals (please see D2.2 "Design and Architecture for the Airport and Train (AT) pilot", D3.2 "Design and Architecture for the S2C Pilot" and D4.2 "ISAC Pilot Architecture").

**Table 2 Contribution of the analytics to the pilots' block diagrams**

| Pilot | Block Diagram | Related Use Cases | Required Analytics | Contribution of the Analytics to the Block Diagram Goals |
|---|---|---|---|---|
| Airport- train (AT) | AT-BD-01 PRM Passenger Assistance | AT-UC-01, AT-UC-06, AT-UC-14 | E-CORRIDOR-IAI-PBI, E-CORRIDOR-IAI-FR, E-CORRIDOR-IAI-PL | To properly support People with Reduced Mobility (PRM), the analysis of camera feeds to understand the status of the train/station (e.g., crowded), the localization of assistive devices (e.g., wheelchair) and of the available assistants from each mode of transportation are required. Moreover, by adopting a biometric-based authentication, the handling of the travel documents may be simplified. |
| | AT-BD-02 Multi-Biometric Passenger Authentication and Baggage Monitoring | AT-UC-02, AT-UC-03, AT-UC-04, AT-UC-05, AT-UC-06, AT-UC-13 | E-CORRIDOR-IAI-FR, E-CORRIDOR-IAI-AR, E-CORRIDOR-IAI-GA, E-CORRIDOR-IAI-PL, E-CORRIDOR-IAI-PBI | To provide a robust and seamless authentication of the passengers, data from multi biometric-based sensors are advocated. Analysis of the camera feeds can be useful in linking passengers (perhaps travelling under the same Passenger Name Record, PNR) with their baggage and identify any potential anomaly. |
| | AT-BD-03 Frictionless Access to Multi-Modal Services | AT-UC-06, AT-UC-07, AT-UC-08 | E-CORRIDOR-IAI-GA, E-CORRIDOR-IAI-PL | To support a frictionless access to multi-modal services a wide set of authentication solutions (described in Sec 2) are combined (see activities in WP8). An enhanced passenger experience may be reached thanks to the adoption of BYOD (Bring Your Own Device) solutions. |

| | | | | |
|---|---|---|---|---|
| | AT-BD-04 Controlled Data Sharing for Service Prediction, Optimization and Security | AT-UC-09, AT-UC-10, AT-UC-11, AT-UC-12, AT-UC-13 | E-CORRIDOR-IAI-FHEC, E-CORRIDOR-IAI-FHEIDS, E-CORRIDOR-IAI-PL, E-CORRIDOR-IAI-PBI, E-CORRIDOR-IAI-AR | A multitude of data are collected by cyber and physical sensors deployed in the transportation domains. By anonymizing (through Data Manipulation Operation, DMO), sharing in a controlled fashion (through Data Sharing Agreements, DSAs) and analyzing such data it is possible to extract insights concerning service performance and identify security threats. |
| Smart cities and car sharing (S2C) | S2C-BD-01 Sign-in | S2C-UC-01 | This block diagrams refers to the eWallet-based registration of the traveler's information and therefore requires only the ISI and ASI services. | |
| | S2C-BD-02 Micro-subsidies | S2C-UC-02 | E-CORRIDOR-IAI-MSA | Based on the trip request characteristics (e.g., origin, destination, time of departure and arrival, modes of transportation) - expressed in the itinerary planning E-CORRIDOR-IAI-MMIP - and the user profile saved in the eWallet (including age, profession, address, etc.) the eligibility to receive a micro subsidy and the corresponding amount are calculated. The result is returned to the itinerary planning tool. |
| | S2C-BD-03 Trip planning and carbon footprint analytics | S2C-UC-03 | E-CORRIDOR-IAI-CFA, E-CORRIDOR-IAI-MMIP | Allows the user (with an eWallet profile) to search for a multimodal trip taking into consideration trip preferences and the estimated carbon footprint for the suggested trips. Furthermore, it is connected to the socio-geographic micro-subsidies analytics. |
| | S2C-BD-05[1] Security analytics services | S2C-UC-06 | E-CORRIDOR-IAI-CANIPS, E-CORRIDOR-IAI-FHEIDS, E-CORRIDOR-IAI-CANIDS | Intrusions to the car system, the IoT network running on the car fleet, or the infrastructure are detected and notified. |
| | S2C-BD-06 Privacy-aware | S2C-UC-07 | E-CORRIDOR-IAI-FHEC | The existence of a driving license number is checked for |

