D5.1 AI-based Framework for Green & Trustworthy Operations Intermediate Version

Document Identification			
Status	Final	**Due Date**	31/01/2024
Version	1.0	**Submission Date**	31/01/2024

Related WP	WP5	**Document Reference**	D5.1
Related Deliverable(s)	D2.1, D2.2, D2.3	**Dissemination Level**	PU
Lead Participant	ATOS	**Lead Author**	Pariente Lobo, Tomas
Contributors	ATOS, EXUS, UOG, UOM, FUJ_LU, SQD, XLAB	**Reviewers**	Athanasios Stratikopoulos (UOM)
			Kaitai Liang (DUT)

Keywords:
Artificial Intelligence, AI, Machine Learning, ML, Deep Learning, DL, Exploratory Data Analysis, Federated Learning, MLOps, AutoML, Risk Analysis, Explainable AI, Energy Efficiency

Disclaimer for Deliverables with dissemination level PUBLIC

This document is issued within the frame and for the purpose of the TANGO project. This project has received funding from the European Union’s Horizon Europe Framework Programme under Grant Agreement No. 101070052. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the European Commission.

The dissemination of this document reflects only the author’s view and the European Commission is not responsible for any use that may be made of the information it contains. This deliverable is subject to final acceptance by the European Commission.

This document and its content are the property of the TANGO Consortium. The content of all or parts of this document can be used and distributed provided that the TANGO project and the document are properly referenced.

Each TANGO Partner may use this document in conformity with the TANGO Consortium Grant Agreement provisions.

Document Information

List of Contributors	
Name	Partner
Sergio Salmerón	ATOS
Iván Martínez	ATOS
Konstantinos Kentrotis	EXUS
Nahime Torres	EXUS
Maria Plakia	EXUS
Sofiane Lagraa	FUJ_LU
Moussa Ouedraogo	FUJ_LU
Sven Rasmusen	FUJ_LU
Sakshyam Panda	UOG
Manos Panaousis	UOG
Athanasios Stratikopoulos	UOM
John Zaras	SQD
Jakob Jenko	XLAB

Document History			
Version	Date	Change editors	Changes
0.1	15/11/2023	Tomás Pariente (ATOS)	Table of Contents to be approved
0.2	04/01/2024	Tomás Pariente (ATOS)	Additions from all editors collected in a single document
0.3	12/01/2024	Tomás Pariente (ATOS)	Added ES, section 1, 2 and 4. Some updates in the rest of the sections (2nd round of contributions)
0.4	19/01/2024	Tomás Pariente (ATOS)	Addressed some comments internal reviewers
0.5	23/01/2024	Tomás Pariente (ATOS)	Addressed comments in section 3.4. Sent for final QA before submission
1.0	30/01/2024	Tomás Pariente (ATOS)	FINAL VERSION TO BE SUBMITTED

Quality Control		
Role	Who (Partner short name)	Approval Date
Deliverable leader	Tomás Pariente Lobo (ATOS)	23/01/2024
Quality manager	Jürgen Neises (FUJ_GE)	26/01/2024
Project Coordinator	Tomás Pariente Lobo (ATOS)	30/01/2024

Table of Contents

List of Tables

List of Figures

List of Acronyms

Abbreviation / acronym	Description
AI	Artificial Intelligence
API	Application Programming Interface
CNN	Convolutional Neural Network
DL	Deep Learning
Dx.y	Deliverable number y and belonging to WP x
EC	European Commission
EDA	Exploratory Data Analysis
EDAE	Exploratory Data Analysis Engine
FL	Federated Learning
GDPR	General Data Protection Rule
GUI	Graphical User Interface
IoT	Internet of Things
KPI	Key Performance Indicator
ML	Machine Learning
MLOps	Machine Learning Operations
NIST	National
OS	Operative System
PAT	Privacy Assurance Tool
PEC	Privacy Enhancing Component
PCE	PEC Compliance Engine
PDI	PEC Data Initialiser
PRE	PEC Recommendation Engine
PRI	PEC Risk Identifier
PSE	PEC Scoring Engine
RENOPS	Renewable Energy Forecast Production Service
SOTA	State Of The Art
Tx.y	Task number y belonging to WPx
WPx	Work Package x
X-AI	Explainable Artificial Intelligence

Executive Summary

This deliverable is the first iteration of the software components developed in the scope of TANGO WP5 (AI-based Framework for Green & Trustworthy Operations). The document describes the work carried out in the six tasks of the work package focusing on the implementation aspects and explained the main components and features implemented so far. The document is therefore accompanied by software artifacts and demos to be integrated in the scope of WP6 in the different architecture layers of the TANGO platform.

Firstly, the document provides an overview of the main vision of the work package and the objectives of the release, mostly focusing on providing a first set of components with the aim of being tested in the different pilots.

Secondly, a detailed description per task and component is provided, broken down into the description of the component, its internal architecture, features implemented so far, the expected support for pilots, a description of the software artifacts, and potential future work for further releases. The document provides this information for all the tools developed in WP5: the Exploratory Data Analysis Engine (EDAE); the support for MLOps, Federated Learning and AutoML; the TornadoVM for dynamic intelligent execution on heterogeneous systems; the Privacy Enhancing Component (PEC) and Privacy Assurance Tool (PAT); the X-AI library for explainability; and finally, the RENOPS tool for allocating the execution of heavy processing tasks when the availability of renewal energy is more favourable. Pointers to software and demos are provided if appropriate, although it is expected that the tools will be integrated into the TANGO platform, and will therefore be accessible for other components, pilots and end users from there.

Finally, the document concludes with a summary of the work done so far and the work to be done for the integration of the tools and the remaining aspects to be covered in following releases.

1 Introduction

1.1 Purpose of the document

The goal of this deliverable is to provide an intermediate release consisting of working implementations of components developed in the scope of *WP5 - AI-based Framework for Green & Trustworthy Operations*. This deliverable is of type demonstrator. Therefore, the document provides the initial outputs in terms of implementation of all the tasks of WP5 accompanied by the software prototypes developed until M17 of the project (January 2024).

Within the document, all tasks are reporting the outcomes in the same way in order to provide a quick overview of their results and facilitate to the reader the possibility of finding easily the features implemented so far, along with extra technical information and relation to pilots. The document explains the current view in terms of future work for each of the components, a view that may require adaptation based on the feedback from the integration and pilots in the coming months.

1.2 Relation to other project work

This document is related to the outputs of T2.1 – GAP Analysis in Distributed Data Management, Processing & Storage, T2.2 – User Needs and Requirements for Data Management, Processing & Storage and T2.3 – Use Case Scenarios & KPIs Definition, where the GAP analysis in terms of technologies, the TANGO offerings, including those of WP5, and the mapping with the user requirements from pilots have been thoroughly described. In this sense, deliverables D2.1 (TANGO D2.1, 2023), D2.2 (TANGO D2.2, 2023) and D2.3 (TANGO D2.3, 2023) are important to understand the current document.

The document is also related to the work to be carried out in WP6 in terms of architecture, integration and testing in tasks *T6.1 -Continuous Integration and Delivery* and *T6.2 Functional Testing and Monitoring*. In the case of T6.1 the relation is bidirectional, as the integration will require the outputs of this deliverable in terms of software artifacts, as well as the architectural choices discussed in WP2 and WP6 have an impact in the way of implementing in the most effective way the results to facilitate the integration.

Finally, although there is no absolute dependency, the work carried out in *WP3 - Distributed Privacy-preserving Data Management and Storage* and *WP4 - Distributed Trust Management Framework* in terms of implementing key aspects for data sharing affect the input of most of the tasks of WP5, as most of them are applying AI or different techniques over the available data. Therefore, it is key to keep a close eye on all the developments in other technical work packages.

1.3 Structure of the document

This document is structured in 4 major sections:

- **Section 1** introduces the document, its purpose and structure.
- **Section 2** provides an overview of WP5, including the vision and objectives of the current release.
- **Section 3** is the main chapter of the document, presenting for each of the WP5 tasks the current state of the components implemented for the intermediate release.
- Finally, **Section 4** concludes the document with a summary of the main achievements and future work.

2 Overview of WP5

In the scope of TANGO, WP5 (AI-based Framework for Green & Trustworthy Operations) is mainly related to provide support to pilots and users in AI-related tasks. The work package has the following main objectives:

- To offer tools to reveal patterns and features that will enable the comprehension, analysis and modelling of data.
- To provide support to automate processes and relatively repetitive tasks of the AI training and serving life cycle.
- To increase trustworthiness and privacy through the identification of privacy risks inflicted by AI mechanisms and processes.
- To provide mechanisms to help AI developers to increase the trustworthiness and explainability of their models.
- To help on energy efficiency aspects related to the training and execution of AI, such as i) optimising energy consumption during the execution of an AI task on a single node; ii) minimising the energy footprint of some of the AI components; or iii) enabling an intelligent allocation of processing time for heavy energy consumption tasks when renewal energy is available.

The WP is not intending to provide a complete framework for AI. There are already many tools and frameworks that can be used to implement and deploy AI models that can be used for that purpose. In fact, users may select tools (e.g., Keras[1] or other frameworks) to develop their models complementing what is available in TANGO. The focus of WP5 is therefore to provide a selected variety of components to be integrated in the TANGO platform, which can be complemented with other AI tools to cover aspects not tackled in WP5. This will help users and pilots to make sense of the data shared and will open the possibility to develop and run analysis on top of the data in a trustworthy, privacy-preserving, explainable and energy efficient manner.

Most of the tools developed in WP5 need data to perform aspects such as pre-processing, training AI models, running the models, assessing privacy and explaining how the models have been developed. In this sense, most of the components of this work package can be considered as value-added tools to exploit the data being shared or belonging to the end users (e.g., our pilots). In this sense, most of the tools will be placed closer to where the data is produced or ending after the data sharing, which means that in most of the cases the components will be placed in the TANGO Connector at the user layer of the architecture. More details of the placement will be provided for each component in their respective subsections.

Figure 1 shows graphically the main tasks and components implemented in WP5 in support of the AI life cycle.

[1] https://keras.io/

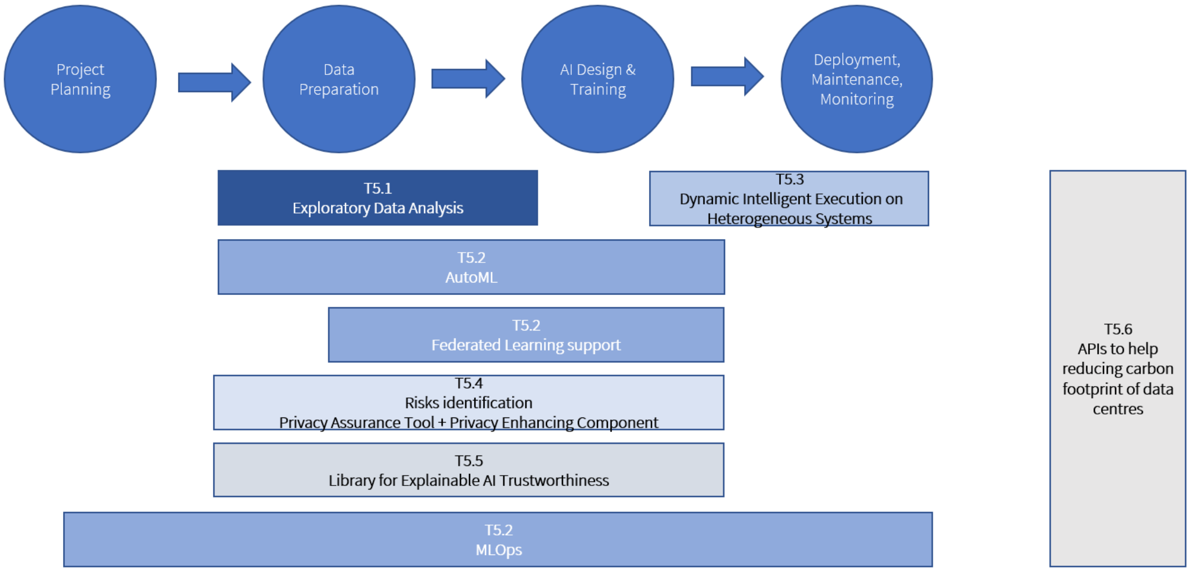

Project Planning
Data Preparation
AI Design & Training
Deployment, Maintenance, Monitoring
T5.1 Exploratory Data Analysis
T5.3 Dynamic Intelligent Execution on Heterogeneous Systems
T5.2 AutoML
T5.2 Federated Learning support
T5.4 Risks identification Privacy Assurance Tool + Privacy Enhancing Component
T5.5 Library for Explainable AI Trustworthiness
T5.2 MLOps
T5.6 APIs to help reducing carbon footprint of data centres

Figure 1. WP5 components relation

A brief description of the components depicted above is provided in Table 1.

Table 1: WP5 components

Task	Component name	Short description
T5.1	EDAE	The exploratory data analysis engine (EDAE) aims at pre-processing data and helping on the identification of patterns or features of interest. It provides ways to identify correlations in data, feature selection and provides several analysis and visualisations both generic and tailored to specific pilots' needs.
T5.2	MLOps	MLOps provides the possibility of managing several steps of the ML life cycle, from their conception to the training, serving, and monitoring phase. ML developers will be able to track experiments and data used for training, catalogue the different versions of the models, and perform model serving if required.
	FL	The Federated Learning (FL) component provides means and techniques to train models with data pools which are not shared among the stakeholders. The models are trained locally with the data at hand, and these partial models are then aggregated in order to come up with a more accurate model that can be shared then with the stakeholders.
	AutoML	AutoML provides support to automate some tasks of the training ML models, such as hyperparameter optimisation or selection of the right model.
T5.3	TornadoVM	TornadoVM is a programming framework for hardware acceleration. It provides APIs for developers to include functions meant to be compiled for hardware acceleration.
T5.4	PEC	The Privacy Enhancing Component (PEC) is a decision-support tool that advices the end-users on their privacy risks and how the privacy risks can be effectively managed. It computes privacy risk of a cyber threat, identifies measures to mitigate the privacy risks and provides guidance on GDPR compliance.

Task	Component name	Short description
	PAT	The Privacy Assurance Tool (PAT) is a secure framework designed for compliance checking, data exchange and record system. It offers the possibility of storing encrypted information about privacy preservation and operates on a consent-based access control mechanism for secure and privacy-preserving data exchange.
T5.5	X-AI	The Explainable AI (X-AI) component is a library that helps to describe an AI model, its expected impact and potential biases. The result is a characterisation of model accuracy, fairness, transparency and outcomes in AI-powered decision making.
T5.6	RENOPS	The Renewable Energy Forecast Production Service (RENOPS) provides functionality to assess and predict the availability of renewable energy sources based on solar irradiation power forecast, and predictions on energy price. It provides means for data centres or other tools to switch the processes that require heavy energy consumption (such as AI model training) to more favourable hours.

In the next sections, the document provides a detailed view of the current status of the implementation of these components for the intermediate release.

3 Description of components

3.1 Exploratory Data Analysis Engine [T5.1]

3.1.1 Introduction

The Exploratory Data Analysis Engine (EDAE) is a module designed to perform advanced data analytics, specifically targeting the discovery of patterns and features within extensive datasets, as well as the preparation of training datasets for other ML components within the TANGO platform. EDAE's primary functionality is to enable identification of correlations and selection of correlation-based features among heterogeneous real-world data, thereby facilitating the preparation and pre-processing of data for further analysis or machine learning applications.

The tool integrates a series of established data analysis techniques alongside AI-driven methods to automatically execute correlation analysis. EDAE’s process is thorough, consisting of variable identification, univariate and bi-variate analysis, treatment of missing values, outlier detection, variable transformation, and feature creation and selection. These capabilities are essential for providing or enabling a detailed exploration and visualisation of data characteristics, which are crucial for the TANGO project's diverse use cases.

EDAE’s architecture is structured into three main sub-modules to ensure a streamlined workflow. The data analytics sub-module is responsible for generating descriptive statistics and performing initial data analyses. The second sub-module manages data quality issues, including the handling of missing or duplicate data. The third sub-module prepares data for subsequent use in AI/ML training within the TANGO ecosystem. Python 3 serves as the foundation for the development of EDAE, with various libraries and functions employed for the implementation of its methods and techniques.

By providing these functionalities, EDAE offers the necessary analytical capabilities to transform raw data inputs into structured, analysed, and visualised datasets. These datasets are then ready to be utilised for the predictive modelling and analytical tasks that drive the project's objectives.

3.1.2 EDAE

3.1.2.1 Short description of the component

The exploratory data analysis engine (EDAE) will be placed at the ‘User Plane’ of the TANGO architecture, being fed with data deriving from the end users aiming to perform data analytics with a focus on the identification of patterns or features of interest among the data. EDAE will be capable of a) identifying correlations among the heterogeneous real-world data that will be inserted into the system and b) selecting correlation-based features. A combination of well-established data analysis techniques will be exploited in order to develop an AI-driven method to automatically perform the correlation analysis, visualising the generated results. The module’s envisioned process consists of the following steps: 1) Variable Identification, 2) Uni-variate Analysis, 3) Bi-variate Analysis, 4) Missing values treatment, 5) Outlier treatment, 6) Variable transformation and 7) Variable creation and selection.

The exploratory data analysis (EDA) applied in the context of the TANGO project, constitutes an operative approach to data analysis, aiming to improve the understanding and accessibility of results. In more detail, EDA builds upon the soundness of statistical models and hypothesis formulation, which are also used in the classical approach, to reveal hidden and unknown information from data in a form that enables analysts to obtain an immediate, direct and easy-to-understand representation of it. Hence, visual graphs are usually generated when this approach is followed, so that a more direct and trustworthy interpretation of similarities, differences, trends, clusters, and correlations are obtained through a picture, rather than through a series of numbers. In reality, EDA forms an analytical framework where the visual examination of data sets, by means of statistically significant representations, plays the pivotal role to support the formulation of hypotheses that could be tested on new data sets. The ability of comparison between two concepts, for instance the dynamic experimentation on data (e.g., evaluating the results on different subsets of a same data set, under different pre-processing conditions), along with the exhaustive visualisation capabilities, empower researchers to identify outliers, trends and patterns in data, upon which new theories and hypotheses can be built.

3.1.2.2 Internal architecture

The EDAE consists of 3 main sub-modules: 1) A data analytics sub-module that is responsible for studying the relationship between variables among the data provided by generating descriptive statistics, data type summary, extracting the data characteristics, etc.; the sub-module will be also capable of performing univariate, bivariate and multivariate analysis. 2) The second sub-module is responsible of handling missing or duplicate data, providing the option to the user to decide how the platform will treat this data or asking him to confirm whether or not they would like to proceed with the proposed treatment methods. 3) The third sub-module aims to prepare the data to be used by other TANGO components in order to train their AI/ML algorithms. The functionalities of this sub-module include the semi-automated data transformation, the implementation of dimensionality reduction techniques (AI may also be implemented to assess what is the optimal dimensionality), variable selection and feature engineering as well as outlier identification and treatment (if needed and upon confirmation by the user).

The methods/techniques that have been implemented as part of the EDAE tool have been developed in Python 3, and all the references to different libraries or functions are correlated with the Python programming language. In the next paragraphs, more details about the implementation of each sub-module are presented.

The results of the **data analytics sub-module** (**1st sub-module**) inform the reasonable next steps you can take with the data, what corrective actions you will need to perform, and generally help better understand the range of values and data types that you will be working with during a machine learning project.

Under descriptive analysis, there have been created a number of functions to display useful characteristics of the data. These characteristics include:

- The 'basic characteristics' of the DataFrame: the number of rows and columns in the raw data, and how much space it takes in memory.
- Missingness analysis: the number and proportion of missing values in each column in the data set. (Missingness is further explored in the Missing Value Handling section)
- Data type summary: the inferred data type of each column. Here, we aggressively try to identify each data type, including date-time variables (although this can be turned off while using the function)
- Descriptive statistics: a table of the mean, minimum, maximum values in each column as well as the 25th, 50th and 75th percentiles of the data (clearly, this only applies for numeric variables)
- Duplicate and constant analysis: the number of rows and/or columns that are duplicates of another row or column, and all columns that contain a constant value.
- The descriptive analysis functions also include the ability to plot the numeric and categorical variables in the data. Histograms are plotted for the numeric variables, to identify their overall distributions; frequency charts are created for categorical variables (when less than ten distinct categories are present).

Additionally, a separate set of functions further explores the behaviour of variables in the data set and their relationships to one another. Under the univariate analysis functions, plots of the kernel density estimations for each of the numeric variables are generated to better understand the underlying distribution function that may be generating the observed variables. Under the multivariate analysis functions, the correlation matrix between the variables is printed (using Pearson correlation). Plots with the relationship between all continuous variables and each date time variable in the data set are also generated. The idea here is to help the user visually identify any patterns in the numeric variables that emerge over time.

The **2nd sub-module** is responsible for **handling missing data**. Missing values are almost inevitable in any reasonably sized data collection effort. Many algorithms fail in the presence of missing values, and they must be dealt with in some way or another. Values can be missing from the data set for a variety of reasons: it can be random, or it can correlate other values in the data or characteristics of the data subject. EDAE does not deal with the different reasons for why data can be missing in this set of functions. Instead, EDAE presents a further set of charts to better explore the missingness in the data set, as well as a variety of methods to handle them. Three charts are plotted to visualise the missing values in the data set:

- The missingno matrix: presents a bird's eye view of the missing data in the data set, and allows for comparisons to be made between columns and identify the rows with the highest levels of missingness.
- The missingno bar chart: shows how much data is present as a percentage in each column.
- The missingno heat map: shows the missingness correlation between columns; this is, it shows which columns are more likely to both have missing values, which have a negative relationship and which columns' missing values are independent of one another.

The techniques provided to handle missing values are as follows:

- delete: remove all rows that have a missing value.
- mean: fill all missing values with the mean of that column.
- median: fill all missing values with the median of that column.
- mode: fill all missing values with the mode of that column.
- K nearest neighbours: identify the nearest 1\% of the data set and fill missing values with the mean of the nearest neighbours

The **3rd sub-module** is responsible for several tasks, namely a) outlier treatment, b) dimensionality reduction and c) variable importance/feature selection, all further detailed in D2.3. Overall, it aims to prepare the data to be used by other TANGO components in order to train their AI/ML algorithms.

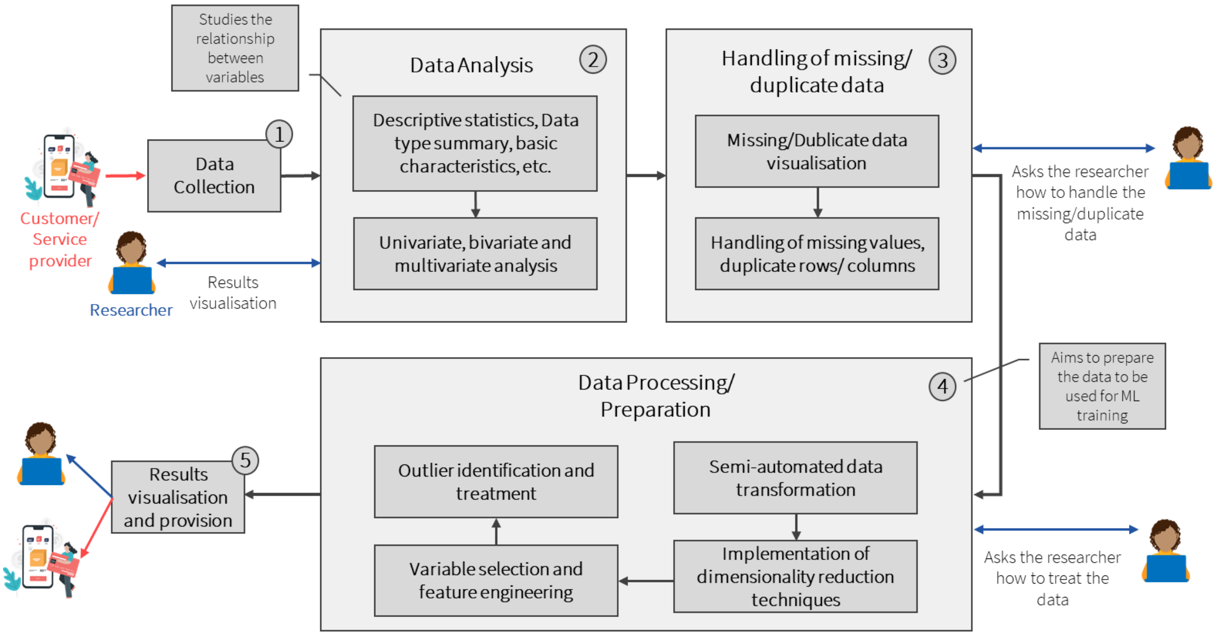
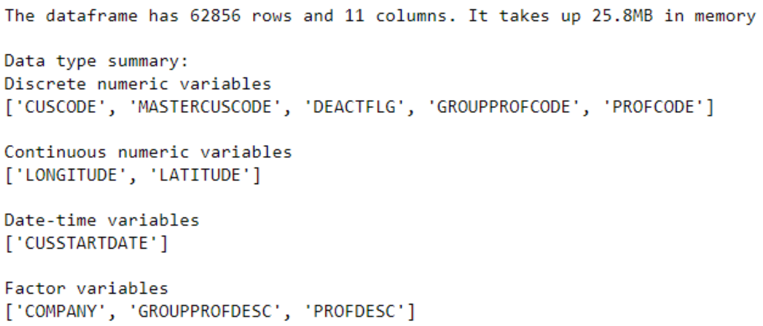

Studies the relationship between variables
Data Analysis
2
Descriptive statistics, Data type summary, basic characteristics, etc.
Univariate, bivariate and multivariate analysis
Handling of missing/ duplicate data
3
Missing/Dublicate data visualisation
Handling of missing values, duplicate rows/ columns
1
Data Collection
Customer/ Service provider
Researcher
Results visualisation
Asks the researcher how to handle the missing/duplicate data
Aims to prepare the data to be used for ML training
Data Processing/ Preparation
4
Outlier identification and treatment
Semi-automated data transformation
Variable selection and feature engineering
Implementation of dimensionality reduction techniques
5
Results visualisation and provision
Asks the researcher how to treat the data

Figure 2. EDAE's High level Architecture

3.1.2.3 Features implemented

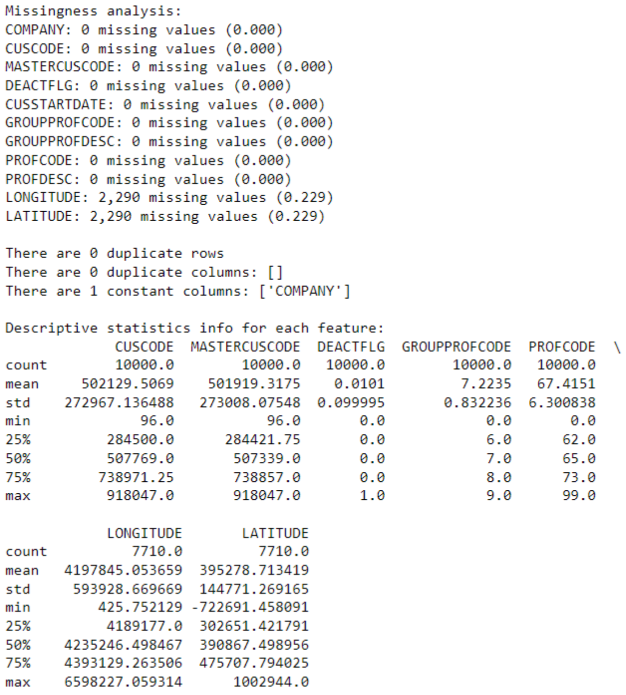
- **Descriptive Analysis:** EDAE's descriptive analysis capabilities are extensive, offering a fundamental understanding of the dataset at hand. It includes:
    - **Basic Data Characteristics:** Quickly assesses the size and memory footprint of the dataset, providing an immediate sense of scale and complexity.
    - **Missingness Analysis:** Evaluates and quantifies missing data, enabling informed decisions on treatment methods.
    - **Data Type Summary:** Aggressively infers data types, including date-time variables, to ensure correct processing and analysis.
    - **Descriptive Statistics:** Provides a statistical summary of the data, including measures of central tendency and variability for numerical variables.
    - **Duplicate and Constant Analysis:** Identifies redundant or static features within the data, which are critical for ensuring the quality of subsequent analyses.
- **Univariate and Multivariate Analysis:** The engine facilitates both univariate and multivariate analyses, allowing users to explore single variables as well as the relationships between them:
    - **Variable Densities and Frequencies:** Through kernel density estimations and frequency charts, EDAE reveals the underlying distributions and commonalities within the data.
    - **Correlation Matrix:** Utilizes Pearson correlation to highlight relationships between variables, providing a foundation for deeper bivariate and multivariate examinations.
    - **Violin Plots and Pie Charts:** Offers visual representations for continuous and categorical data, respectively, enhancing the interpretability of the dataset's features.
- **Missing Value Treatment:** EDAE empowers users to dictate how missing values are addressed, with methods such as:
    - **Deletion:** Removes incomplete records, ensuring analyses are performed on fully populated datasets.
    - **Imputation:** Utilises statistical measures like mean, median, mode, or more sophisticated techniques like KNN imputation to fill in gaps within the data.
- **Outlier Treatment:** Recognising and managing outliers is crucial for maintaining data integrity:
    - **Detection:** Identifies anomalies using visualisation and statistical methods.
    - **Treatment:** Offers solutions such as removal, statistical imputations, or algorithmic predictions to address outliers effectively.
- **Dimensionality Reduction:** To combat the curse of dimensionality and enhance model performance, EDAE includes methods like PCA and UMAP, facilitating a more manageable representation of the data while preserving essential information.
- **Feature Importance and Anomaly Detection:** Determining variable significance is streamlined through various algorithms, providing insights into the data's predictive power and revealing anomalous patterns that may warrant further investigation.
- **Simple Automatic Clustering:** This feature helps in discerning inherent groupings within the data, which can be pivotal for segmentation and pattern recognition.

The following figures depict some of the aforementioned functionalities implemented to date.



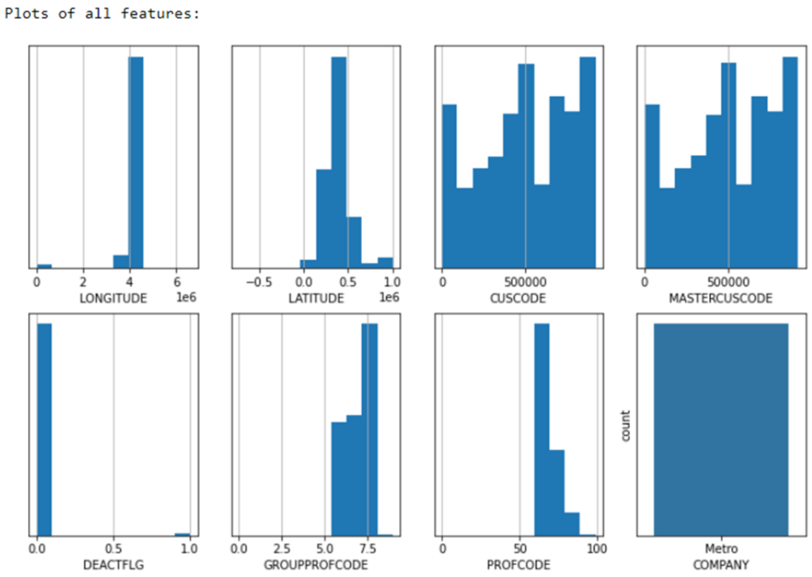
```
The dataframe has 62856 rows and 11 columns. It takes up 25.8MB in memory

Data type summary:
Discrete numeric variables
['CUSCODE', 'MASTERCUSCODE', 'DEACTFLG', 'GROUPPROFCODE', 'PROFCODE']

Continuous numeric variables
['LONGITUDE', 'LATITUDE']

Date-time variables
['CUSSTARTDATE']

Factor variables
['COMPANY', 'GROUPPROFDESC', 'PROFDESC']
```

Figure 3. Data type summary of the dataset



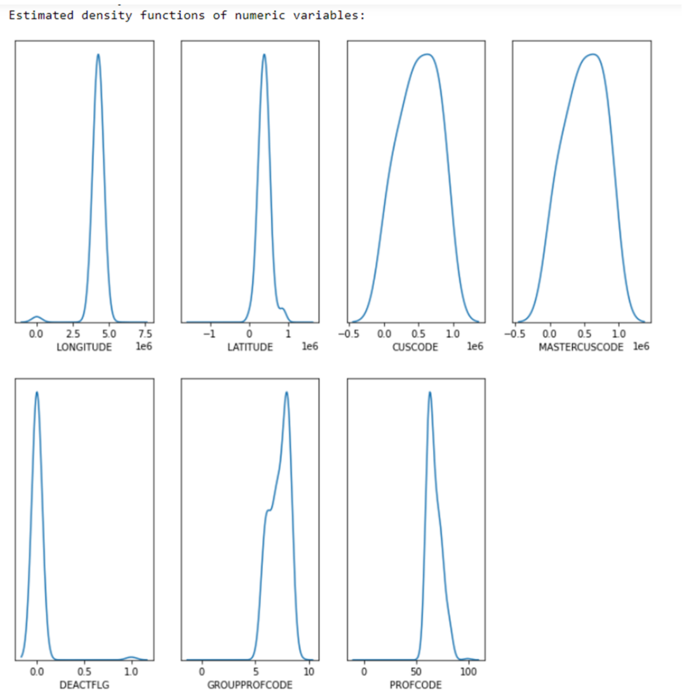
```
Missingness analysis:
COMPANY: 0 missing values (0.000)
CUSCODE: 0 missing values (0.000)
MASTERCUSCODE: 0 missing values (0.000)
DEACTFLG: 0 missing values (0.000)
CUSSTARTDATE: 0 missing values (0.000)
GROUPPROFCODE: 0 missing values (0.000)
GROUPPROFDESC: 0 missing values (0.000)
PROFCODE: 0 missing values (0.000)
PROFDESC: 0 missing values (0.000)
LONGITUDE: 2,290 missing values (0.229)
LATITUDE: 2,290 missing values (0.229)

There are 0 duplicate rows
There are 0 duplicate columns: []
There are 1 constant columns: ['COMPANY']

Descriptive statistics info for each feature:
             CUSCODE  MASTERCUSCODE  DEACTFLG  GROUPPROFCODE  PROFCODE  \
count        10000.0        10000.0   10000.0        10000.0   10000.0
mean     502129.5069    501919.3175    0.0101         7.2235   67.4151
std    272967.136488   273008.07548  0.099995       0.832236  6.300838
min             96.0           96.0       0.0            0.0       0.0
25%         284500.0      284421.75       0.0            6.0      62.0
50%         507769.0       507339.0       0.0            7.0      65.0
75%        738971.25       738857.0       0.0            8.0      73.0
max         918047.0       918047.0       1.0            9.0      99.0

            LONGITUDE       LATITUDE
count          7710.0         7710.0
mean   4197845.053659  395278.713419
std     593928.669669  144771.269165
min        425.752129 -722691.458091
25%         4189177.0  302651.421791
50%    4235246.498467  390867.498956
75%    4393129.263506  475707.794025
max    6598227.059314      1002944.0
```

Figure 4. Output of the missingness, descriptive and duplicate or constant rows or columns analysis



Plots of all features:

Figure 5. Histograms of the numerical variables in the provided dataset



Estimated density functions of numeric variables:

Figure 6. Density functions of the numerical variables in the provided dataset

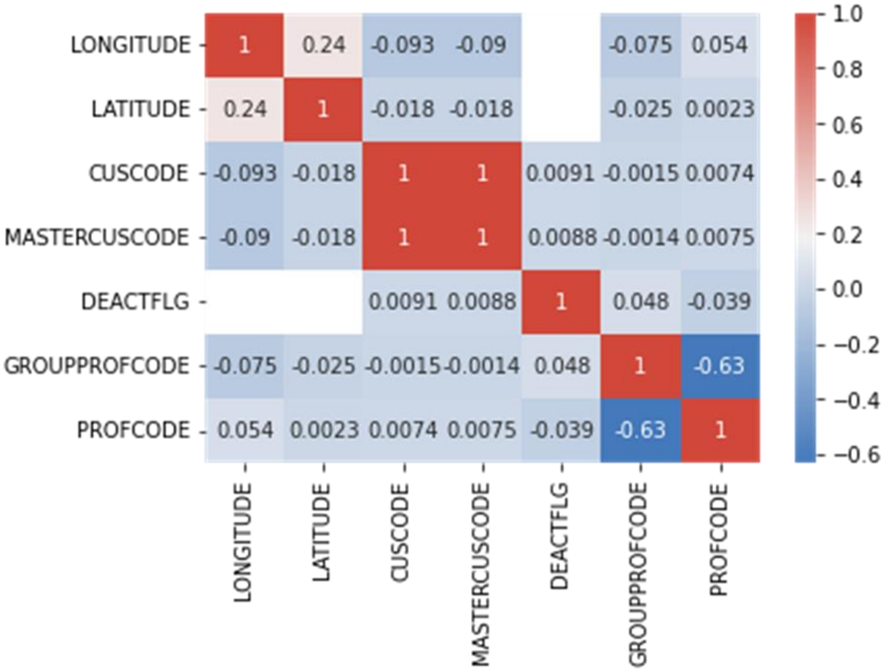


Figure 7. Correlation matrix of numeric columns

For the 2nd sub-module which is responsible for the identification and handling of missing data. In Figure 6 and Figure 7 the user observes the existence of missing values in the different columns of the dataset. While observing the missingness heatmap the user is able to identify correlation among the columns in terms of missing values. For example, in Figure 8, each time a row appears to have a missing value in the longitude column, the same row will have a missing value in the latitude column.

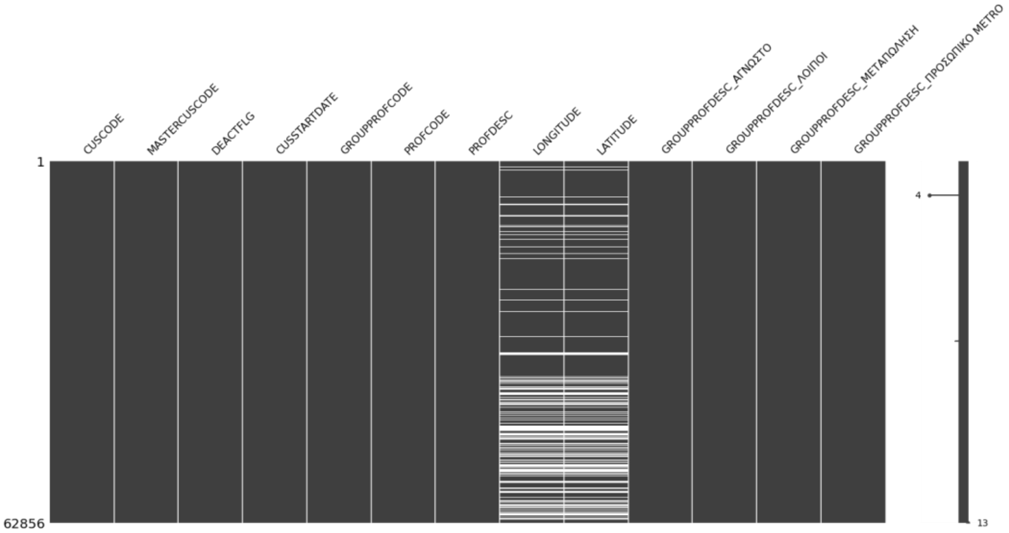


Figure 8. Missingness matrix

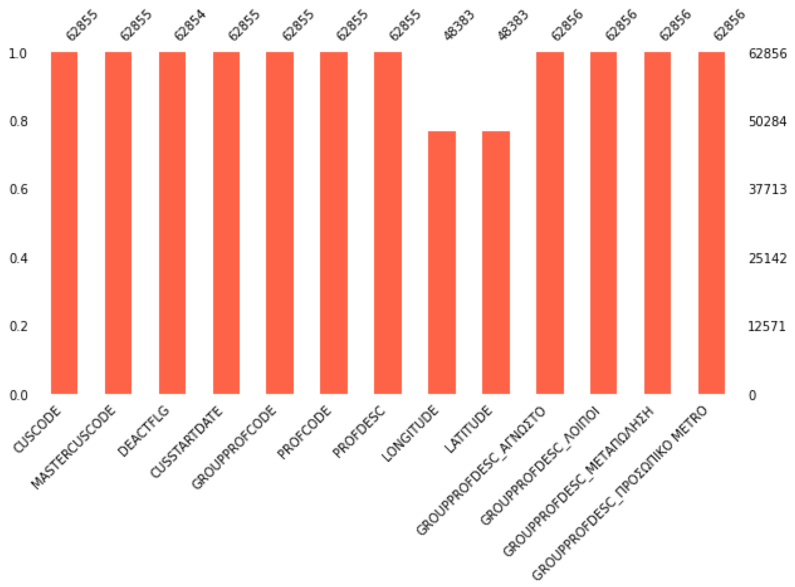


Figure 9. The bars show what proportion of data is present

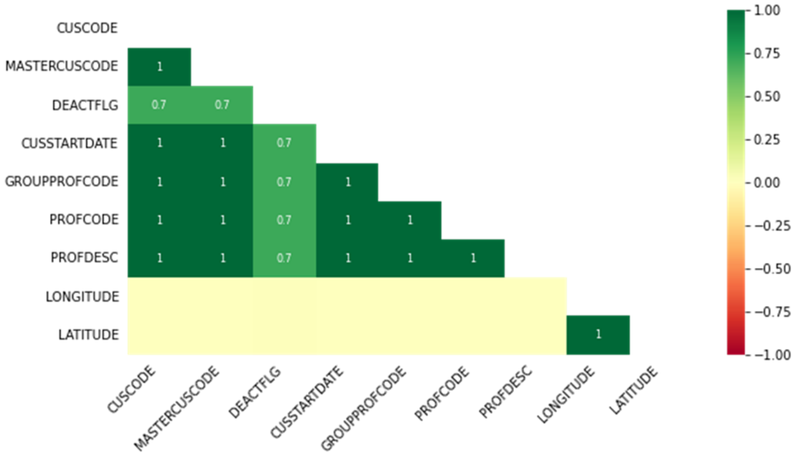
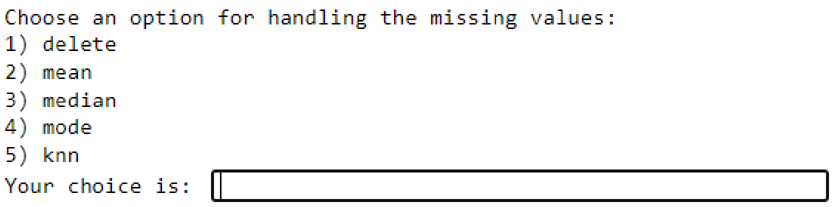
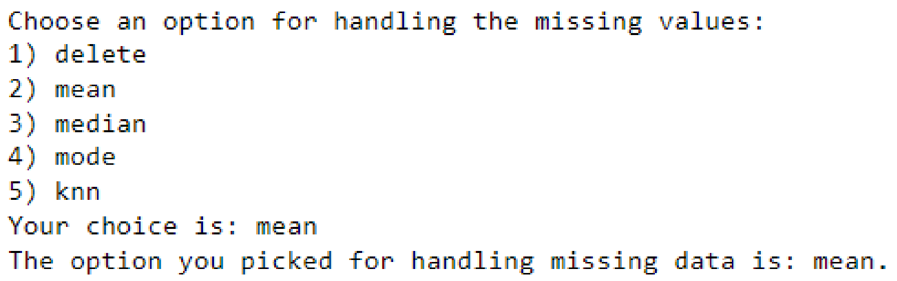


Figure 10. The correlation plot shows which variables are more likely to be missing together

For the handling of the missing values, the following choices presented in Figure 11a considering the deletion of them or filling with mean, median and other statistical methods.

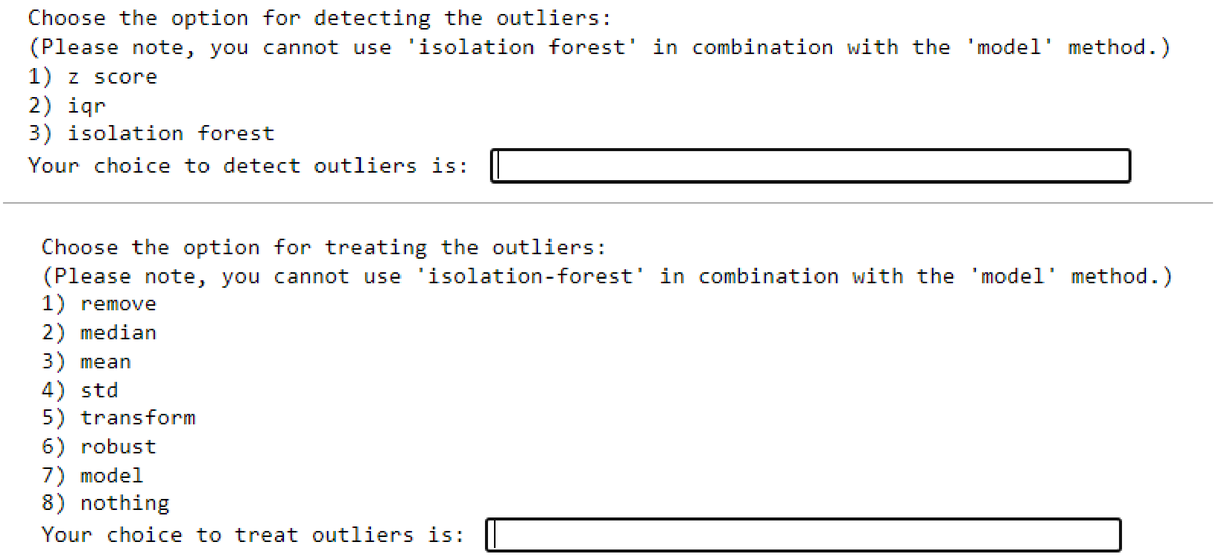


```
Choose an option for handling the missing values:
1) delete
2) mean
3) median
4) mode
5) knn
Your choice is:
```



```
Choose an option for handling the missing values:
1) delete
2) mean
3) median
4) mode
5) knn
Your choice is: mean
The option you picked for handling missing data is: mean.
```

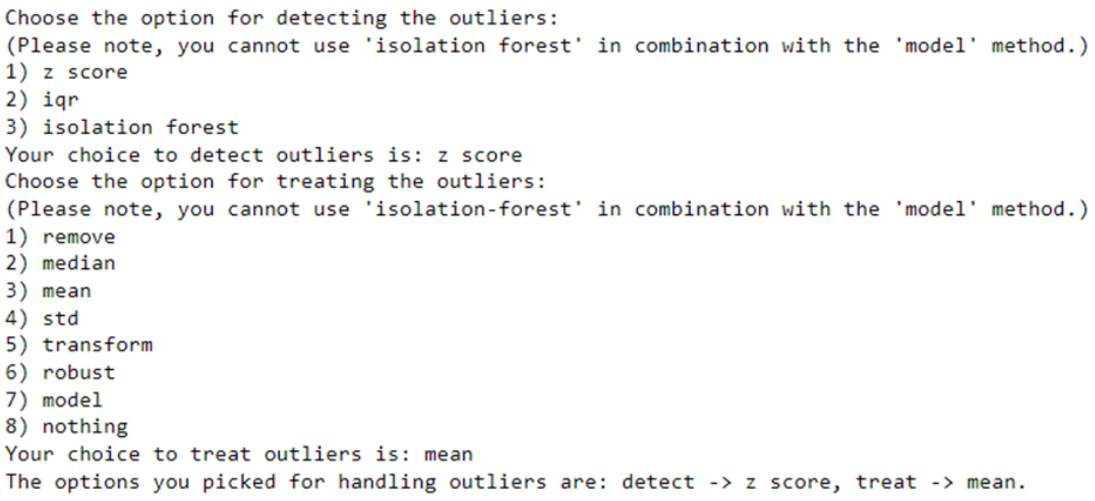
Figure 11. a) Menu for choosing how to handle missing values and b) example of choosing one of the options



```
Choose the option for detecting the outliers:
(Please note, you cannot use 'isolation forest' in combination with the 'model' method.)
1) z score
2) iqr
3) isolation forest
Your choice to detect outliers is:
```

```
Choose the option for treating the outliers:
(Please note, you cannot use 'isolation-forest' in combination with the 'model' method.)
1) remove
2) median
3) mean
4) std
5) transform
6) robust
7) model
8) nothing
Your choice to treat outliers is:
```

Figure 12. a) Menu presenting options to detect outliers in the dataset and b) for treating outliers



```
Choose the option for detecting the outliers:
(Please note, you cannot use 'isolation forest' in combination with the 'model' method.)
1) z score
2) iqr
3) isolation forest
Your choice to detect outliers is: z score
Choose the option for treating the outliers:
(Please note, you cannot use 'isolation-forest' in combination with the 'model' method.)
1) remove
2) median
3) mean
4) std
5) transform
6) robust
7) model
8) nothing
Your choice to treat outliers is: mean
The options you picked for handling outliers are: detect -> z score, treat -> mean.
```

Figure 13. Users' selections example for the identification and treatment of the outliers

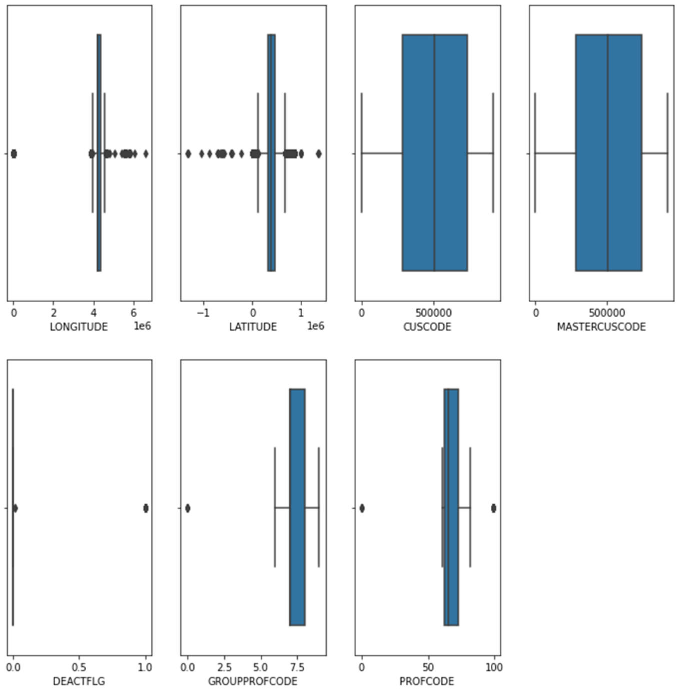


Figure 14. Boxplots presenting the outliers in the numeric values of the dataset identified based in user's selection of available methods

3.1.2.4 Support for pilots

For the **Smart Hospitality (Pilot 1)**, EDAE integrates with the NADIA hotel system, focusing on the preparation of customer data. It provides hotel managers with statistical analyses of customer traits and preferences, revealing associations, trends, and patterns crucial for personalised service delivery. Receptionists and managers gain insights into customer behaviours, facilitating the delivery of tailored experiences. Additionally, EDAE incorporates anonymisation techniques to ensure sensitive information, such as nationality or gender, is processed in compliance with privacy regulations. Key functionalities include:

- Customer categorisation based on shopping frequency, amounts spent, and recency (FR-AI-EDAE-001).
- Identification of patterns, trends, and associations in customer traits and preferences (FR-AI-EDAE-008, FR-AI-EDAE-009).
- Anonymisation of sensitive data to protect individual privacy.

In **Smart Manufacturing (Pilot 3)**, EDAE's module assists technicians by automating the generation of reports on failed prints. It supplies pre-training statistics on print data and key performance indicators for managerial oversight, including the identification of outliers and failure trends. This enables a proactive approach to quality control and operational efficiency in the manufacturing process. Key functionalities include:

- Analysis of print data to identify outliers and failure trends (FR-AI-EDAE-010).
- Key Performance Indicators (KPIs) for printing tasks and quality issues data for data scientists (FR-AI-EDAE-011).