[1] The submitted D3.2 contains a typo in numbering the block diagrams, skipping the SC-BD-04. To be consisted with D3.2, the same numbering is preserved here.

| | | | | |
|---|---|---|---|---|
| | interest-based sharing | | | validity in a privacy-preserving way against multiple lists. |
| | S2C-BD-07 Driving behavior recognition | S2C-UC-08 | E-CORRIDOR-IAI-SR | Driving style is computed and compared against a population of other drivers. Suggestions can be provided to the drivers (e.g., if a route is more dangerous than others). |
| Multi-Modal Transportation Information Sharing and Analysis Center (ISAC) | ISAC-DB-01 Collecting public data | ISAC-UC-01 | E-CORRIDOR-IAI-CDLA | Gathering cyber-threat information from several sources increases the accuracy of a security system by predicting (emerging threats) and timely reacting (by learning from previous events) to any attack. |
| | ISAC-BD-02 Collecting private data | ISAC-UC-02, ISAC-UC-03, ISAC-UC-06 | E-CORRIDOR-IAI-CDV | Data coming from users and transportation operators play a key role in identifying and preventing threats. The data are converted and shared with other stakeholders according to strict security and access policies set by the data owner. |
| | ISAC-BD-03 Running analytics and visualization | ISAC-UC-04, ISAC-UC-07, ISAC-UC-08 | E-CORRIDOR-IAI-CDV<br>E-CORRIDOR-IAI-CDA | Data can be exploited by running analytics and visualizing results in terms of graphs, statistics, and data aggregates. The analysis on data can discover new threats and vulnerabilities on the system, while the visualization tools help to exploit this information by providing overview on the results. |
| | ISAC-BD-04 Cyber-threat notification | ISAC-UC-05 | E-CORRIDOR-IAI-CDA | The sharing of the analysis results allows to implement a notification service able to inform each transportation service provider about new threats and vulnerabilities discovered, along with countermeasures to mitigate a potential or an ongoing attack. |

As summarized in the above Table 2, analytics in the toolbox cover a critical role in satisfying the pilot requirements. When the analytics are used along with all the other features provided by the E-CORRIDOR framework (e.g., data sharing and advanced authentication) these building blocks can achieve their full potential.

# 9. Contributions to the E-CORRIDOR objectives at M24

The data analytics discussed in the previous sections contribute to the fulfillment of some of the E-CORRIDOR objectives (Obj.), namely:

- Objective 2: E-CORRIDOR will define edge enabled data analytics and prediction services in a collaborative, distributed and confidential way
- Objective 3: E-CORRIDOR will define a secure and robust platform in holistic manner to keep the communication platform safe from cyber-attacks and ensure service continuity
- Objective 4: E-CORRIDOR will improve, mature, and integrate several existing tools provided by E-CORRIDOR partners and will tailor those to the specific needs of the E-CORRIDOR platform and Pilots
- Objective 5: E-CORRIDOR will provide mechanisms for seamless access to multimodal transport
- Objective 6: the framework and the services developed will be used to deliver three pilot products

For each of the above objectives, Table 3 summarizes how data analytics and WP7 activities contribute, the goals reached at the end of the second year and the ongoing effort.

**Table 3 Contribution of the WP7 activities to the project goals at M24**

| Obj. | Data analytics contribution | Reached goals | Next steps |
|---|---|---|---|
| Objective 2 | Prosumers can share data in a controlled and privacy-aware fashion by specifying appropriate DSA. Analytics in the IAI toolbox are edge-enabled and can then exploit the information sharing capability of the E-CORRIDOR framework. | By considering performance requirements and resource constraints expressed by the pilots, analytics run on lightweight devices, smartphones, the cloud or in a hybrid edge-cloud fashion according to the scenario. | Finalize the integration of all the data analytics in the E-CORRIDOR framework in such a way that the analytics toolbox can exploit the data shared through the ISI subsystem. |
| Objective 3 | Analytics provide identity, privacy, and security services. In some cases, these are also used as building blocks for defining more complex cyber-security services performed by the ASI subsystem. | Three security analytics such as IP blacklist checker, intrusion detection and intrusion prevention systems have been designed and developed. | Increase the maturity of the tools and evaluate their use in the pilots' infrastructure to strengthen their cyber-security |