For the **Retail (Pilot 6)** pilot, EDAE delves into sales, product, and customer data to provide a comprehensive analysis that underpins business strategies. It segments customers based on loyalty metrics derived from their shopping behaviours and preferences, providing a nuanced understanding of different customer demographics. The insights garnered are pivotal for business analysts and store managers, who can identify and bolster underperforming areas, align inventory with consumer tastes, and tailor marketing strategies to regional preferences. Key functionalities include:

- Detecting and reporting shopping patterns and preferences across different countries (FR-AI-EDAE-002).
- Exposing data to analysts for identifying new market opportunities and weak areas in stores for revenue optimisation (FR-AI-EDAE-003, FR-AI-EDAE-005).
- Exploring large datasets to indicate regional taste preferences and customer behaviour insights (FR-AI-EDAE-006).

Additionally, EDAE must be compatible with all TANGO components that consume its data (NFR-COMP-EDAE-001) and dockerised for integration ease (NFR-PORT-EDAE-001). The engine also generates notifications about the status of all EDAE processing tasks (FR-AI-EDAE-013), improving user experience.

Additionally, EDAE must be compatible with all TANGO components that consume its data (NFR-COMP-EDAE-001) and dockerised for integration ease (NFR-PORT-EDAE-001). The engine also generates notifications about the status of all EDAE processing tasks (FR-AI-EDAE-013), improving user experience.

An example of custom report for the Retail pilot is presented in Figure 15. This analysis provides insights into customer purchasing patterns and behaviour (FR-AI-EDAE-006) focussing on the variation in these patterns at different times throughout the day, highlighting the distinctions between weekdays and weekends.

Understanding the mean hourly transactions volume helps recognising peak business hours, which is crucial for the strategic planning of promotions, marketing initiatives, and overall store operations. For instance, it enables to make well-informed decisions regarding staffing and the allocation of resources, ensuring that the store is optimally prepared to meet customer demand, thereby enhancing customer satisfaction and operational efficiency.

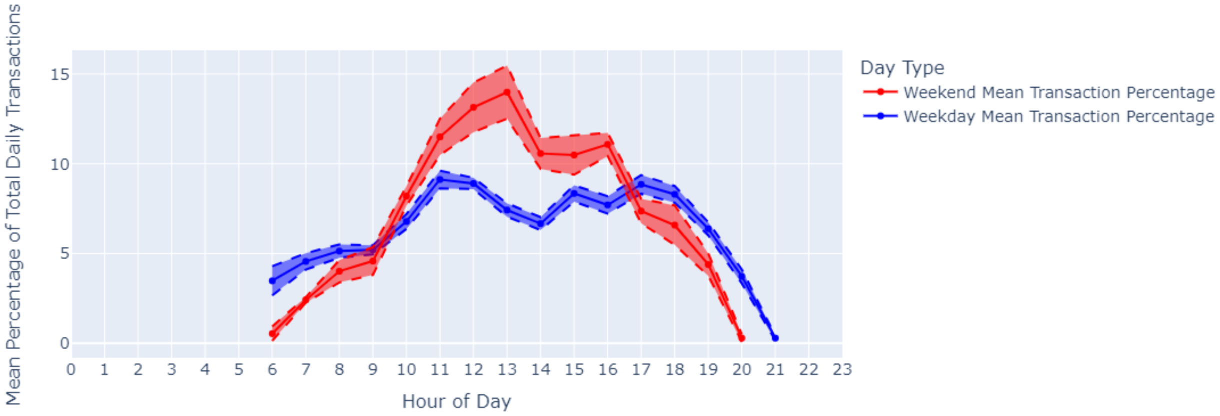


Figure 15. Hourly Transactions Volume Distribution Analysis by Day Type (Weekdays vs. Weekends)

3.1.2.5 Software artifacts

A dockerised version of the EDAE will be deployed on the provided server. In order to explore the TANGO CI/CD system automated testing capabilities, we will push a Docker image to registry from external source (due to restrictions on sharing code directly).

3.1.2.6 Future work

Up to M17 (January 2024), the development efforts were concentrated on enhancing EDAE’s functionalities to suit the unique needs of each use case. The retail pilot, being the most advanced, along with the smart manufacturing use case, received focused attention for feature development and optimisation. Coinciding with the delivery of D5.1, the beta version of EDAE is successfully launched. This milestone marked approximately 60% completion of the development, showcasing a significant leap in the engine's capability and readiness for practical application.

The next phase of development will be dedicated to completing the remaining portion of EDAE’s functionalities. This involves finalising and fine-tuning features specific to each use case, ensuring comprehensive coverage and effectiveness across the TANGO project. A pivotal part of the upcoming development stage is the integration of EDAE with user interfaces (T6.3) and other relevant components within the TANGO ecosystem (T5.2, T5.3, T5.5, T5.6). This integration is crucial for enhancing the usability of EDAE, making its analytical power readily accessible and user-friendly within the project’s operational framework.

3.2 Energy efficient AI model training [T5.2]

3.2.1 Introduction

Task T5.2, energy efficient AI model training, provides a set of tools that helps to train Machine Learning models. As there are many tools in the market supporting different aspects of the ML life cycle, this task does not intend to cover all aspects, but focuses on some specific elements that may impact in solving specific problems of data sharing:

- applying federated learning techniques to minimise the needs of sharing data while maximising the accuracy of the trained models and preserving data privacy;
- providing methods to help data scientist and engineers to ease the process of training and selecting the best candidate model to solve specific predictive problems using AutoML tools;
- and giving the possibility to track the ML life cycle and serving the models using a MLOps.

Deliverable D2.3 provided an in-depth overview of the three components in Section 5.2. Figure 16, borrowed from Section 5.2.1 of D2.3, shows the three components and the expected relationship among them.

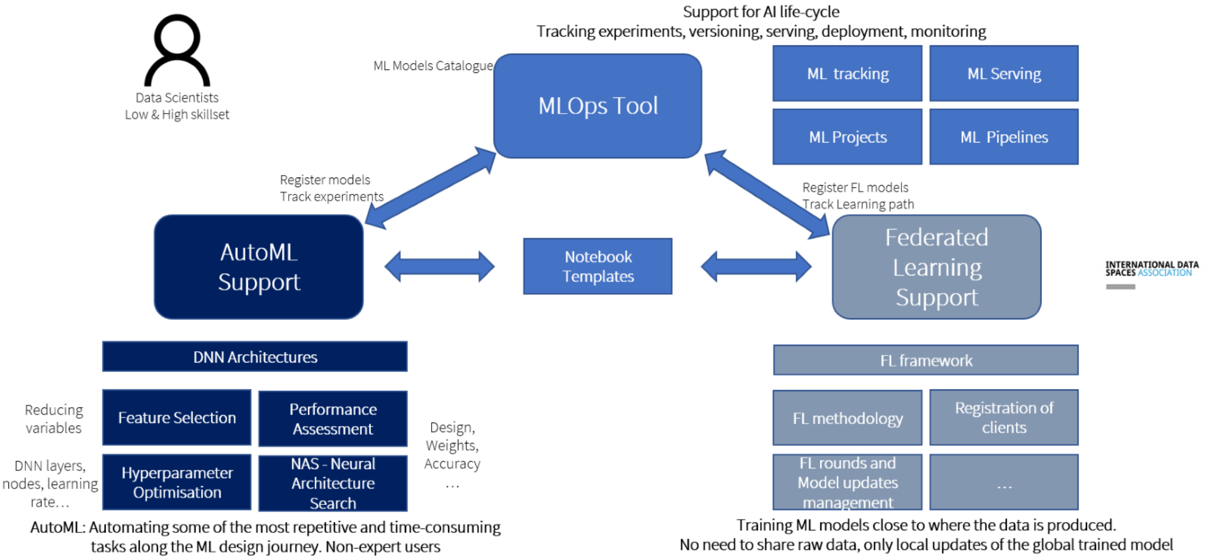

Support for AI life-cycle
Tracking experiments, versioning, serving, deployment, monitoring
Data Scientists
Low & High skillset
ML Models Catalogue
MLOps Tool
ML tracking
ML Serving
ML Projects
ML Pipelines
Register models
Track experiments
Register FL models
Track Learning path
AutoML Support
Notebook Templates
Federated Learning Support
INTERNATIONAL DATA SPACES ASSOCIATION
DNN Architectures
Reducing variables
Feature Selection
Performance Assessment
Design, Weights, Accuracy …
DNN layers, nodes, learning rate…
Hyperparameter Optimisation
NAS - Neural Architecture Search
FL framework
FL methodology
Registration of clients
FL rounds and Model updates management
…
AutoML: Automating some of the most repetitive and time-consuming tasks along the ML design journey. Non-expert users
Training ML models close to where the data is produced.
No need to share raw data, only local updates of the global trained model

Figure 16. T5.2 Main Components and their Interactions (from D2.3)

MLOps component has a central role as main driver of the ML life cycle. The models developed either using AutoML or Federated Learning could be placed in the MLOps tool and the integration is expected to happen via Notebooks or APIs. However, for this first iteration all these components are provided with a minimal set of features with no integration among them or with other components of the TANGO framework whatsoever. The components are expected to be mainly deployed at the client side (TANGO Connector) of the architecture, although other configurations might be useful to be explored in future releases.

3.2.2 Support for MLOps

The MLOps component, which was first described in D2.3, enables TANGO to handle the different ML models that are produced. Models' entire life cycle (from conception to the serving, training, and monitoring stages) is covered by this component. AI system engineers as well as ML and AI developers are supposed to use it. The particular feature set of this component is described in detail in Section 3.2.2.3.

3.2.2.1 Short description of the component

From a functional perspective the MLOps component allows the tracking of ML model generation experiments. The MLOps component is oriented to advanced-AI/ML users in order to allow them to keep track of the experiments carried out generating ML models, easing:

- Models and their results (or other artifacts) storage.
- Model management & deployment.

In the TANGO ecosystem the MLOps component will be placed into the TANGO Connector in the architecture defined in deliverable D2.3. This component provides to TANGO the capacity of managing the different ML models that will be developed.

As added value for the TANGO platform, the Data scientists and ML system engineers will may now receive help for the whole life cycle of machine learning, including standardized and traceable model building and administration, thanks to MLOps technology. Training and experiments in machine learning are documented, which lends credibility to the model.

3.2.2.2 Internal architecture

The main building blocks of the MLOps component are shown in Figure 17 from a high-level, cross-technology viewpoint. This figure illustrates the many components that make up the anticipated solution. Functionalities like ML experiment tracking will be made possible by MLOps, registering model performance across several iterations generated (via training and testing) on the model building platform or development environment. This component also allows for model serving or deploying the model for usage in a production environment, as well as the archiving of models and other artifacts related with the model.

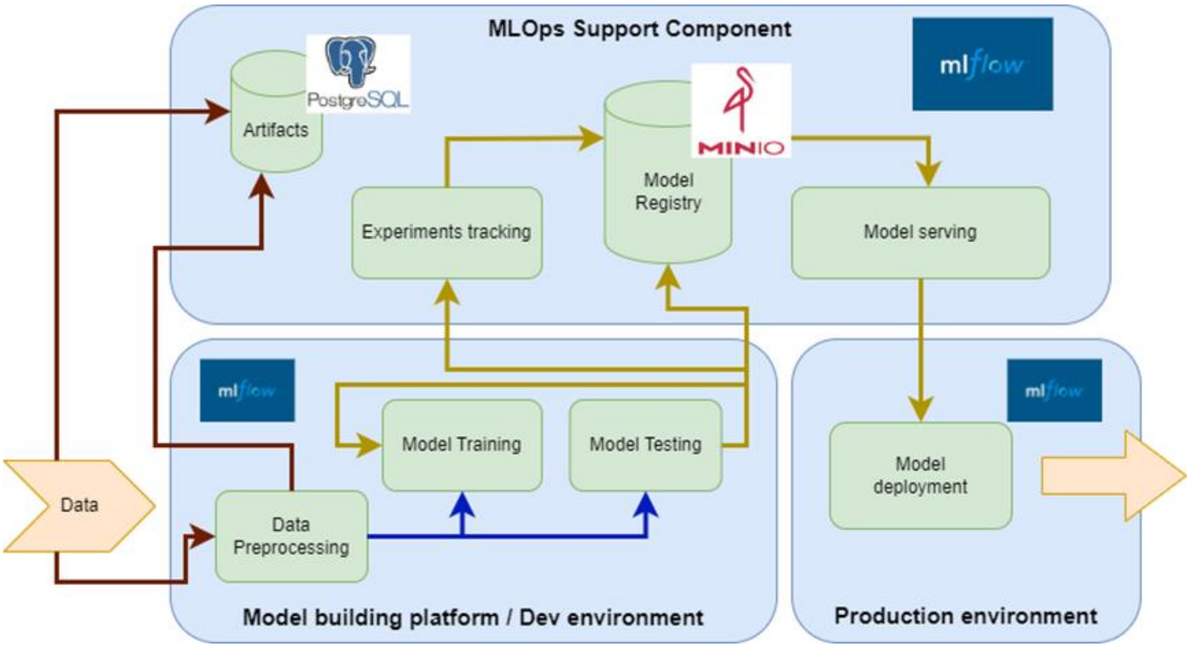


Figure 17. MLOps component internal architecture.

The MLOps component might be installed in either the central TANGO platform or the pilot side of the platform. The former deployment would allow pilots to share ML models, but it would raise privacy and security concerns if the ML models were not made public. After reviewing the pilots demands as indicated in deliverable D2.2, it was discovered that the use cases only need the sharing of ML models among pilot participants. As a result, the project does not require a central ML Model Registry for all pilots, but rather decentralized registries for the individual pilot use cases.

Figure 17 illustrates not only the primary building parts of the MLOps component but also their mapping with supporting technologies. These technologies mostly include MLFlow[2], PostgreSQL[3], and MinIO[4]. Because of its lightweight nature, MLFlow only needs a small amount of RAM and processing power in addition to certain deployment requirements. This is important from the point of view of energy efficiency of the solution, as other tools (e.g., Kubeflow[5]) have more requirements in terms of energy footprint, deployment weight, etc. Furthermore, nothing prevents the deployment of MLFlow in the TANGO platform central node if necessary to oversee the requirements of upcoming pilots or even to assist with the ML life cycle of particular platform components. Since there are currently no needs for model sharing among pilots, it is more likely that the MLOps component will be deployed at the user side of the TANGO platform.

3.2.2.3 Features implemented

The MLOps component provides support for the following main core features:

- Model training and tuning: to train and enhance model performance.
- Model governance and review: to keep track of model versions and lineage and oversee the life cycle management of model artifacts and transitions. MLFlow is an open source MLOps platform that facilitates collaboration, discovery, and sharing among ML models.

[2] https://mlflow.org/
[3] https://www.postgresql.org/
[4] https://min.io/
[5] https://www.kubeflow.org/

- Model inference and serving: to control the frequency of model refreshes, the duration of inference requests, and other related production-specifics in testing and quality assurance.
- Model deployment and monitoring: to productionize registered models, automate cluster setup and permissions.

3.2.2.4 Support for pilots

The MLOps component aims to support mainly the following TANGO pilots:

- Smart Manufacturing pilot (FMAKE):
  - Providing support to the ML model creation.
  - Using data on the effectiveness of each trained model, the MLOps component may assist with versioning the model in order to monitor its evolution at various phases.
- Banking pilot:
  - Owing to the substantial volume of data exhibited in the banking use case, machine learning models are suggested as a means of supplying certain functionality. The MLOps component may be useful to track various models and their produced results.

Finally, Figure 18 shows a tentative deployment scenario in which the MLOps component is used in UC1_1 node to run some experiments and tuning the model produced. UC1_2 will be in charge of model governance to keep track of model versions provided by UC1_1 (ML_Model1, ML_Model2 and ML_Model3 in Figure 3). Finally, the UC1_2 will be serving the ML_Model2 to UC1_3 to deploy the model in a real production environment enabling the model to be consumed by the pilot end users.

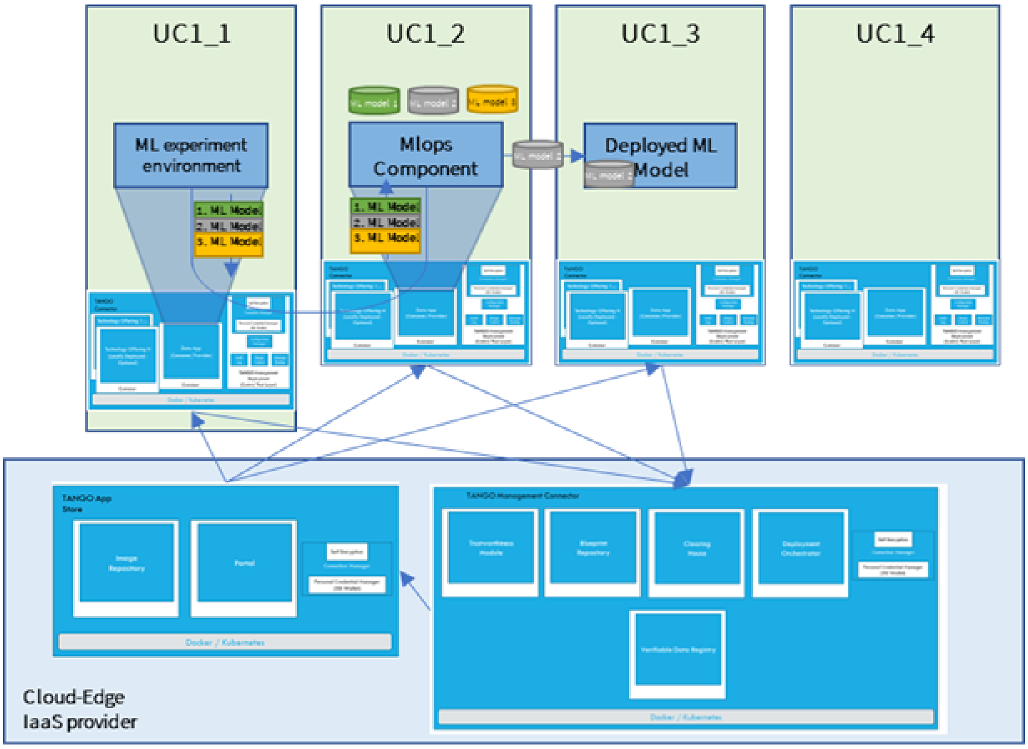


Figure 18. MLOps pilot deployment scenario.

3.2.2.5 Software artifacts

The MLOps component offers user interfaces from the chosen tool to handle the various stages of the life cycle of the ML models (MLFlow has a separate GUI and API). Use should be rather simple for ML developers (data scientists, ML system engineers). The MLOps component offers the foundation for creating, training, and providing ML models, and the user would interact via Notebooks (such as Jupyter Notebooks), Python code, or equivalent.

A deployment vision of the MLOps component is provided in Figure 19.

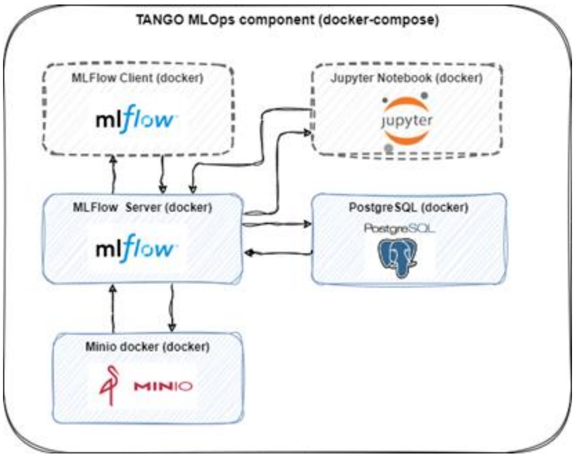


Figure 19. MLOps component docker approach.

By logging experiments in the same location, MLFlow allows multiple ML developers (data scientists, ML system engineers) using different machines to collaborate. Instead of keeping the information from MLFlow Tracking in a local mlruns folder, we save the parameters and metrics in a PostgreSQL database and the artifacts in MinIO object storage.

Finally, in Figure 19, the dotted lines represent two possibilities for interacting with the MLFlow server that might possibly extend the MLOps deployment: the MLFlow client and the Jupyter notebook.

3.2.2.6 Future work

The initial version of the MLOps component will provide support for experiment tracking and artifact management mainly.

Regarding to the functionalities expected for the second half of the project, the focus will be the integration with the authentication and authorisation approaches defined for TANGO ecosystem. Other functionalities to be taken into consideration could be the followings:

- Provision of an initial set of MLFlow Recipes.
- Jupyter notebook templates supporting a kind of generic pilots ML pipelines.
- Potential integration with other tools (i.e., with the results of the AutoML and Federated Learning components, the XAI library, the RENOPS or others).

3.2.3 Support for Federated Learning

In the field of AI, Federated Learning (FL) has emerged offering a transformative approach to training AI models while safeguarding data privacy and security. Traditional centralized ML approaches, where data is gathered and analysed at a central server, raise concerns about data breaches, misuse, and potential privacy violations. FL addresses these challenges by shifting the training process to individual devices or servers, ensuring that sensitive data remains decentralized and confidential.

Federated learning breaks free from the limitations of centralized ML, enabling the collaborative training of AI models on decentralized data. Unlike conventional methods, FL prohibits the direct sharing of raw data between participants. Instead, individual devices or servers train local copies of the model using their own data. These updated models are then exchanged with a central server, which aggregates the information and applies it to the global model. This process ensures that data remains securely encrypted on individual devices, preventing unauthorized access and data breaches.

Federated learning offers a plethora of advantages that extend beyond data privacy protection. It fosters collaboration among organizations or individuals while preserving data confidentiality, enabling the development of AI models with broader applicability. The decentralized nature of FL also facilitates the utilization of diverse datasets from multiple sources, enhancing model robustness and generalizability. This capability proves particularly valuable in healthcare, where sensitive patient data is often siloed across healthcare providers, hindering the development of comprehensive AI-powered solutions.

In the realm of finance, FL empowers fraud detection, credit scoring, and customer segmentation without compromising sensitive financial information. For retailers, FL enables personalized product recommendations, targeted advertising, and inventory management, all while upholding customer data privacy. The Internet of Things (IoT) stands to benefit from FL's ability to train models on edge devices, facilitating real-time predictive maintenance, anomaly detection, and traffic optimization. In the social media domain, FL revolutionizes language translation, sentiment analysis, and personalized recommendations, all while respecting user privacy concerns.

Federated learning stands as a transformative force in the AI landscape, providing a secure and collaborative framework for training AI models. As FL technologies mature, we can expect to see its widespread adoption across various industries, paving the way for a future where AI thrives without compromising data privacy and security.

3.2.3.1 Short description of the component

The component to be deployed in TANGO to give support for the Federated Learning functionality (presented in deliverable D2.3) is based on a new solution under development by the ATOS team. This new solution, named FLEVIDEN aims at providing Federated Learning capabilities offering a modular approach that allows a rich level of flexibility when creating FL-based scenarios.

This tool should be deployed in different client nodes, each one with their training dataset, and one that will serve as a central node (performing the aggregation of the weights of the different models trained. During the design of the federated learning scenario, the different containers deployed should be provided not only their corresponding training dataset, but also the reachable addresses of the different nodes they should communicate with and some model definition (to be shared by all the nodes involved).

3.2.3.2 Internal architecture

In the following diagram (Figure 20), the internal architecture of the Federated Learning component is depicted. As we can see, at the time of the delivery of this deliverable, two main offerings are to be provided: the server node (in green) and the client node (in orange). These nodes can be instantiated as many times as needed (one server node for several client nodes), and their reachable addresses should be provided as input parameter on deployment time.

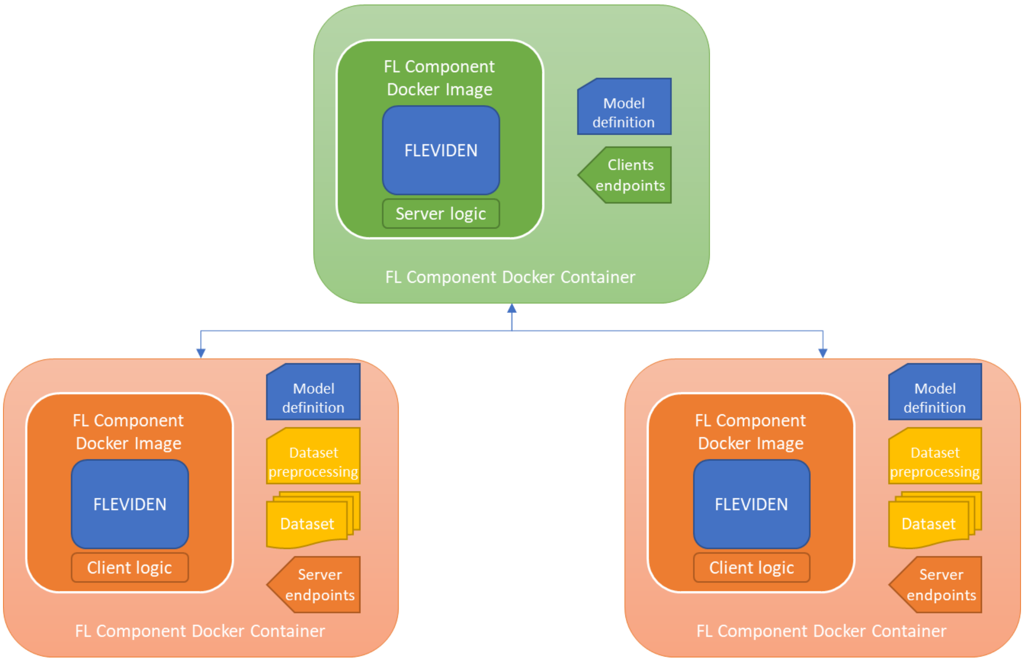


Figure 20. Federated Learning component internal architecture.

Figure 20 shows the different nodes involved in Federated Learning scenario in TANGO. All nodes should share the same model definition (depicted in blue) in order to perform the training over the same model architecture. Client nodes should include their corresponding dataset each one (in yellow) as well as the pointer to the server endpoint. Regarding the server node it should have the pointer to all the different clients in order to act as an aggregator for all the models independently trained in the scenario.

3.2.3.3 Features implemented

In order to give support for the FL feature in the context of TANGO, a solution based on a tool under development by the ATOS Research & Innovation area is being deployed in the project. This tool, named FLEVIDEN, aims to provide FL capabilities in a modular and encapsulated way, aiming to ease the creation of model training scenarios in the scope of FL.

Current version of the Federated Learning component offered in TANGO, allows the execution of a Federated Learning scenario where different nodes can stablish connection though HTTP protocol. Supported frameworks for model creation include Keras and, still in beta version, PyTorch.

3.2.3.4 Support for pilots

The Federated Learning is expected to be deployed on the following TANGO pilots:

- **Smart Manufacturing** pilot: (FMAKE): Allowing the distributed model generation from data obtained from different machines.
- **Banking** pilot: Allowing the model generation with no need of dataset sharing between involved entities.

3.2.3.5 Software artifacts

This component is expected to be offered as docker images to be deployed given the parameters mentioned in section 3.2.3.2. The images, one offered as server node and another as client node, are to be deployed as containers in different nodes with HTTP visibility between them. In the current version, the client node image-based containers should be deployed initially, and the server node image-based container last, as it will expect to perform the connection with the client nodes and maintain it until the training has finished in the different nodes.

3.2.3.6 Future work

As aforementioned, the Federated Learning component in TANGO is based on the FLEVIDEN core component, which is under development. Further step to carry out in the creation of this solution include:

- Provide further support for PyTorch-based models.
- Support other communication protocols.
- Allow further configurations in the communication protocols.
- Evaluate further support for newer pilot requirements.
- Potential integration of the results of the ML training cycles with the MLOps tool.

3.2.4 Support for AutoML

Automated machine learning (AutoML) is a revolutionary approach to machine learning that simplifies the process of building and training predictive models. This approach automates some key steps in the creation of a predictive model. By automating these intricate steps, AutoML allows even non-experts to get benefits from the power of machine learning, democratizing its accessibility and application.

This AutoML approach offers a set of advantages that make it an invaluable asset for individuals and organizations alike. Firstly, it significantly streamlines the machine learning workflow, saving developers valuable time and resources. AutoML algorithms efficiently handle data pre-processing, feature extraction, and model selection, reducing the manual effort required to build effective machine learning solutions. Secondly, AutoML consistently outperforms manually crafted models, leading to enhanced accuracy and predictive power. Additionally, AutoML broadens the reach of machine learning, empowering a wider range of users to tap into its capabilities. This democratization of machine learning fosters innovation and drives progress in various domains, from healthcare and finance to marketing and manufacturing.

3.2.4.1 Short description of the component

The TANGO AutoML component aims at providing AutoML capabilities for easing the model training in the generation of predictive models. This component will help in the set-up phase of a model by means of automating the initial model creation phase, comparing several ML-based model approaches in order to decide the one offering the best results. Current version of the component allows to select a subset of models to evaluate depending on some major characteristics of them (i.e. explainability, performance or all). Additionally, the evaluation metric can be selected in order to perform the model selection based on that metric.

For the component to work, the end-user should provide a csv file with the training data, select the features that will be taken into account, and the target label (the feature to be predicted). After that, the end user will be able to launch the AutoML process that will provide a model as result as well as the results of the different models evaluated.

3.2.4.2 Internal architecture

The AutoML component provided in TANGO is based on the functionalities of the MLJAR[6] library in order to ease the initial evaluation of several ML-based models. This tool eases the automation of these initial model selection and tests, offering, as result, a report with the comparison of the different models evaluated. In order to ease the use of the tool, this initial version is being served as a webApp by means of using the Mercury framework[7]. This framework allows the "translation" of python notebooks to webapps, allowing the end user to interact with these notebooks by means of a GUI. Different frameworks will be assessed for the next iterations.

[6] https://mljar.com/

[7] https://mercury.mljar.com/

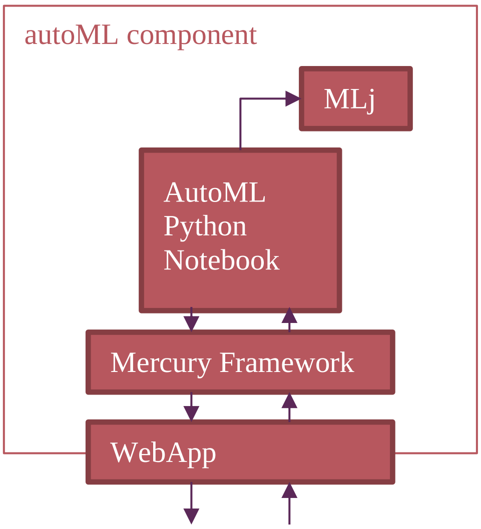


Figure 21. AutoML component internal architecture

As can be seen in Figure 21, AutoML functionality provided by MLJAR is defined by an AutoML notebook. The mercury framework imports the notebooks on deployment time and serves them as webapps to be accessed by the end user. These apps ease the interaction with the notebook, allowing the execution of the AutoML process interacting with the webapp GUI and without needing to edit a line of code.

3.2.4.3 Features implemented

The deployed solution based on MLJAR and Mercury allows:

- Performing initial training on a user-provided dataset using different ML-based models.
- Training data must be provided as a CSV file.
- Get the model comparison report in order to check the different performance of each model.
- Perform predictions using a model generated by the AutoML component and new data entries.

3.2.4.4 Support for pilots

This component can be potentially used by any pilot aiming at performing supervised learning-based model prediction. Although the current implementation is more oriented for pilots handling non-large-scale datasets, this tool is very appropriate to perform an initial approach to evaluate the fitness and performance of a ML-based model on a given initial dataset.

3.2.4.5 Software artifacts

The solution provided will be deployed using docker. This ensures the proper working of the component in different environments and eases its deployment. The current solution deployment will require to have an open port (by default 80) to be accessed via web browser. Additionally, the machine where the solution will be deployed requires to be accessible from the device where the end user will access the component.

3.2.4.6 Future work

Due to the lack of specific requirements for AutoML from the TANGO pilots, the solution provided in this initial version of the component is a proof of concept to be evaluated and improved in the next releases based on the feedback of the users. The following issues have been identified to be addressed in next versions of the component:

- Improve deployment strategy (by means of improving scalability of the solution) within the docker container (or by means of using docker compose in order to serve the deployment of the component as a set of different containers).
- Perform customized development on the notebooks served based on pilots' needs.
- Evaluate other approaches or frameworks to extend the AutoML functionality.
- Evaluate other approaches for serving the AutoML solutions (apart from Mercury framework).
- Potential integration of the results with the MLOps tool.

3.3 Dynamic Intelligent Execution on Heterogeneous Systems [T5.3]

3.3.1 Introduction

Hardware acceleration has emerged as a mean to increase performance and power efficiency of software applications. In the context on TANGO, Task 5.3 aims to deliver automated execution plans on the hardware accelerator devices that will be available on a single node. The work entailed in this task is focused on TornadoVM; an open-source software technology originated from the University of Manchester to bridge the gap between managed programming languages (e.g., Java) and hardware acceleration.

3.3.2 TornadoVM

This paragraph aims to describe the core technology that is extended in Task 5.3 to deliver the ultimate objective of devising automated execution plans for hardware accelerators. TornadoVM is a Java software library that exposes a Java API for developers to define the relevant functions that are meant to be compiled for hardware acceleration. Hence, it contains a compiler that extends the Graal compiler and contains at its backbone three backend that generate kernels of two types: OpenCL (source code), PTX (source code), and SPIR-V (binary code). The generated artifacts can be generated on different hardware accelerators' architectures, such as multi-core CPUs, GPUs and FPGAs. The architecture of TornadoVM has been already introduced in Deliverable D2.3 (Section 5.3).

3.3.2.1 Short description of the component

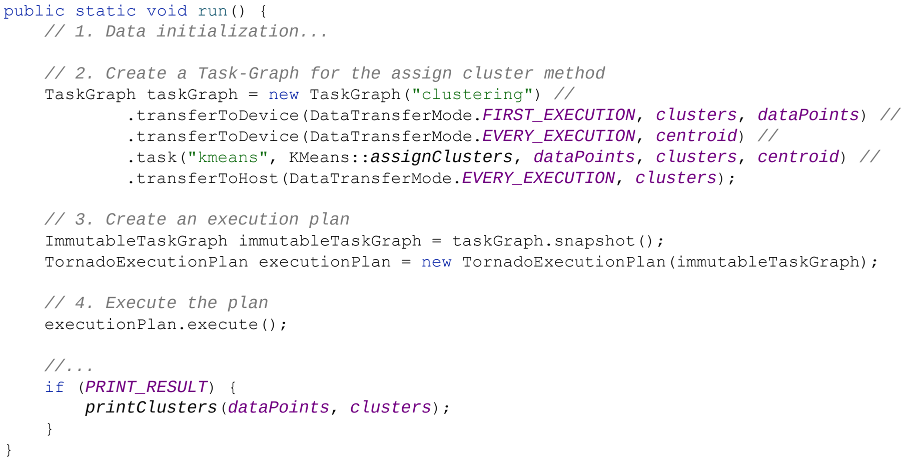
The component of Task 5.3 utilizes two key sub-components:

- TornadoVM: to express in Java the functions that are going to be supported for hardware acceleration.
- GraalPython: the GraalVM implementation of Python that is shipped as a standalone toolkit that can interoperate with TornadoVM via JVM in order to deliver functionality implemented in Java classes to Python programs.

The component of Task 5.3 is going to be added in the TANGO architecture as a docker container that will accept as input a function, and in turn it will execute it on the most efficient hardware accelerator (i.e., multi-core CPU, GPU, etc.) that is available on the system.

3.3.2.2 Internal architecture

A more elaborate overview of the software component that will offer the dynamic and intelligent execution on heterogeneous systems is shown in the figure below. The containerized software encompasses the accelerated versions of the functions (namely, Accelerated Library) that can be invoked by the EXUS EDAE. At this stage, a first function that we have integrated for the first prototype is k-means; the most popular unsupervised machine learning algorithm. The accelerated version of k-means is implemented in Java (my.accelerated.library.Kmeans) and employs the TornadoVM API to express the parallel implementation of the algorithm. We followed the guidelines of how to use the API as described in the TornadoVM documentation[8]. The Java class contains a method run, which is exposed also to the Python program to invoke the execution of k-means. The implementation of the run method is as depicted in Figure 22:



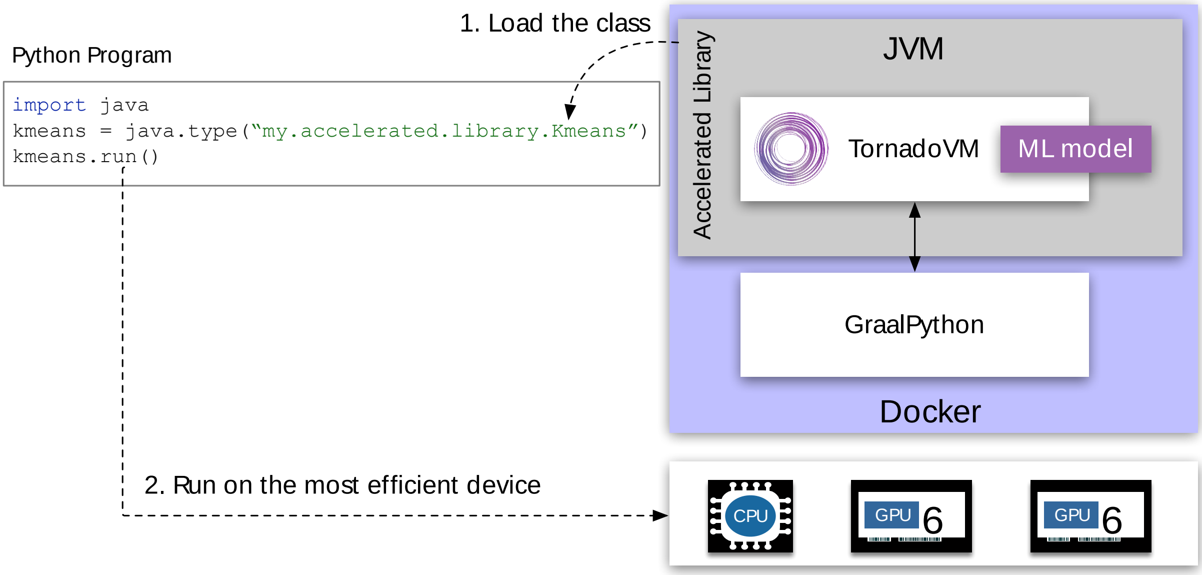
```
public static void run() {
    // 1. Data initialization...

    // 2. Create a Task-Graph for the assign cluster method
    TaskGraph taskGraph = new TaskGraph("clustering") //
            .transferToDevice(DataTransferMode.FIRST_EXECUTION, clusters, dataPoints) //
            .transferToDevice(DataTransferMode.EVERY_EXECUTION, centroid) //
            .task("kmeans", KMeans::assignClusters, dataPoints, clusters, centroid) //
            .transferToHost(DataTransferMode.EVERY_EXECUTION, clusters);

    // 3. Create an execution plan
    ImmutableTaskGraph immutableTaskGraph = taskGraph.snapshot();
    TornadoExecutionPlan executionPlan = new TornadoExecutionPlan(immutableTaskGraph);

    // 4. Execute the plan
    executionPlan.execute();

    //...
    if (PRINT_RESULT) {
        printClusters(dataPoints, clusters);
    }
}
```

Figure 22. Implementation of the run method in TornadoVM

The first step contains the definition of the TornadoVM data structures for the representation of the data points, the centroid and the clusters. The second step creates the TornadoVM TaskGraph, which is an object that defines which data should be moved on the accelerator (i.e., transferToDevice, transferToHost) along with which functions (namely, tasks) should be offloaded on the accelerator. The next step creates a snapshot of the current function implementation in order to be able to compile it to a GPU version without any mutations triggered during compilation. At this stage a snapshot is also associated with an execution plan, which can contain one or multiple snapshots of TaskGraphs. Finally, the execution plan is executed, which will trigger the compilation from Java to OpenCL (the first time), and the actual execution on an accelerator (e.g., a GPU).

Additionally, the container encompasses the python implementation of GraalVM (namely, GraalPython) which enables the interoperability between Java classes and Python programs. Thus, a Python program that runs in the provided container is able to load the Java classes of the Accelerated Library, as shown in the figure below. Once the class is loaded, GraalPython enables the execution of the methods defined in the Java classes from the Python runtime implementation that runs on top of the JVM. At this point, it is important to mention that to achieve interoperability between TornadoVM and GraalPython, both software should be compatible and support the same Graal compiler (Graal 23.1.0, at the moment). The steps are depicted in Figure 23.

[8] https://tornadovm.readthedocs.io/en/latest/programming.html



```
import java
kmeans = java.type("my.accelerated.library.Kmeans")
kmeans.run()
```

Figure 23. Steps in TornadoVM to run accelerated code from Python code

3.3.2.3 Features implemented

Since the beginning of Task 5.3 we have undertaken several fundamental features that are vital for the successful delivery of the defined functionality. Some of the features have been up-streamed in the main TornadoVM repository:

- The Multiple Tasks on Multiple Devices feature, which enables the concurrent execution of various functions on all available heterogeneous accelerators (CPUs, GPUs, FPGAs).
- The integration with the Graal 23.1.0 version which makes TornadoVM compatible with JDK 21 and the polyglot runtime implementations, such as GraalPython, GraalJS, TruffleRuby, etc.

Additionally, we have collaborated in the joined discussions in the context of other work packages. In particular, we have been in contact with the pilots in order to comprehend the requirements of the use cases and seek the means to deliver our technology to them. At the same time, we have participated in discussions for the definition of the system requirements for the TANGO architecture as well as how our component will be integrated and interoperate with other software components in the TANGO architecture (see Section 5.3 in D2.3).

3.3.2.4 Support for pilots

The technology offering of Task 5.3 will be indirectly consumed by TANGO pilots that employ the EXUS Data Analytics Engine offered in Task 5.1. The reason that the component of Task 5.3 is coupled with the offering of T5.1 is that the requirements of the TornadoVM component are not satisfied by the characteristics of the TANGO use cases pilots. More specifically:

1. No TANGO use case presents significant data processing that can justify the employment hardware acceleration to increase performance/power efficiency.
2. No TANGO use case uses Java as the main programming language.

Thus, it has been decided to expose the scientific objectives of Task 5.3 in a different programming language Python; by integrating TornadoVM with GraalPython. This decision enabled the consortium to pivot and demonstrate the offering of Task 5.3 for the acceleration of various data analytics functions shipped with the EXUS EDAE. Hence, Task 5.2 may be indirectly by the use cases that are going to be supported by EDAE (see Section 3.1.2.4).

3.3.2.5 Software artifacts

The current state of the docker container that is used to develop and test the component for Task T5.3 is available in dockerhub, and will be ported to the project docker registry:

- https://hub.docker.com/repository/docker/beehivelab/tornadovm-polyglot-graalpy-23.1.0-nvidia-opencl-container/general.

The docker container is built based on two specific artifacts that are open source:

- TornadoVM: https://github.com/beehive-lab/TornadoVM
- GraalPython: https://github.com/oracle/graalpython/tree/graal-23.1.0

3.3.2.6 Future work

In the following months, we plan to extend TornadoVM with new functionality that includes:

- The integration of power metrics in the runtime layer used in our task (TornadoVM Runtime). This step will enable TornadoVM to monitor the energy class of each hardware accelerator.
- Initiate the design of a ML model that will be able to predict which heterogeneous device will be the most energy efficient. This model is planned to be integrated within the TornadoVM Runtime at the second half of the project.
- Extend the coverage to support more functions for the EXUS EDAE that is described in T5.1. The selection of the candidate functions that can be accelerated on heterogeneous hardware will be decided in collaboration with EXUS.

3.4 Privacy Threat Modelling & Identification for Trustworthy AI [T5.4]

3.4.1 Introduction

As Task 5.4 harbours two components, Privacy Enhancing Component (PEC) by UOG and Privacy Assurance Tool (PAT) by FUJ_LU, this section provides an overview of each component separately highlighting the problem that a component aims to solve along with the need for solving the problem.

Privacy Enhancing Component (PEC):

Privacy Risk Assessment, alternatively known as Privacy Impact Assessment, is a systematic process used to identify, evaluate and address privacy risks associated with the collection, use and handling of personal and sensitive data. Privacy risk assessment can help organisations to take decisive actions against privacy-related concerns. Addressing these concerns ensure businesses comply with privacy regulations and accommodate data privacy requests from consumers. Thus, organisations that conduct privacy risk assessments are more likely to avoid legal and business implications of non-compliance and ensure a long-term trustworthy relationship with their customers. Last but not the least, privacy risk assessment can significantly complement data protection impact assessments which is a requirement from GDPR for businesses using Personal Identifiable Information. The Privacy Enhancing Component (PEC) aims to offer a way of performing privacy impact assessments while recommending a list of safeguards to mitigate potential privacy threats.

Privacy Assurance Tool (PAT):

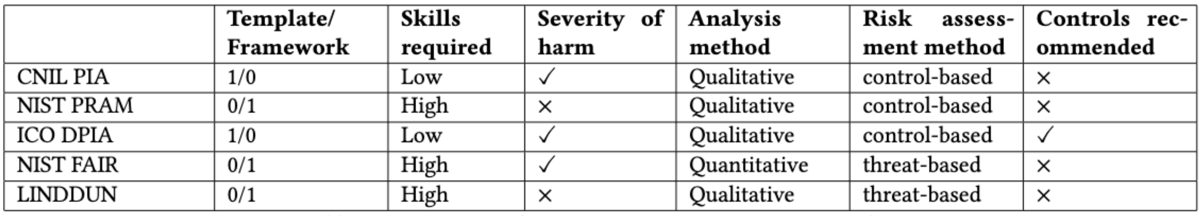
Nowadays, global organizations are transitioning to electronic records. This shift offers a dependable and easy method for these organizations to share and access information about data subjects. Therefore, it's essential to establish systems that allow data subjects to securely exchange their information, while also giving them the autonomy to choose the type of data they wish to share. This highlights the necessity for a secure and privacy-preserving data exchange and record system that aligns with GDPR regulations, bridging the gap between a data subject and a data requestor. The challenge is to explore the ways in which organizations can guarantee that the data shared by the data subject is managed respectfully and in accordance with GDPR regulations.

PAT a broker framework which protects data subjects' information during data exchanges. To achieve security and privacy-preservation for information exchange between an organisation and data subject. The PAT includes a consent-based access control (CBAC) mechanism. A consent is created by an organisation and sent to the data subject. A consent is an authorization initiated by a data subject for an intended data requester via an agreement between them. After obtaining the consent from the data subject, an organization can gain access to the data, which is encrypted by a PAT. Data subject data is protected against access of unauthorized parties. Furthermore, PAT assures the authentication and privacy-preserving, consent-preserving, consent compliance by provide access control, rules engines, and audit logs.

3.4.2 Privacy Enhancing Component (PEC)

The Privacy Enhancing Component (PEC) is main tool of the TANGO platform that computes privacy risk of a cyber threat, identifies measures to mitigate the privacy risks and provides guidance on GDPR compliance. PEC is a decision-support tool that advices the end-users on their privacy risks and how the privacy risks can be effectively managed. It is supported by a Core Engine and a Dashboard. The core engine identifies privacy threats, determines possible impacts, computes privacy scores, defines measures and evaluates compliance. While the Dashboard visualises the privacy risks and provides guidance on mitigating privacy threats. Table 2 presents an overview of existing privacy impact assessment approaches highlighting their general characteristics.

Table 2: Overview of Privacy Impact Assessment Methods



	Template/ Framework	Skills required	Severity of harm	Analysis method	Risk assessment method	Controls recommended
CNIL PIA	1/0	Low	✓	Qualitative	control-based	×
NIST PRAM	0/1	High	×	Qualitative	control-based	×
ICO DPIA	1/0	Low	✓	Qualitative	control-based	✓
NIST FAIR	0/1	High	✓	Quantitative	threat-based	×
LINDDUN	0/1	High	×	Qualitative	threat-based	×

PEC builds upon the existing concepts presented here to provide a more comprehensive approach to performing privacy impact assessment. It enhances the existing approaches by considering cyberattacks for privacy risk assessment while linking it to privacy impact categories, enterprise loss and possible harm to individual. This consideration addresses a major gap in existing practices which is to integrate privacy risk assessment with cyber risk management. Using PEC, organisations will be able to better identify and manage their privacy threats. Further, viewing cyber threats from a privacy lens can help organisations understand and prioritise risks promoting better resource allocation and decision making.

3.4.2.1 Short description of the component

PEC aims to facilitate a way of documenting processes that handles personal and sensitive data. In the process of supporting the documentations, it allows users to perform privacy impact assessment of data processes handling personal and sensitive data. The privacy impact assessment highlights threats that could lead to privacy breach and potentially harm individuals. The main functionalities of PEC include privacy impact assessment leading to identification of privacy threats and privacy impact scores, assessment of privacy safeguards to mitigate the identified privacy threats, and insights into compliance with GDPR through implementation of the recommendation safeguards.

PEC will reside in the TANGO App Store that allows TANGO stakeholders, users and providers to discover and access the technology offering. PEC will be offered as a web application allowing user to enter requisite data (data actions, processes, type of data etc) to perform privacy impact assessment and visualise the privacy risks and suggested recommendation through the dedicated PEC dashboard. PEC will be served as an essential graphical user interface for TANGO users, allowing them to conveniently record and inspect their privacy impact assessment reports. This will be achieved by deploying the PEC container image on the TANGO App Store allowing the users to access the PEC web application at will.

Various key performance indicators will be assessed to validate the applicability of PEC. These KPI include:

- Assessing how satisfied the users are with using the software through usability testing and feedback.
- Assess the performance of the software in terms of response time.
- Implementation of security features to reduce common security risks and enforce better coding practices.
- Unit testing, functional testing and integration testing to identify logical and functional bottlenecks.

3.4.2.2 Internal architecture

PEC consists of five components. These components work in synergy to deliver the outcome of PEC which is a list of **privacy risk scores**, privacy **impacts** and a set of **recommendations** to mitigate the privacy risks. The outcome is visualised to the user through the Dashboard to gain insights into the privacy risks and actionable advice. The PEC components, presented in Figure 24, are:

- PEC Data Initialiser (PDI)
- PEC Risk Identifier (PRI)
- PEC Scoring Engine (PSE)
- PEC Recommendation Engine (PRE)
- PEC Compliance Engine (PCE)

A user initiates the privacy risk assessment process by providing the requisite data to the **PEC Data Initialiser (PDI)**. PDI assists the user in determining whether a privacy impact assessment is necessary based on the type of data the user aims to collect, process or consume. Once the context around the use of data has been identified - in the form of data actions and associated components, and the need of a privacy risk assessment, as required by GDPR, is assessed, the **PEC Risk Identifier (PRI)** identifies privacy threats and possible impact categories associated with the provided data action. For example, the privacy impact categories considered for the Autonomous Vehicle use case are: (i) Identify theft and financial fraud; (ii) Unauthorised access to personal and sensitive data; (iii) Interception of communication; (iv) Vehicle tracking and location history access, and (v) Alteration of non-critical vehicle settings.

PRI also established a mapping between the identified privacy threats and the NIST Privacy Engineering Objectives which further assist with integration with the NIST Privacy Risk Assessment Methodology (PRAM). NIST IR 8062[9] presents three privacy engineering objective which are *Predictability*, *Manageability* and *Disassociability*. Predictability as the ability that enables reliable assumptions about individuals, owners, and operators based on the processing of personal information. Manageability as the ability that enables granular administration of personal information including alteration, deletion, and selective disclosure. While Disassociability is enabling the processing of personal information or events without associating them to individuals or devices beyond the operational requirements. These three objectives are considered as the key reference points of privacy in PEC which must be met to reduce privacy risks and protect privacy at scale.