| | | | |
|---|---|---|---|
| **Objective 4** | Data analytics arise from research activities of several partners and input collected by the pilots. The goal is to increase the maturity of such technologies. | Initial solutions already available by the technology providers have been customized to cover a wide set of scenarios presented by the pilots (see Section 8). So far, two data analytics have been fully integrated in the E-CORRIDOR framework and their entire workflow has been tested. Other two analytics are currently under integration. | Test, validate and refine the analytics in realistic environments provided by the pilots. Collect feedback from the pilots, experts in the multi-modal transportation domain. |
| **Objective 5** | Analytics in the IAI toolbox offers a set of passenger and driver identification and authentication capabilities targeting the peculiarity of multimodal transportation environments | A set of identification and authentication analytics has been built (see Section 2). Analytics exploit a variety of sensors and authentication procedures to accommodate a rich set of user preferences and security requirements. | Exploit the available analytics to build more advanced services (in the ASI subsystem) to allow an enhanced user experience and access to services provided by the transportation service providers. |
| **Objective 6** | Collaboration between technology providers and pilots to understand the criticalities of the multimodal transport and design solutions targeting the needs of their environments. | The analytics have been designed through a strong collaboration between technology and domain experts. Results of this collaboration have determined the match between the expressed scenarios and the analytics (reported in Section 8). | Connect data sources and solutions already available in the pilots' environment. Test and evaluate the benefit of the E-CORRIDOR framework when the analytics are connected to a realistic pilot ecosystem to further mature the products. |

# 10. Conclusion

A wide set of analytics have been deemed relevant by the multi-modal pilots to successfully fulfil their scenarios. This document reports on the status of the data analytics components developed in E-CORRIDOR along some initial tests. With respect to the analytics originally identified at M12 and reported in D7.1, a few components have been further customized to accommodate emerging considerations from the pilots (e.g., privacy concerns). For some of the technologies, another pilot has expressed interest in evaluating a solution targeting its own scenarios (this is the case of the interest-based services at the airport information kiosks).

The relevance of the proposed analytics in achieving a frictionless passenger experience and improving the cyber-security toward a really integrated pan-European multi-modal transportation environment has been highlighted in Section 8. Instead, the contributions to the E-CORRIDOR objectives, with completed tasks and next steps summarize the current progress and show a clear path for the ongoing efforts related to the WP7 activities.

All in all, the analytics have a various degree of maturation also considering that some partners preferred to adopt a more agile approach (starting the tests earlier with fewer features) whereas other adopted a more traditional approach (building the bulk of the features before intensifying the experiments). Currently, tests have been mainly focused on labs, and activities are on track as all the analytics already run as stand-alone components. About 50% of the requirements marked as “must” concerning development and integration efforts have been fulfilled. For the “should” and “could” requirements the competition is of about 30%. For the remaining there is no identified risk, and these requirements are currently partially fulfilled or expected to be deployed in the next months. Moreover, results for some of the analytics have been shared with the security community and/or published in international conferences and industrial events.

The next months will be devoted, in addition to the finalization of the analytics solutions, to their integration in the E-CORRIDOR framework and to validation and final maturation. In collaboration with WP6, a few components have been already fully integrated in the analytics toolbox and their whole workflow tested. On the other hand, the experienced collaboration and extensive discussions between partners should assist the maturation of all the analytics.

# 11. References


[1] G. Costantino, F. Martinelli, P. Santi and I. Matteucci, "A Privacy-Preserving Infrastructure for Driver's Reputation Aware Automotive Services," in *Proceedings Workshop on Socio-Technical Aspects in Security (STAST)*, Luxembourg City, Luxembourg, 2019.

[2] E-CORRIDOR, "D7.1: Data Analytics Techniques Requirements and Architecture," 2021.

[3] A. Bochkovskiy, C.-Y. Wang and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv e-prints*, 2020.

[4] N. Wojke, A. Bewley and D. Paulus, "Simple Online and Realtime Tracking with a Deep Association Metric," *IEEE International Conference on Image Processing (ICIP)*, 2017.

[5] G. Giorgi, A. Saracino and F. Martinelli, "Using Recurrent Neural Networks for Continuous Authentication through Gait Analysis," *Pattern Recognition Letters*, vol. 147, pp. 157-163, 2021.