Next, on receiving the inputs from PRI, the **PEC Scoring Engine (PSE)** has two main tasks: the first is to identify problematic data actions and problems for individual from the NIST PRAM associated with the identified privacy threats, and second is to compute the privacy risk scores considering the likelihood and potential impact of a privacy threat. NIST PRAM identifies nine problematic data actions that could lead to potential harm to an individual. The problematic data actions are: (i) *Appropriation* (AP) includes scenarios in which data is used in ways that exceed individual's expectation or authorisation; (ii)

[9] https://csrc.nist.gov/pubs/ir/8062/final

Distortion (DI) refers to the use or dissemination of inaccurate or misleading data; (iii) *Induced Disclosure* (ID) refers to scenarios in which individuals feel compelled to provide information disproportionate to the purpose or outcome of the transaction. Induced disclosure can include leveraging access or rights to an essential (or perceived essential) service; (iv) *Data Insecurity* (IN) resulting in a breach of confidentiality and integrity of personal data; (v) *Re-identification* (RE) refers to scenarios where data from multiple sources can be associated or identified to a specific individual; (vi) *Stigmatisation* (ST) refers to the scenario in which data is linked to an actual identity in such a way as to create a stigma; (vii) *Surveillance* (SU) refers to scenarios in which data, devices and individuals are tracked or monitored in a manner disproportionate to the purpose leading to an adverse situation for individuals or groups; (viii) *Unanticipated Revelation* (UR) refers to situations in which data in revealed or exposed in unexpected ways; (ix) *Unwarranted Restriction* (WR) includes not only blocking access to data or services. but also limiting awareness of the existence of data or its use in ways that are disproportionate to operational purposes.

The harm associated with these problematic data actions include: (i) *Dignity Loss* that includes embarrassment and emotional distress; (ii) *Discrimination* that covers unfair or unethical differential treatment of individuals or at-risk groups arising from the processing of data; (iii) *Economic Loss* that includes direct financial losses as the result of identity theft or the failure to receive fair value in a transaction; (iv) *Loss of Autonomy* that includes losing control over determinations about information processing or interactions with systems, products or services, as well as needless changes in ordinary behaviour, including self-imposed restrictions on expression or civic engagement; (v) *Loss of Liberty* that covers impacts from incomplete or inaccurate data which can lead to improper exposure to arrest or detainment and/or improper exposure or use of information to abuse governmental power; (vi) *Physical Harm*; and (vii) *Loss of Trust* that includes the breach of implicit or explicit expectations or agreements about the processing of data which could lead to diminishing morale or leave individuals reluctant to engage in future transactions potentially creating larger economic or civic consequences.

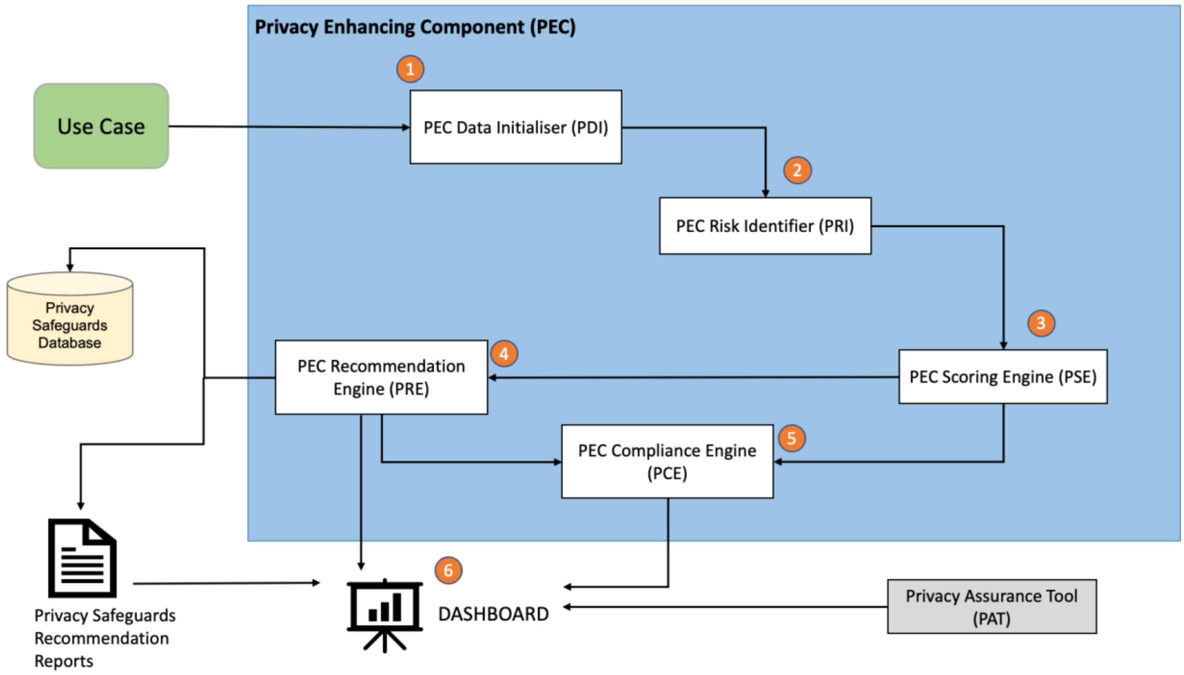

Privacy Enhancing Component (PEC)
Use Case
1
PEC Data Initialiser (PDI)
2
PEC Risk Identifier (PRI)
3
PEC Scoring Engine (PSE)
4
PEC Recommendation Engine (PRE)
5
PEC Compliance Engine (PCE)
Privacy Safeguards Database
Privacy Safeguards Recommendation Reports
6
DASHBOARD
Privacy Assurance Tool (PAT)

Figure 24. PEC Architecture with internal components

To compute the privacy risk score, PSE builds upon existing databases, industry reports and in-house knowledge bases to define the likelihood and loss magnitude of an attack against the privacy impact categories. The Likelihood of Impact (FI) represents the probability of a successful event leading to the violation of privacy and causing specific harm to the individual or group. It is defined in the range of 1 to 5, where 1 be less likely and 5 being most likely. Whereas Loss Magnitude (L) expresses the potential business impact from an adverse event. It is composed of three categories of loses: *(i) Legal Fines; (ii) Financial Cost, and (iii) Reputation Cost*. The loss magnitude is a quantitative value mapped to a 3-tier scale value ranging between "Low = 1", "Medium = 2" and "High = 3". The Loss Magnitude (L) is obtained by adding the factors altogether.

$$L = \sum \left\{ \text{Legal fines, Financial Cost, Reputation Cost} \right\}$$

The risk score presents the privacy risk which is the likelihood of impact (FI) times loss magnitude (L). Mathematically, we define risk as the inner product of these two factors. The final score will be normalised to 0-10 range.

$$\begin{aligned} \textbf{Privacy Risk} &= \langle FI \cdot L \rangle \\ &= [FI_1 \times L_1, FI_2 \times L_2, \ldots, FI_r \times L_r] \end{aligned}$$

Table 3 presents the privacy risk range used to compute the scores.

Table 3: Privacy Risk Range

Privacy Risk Range	Qualitative Value
0 – 2	Very Low
2 – 4	Low
4 – 6	Medium
6 – 8	High
8 – 10	Very High

Once the privacy risk scores have been computed, the **PEC Recommendation Engine (PRE)** identifies the list of safeguards form the Privacy Safeguards Database to mitigate the privacy risks. The **Privacy Safeguards Database** holds a list of safeguards that the PEC uses to mitigate the identified privacy risks. For TANGO, we have considered the NIST Privacy Controls as the list of safeguards considered for PEC. The final recommendations that the PEC provides are derived from the list of safeguards that could be used to mitigate the privacy threats tailored to a pilot use case. Along with a list of safeguards that could be implemented to address the privacy risks, the PRE recommendations provide insights on possible harms that the risk event could lead to and how implementing the recommended safeguard would mitigate the privacy impacts. Finally, the **PEC Compliance Engine (PCE)** collates and analyses the recommendations from PRE to generate insights and correlation between GDPR and the list of recommendations.

The **PEC Dashboard** is a graphical user interface that provide a means of visualising the outcomes of a privacy impact assessment process. The digital dashboard permits users to assess their privacy risks and safeguards required to mitigate the potential privacy impacts. The PEC Dashboard is mainly used for three purposes:

- To communicate to the end user the privacy risk, possible privacy impacts and recommendations to mitigate the privacy risks,
- To record the privacy impact assessment process and visualise and prioritise the privacy risks, and
- Provide insights and guidance on complying with GDPR.

The technologies used to build the PEC Dashboard includes:

- **Django Framework**: Django is a high-level web framework that enables development of secure and maintainable web applications. Django follows the Model-Template-View (MTV) architectural pattern, which is like the Model-View-Controller (MVC) architecture. It separates data handling (Model), user interface (Template), and business logic (View), promoting clean, readable, and reusable code. Django's Object-Relational Mapping (ORM) allows one to interact with the database using Python code instead of SQL. This abstraction enables developers to work with various databases and perform database operations without writing backend-specific SQL. Django automatically generates a powerful and production-ready administrative interface that can be used to manage the content of your site. Finally, Django emphasizes security and helps developers avoid many common security mistakes, such as SQL injection, cross-site scripting, cross-site request forgery, and clickjacking. Its user authentication system provides a secure way to manage user accounts and passwords. These features made Django an ideal choice to develop PEC for TANGO. Another consideration is that Django provides efficient ways of building RESTful APIs making it a great choice for backend development in web services and mobile applications.
- **PostgreSQL**: PostgreSQL, often simply called Postgres, is an open-source, advanced, and feature-rich relational database management system (RDBMS). It's known for its robustness, scalability, and performance and is used widely across various types of applications. Postgres is used to store the PEC models and user inputs.
- **HTML, CSS and JavaScript**: HyperText Markup Languages, Cascading Style Sheets and JavaScript are vital elements for design and production of the dashboard. Within the PEC, these are used as the building blocks for styling, visualisation (charts, graphs etc) and functionalities.
- **Bootstrap**: Bootstrap is a free and open-source front-end framework for developing responsive, mobile-first websites and web applications. It's one of the most popular HTML, CSS, and JavaScript frameworks for creating dynamic and visually appealing user interfaces. PEC uses Bootstrap to build a responsive design ensuring that the web pages and dashboard adjust smoothly to different screen sizes and devices, are compatible with modern browsers, and are customisable to meet the project visual requirements.
- **Large Language Models (LLMs)**: LLMs are advanced artificial intelligence systems designed to understand, generate and interact with human languages. This project uses LLMs to understand context and process natural language to derive threat scenarios and to support the overall functionality of PEC.

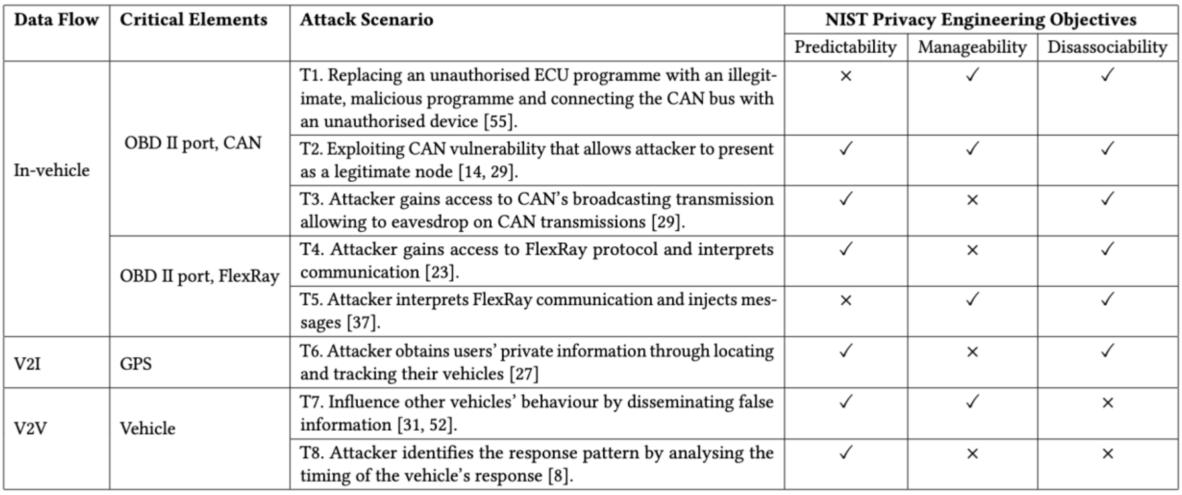
3.4.2.3 Features implemented

The scope of PEC for this submission includes two of the three core features to be developed during TANGO. These functionalities assist to assess privacy impact from an organisation’s handling of personal and sensitive data. The developed features include:

1. Perform Privacy Impact Assessment: Allowing users to enter contextual data on their organisation’s processing and handling of data to perform privacy impact assessment. On receiving the user data, PEC identifies potential threats that could lead to privacy harm, computes privacy scores for the potential threats with insights on potential problem for individuals that this would lead to.

Here, we present a small working example of the PEC for the Autonomous Vehicle uses case with underline mapping and a sample application of PEC to compute privacy threats, privacy scores and possible impacts. We assume that the user has provided the requisite pre-context information regarding the processing of data in this use case. Using the information PEC derives possible threat scenarios and performs a mapping of each threat scenario to the NIST Privacy Engineering Objectives. A version of the result is presented in Table 4.

Table 4: Data flow elements with attack scenario mapped to NIST Privacy Engineering Objectives

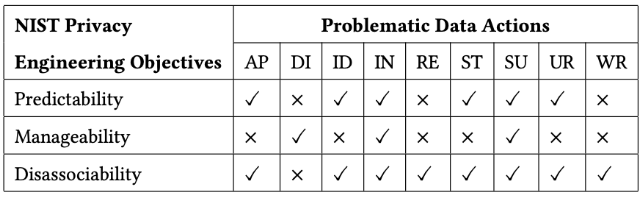


Data Flow	Critical Elements	Attack Scenario	NIST Privacy Engineering Objectives		
			Predictability	Manageability	Disassociability
In-vehicle	OBD II port, CAN	T1. Replacing an unauthorised ECU programme with an illegitimate, malicious programme and connecting the CAN bus with an unauthorised device [55].	×	✓	✓
		T2. Exploiting CAN vulnerability that allows attacker to present as a legitimate node [14, 29].	✓	✓	✓
		T3. Attacker gains access to CAN's broadcasting transmission allowing to eavesdrop on CAN transmissions [29].	✓	×	✓
	OBD II port, FlexRay	T4. Attacker gains access to FlexRay protocol and interprets communication [23].	✓	×	✓
		T5. Attacker interprets FlexRay communication and injects messages [37].	×	✓	✓
V2I	GPS	T6. Attacker obtains users' private information through locating and tracking their vehicles [27]	✓	×	✓
V2V	Vehicle	T7. Influence other vehicles' behaviour by disseminating false information [31, 52].	✓	✓	×
		T8. Attacker identifies the response pattern by analysing the timing of the vehicle's response [8].	✓	×	×

Next is to map NIST Privacy Engineering Objectives to NIST PRAM Problematic Data Actions such as Appropriation (AP) in the context of privacy refers to the unauthorised use of an individual's data for purpose other than those for which the data was originally collected. Unauthorised use of data can allow the attacker to have reliable assumptions about the entity as well as associating events and actions to an entity. Appropriation, thus, can affect predictability and disassociability. Unanticipated revelation (UR) in the context of privacy refers to the unexpected disclosure or exposure of information that was not meant to be shared. The unexpected revelation of information can allow the attacker to have reliable assumptions about the entity's behaviour or characteristics as well as can associate actions with an entity affecting predictability and disassociability.

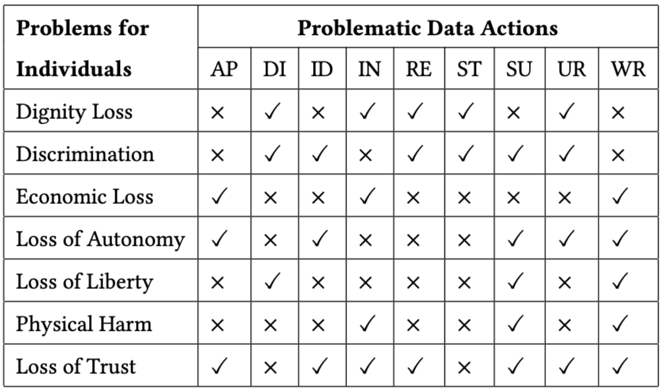
On the other hand, distortion (DI) which refers to the manipulation or modification of information will affect the manageability metric. A similar assessment is performed for all the rest of the problematic data actions and the mapping is presented in Table 5, while Table 6 presents the mapping between the problematic data actions and problems for individual.

Table 5: Mapping NIST Privacy Engineering Objectives and NIST Problematic Data Actions



NIST Privacy Engineering Objectives	Problematic Data Actions								
	AP	DI	ID	IN	RE	ST	SU	UR	WR
Predictability	✓	×	✓	✓	×	✓	✓	✓	×
Manageability	×	✓	×	✓	×	×	✓	×	×
Disassociability	✓	×	✓	✓	✓	✓	✓	✓	✓

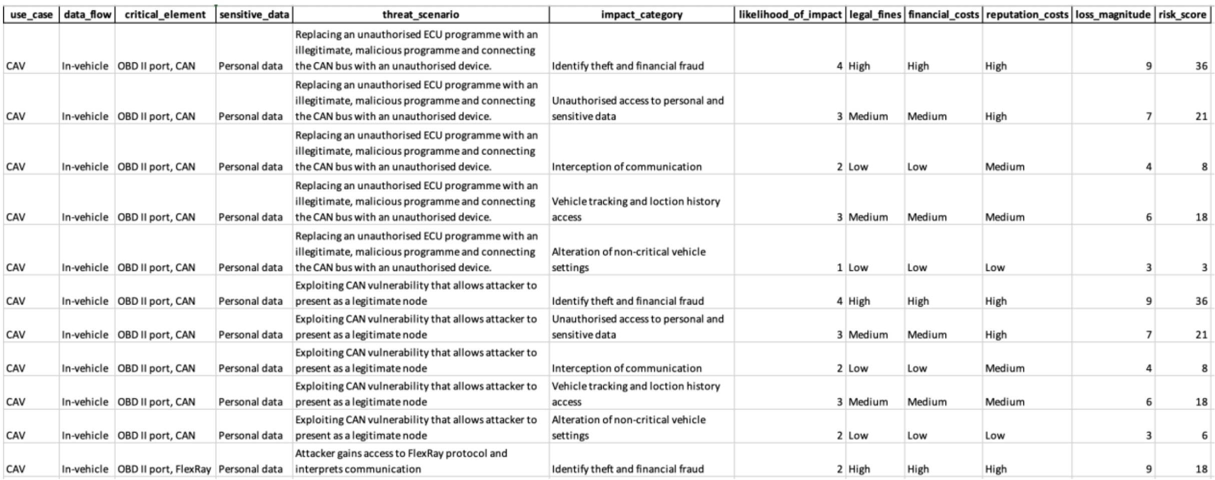
Table 6: Mapping NIST problems for individual to NIST Problematic Data Actions



Problems for Individuals	Problematic Data Actions								
	AP	DI	ID	IN	RE	ST	SU	UR	WR
Dignity Loss	×	✓	×	✓	✓	✓	×	✓	×
Discrimination	×	✓	✓	×	✓	✓	✓	✓	×
Economic Loss	✓	×	×	✓	×	×	×	×	✓
Loss of Autonomy	✓	×	✓	×	×	×	✓	✓	✓
Loss of Liberty	×	✓	×	×	×	×	✓	×	✓
Physical Harm	×	×	×	✓	×	×	✓	×	✓
Loss of Trust	✓	×	✓	✓	✓	×	✓	✓	✓

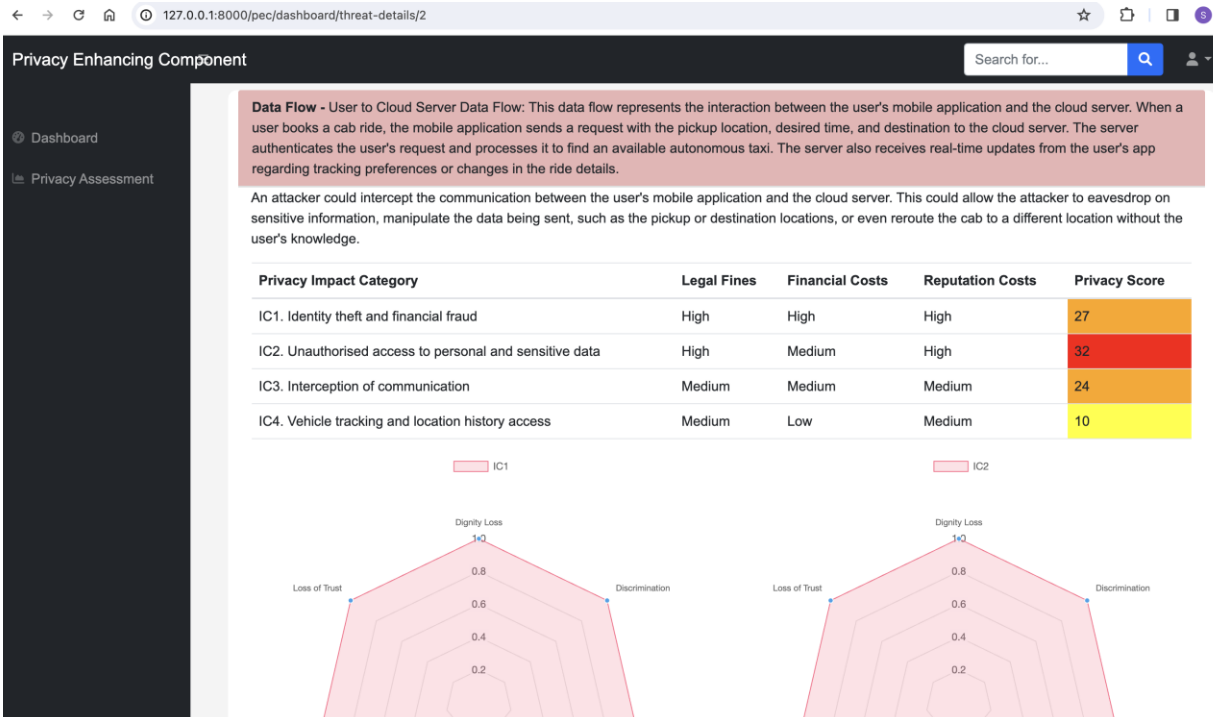
The PEC Scoring Engine aggregates all the information from other components (see section 3.4.2.2) and undertakes the calculation of the privacy scores. The score is calculated per privacy impact category against each loss category (see the five impact categories and three loss categories listed in section 3.4.2.2). Note that the privacy score is bounded between 0 and 50 which is obtained by multiplying the likelihood of impact to the loss magnitude. The final privacy score is normalised between the range 0 and 10. Table 7 presents a sample of the final risk score computed using PEC.

Table 7: Final risk score of a threat scenario with values for impact categories and loss categories



use_case	data_flow	critical_element	sensitive_data	threat_scenario	impact_category	likelihood_of_impact	legal_fines	financial_costs	reputation_costs	loss_magnitude	risk_score
CAV	In-vehicle	OBD II port, CAN	Personal data	Replacing an unauthorised ECU programme with an illegitimate, malicious programme and connecting the CAN bus with an unauthorised device.	Identify theft and financial fraud	4	High	High	High	9	36
CAV	In-vehicle	OBD II port, CAN	Personal data	Replacing an unauthorised ECU programme with an illegitimate, malicious programme and connecting the CAN bus with an unauthorised device.	Unauthorised access to personal and sensitive data	3	Medium	Medium	High	7	21
CAV	In-vehicle	OBD II port, CAN	Personal data	Replacing an unauthorised ECU programme with an illegitimate, malicious programme and connecting the CAN bus with an unauthorised device.	Interception of communication	2	Low	Low	Medium	4	8
CAV	In-vehicle	OBD II port, CAN	Personal data	Replacing an unauthorised ECU programme with an illegitimate, malicious programme and connecting the CAN bus with an unauthorised device.	Vehicle tracking and loction history access	3	Medium	Medium	Medium	6	18
CAV	In-vehicle	OBD II port, CAN	Personal data	Replacing an unauthorised ECU programme with an illegitimate, malicious programme and connecting the CAN bus with an unauthorised device.	Alteration of non-critical vehicle settings	1	Low	Low	Low	3	3
CAV	In-vehicle	OBD II port, CAN	Personal data	Exploiting CAN vulnerability that allows attacker to present as a legitimate node	Identify theft and financial fraud	4	High	High	High	9	36
CAV	In-vehicle	OBD II port, CAN	Personal data	Exploiting CAN vulnerability that allows attacker to present as a legitimate node	Unauthorised access to personal and sensitive data	3	Medium	Medium	High	7	21
CAV	In-vehicle	OBD II port, CAN	Personal data	Exploiting CAN vulnerability that allows attacker to present as a legitimate node	Interception of communication	2	Low	Low	Medium	4	8
CAV	In-vehicle	OBD II port, CAN	Personal data	Exploiting CAN vulnerability that allows attacker to present as a legitimate node	Vehicle tracking and loction history access	3	Medium	Medium	Medium	6	18
CAV	In-vehicle	OBD II port, CAN	Personal data	Exploiting CAN vulnerability that allows attacker to present as a legitimate node	Alteration of non-critical vehicle settings	2	Low	Low	Low	3	6
CAV	In-vehicle	OBD II port, FlexRay	Personal data	Attacker gains access to FlexRay protocol and interprets communication	Identify theft and financial fraud	2	High	High	High	9	18

Figure 25 visualises the privacy score of a threat scenario for a data flow. It also provides insights on possible privacy harm on individuals based on the nine categories defined in NIST PRAM and also lists possible countermeasures to mitigate the risk.



127.0.0.1:8000/pec/dashboard/threat-details/2
Privacy Enhancing Component
Search for...
Dashboard
Privacy Assessment

Data Flow - User to Cloud Server Data Flow: This data flow represents the interaction between the user's mobile application and the cloud server. When a user books a cab ride, the mobile application sends a request with the pickup location, desired time, and destination to the cloud server. The server authenticates the user's request and processes it to find an available autonomous taxi. The server also receives real-time updates from the user's app regarding tracking preferences or changes in the ride details.

An attacker could intercept the communication between the user's mobile application and the cloud server. This could allow the attacker to eavesdrop on sensitive information, manipulate the data being sent, such as the pickup or destination locations, or even reroute the cab to a different location without the user's knowledge.

Privacy Impact Category	Legal Fines	Financial Costs	Reputation Costs	Privacy Score
IC1. Identity theft and financial fraud	High	High	High	27
IC2. Unauthorised access to personal and sensitive data	High	Medium	High	32
IC3. Interception of communication	Medium	Medium	Medium	24
IC4. Vehicle tracking and location history access	Medium	Low	Medium	10

Figure 25. Privacy impact categories, associated costs and privacy score for a threat

2. Identify safeguards to mitigate the privacy impacts: Once the privacy scores and possible impact categories have been assessed, PEC identifies a list of safeguards from the NIST Privacy Controls to mitigate the privacy threats. These safeguards are provided to the users in form of a recommendations with additional insights on a safeguard's effect on the privacy threat.

3.4.2.4 Support for pilots

The first version of PEC, presented in this document, will support the Autonomous Vehicles and Retail use cases. Note that this version of PEC only supports privacy impact assessment for processing, handling and using of personal and sensitive data. While the next version, to be submitted in M30, will also cover privacy impact assessment for the use of AI/ML services in the chosen use cases. Additional details on the functional requirements, user requirements and descriptions and applicability of PEC to the pilot use cases can be found in TANGO deliverables D2.2 (TANGO D2.2, 2023) and D2.3 (TANGO D2.3, 2023).

Table 8: Pilot user requirements and PEC functionality

Functional Requirement	Related User Requirements	Description of user requirements	PEC's contribution
PEC_privacy_assessment	UR-PCY-AV-003	The platform must protect the data flows from cyberattacks during the ride.	PEC will allow users to perform and record privacy impact assessment by identifying processes that handle sensitive data, accessing threats that could lead to privacy issues and determining possible privacy impacts that the threats would lead to. Insights into these would enable organisations to be aware of possible privacy threats, prioritise and address them mitigating their impacts and improving the overall data security and thus ensuring better privacy.
	UR-PCY-AV-004		
	UR-PCY-RT-018	The platform must protect the data flows from cyberattacks.	
	UR-PCY-RT-019	The system must protect the data from cyberattacks.	
PEC_recommendation	UR-GCO-AV-006	The data must be managed based on GDPR and all related data protection regulations.	PEC provides a list of recommendations to mitigate the identified privacy concerns. Implementation of these recommendations will ensure that the organisation's data protection measures are improved, will meeting some/most of the recommended data protection guidelines. These user requirements will be further supported by the third functional requirement to be developed in the second phase.
	UR-GCO-RT-017		

Beside the listed user requirements in Table 8, PEC could also partially support user requirements related to data protections data protection in Autonomous Vehicles such as UR-DIN-AV-002 that addresses data accuracy and integrity, UR-DAC-AV-007 that aims to uphold data access controls, and UR-PCY-AV-012 that aims to protect from cyberattacks and unauthorised access. Similarly, user requirements related to data management in the Retail use case could be partially supported by the PEC recommendations.

3.4.2.5 Software artifacts

To simplify the deployment process, a dockerised version of PEC will be used for continuous integration and continuous deployment activities. Docker is a container technology allowing compartmentalisation of various components into controlled execution environments. Docker can be installed on all major operating systems and facilitates running micro-services agnostic to specific system requirements (OS, software, hardware etc). The minimum software requirements to deploy PEC include:

- Operating system (Windows OS, Linux, CentOS)
- Dockerisation of PEC
- Web browser such as Internet Explorer, Chrome, Safari etc.

Whereas the minimum requirements include:

- Computer system with at least i3 core, 230GB of storage and 8GB of RAM
- CPU 1.7Hz (at least)
- Mobile devices with iOS or Android

Table 9 and Table 10 present additional technical details on the frontend and backend components of PEC.

Table 9: Technical details of PEC Frontend

DEVELOPMENT – PEC FRONTEND	
Type (Software/Hardware)	Software
Operating Systems	Windows 11 (64-bit), iOS, Linux
GUI	Web-based application using a web browser (Microsoft Edge, Chrome, Safari)
Programming language	HTML, CSS, JavaScript
Development Environment	Google Chrome, Microsoft Edge
Software Requirements	Recent web browser

Table 10: Technical details of PEC backend

DEVELOPMENT – PEC BACKEND	
Type (Software/Hardware)	Software
Operating Systems	Windows 11 (64-bit), iOS, Linux
Database and databases tools	PostgreSQL
Programming language	Python
Development Environment	Django, GitHub
INSTALLATION AND DEPLOYMENT	
Software Requirements	Django and related dependencies
Containerization	YES: Docker

Additional details such as screenshots of the PEC UI are presented in Annex 6.1.

3.4.2.6 Future work

The final version of PEC (to be submitted in M30) will extend the privacy risk assessment for AI/ML services. It will allow the users to perform privacy impact assessment on AI/ML-powered services. The assessment process will follow a similar structure as detailed in section 3.4.2.3 but specifically focus on capturing details related to the AI/ML lifecycle, identifying possible threat scenarios, assessing their privacy impacts and justifying safeguards to mitigate the impacts. Moreover, PEC will also provide guidance on complying with GDPR as an additional feature.

Besides these features, the final version of PEC will have:

- Complete integration with the PAT component,
- Provide API endpoints to communicate when needed with other TANGO components such as the Exploratory Data Analysis Engine, Trustworthy Data Sharing component, X-AI module and MLOps component. For more details on possible interactions of PEC with other TANGO components refer to deliverable D2.3,
- Full integration and deployment of PEC in the TANGO architecture including meeting the technical requirements such as authentication, testing and user evaluation, and
- Extend the analysis to other pilot use cases, where possible if time permits.

3.4.3 Privacy Assurance Tool (PAT)

3.4.3.1 Short description of the component

The Privacy Assurance Tool (PAT) is a secure framework designed for data exchange and record system. It offers encrypted data information for privacy preservation to the data requester, typically an organization. PAT operates on a consent-based access control mechanism for secure and privacy-preserving data exchange. It also includes a compliance engine to assure and check whether or not an organization complies with the data subject's consent. The PAT system is composed of three main components: a user interface, a server, and a NoSQL database. It is designed to adhere to GDPR regulations, ensuring secure and compliant data sharing practices. In Tango architecture, PAT as a docker container could be placed in the Tango connector in the organisation side.

The main characteristics and contributions of the PAT are threefold.

- **Secure Framework**: PAT is built with robust security measures in place. It uses encryption techniques to protect the data during transmission and storage, ensuring that the data is not accessible or readable by unauthorized entities. This makes it a secure framework for data exchange. This also ensures the confidentiality of the data.
- **Trusted and Authentic Record**: PAT maintains the accuracy and consistency of the data. It ensures that the data is not altered in transit and that it arrives at the destination in the same state as it was sent. This ensures the integrity of the data.
- **Consent-Based Access**: PAT operates on a consent-based access control mechanism. This means that other entities are unable to access the data without prior consent from the data subject. The data subject has full control over who can access their data, for what purpose, and for how long. This not only ensures the privacy of the data subject but also complies with data protection regulations like the GDPR.

3.4.3.2 Internal architecture

The PAT consists of three main components. The first component is the user interface, for a data subject and organization users to interact with the PAT. The user interface is a web application developed using the Flutter cross-platform application development framework supported by the Dart programming language.

The second component of the PAT is the server. The server is developed with the Java programming language and uses the Spring Boot framework to create a RESTful API to handle GET, PUT, POST, and DELETE requests. The Flutter web application uses the provided RESTful API to perform actions and store data into the PAT's third component, the PAT NoSQL database (such as MongoDB). The results from each API request to the Java Server are returned in JSON format.

The PAT NoSQL database is used to store consents and data subject data. Additionally, the connection between the Java Server and the PAT database is established through the integrated API or drivers through which the Java Server can perform queries to manipulate the stored data. The result of the queries is then sent back to the Java Server in JSON format. Figure 26 provides an illustration of the overall PAT architecture.

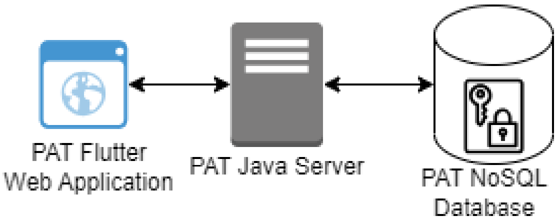


Figure 26. PAT Manager's overall architecture

Figure 27 briefly elaborates on the methodology used to develop the PAT components. The methodology is inspired by the agile development methods, allowing the PAT's development process to be iterative and flexible for any necessary changes that might need to be performed in the future.

Before any implementation starts, the user requirements are analysed from the D2.2 an D2.3, so that the PAT fulfils its functionalities for the Tango platform. Following, a common data structure is required. GDPR-compliant consents are complex data structures and can therefore be modelled using ontologies as a data structure. Once the requirement information for consent construction has been created, the implementation of the PAT's Java server and the PAT's Flutter web application commence. Both the database, Java server and Flutter web (for both data subject and organizations sides) application are developed in parallel and later integrated with each other through the Java server's REST API. As soon as the integration is successful the next functionality can be implemented and tested. The testing for the Flutter web application and the Java Server can be done simultaneously through the REST API.

Furthermore, when a functionality has been tested, then the next functionalities can be implemented to further extend the PAT's capabilities. Lastly, a review on the PAT's state is done and any missing functionalities can be added, resulting in an iterative development process for the PAT.

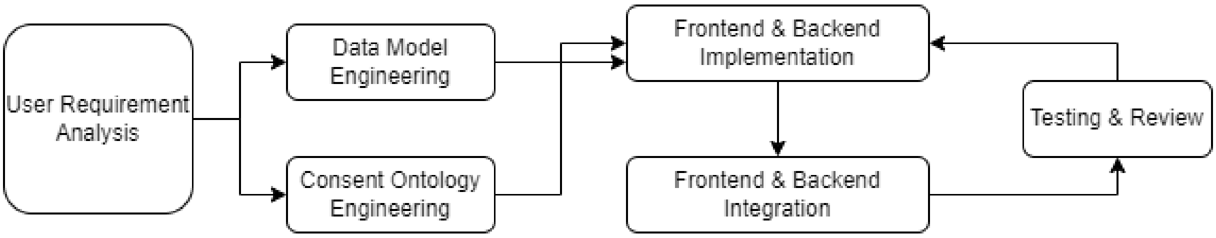

User Requirement Analysis
Data Model Engineering
Consent Ontology Engineering
Frontend & Backend Implementation
Frontend & Backend Integration
Testing & Review

Figure 27. The methodology used to develop the PAT.

The PAT architecture contains data subject or user (data subject) and organization sides, see Figure 28. Data and consent creator and collector are used for data privacy monitoring. The user is aware about the shared data and its policy. The organization is transparent about the data purposes.

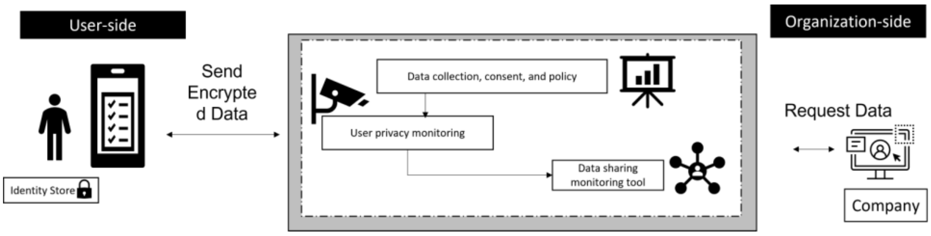
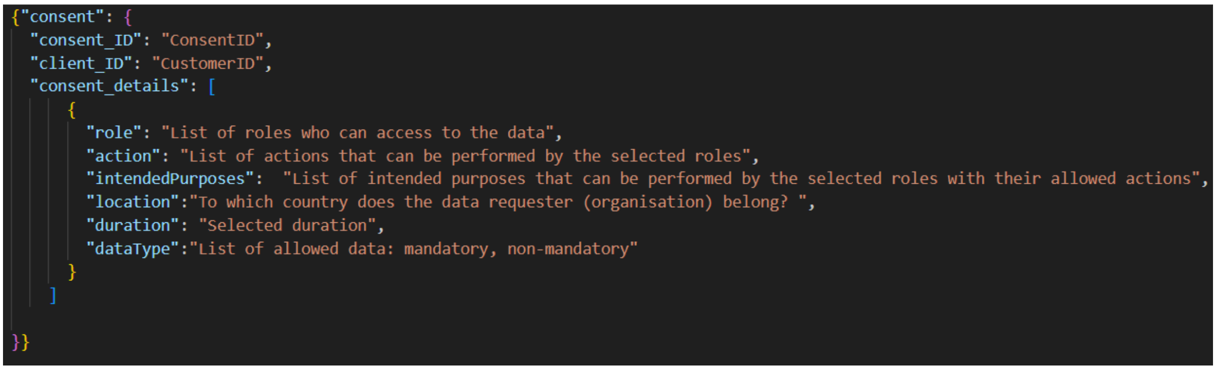


Figure 28. PAT Architecture

The PAT offers a range of features designed to enhance data privacy and security. It facilitates data sharing between a user and an organization by providing assurance to the user about the safety of their data. PAT also enables privacy monitoring, allowing both the user and the organization to monitor data sharing, consent, and any privacy risks related to the requested data throughout the data lifecycle. Transparency is a key feature of PAT, ensuring that end-to-end data sharing is clear to the user and that they are fully aware of how their data is being used. In terms of compliance, PAT supports organizations in fulfilling their obligations under the General Data Protection Regulation (GDPR). Finally, PAT is designed with security and privacy at its core, adhering to the principles of secure and privacy by design.

Functional requirements. The PAT is designed with several functional requirements to ensure secure and compliant data handling. It can collect a client's personal data and preferences directly through the web, which not only guarantees the collection of accurate client information but also enables the provision of personalized services. PAT also ensures that data sharing is only enabled among relevant stakeholders or business partners, enhancing data security. Adherence to GDPR requirements is a key feature of PAT, allowing for the management of data in a compliant manner. To ensure secure access, PAT supports trusted user authentication for applicants. Lastly, PAT provides secure data storage, ensuring that all collected data is stored in a trustworthy manner, further enhancing the security and privacy of the system.

Consent creation and request: In general, consent is defined as permission for something to happen or agreement to do something. It refers to the compliance in or approval of what is done or proposed by another. A consent refers to a data subject granting access to specified details stored on a record system. It is a voluntary agreement given by an individual prior to a targeted procedure (such as booking a room in the smart hospitality use case). This implies that the data subject has been fully informed about the procedure, understands the information, and agrees to undergo the procedure. Consent is fundamental in any information sharing as it respects the data subject's autonomy and privacy. The PAT offers the autonomy to the data subjects to choose the different options in the consent and help them to understand the content of the consent. Figure 29 shows the details of the consent JSON format in the backend side.



```
{"consent": {
  "consent_ID": "ConsentID",
  "client_ID": "CustomerID",
  "consent_details": [
      {
        "role": "List of roles who can access to the data",
        "action": "List of actions that can be performed by the selected roles",
        "intendedPurposes":  "List of intended purposes that can be performed by the selected roles with their allowed actions",
        "location":"To which country does the data requester (organisation) belong? ",
        "duration": "Selected duration",
        "dataType":"List of allowed data: mandatory, non-mandatory"
      }
    ]

}}
```

Figure 29. Consent in JSON format

The consent contains the following properties or attribute:

- **Role**: a role refers to the designation or position of an individual or entity in an organization. In the case of the smart hospitality use case, this could be a receptionist, hotel manager, group hotel manager, etc. The role determines what level of access they have to the data and what actions they can perform.
- **Action**: This refers to the operations that the individual or entity (based on their role) can perform on the data. Actions can include reading, writing, updating, deleting, etc.
- **Purpose**: This is the intended use of the data by the individual or entity. The purpose should be clearly defined and aligned with the organization's data usage policies. For example, the purpose could be the check-in-in and check-out for the receptionist.

- **Location**: This refers to the geographical location of the organization or individual requesting access to the data. Data protection laws vary by country, so the location can affect what data can be accessed and how it can be used.
- **Duration**: This is the length of time for which the data access is granted. After the duration expires, the individual or entity would no longer have access to the data unless the access is renewed.

Audit log collector for user's consent compliance. Audit logs are created from the java sever, generated, and stored in a database. The logs represent the behaviour of a user organisation when connecting and requesting the data. The event log contains the following information: timestamp, the user's email, IP address, the accessed data, and the type of operations.

3.4.3.3 Features implemented

The Privacy Assurance Tool (PAT) aims to monitor privacy related to the consent and data policy between data subjects and organisations. PAT has the following main features:

- **Consent-based access control (CBAC) mechanism for secure and privacy-preserving compliant:** The access to the data is assured by the data subject's consent. If the organisation user attributes match the consent, then the user has access to the requested data.
- **Automated privacy monitoring** is a process that facilitates all tasks involving data subjects' personal data with little to no human involvement. It's designed to save time and energy, and it includes the following task:
  - Handling data subjects' data privacy requests: This could include creation of requests.
  - Government and regulatory compliance: Ensuring that data handling practices are in line with the GDPR laws.
  - Data mapping: Identifying what data is stored where within the organization.
  - Consent management: Managing user consent for data collection.
- **GDPR compliance**: Organizations exhibit no apprehension regarding the GDPR compliance of their systems or data sharing protocols. Indeed, the PAT is inherently designed to adhere to GDPR regulations. Consequently, inquiries regarding their alignment with GDPR standards are provided by the PAT. Thus, it allows the organizations to focus on their main business.
- **Secure data sharing**: the data subject data is encrypted in the database. The decryption is on-demand based on a set of rules detailed in the next paragraph.

All these features are fulfilled by the PAT compliance checker composed by three types of rules:

- **Organisation users' access-based rules**: When a data subject consents to share their data, a rule is initiated to grant access based on certain conditions. These conditions include the role of the organization's user, its location and the permitted operations on the data. For example, if a data subject specifies that only a receptionist can access their data, the receptionist is granted READ and UPDATE permissions knowing that the receptionist has the following operations by default READ, UPDATE, WRITE, DELETE. This access is provided through a rule that is initiated based on the data subject's consent. The PAT is responsible for managing all data subject consents and their data access by the organization's users.
- **Consent-based rules**: A set of rules is established to verify the compliance and validity of consents. For example, if a data subject selects a duration of 6 months for data storage on the organization's side, the PAT checker routinely verifies the consent's validity. If the consent is no longer valid, the PAT removes the data subject's data from the database.
- **Audit logs-based rules**: A rule set is in place to scrutinize the audit logs generated, which pertain to permissible data operations, taking into account the given consents. In its current version, these rules are designed to ensure that data access by organizations is compliant. For example, if the consent specifies Spain as the location, it implies that only users from Spanish organizations can access data subject data. The event log records the behaviour of the

organization's user, primarily their IP address. This enables the initiation of a rule to verify whether the user's data access is from within or outside Spain.

The main core features of PAT are implemented and described for common the targeted pilot use cases (smart hospitality and public administration). In the current version, only specific features are implemented for the smart hospitality use case. These features will be reported in the next subsection (Support for pilots). The implemented features of the PAT composed of both frontend and backend described as follows:

PAT operations in the data subject (data subject side) frontend. The PAT operations are implemented on the frontend of the data subject, which is the data subject side. The PAT facilitates the entry of data requested by an organization and modifies the consent according to the data that the individual wants to share with the organization. This data can be either mandatory or non-mandatory. The data subject has the ability to take a fine-grained decision to accept, update each piece of information in the consent and requested dataset.

PAT operations in the data requester (organization) frontend. The PAT operations are carried out on the frontend of the organization and consist of two types of operations. Firstly, the organization has the capability to create a consent form, filling in the necessary information based on GDPR requirements. This includes both mandatory and non-mandatory data, as well as details in the consent that a data subject can either not change or change freely. Secondly, the organization has the ability to update an existing consent at any time. However, these updates should be made prior to any acceptance of the consent by a data subject (data subject). If the data subject has already accepted the consent, the data requester (organization) cannot update the consent. In addition, the data requester has the option to remove a consent, especially in situations where the consent has not been signed by a data subject.

Consent visualisation in the data subject (data subject side) frontend. It's essential to consider how the consent is displayed to the end users. This could be via a user interface (UI) or a graphical visualization. It's vital that this display is informative. Users frequently lack knowledge about data sharing and the related privacy risks. They might not completely grasp the ramifications of providing consent. Therefore, we suggest a consent user interface to ensure that a user's informed consent always guarantees that they fully understand the consequences of their actions. The pursuit of complexity, slowness, and challenging interactions often encourages individuals to pay careful attention to crucial consent information, thoroughly contemplating the requirements and fully grasping the potential results before agreeing.

PAT backend for both data subject and organisations. The PAT backend in developed in Java based on Spring boot framework to create API with the frontend as well as the future integration with other components proposed in Tango. The PAT endpoints are described in Annex 6.1. The API endpoints are the points of interaction and communication between an API and a serve, and NoSQL database. Moreover, the PAT generates audit logs related to requested and accessed data in the database. The audit logs are also stored in a database. Specific rules are implemented on the audit logs to detect a non-authorised access or non-compliant use of the data subject data. In addition, other rules-based consents are implemented such as the validity of the consent.

PAT Ontology: The ontology is developed in the beginning of the project in order to modelized the concepts of the consent and their relationship using Protégé. The ontology represents a consent model for GDPR compliance. The ontology's main purpose is to help us to capture the knowledge about the characteristics of a consent and data sharing domains. In addition, it provides a set of machine-readable statements.

3.4.3.4 Support for pilots

PAT targets two pilot use cases: smart hospitality and public administration (visa request). For the first version of the deliverable, PAT will target the smart hospitality use case.

PAT will directly interact with pilots: It means that the tool will be used by the organization leading the pilots for requesting data and its consents. It will be a pilot-centred approach where the privacy assurance is managed during the data-life cycle and end-to-end data workflow: from data collection to organization.

In both smart hospitality and public administration use cases, here are the following functionalities:

In customer side:

- Fill-in the requested data (ex. personal data)
- Preferences (activities, drinks, foods, etc. in smart hospitality use case),
- Data policies and privacy awareness.
- Monitoring privacy risks of the requested data

In organization side:

- Request the data (ex. personal data)
- Data policies
- Data policies and privacy awareness.
- Monitoring privacy risks

Example.

Consider a user named "John" who works as a receptionist in the hotel. John is trying to access a mandatory client data (the resource) on the hotel's server. The action he wants to perform is "read". The access request is made from a company-registered device.

In an CBAC model, the following attributes would be evaluated:

- Subject attributes: User ID, job role (receptionist)
- Resource attributes: Type of resource (mandatory client data)
- Action attributes: Type of action (read, update)
- Environment attributes: location of access (city or country)

The CBAC system would then check these attributes against the clients' access control policies provided in the consents. For instance, one policy could be: an employee in the hotel in Malaga with a role receptionist can read and update mandatory client data from a company device. If John's attributes match this policy, he would be granted access to the mandatory client data. This is a simplified example, but it illustrates how CBAC can provide fine-grained, dynamic access control based on a variety of factors.

3.4.3.5 Software artifacts

Table 11, Table 12 and Table 13 represent the technical characteristics of the implementation of PAT fronted, backend, and server, respectively.

Table 11: Technical details of PAT frontend

DEVELOPMENT – PAT FRONTEND	
Type (Software/Hardware)	Software
Operating Systems	Windows 11 (64-bit)
Database and databases tools	None
GUI	Web-based application using a web browser (Microsoft Edge, Chrome)
Programming language	Dart
Development Environment	Visual Studio Code, Microsoft Edge
INSTALLATION AND DEPLOYMENT	

Software Requirements	Flutter (3.7.11)
Hardware Requirements	None
Containerization	YES: Docker
Communications	It communicates with the PAT server.

Table 12: Technical details of PAT backend

DEVELOPMENT – PAT BACKEND	
Type (Software/Hardware)	Software
Operating Systems	Windows 11 (64-bit)
Database and databases tools	NoSQL database (MongoDB)
GUI	None
Programming language	Java
Development Environment	IntelliJ, Protégé, GitHub
INSTALLATION AND DEPLOYMENT	
Software Requirements	Java 20 JDK, GraphDB
Hardware Requirements	None
Containerization	YES: Docker
Communications	It communicates with the PAT server.

Table 13: Technical details of PAT Server

DEVELOPMENT – PAT Server	
Type (Software/Hardware)	Software
Operating Systems	Windows 10 (64-bit)
Database and databases tools	None
GUI	None
Programming language	Java, Sprint boot 2
Development Environment	IntelliJ, GitHub
INSTALLATION AND DEPLOYMENT	
Software Requirements	Java 20 JDK
Hardware Requirements	None
Containerization	YES: Docker
Communications	The server allows to connect the frontend and backend. It offers a REST API for different endpoints such as getting and posting consent and data.

Additional details on end points and mock-up are presented in Annex 6.1

3.4.3.6 Future work

The future work is described as follows:

- Validate the proposed data model for the smart hospitality use case.

- Develop the data model for the public administration use case in the backend side while developing its frontend in the user side.
- Enhance the front-end interfaces for both data subjects and organizations.
- Add and implement more rules-based compliance based on the consent, audit logs, and organization user's attributes.
- Integrate the identity management tool in the organization side for managing the user of the organizations. For instance, the employees of the hotel in the smart hospitality use case, or the employees in the public administration use case.
- Improve the privacy of PAT by enforcing the consent.
- Convert the mobile UI design to a desktop UI design for facilitating the integration with other Tango components.
- Based on the deliverables of WP3, 4, and 5 and their outcomes, the challenge is to potentially connect PAT with PEC or other components targeting the same use cases in the WP5 or/and WP3/WP4 such as the use of SSI (Task 4.1) for PAT authentication and the use of the blockchain (Task 3.1) for storing the consents by offering more transparency and regulation. In addition, the challenge is to connect PAT to Tango connector.

3.5 X-AI for Privacy and Trust Enhancement [T5.5]

3.5.1 Introduction

The purpose of this task is to provide explainability to the models that are trained with the help of Federated Learning within the TANGO project by leveraging explainable artificial intelligence (XAI) techniques.

XAI is a set of processes and methods that allows human users to comprehend and trust the results and output created by machine learning algorithms. It is crucial for an organization in building trust and confidence when putting AI models into production and it also helps the organization adopt a responsible approach to AI development.

3.5.2 X-AI Library

3.5.2.1 Short description of the component

The XAI module will be used to describe an AI model, its expected impact and potential biases. It will help characterize model accuracy, fairness, transparency and outcomes in AI-powered decision making. Apart from helping build trust in the predictions the Machine Learning models produce, it will also help debug the models in cases where bias in the decision-making process is discovered.