[6] CNR IIT Cybersecurity, "GaitAuth," [Online]. Available: https://github.com/iitcybersecurity/ActivityRecognition/tree/gaitauth.

[7] S. Z. Li and A. K. Jain, Handbook of face recognition, Springer, 2011.

[8] A. Geitgey, "Face Recognition," April 2022. [Online]. Available: https://pypi.org/project/face-recognition/.

[9] H.-W. Ng and S. Winkler, "A data-driven approach to cleaning large face datasets," *IEEE international conference on image processing (ICIP)*, pp. 343-347, 2014.

[10] O. M. Parkhi, A. Vedaldi and A. Zisserman, "Deep Face Recognition," *BMVC*, 2015.

[11] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, 2016.

[12] T. Drutarovsky and A. Fogelton, "Eye blink detection using variance of motion vectors," *European conference on computer vision*, pp. 436-448, 2014.

[13] G. Pan, L. Sun, Z. Wu and S. Lao, "Eyeblink-based anti-spoofing in face recognition from a generic webcamera," *IEEE 11th international conference on computer vision*, pp. 1--8, 2007.

[14] "Pose Mediapipe," April 2022. [Online]. Available: https://play.google.com/store/apps/details?id=com.farproc.wifi.analyzer&hl=en&gl=US.

[15] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, USA: Prentice Hall Press, 2009.

[16] "Rideal the platform for rider incentive programs," Mobi, [Online]. Available: https://rideal.mobi/. [Accessed 17 May 2022].

[17] D. Kroening and E. Clarke, *CBMC Bounded Model Checker - https://www.cprover.org/cbmc/*, Carnegie Mellon, 2016.

[18] S. G. Akl and P. D. Taylor, "Cryptographic Solution to a Multi Level Security Problem," in *Advances in Cryptology*, 1982.

[19] Y. Chevalier, F. Fenzl, M. Kolomeets, R. Rieke, A. Chechulin and C. Kraus, "Cyberattack detection in vehicles using characteristic functions, artificial neural networks, and visual analysis," *Informatics and Automation*, vol. 20, p. 845–868, 8 2021.

[20] F. Fenzl, R. Rieke and A. Dominik, "In-vehicle detection of targeted CAN bus attacks," in *The 16th International Conference on Availability, Reliability and Security*, 2021.

[21] L. Buschlinger, R. Rieke, S. Sarda and C. Kraus, "Decision Tree-Based Rule Derivation for Intrusion Detection in Safety-Critical Automotive Systems," in *2022 30th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, 2022.

[22] C. Plappert, D. Zelle, H. Gadacz, R. Rieke, D. Scheuermann and C. Kraus, "Attack Surface Assessment for Cybersecurity Engineering in the Automotive Domain," in *2021 29th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*, 2021.

[23] D. Zelle, C. Plappert, R. Rieke, D. Scheuermann and C. Krauß, "ThreatSurf: A method for automated Threat Surface assessment in automotive cybersecurity engineering," *Microprocessors and Microsystems*, vol. 90, p. 104461, 4 2022.

[24] C. Plappert, J. Stancke and L. Jager, "Towards a Privacy-Aware Electric Vehicle Architecture," in *2022 30th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, 2022.

[25] C. Plappert, L. Jäger and A. Fuchs, "Secure Role and Rights Management for Automotive Access and Feature Activation," in *Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security*, 2021.

[26] D. Zelle, T. Lauser, D. Kern and C. Krauß, "Analyzing and Securing SOME/IP Automotive Services with Formal and Practical Methods," in *The 16th International Conference on Availability, Reliability and Security*, 2021.

[27] C. Plappert, F. Fenzl, R. Rieke, I. Matteucci, G. Costantino and M. D. Vincenzi, "SECPAT: Security Patterns for Resilient Automotive E / E Architectures," in *2022 30th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, 2022.

[28] United Nations Economic Commission for Europe, *UN Regulation No 155 – Uniform provisions concerning the approval of vehicles with regards to cybersecurity and cybersecurity management system - https://unece.org/sites/default/files/*, Geneva, CH: Regulation Addendum 154, 2021.