Main features:

- Provide Machine Learning and Deep Learning model explainability.
- Present the explainability results in an easy-to-understand visual way.

Since Federated Learning will be leveraged for the model training, the XAI module will be packed in a Docker container and will be deployed on the pilots' premises. This will allow the pilots to have full control over their data without the need to share sensitive information that they may contain. An added benefit is that there will be no unnecessary transferring of large amounts of data, thus reducing energy consumption.

3.5.2.2 Internal architecture

The Explainable Artificial Intelligence module for Machine Learning and Deep Learning model explainability is written in Python. The internal architecture of the module is:

- **Input data**: The input data that could be of a specific class (e.g., tabular, images, time-series)

- **Processing pipeline**: The processing pipeline that the input data is inserted to be transformed via specified functions, resulting to processed data.
- **Processed Data**: The output data of the processing pipeline is data transformed in the appropriate format required by the explainer of the XAI pipeline.
- **Model Storage**: The model storage that the necessary models are stored in.
- **Model**: This is the trained model. It is required in most of the XAI methods. In some methods (model-agnostic), other techniques will be used to compensate for the model's absence, like training a new model based on the input data.
- **XAI Pipeline**: Runs the explainer pipeline that will be used depending on the type of the data and model. For tabular data, libraries such as LIME or SHAP and for images Grad-CAM can be used.
- **The Explain-ability report**: The report that contains the results produced by the explainer of the XAI pipeline.

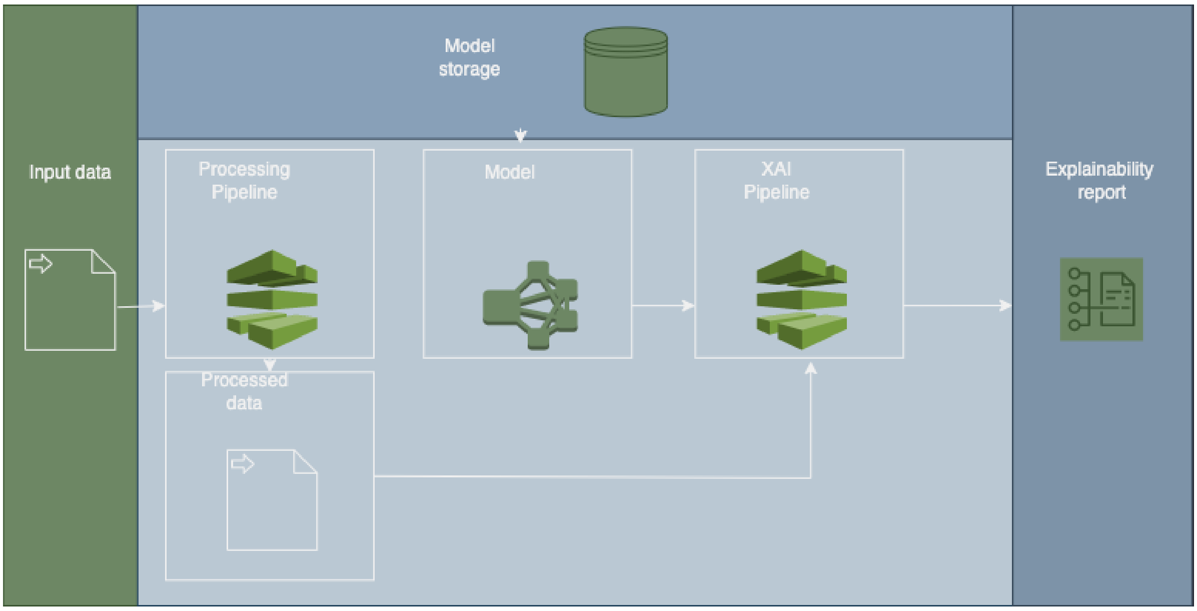


Figure 30. XAI Module Architecture

Depending on the model availability and the data classes, different algorithms can be applied to perform XAI. So far four different algorithms are supported:

- **Local Interpretable Model-agnostic (LIME):** It is a visualization technique that helps explain individual predictions. It is model agnostic so it can be applied to any supervised regression or classification model.
- **Gradient-weighted Class Activation Mapping (Grad-CAM):** It's an algorithm that uses the gradients of any target concept (say 'dog' in a classification network or a sequence of words in captioning network) flowing into the final convolutional layer, to produce a coarse localization map highlighting the important regions in the image for predicting the concept. It works only with Convolutional Neural Networks (CNNs).
- **Layer-wise Relevance Propagation (LRP):** It is a technique that brings such explainability and scales to potentially highly complex deep neural networks. It operates by propagating the prediction backward in the neural network, using a set of purposely designed propagation rules.
- **Anchors:** They are high precision explainers that use reinforcement learning methods to come up with the set of feature conditions (called anchors), which will help explain the observation of interest and also a set of surrounding observations with a high precision (the user is free to choose their minimum precision cut-off).

Depending on whether the models that need explaining are provided or not, the following methods can be followed:

- **Model-specific explainability:** These methods work based on the details of the specific structures of the machine learning or deep learning model which they are applied on. Model-specific XAI algorithms work by analysing the internal workings of a specific ML or DL model to provide insights into how it makes decisions. This approach, although it allows us to get a deeper understanding of the decision since we know the internal structure of the model, it compromises the performance of the model because the model itself will have to be recreated.
- **Model-agnostic explainability:** These methods do not take into account the structure of the model so they can work with any machine learning model, giving more flexibility. A drawback is that these methods treat the model as a black box so they try to explain the output based only on the input whereas model-specific methods can explain the model's behaviour on all its layers.

The model-specific method needs for the trained model to be provided whereas the model-agnostic method does not. It should be mentioned however that even if the model is provided, it is not sure that a model specific approach could be applied, as it is difficult to decompose extremely complex models. All-in-all, despite of which explaining method will be used, the dataset that the model was trained with (training dataset) and the expected results (inferences) are a requirement to successfully produce an explainability report.

Regarding the dependencies with other components, the "Energy efficient AI model training" component by ATOS in T5.2 will facilitate the training of the models and the XAI module with will be used to explain the results and help with the model debugging process.

3.5.2.3 Features implemented

Up to this point, generic pipelines of the XAI module have been implemented. It is needed to apply these pipelines on the final or checkpoints of the models trained for the TANGO project's purposes and test their results.

3.5.2.4 Support for pilots

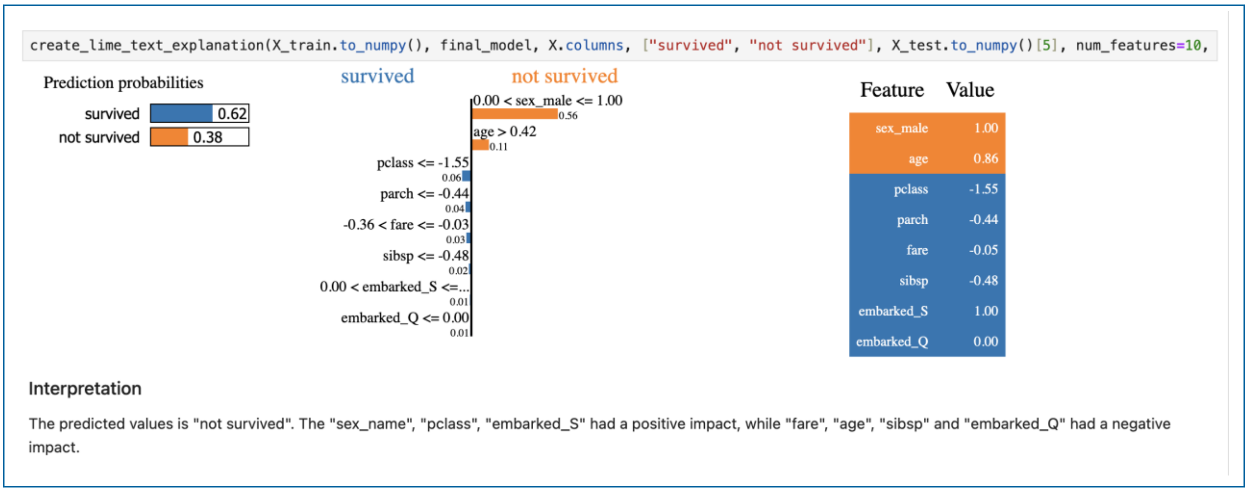
Table 14 provides an overview of the current support for pilots expected for the XAI component.

Table 14: XAI support for pilots

Pilot	Association
Smart Manufacturing FMAKE (Pilot 3)	Data scientists from FMAKE will facilitate the usage of XAI to re-train their current quality models and improve the printing laser's parameters. The debugging of the training process will be much easier and faster since the results of the model will be explained.
In order to achieve the desired results, a model specific approach will be implemented, specifically by developing an XAI algorithm that will be able to identify which parts of an image play a crucial role in order to output and predict the laser's parameters.	
Banking (Pilot 4)	The banking pilot will train Machine Learning models that will be able to detect fraudulent activities. XAI would be great at explaining why a specific activity was characterized as fraudulent and in turn detect flaws in the training process.
The data classes for the fraudulent transaction detection models will be in a tabular form so a model agnostic approach will be implemented, likely by using the LIME or SHAP algorithms. These two algorithms are very good in providing insights in tabular data use cases.	

3.5.2.5 Software artifacts

Jupyter notebooks were created to demonstrate the use of XAI pipelines:



```
create_lime_text_explanation(X_train.to_numpy(), final_model, X.columns, ["survived", "not survived"], X_test.to_numpy()[5], num_features=10,
```

Feature	Value
sex_male	1.00
age	0.86
pclass	-1.55
parch	-0.44
fare	-0.05
sibsp	-0.48
embarked_S	1.00
embarked_Q	0.00

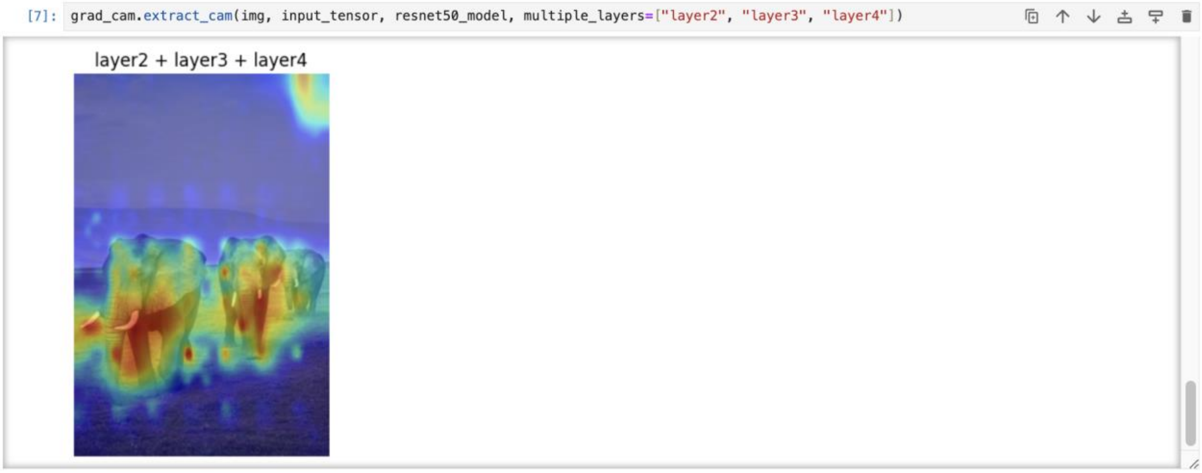
Interpretation

The predicted values is "not survived". The "sex_name", "pclass", "embarked_S" had a positive impact, while "fare", "age", "sibsp" and "embarked_Q" had a negative impact.

Figure 31. Example XAI report using LIME.

In the above example, a dataset[10] describing the survival status of individual passengers on the Titanic is used to show which features had a positive and a negative impact to the predicted value of “Not survived”. Here, the LIME method is used on tabular data, but it also supports text data and picture data as input formats.

Below, three different layers are displayed to indicate, using Grad-CAM and with the help of XAI, which parts of the image the model took into consideration in order to locate the elephants:



```
[7]: grad_cam.extract_cam(img, input_tensor, resnet50_model, multiple_layers=["layer2", "layer3", "layer4"])
```

Figure 32. XAI on image using Grad-CAM for a custom ResNet-50 model.

Each layer of the model has different importance:

[10] https://www.tensorflow.org/datasets/catalog/titanic

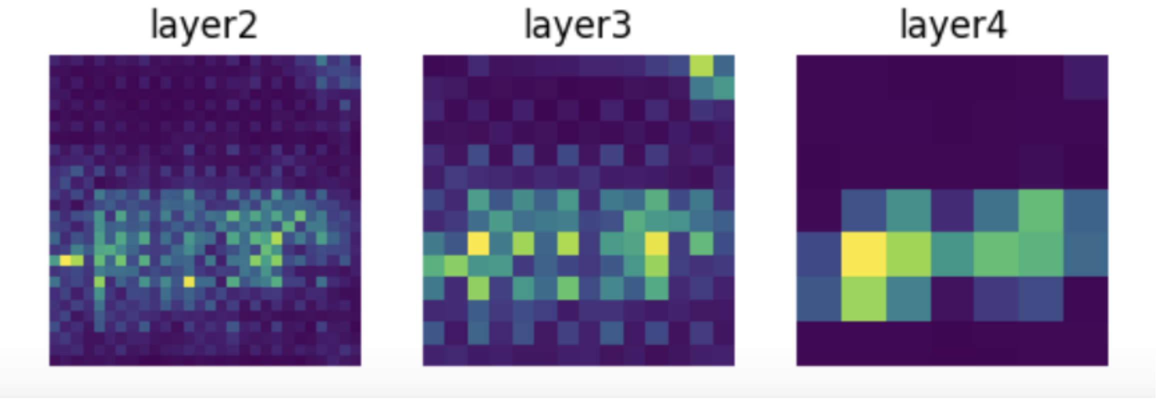


Figure 33. Each layer has different importance.

The repo for the module is available on GitHub:

https://github.com/TANGO-EU-PROJECT/xai-component

3.5.2.6 Future work

The remaining work for the second half of the project will be to:

- Complete the custom XAI algorithm for the Smart Manufacturing pilot.
- Start implementing XAI techniques on the models that will be trained by the banking pilot.
- Design and implement the connection between the XAI module and ATOS's Federated Learning component.
- Dockerise the component and deploy it on the pilot's premises.

3.6 Infrastructure Management based on AI [T5.6]

3.6.1 Introduction

This task aims to address one of the key goals of the TANGO project, which is to increase the energy efficiency of the platform. The increasing number of data centres is increasing the overall carbon footprint of computing. Sources from IAE (iea, 2023) and Rong et. al (H. Rong, may 2016) state that HPCs account for around 1.5 % of annual global energy consumption and when combined with the ICT sector, for around 2 % Greenhouse emissions. A source from the EU Commission (Commission, 2020) states that energy consumption of data centres in EU member states is set to increase from 2.7% of the electricity demand in 2018 to 3.2% by 2030. With an increasing demand for LLMs, the demand for HPC is set to increase even more with some more recent sources (Consumption, 2022) stating that the number could reach up to 4 % by 2023. A more detailed analysis of the impact of HPCs can be found in D2.3, section 5.6 (TANGO D2.3, 2023).

The authors in the paper (Anna Maria Oosthuizen, 2022) state that renewable energy sources are of lower cost than conventional sources. As energy consumption and production must be in sync, shifting the load to times when renewable sources are available should drive down the utility costs and help stabilise the grid. In production demand fall out of sync, the power grid may collapse due to the frequency being too high or too low. To combat this issue, constant adjustments the energy production are required. To avoid expenses of reducing production and issues of storing the energy, we need to strive towards the goal of shifting the energy consumption to the periods when the energy is in abundance. To achieve this, we need to forecast energy production and have the means to consume it. Data centres serve as one such means.

Through TANGO, we plan to provide RENOPS, a renewable energy forecasting service aimed at addressing the mentioned issues. As price is used to indirectly regulate the grid load, accurate price predictions could also play a key role in maximising the use of renewable energy sources. Through TANGO we plan to use state-of-the-art deep learning models to provide accurate price predictions.

3.6.2 RENOPS

3.6.2.1 Short description of the component

RENOPS stands for Renewable Energy Forecast Production Service. Its main task is to predict the availability of renewable energy sources. This can be done directly by using solar irradiation power forecast, or indirectly by predicting the energy price.

The system is composed of three components, the first component being the prediction services i.e., models providing the forecast for renewable availability or price.

The second part is the service that seamlessly provides information about renewable energy potential to the client. Since this information is gathered from multiple models, the service acts as a gateway service between the client and the models. The information it retrieves can be used to build run schedules, automate, or simply display information on a dashboard.

To demonstrate the models' capabilities, we propose the third part of the system, a scheduler. RENOPS scheduler is a program that utilises the forecast from the RENOPS service to find the most optimal window to run energy-intensive tasks. A scheduler can be applied to shift recurrent energy-intensive jobs such as AI model training, big data analytics, backups, and various scans. While it can be initiated manually, its main strengths are that it can be used for periodic tasks. In this light, AI model training is an excellent fit, as models in production need to be finetuned to new data on a weekly basis.

3.6.2.2 Internal architecture

The following subsection will focus on describing individual components of the architecture as seen in Figure 34 below. The architecture does not depict the request flow. A request is initiated on the left, on the client side and moves from to the right towards the data sources. The response (dataflow, depicted with grey arrows) flows in the other direction from to right. We will describe the individual component starting with data sources on the right and then moving towards the left following the dataflow.

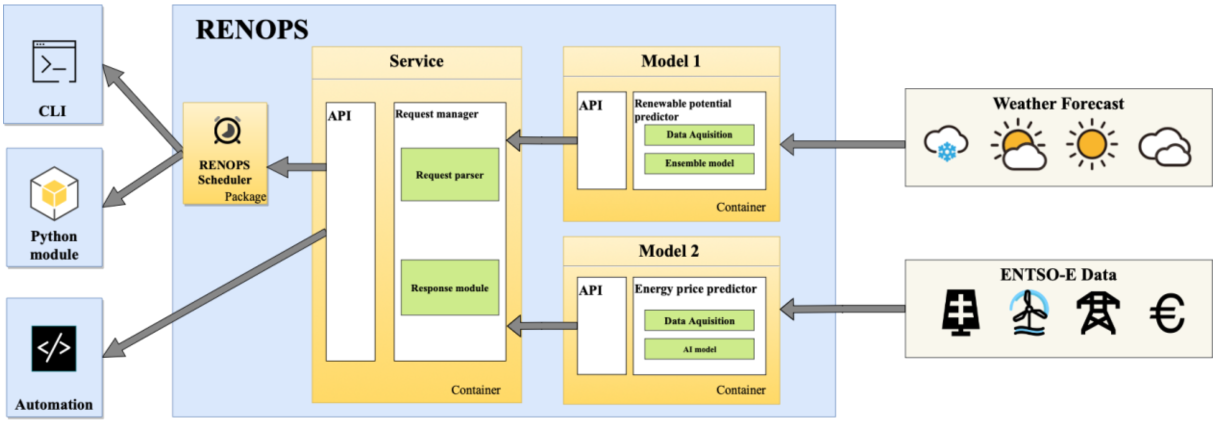


Figure 34. RENOPS architecture

3.6.2.2.1 Data sources

Data sources can be split into two main groups: weather and energy data. Weather data is obtained from Open Meteo (Zippenfenig, 2023). They obtain their data from weather providers such as NOAA, DWD and MeteoFrance. Weather forecast models already provide estimated solar irradiation in their forecast. More detailed metadata for used models can be seen in Table 15.

Table 15: Weather models

Model	Weather Provider	Origin Country	Spatial Resolution	Forecast Length
ICON	DWD	GER	2 - 11 km	7.5 days
GFS	NOAA	US	3 - 25 km	16 days
Arpege & Arome	Meteo France	FR	1 - 40 km	5 days
Ensemble	all	all	1 - 40 km	5 days

The electrical energy data is obtained from ENTSO-E (ENTSO-E, brez datuma). We obtained forecasted and measured load demand as well as day-ahead energy price forecast and energy generation per production type. This will enable our model to better predict the price of the energy.

Table 16: ENTSO-E data

Type	Provider	Temporal Resolution	Spatial Resolution	Forecast Length
Load demand	ENTSO-e	1h	Per EU member state	1 day
Energy price	ENTSO-e	1h	Per EU member state	1 day
Production type	ENTSO-e	1h	Per EU member state	/

As we can observe in Table 16, the spatial resolution of the ENTSO-E per EU member state means that this will be the maximum resolution of our price models as well. On the other hand, the renewable potential model will have spatial resolution between 1 - 40 km.

3.6.2.2.2 Model 1 - Renewable Potential Predictor

To obtain a renewable potential as described in D2.3 (TANGO D2.3, 2023), we need to first obtain accurate solar irradiation forecasts. To achieve this, we utilise an ensemble model, where we use grid search to find the optimal contributions of each weather model. One such example can be observed in Figure 35 below.

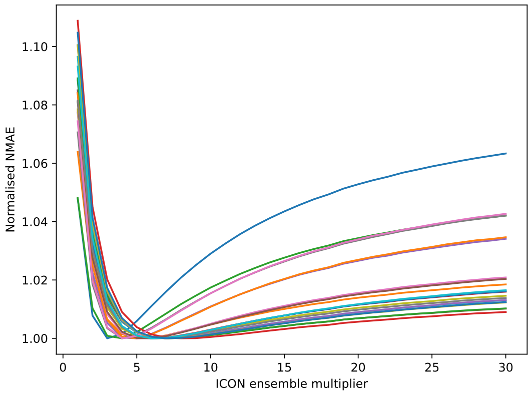


Figure 35. Hyper-parameter grid search for all combinations of 5 locations on train data

Renewable potential predictor uses data from multiple weather models to produce an accurate prediction. We propose an ensemble model as follows:

$$P_{\text{ensemble}} = \frac{\sum_{i=1}^{n} W_i \cdot P_i}{\sum_{i=1}^{n} W_i}$$

Where:

- $P_{enseble}$ is the resulting model.
- W_i is the weight of contribution of the individual model (ICON, GFS, etc.)
- P_i represents the prediction of each weather model.

Using an ensemble model increases robustness against drastic mistakes made by individual models.

3.6.2.2.3 Model 2 – Energy price predictor

Optimally RENOPS would provide the overall presence of renewable energy sources in our grid. To include all renewable sources, we propose to use energy prices to automatically add weight for each source. This enables us to automatically estimate the contribution of individual renewable energy sources for each EU member state.

Overall, price dictates the availability of energy, and it makes sense to run energy-intensive tasks when in abundance. Occasionally, an excess of energy supply can lead to negative energy prices. With an increasing number of renewables in our grid, we can expect these occurrences to become more common. This is beneficial for the consumers and the grid, as production is less volatile and in turn more efficient.

Using grid load, weather forecast, holidays, and energy prices we plan to make accurate predictions. To make the predictions we will utilise a Temporal fusion transformer (Bryan Lim, 2021), a SOTA deep learning model used for time series forecasting. Temporal fusion transformer or TFT uses self-attention methods to focus on the parts of the inputs that have the biggest effect on the predictors.

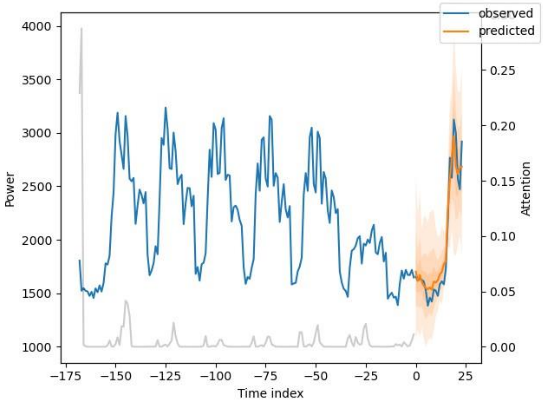


Figure 36. Example of day-ahead load demand prediction with TFT

The training procedure is as follows. Covariates, such as weather, energy consumption, and holidays are passed are passed through a variable selection network (VSN), where their optimal contribution is found. Next, a long short-term memory network (LSTM) is used to find local temporal relations. Here past as well as future covariates are used. Next, a static enrichment layer is used to add information class ID in case of multivariate forecast. Finally, a multi-head attention layer is used to make sense of more global time relations. Using attention, it should learn to focus attention on time when consumption is most similar, as seen in Figure 36.

3.6.2.2.4 RENOPS (service)

RENOPS service is the main entry point into the system. It acts as a gateway to make further requests and actions to obtain results for the client. Interaction is performed using a REST API, described in TANGO D2.3. Endpoints can be observed in Figure 39 within the OpenAPI specifications. Its main parameters are forecast location and forecast type (either potential or energy price). Based on parameters it runs new queries to obtain requested predictions. In case the given model does not have the requested 3-day ahead forecast for the given bidding zone, it makes a direct request to ENTSO-e, to obtain day-ahead predictions. Overall, it's designed to seamlessly provide data to the user.

3.6.2.2.5 RENOPS scheduler

As seen in Figure 34 we can energy price and renewable potential can be directly accessed via RENOPS API and can be used to run various automation or simply display the information in a UI. To demonstrate the possible use-case of such service, we have developed a scheduler. The scheduler communicates with RENOPS to obtain renewable potential or price for the current location. Location can be either specified manually or can be automatically deducted based on source IP address. It identifies an optimal execution window based on user-defined runtime and deadline. Once found, the program enters a loop until reaching this time. Automating this process enables periodic runs without user supervision. This service yields various benefits, from utility cost savings to reducing carbon footprints, as described in the introduction.

3.6.2.3 Features implemented.

While features were already described in D2.3, the main implemented features of our service include:

- Predictions.
  - A 5-day ensemble forecast of renewable energy availability for any location.
  - A Day ahead electrical energy price forecast for all EU member states.
  - A 3-day ahead price forecast for Slovenia.
- Temporal job scheduling and orchestration based on renewable energy availability.
  - Command line interface program.
  - Python package.
- A REST API access to renewable energy and price information.
  - Basic authentication using an API token.
  - Request IP limiter to mitigate DDoS attacks.
- Dockerised services and packaged services.
- CI/CD development pipeline.