[29] Y. Chevalier, F. Fenzl, R. Rieke, C. Kraus, M. Kolomeets and A. Chechulin, "yberattack detection in vehicles using characteristic functions, artificial neural networks, and visual analysis," *Informatics and Automation*, vol. 20, no. 4, 2021.

# A. Appendix

## *A.1 Definitions and Abbreviations*

| Term | Meaning |
|---|---|
| **AAL** | Ambient Assisted Living |
| **AI** | Artificial Intelligence |
| **API** | Application Programming Interface |
| **ASI** | Advanced Security Services Infrastructure (E-CORRIDOR framework) |
| **AT** | Airport-Train (E-CORRIDOR pilot) |
| **AUC** | Area Under the Curve |
| **AUTOSAR** | Automotive Open System Architecture |
| **BLE** | Bluetooth Low Energy |
| **BYOD** | Bring Your Own Device |
| **CAN bus** | Controller Area Network |
| **CNN** | Convolutional Neural Networks |
| **CO2** | Carbon dioxide – air pollutant |
| **CSI** | Common Security Infrastructure (E-CORRIDOR framework) |
| **CTI** | Cyber-Threat Information |
| **CVE** | Common Vulnerabilities and Exposures |
| **CVSS** | Common Vulnerability Scoring System |
| **DB** | Database |
| **DBC** | Database CAN file |
| **DLI** | DSA Lifecycle Infrastructure (E-CORRIDOR framework) |
| **DMO** | Data Manipulation Operation |
| **DNA** | Here, as an analogy to the molecule that carries the genetic instructions |
| **DPO** | Data Protected Object |
| **DSA** | Data Sharing Agreement |
| **EAR** | Eye Aspect Ratio |
| **ECU** | Electronic Control Unit |
| **EU** | European Union |
| **FHE** | Fully Homomorphic Encryption |
| **eWallet** | Digital wallet |
| **GBFS** | General Bikeshare Feed Specification |
| **GPS** | Global Positioning System |
| **GTFS** | General Transit Feed Specification |
| **GUI** | Graphical User Interface |

| | |
|---|---|
| **HAR** | Human Activity Recognition |
| **IAI** | Information Analytics Infrastructure (E-CORRIDOR framework) |
| **IDS** | Intrusion Detection System |
| **IMU** | Inertial measurement unit |
| **IoT** | Internet of Things |
| **IP** | Internet Protocol |
| **IPS** | Intrusion Prevention System |
| **IR** | Infrared |
| **ISAC** | Information Sharing and Analytics Center |
| **ISI** | Information Sharing Infrastructure (E-CORRIDOR framework) |
| **IVI** | In-Vehicle Infotainment (IVI) |
| **JSON** | JavaScript Object Notation |
| **LFW** | Labeled Faces in the Wild |
| **LSTM** | Long Short-Term Memory |
| **ML** | Machine Learning |
| **MMT-ISAC** | Information Sharing and Analysis Center for Multi-Modal transport (E-CORRIDOR pilot) |
| **MoSCoW** | Must have, Should have, Could have, and Won't have but would like |
| **MPC** | Secure Multi-Party Computation |
| **NB** | Naïve Bayes |
| **NN** | Neural Network |
| **NVD** | National Vulnerability Database |
| **OBD-II** | On-board diagnostics - standard version 2 |
| **OEM** | Original Equipment Manufacturer |
| **OS** | Operating System |
| **PID** | Parameter Identification |
| **PNR** | Passenger Name Record |
| **PoI** | Point of Interest |
| **PRM** | People with Reduced Mobility |
| **PSR** | Private Secure Routine |
| **R&D** | Research and Development |
| **REST** | Representational state transfer |
| **RF** | Random Forest |
| **RGB** | Red, green, and blue – color model |
| **RGB-D** | Red, green, blue and depth |
| **RNN** | Recurrent Neural Network |

| | |
|---|---|
| **RPM** | Revolutions per minute |
| **RSSI** | Received Signal Strength Indicator |
| **STIX** | Structured threat information expression |
| **S2C** | Smart cities and car sharing (E-CORRIDOR pilot) |
| **SVM** | Support Vector Machine |
| **TRL** | Technology Readiness Level |
| **2PC** | Secure Two-Party Computation |
| **UC** | Use case |
| **US** | User story |
| **USA** | United States of America |
| **UUID** | Universally Unique IDentifier (UUID) |

## *A.2 List of Figures*