3.6.2.4 Support for pilots

Smart Manufacturing FMAKE (Pilot 3): To perform quality estimation of their 3D prints, FMAKE utilizes computer vision models. Larger language and computer vision models usually consume a significant amount of energy in the training and inference phase. Such models must be retrained or finetuned on weekly basis to take into an account newly collected data. To address the energy efficiency aspect of TANGO we plan to utilize the RENOPS system to find and schedule the training phase to the most optimal time. This should reduce the carbon footprint and utility cost of AI model training.

Pilots such as Smart hospitality, Banking and Retail will be addressed indirectly through T5.1 Exploratory data analysis and T5.2 Energy efficient model training.

3.6.2.5 Software artifacts

3.6.2.5.1 Models

RENOPS will utilise two models. One for predicting the renewable potential and the other for energy price. The two models are very different from one another. For example, if we observe Figure 37, we can see that the weather model is a function of latitude and longitude: $X = F(lat, lon)$).

The output is then passed through renewable potential predictor to obtain prediction Y. In simple terms, we can utilise one function for every location on Earth. Such a model is easy to maintain, as it does not require permanent data storage and model finetuning.

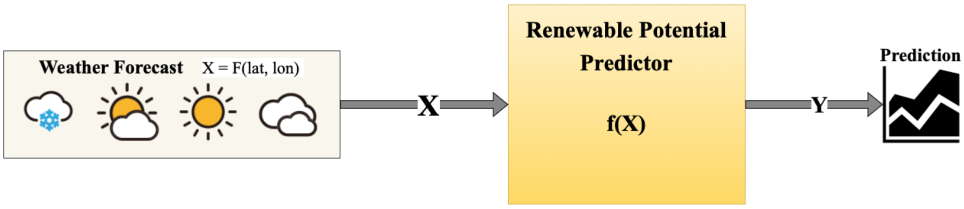


Figure 37. Renewable potential prediction model

On the other hand, when observing Figure 38, we are adding energy data, which is compared to the weather discrete. The spatial resolution of energy data is per-county, as we can observe in Table 16, meaning that we would need one model for every member state. Besides having more models, we need to update the model based on new data on a weekly basis.

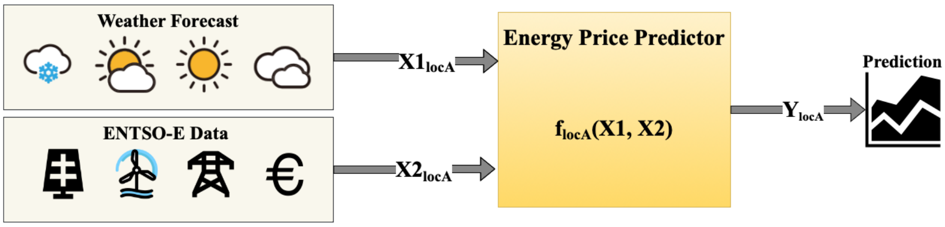
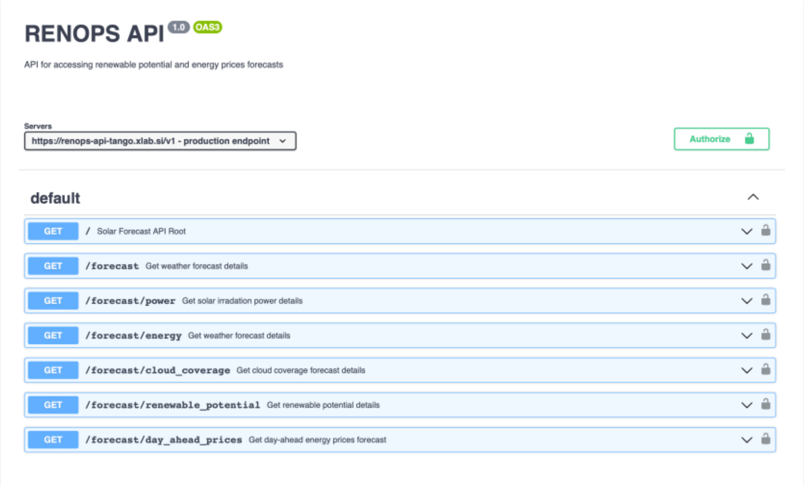


Figure 38. Energy price prediction model

The model uses PyTorch-forecasting (Beitner, 2020) implementation of the temporal fusion transformer. The model will be deployed as a docker container and will use REST API to communicate with the gateway. A REST API call to *forecast/price*, will return the prediction to the gateway, which will then process and forward to the user.

3.6.2.5.2 Service

RENOPS service serves as the gateway to the rest of the system. To communicate we use REST API, using the OpenAPI specifications as seen in Figure 39 below.



RENOPS API 1.0 OAS3

API for accessing renewable potential and energy prices forecasts

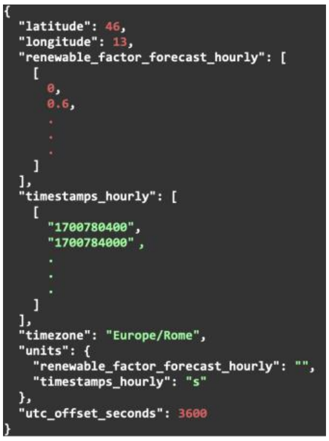
Servers
https://renops-api-tango.xlab.si/v1 - production endpoint

default

GET / Solar Forecast API Root
GET /forecast Get weather forecast details
GET /forecast/power Get solar irradation power details
GET /forecast/energy Get weather forecast details
GET /forecast/cloud_coverage Get cloud coverage forecast details
GET /forecast/renewable_potential Get renewable potential details
GET /forecast/day_ahead_prices Get day-ahead energy prices forecast

Figure 39. OpenAPI specifications (available in GitHub repo)

For example, a request to *forecast/renewable_potential* is expected to return a response as seen in Figure 40 below. As we can see, the response is in JSON format and includes all the necessary entries to further parse the results. Among other metadata, we provide units that can be matched with data entries.



```
{
  "latitude": 46,
  "longitude": 13,
  "renewable_factor_forecast_hourly": [
    [
      0,
      0.6,
      .
      .
      .
    ]
  ],
  "timestamps_hourly": [
    [
      "1700780400",
      "1700784000" ,
      .
      .
      .
    ]
  ],
  "timezone": "Europe/Rome",
  "units": {
    "renewable_factor_forecast_hourly": "",
    "timestamps_hourly": "s"
  },
  "utc_offset_seconds": 3600
}
```

Figure 40. API response for forecast/renewable_potential

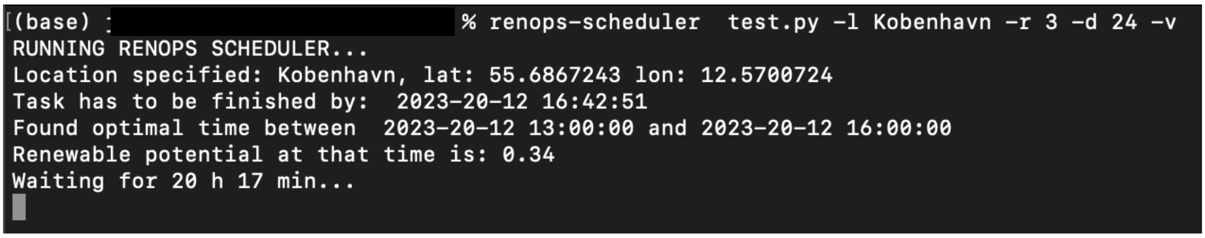
Furthermore, one may test the API by visiting:

- Renewable potential for Ljubljana
- Price forecast for Slovenia

The is built using Flask and served on a Gunicorn WSGI server. Additional security measures were added, including basic auth using API keys, along with the implementation of request limiters (flask-limiter) to mitigate DDoS attacks.

3.6.2.5.3 Scheduler

- The scheduler will be delivered as a Python package. Currently, the package and repository are hosted on XLAB premises. Soon, it will be migrated and made available in TANGO GitHub project[11]. The scheduler was written in Python and can be used in directly in CLI to run any type of executable. Alternatively, it can be imported into a Python script and used as a module.
- Examples in Figure 41 and Figure 42 below show how the scheduler is used in the CLI. The first positional argument is the script we want to shift. It can be any script as long as its executable. The next argument is location. We use geocoding API[12] to translate city names to coordinates. This streamlines the process, eliminating the necessity for entering precise location names. Optionally, if we pass the "-l auto" argument, the location will use IP together with geo-coding API to locate the machine. Other arguments such as runtime and deadline are optional, but a user may specify them in case desired.

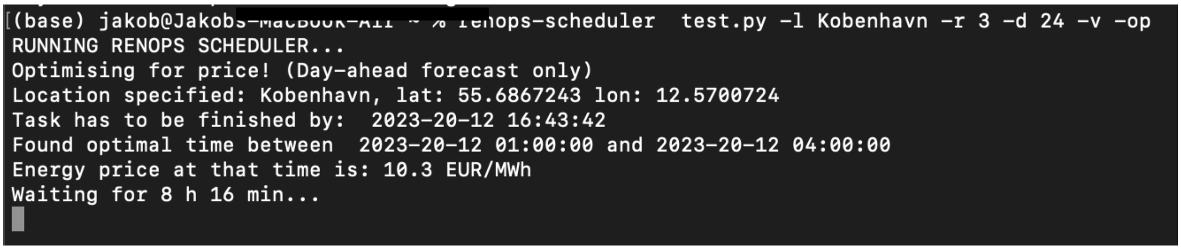


```
[(base) [redacted] % renops-scheduler  test.py -l Kobenhavn -r 3 -d 24 -v
RUNNING RENOPS SCHEDULER...
Location specified: Kobenhavn, lat: 55.6867243 lon: 12.5700724
Task has to be finished by:  2023-20-12 16:42:51
Found optimal time between  2023-20-12 13:00:00 and 2023-20-12 16:00:00
Renewable potential at that time is: 0.34
Waiting for 20 h 17 min...
```

Figure 41. Scheduler CLI example for renewable potential optimisation for Copenhagen

[11] https://github.com/orgs/TANGO-EU-PROJECT/teams/xlab/repositories
[12] https://nominatim.openstreetmap.org

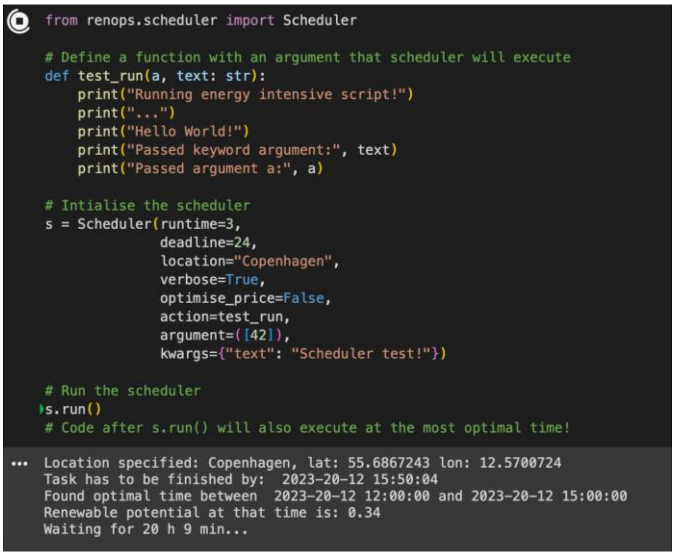
Figure 41 depicts an example when we add the optimise price flag “-op”. In this case, multiday-ahead predictions are not available, so we directly call day-ahead the predictions from ENTSO-e. As mentioned, price predictions are available for the EU only.



```
[(base) jakob@Jakobs[redacted]enops-scheduler  test.py -l Kobenhavn -r 3 -d 24 -v -op
RUNNING RENOPS SCHEDULER...
Optimising for price! (Day-ahead forecast only)
Location specified: Kobenhavn, lat: 55.6867243 lon: 12.5700724
Task has to be finished by:  2023-20-12 16:43:42
Found optimal time between  2023-20-12 01:00:00 and 2023-20-12 04:00:00
Energy price at that time is: 10.3 EUR/MWh
Waiting for 8 h 16 min...
```

Figure 42. Scheduler CLI example for price optimisation for Copenhagen

Figure 42 depicts an example used in a Jupiter notebook. The usage is similar to other schedulers such as Sched[13] integrated in Python 3. With the addition that we incorporate price or renewable potential into the decision.



```python
from renops.scheduler import Scheduler

# Define a function with an argument that scheduler will execute
def test_run(a, text: str):
    print("Running energy intensive script!")
    print("...")
    print("Hello World!")
    print("Passed keyword argument:", text)
    print("Passed argument a:", a)

# Intialise the scheduler
s = Scheduler(runtime=3,
              deadline=24,
              location="Copenhagen",
              verbose=True,
              optimise_price=False,
              action=test_run,
              argument=([42]),
              kwargs={"text": "Scheduler test!"})

# Run the scheduler
s.run()
# Code after s.run() will also execute at the most optimal time!
```

```
Location specified: Copenhagen, lat: 55.6867243 lon: 12.5700724
Task has to be finished by:  2023-20-12 15:50:04
Found optimal time between  2023-20-12 12:00:00 and 2023-20-12 15:00:00
Renewable potential at that time is: 0.34
Waiting for 20 h 9 min...
```

Figure 43. Scheduler in a Jupyter Notebook example for Copenhagen

3.6.2.6 Future work

Future work can be split into two parts: one being further improvement of RENOPS and the other being the integration with other components of the TANGO platform. Currently scheduler supports temporal shifting, the option to shift in geographical space, to the most optimal data centre. If we used temporal shift to find the most optimal time for training AI models, we could use geographical shift to find the optimal data centre for inference. While the training pipeline for the energy price prediction model is implemented and the model trained, additional work must be invested in setting up the MLOps pipeline for continuous improvement of the model with new data. Additionally, we plan to run extensive evaluations of our model’s performance.

On the side of integration with other components, we need to integrate service into other components such as T5.1 Exploratory data analysis engine and T5.2, energy-efficient AI model training where RENOPS could be integrated into their Federated learning or MLOps component.

[13] https://docs.python.org/3/library/sched.html

4 Conclusions and future work

Building upon the work carried out in the scope of WP2 during the first months of the project, this deliverable presents the initial results in the form of an intermediate iteration of the software components developed in the scope of all tasks of TANGO WP5. Besides the vision and objectives of WP5 as a whole, the document explains for each task its current state of the implementation of the tools in support of AI-related activities.

The document presents a detailed explanation of the implementation done so far per task and component. Each component provides its description, internal architecture, features implemented, expected support for pilots, software artifacts, and potential future work, as well as pointers to software/demos where required. A summary of the current release and future work per task and component can be found in Table 17.

Table 17: Summary of current release main features and future work

Task	Component name	Current release main features	Future work
T5.1	EDAE	Data analysis submodule: • Descriptive analysis. • Univariate and multivariate analysis. Handling missing data submodule: • Missing/Duplicate data visualisation. • Handling missing values. Data processing/preparation submodule: • Semi-automated data transformation. • Dimensionality reduction. • Variable selection and feature engineering. • Outlier identification and treatment. • Anomaly detection. • Simple automatic clustering. Visualisation techniques. Support mainly for the METRO pilot requirements so far.	• Ensuring comprehensive coverage and effectiveness across the TANGO project. • Integration of EDAE with user interfaces (T6.3) • Potential integration with other WP5 results. • Finalising and fine-tuning features specific to each pilot.
T5.2	MLOps, FL and AutoML	MLOps component: • Support for model training and tunning. • Support for tracking ML experiments. • Model management and deployment. • ML Model registry.	MLOps component: • Provision of an initial set of MLFlow Recipes. • Jupyter notebook templates supporting a kind of generic pilots ML pipelines.

Task	Component name	Current release main features	Future work
		• Model serving. Federated Learning component: • Execution of simple FL scenarios. • Support for Keras (PyTorch in Beta). AutoML component: • Based on MLJAR, aimed at relatively non-expert users. • Initial training on a user-provided CSV dataset using different ML-based models. • Model selection. Support mainly for the banking and manufacturing (FMAKE) pilots to be tested	• Investigate potential integration with other tools (i.e., AutoML FL, XAI, RENOPS). Federated Learning component: • Further support for PyTorch. • Support other communication protocols. • Allow further configurations in the communication protocols. • Evaluate further support for newer pilot requirements. • Potential integration of the results of the ML training cycles with the MLOps tool. AutoML component: • Improve deployment strategy. • Customized development on the notebooks served based on pilots' needs. • Evaluate other approaches or frameworks to extend the AutoML functionality. • Potential integration with MLOps.
T5.3	TornadoVM	Hardware acceleration software. Automated execution plans on the hardware accelerator devices that will be available on a single node. TornadoVM: To express in Java the functions that are going to be supported for hardware acceleration, including Multiple Tasks on Multiple Devices feature. GraalPython: To deliver TornadoVM functionality implemented in Java classes to Python programs. Integration with EDAE (T5.1) for acceleration of the execution of that component.	• Integration of power metrics in the TornadoVM Runtime to monitor the energy. • Implement a ML model able to predict which heterogeneous device will be the most energy efficient. • Extend the coverage to support more functions for the EDAE
T5.4	PEC & PAT	PEC: Support with several components to deliver a list of privacy risk scores, privacy impacts and a set of recommendations to mitigate the privacy risks: • PEC Data Initialiser (PDI). • PEC Risk Identifier (PRI).	PEC • Extend privacy risk assessment for AI/ML services, identifying possible threat scenarios, assessing their privacy impacts and justifying safeguards to mitigate the impacts. • Guidance on complying with GDPR.

Task	Component name	Current release main features	Future work
		• PEC Scoring Engine (PSE). • PEC Recommendation Engine (PRE). • PEC Compliance Engine (PCE). • PEC Dashboard to visualise the outcomes. PAT: Support for secure and privacy-preserving data exchange based on consent. • Consent-based access control (CBAC) mechanism for secure and privacy-preserving compliant. • Automated privacy monitoring. • GDPR compliance. • Secure data sharing.	• Complete integration with PAT. • API endpoints to allow potential integration with other WP5 components. • Integration and deployment of PEC in the TANGO architecture. • Extend the analysis to other pilots. PAT • Validate the proposed data model for the smart hospitality pilot. • Develop the data model for the public administration pilot. • Enhance the front-end interfaces for both data subjects and organizations. • Add and implement more rules-based compliance based on the consent, audit logs, and organization user's attributes. • Integrate the identity management tool in the organization side for managing the user of the organizations. • Improve the privacy of PAT by enforcing the consent. • Convert the mobile UI design to a desktop UI design for facilitating the integration with other TANGO components. • Investigate integration with other components, especially from WP3 and WP4 and with the TANGO Connector.
T5.5	X-AI	• Library to provide ML and Deep Learning (DL) model explainability. • Four different explainability algorithms supported (LIME, Grad-Cam, LRP and Anchors). • Present the explainability results in an easy-to-understand visual way.	• Complete the custom XAI algorithm for the Smart Manufacturing pilot. • Start implementing XAI techniques on the models that will be trained by the banking pilot. • Design and implement the connection between the XAI module and ATOS's Federated Learning component. • Dockerise the component and deploy it on the pilot's premises.
T5.6	RENOPS	Predictions • A 5-day ensemble forecast of renewable energy availability for any location.	RENOPS • Scheduler to support location shifting besides temporal shifting.

Task	Component name	Current release main features	Future work
		• A Day ahead electrical energy price forecast for all EU member states. • A 3-day ahead price forecast for Slovenia. Scheduler based on renewable energy availability. • Command line interface program. • Python package. A REST API access to renewable energy and price information. • Basic authentication using an API token. • Request IP limiter to mitigate DDoS attacks. Dockerised services and packaged services. CI/CD development pipeline.	• Setting up a MLOps pipeline for improvement of the model with future data. • Run extensive evaluations of the model's performance. Integration with other WP5 components

In general, for all tasks, the delivery of the intermediate release provides a first working prototype to be integrated in the TANGO platform and tested by the users. Most of the components work in isolation, which means that the immediate future work will be related to the closest technical milestone: the integration in M21 (May 2024). Therefore, partners involved in the current release of WP5 software artifacts will be paying attention to the guidelines from WP6 to help in the integration and testing of the components in the architecture of the TANGO platform. Each task provides in their respective sections of the document an initial estimation of the future work towards the final release of WP5, expected in M30 (February 2025), summarised as well in Table 17.

5 References

Anna Maria Oosthuizen, R. I.-L. (2022). The relationship between renewable energy and retail electricity prices: Panel evidence from OECD countries, . *Energy, Volume 238, Part B,* .

Beitner, J. (2020, September 12). *https://towardsdatascience.com/introducing-pytorch-forecasting-64de99b9ef46*. (Medium) Retrieved July 6, 2023, from https://towardsdatascience.com/introducing-pytorch-forecasting-64de99b9ef46

Bryan Lim, S. Ö. (2021). Temporal Fusion Transformers for interpretable multi-horizon time series forecasting, . *International Journal of Forecasting, , 37*(4), 1748-1764.

Commission, E. (2020, 11 9). *Energy Consumption of Data Centres in EU Member States*. Retrieved 12 20, 2023, from https://ec.europa.eu/reports/energy-consumption-of-data-centres-in-eu-member-state

Consumption, G. D. (2022, 12 30). *Global Data Centre Energy Consumption.* Retrieved 12 20, 2023, from https://www.gdec.org/reports/data-centres-and-global-electricity-consumption

ENTSO-E. (n.d.). *transparency.entsoe*. (ENTSO-E) Retrieved 12 20, 2023, from https://transparency.entsoe.eu/dashboard/show

H. Rong, H. Z. (may 2016). Optimizing energy consumption for data centers. *Renewable and Sustainable Energy Reviews, 58*, 674-691.

iea. (2023, 7 1). (iea) Retrieved 12 20, 2023, from https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks

TANGO D2.1. (2023). *D2.1 State-of-the-Art & GAP analysis Distributed Data Management, Processing and Storage.* TANGO Consortium.

TANGO D2.2. (2023). *D2.2 User Needs and Requirements & Use Case Scenarios.* TANGO Consortium.

TANGO D2.3. (2023). *D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment.* TANGO Consortium.

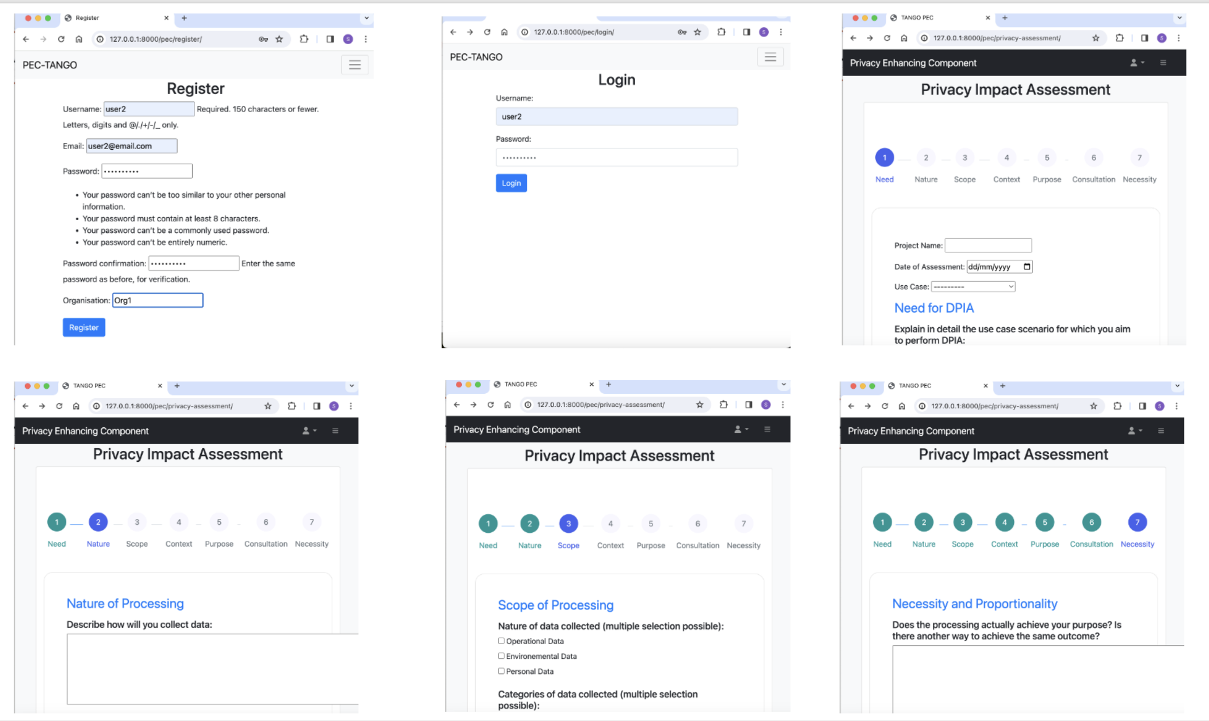
Zippenfenig, P. (2023). Open-Meteo.com Weather API [Computer software]. Zenodo.

6 Annexes

6.1 Annex 1– T5.4 (PEC + PAT)

Privacy Enhancing Component (PEC):

Figure 44 presents the screenshots from the PEC Web Application displaying UI of user registration, login and various snapshots of the privacy risk assessment form used to collect data.


Register
127.0.0.1:8000/pec/register/
PEC-TANGO
Register
Login
127.0.0.1:8000/pec/login/
TANGO PEC
127.0.0.1:8000/pec/privacy-assessment/
Privacy Enhancing Component
Privacy Impact Assessment
1 Need
2 Nature
3 Scope
4 Context
5 Purpose
6 Consultation
7 Necessity

Register

Username: user2 Required. 150 characters or fewer. Letters, digits and @/./+/-/_ only.

Email: user2@email.com

Password: ••••••••••

- Your password can't be too similar to your other personal information.
- Your password must contain at least 8 characters.
- Your password can't be a commonly used password.
- Your password can't be entirely numeric.

Password confirmation: •••••••••• Enter the same password as before, for verification.

Organisation: Org1

Register

Login

Username:
user2

Password:
••••••••••

Login

Privacy Impact Assessment

Project Name:

Date of Assessment: dd/mm/yyyy

Use Case: ---------

Need for DPIA

Explain in detail the use case scenario for which you aim to perform DPIA:

Nature of Processing

Describe how will you collect data:

Scope of Processing

Nature of data collected (multiple selection possible):

☐ Operational Data
☐ Environemental Data
☐ Personal Data

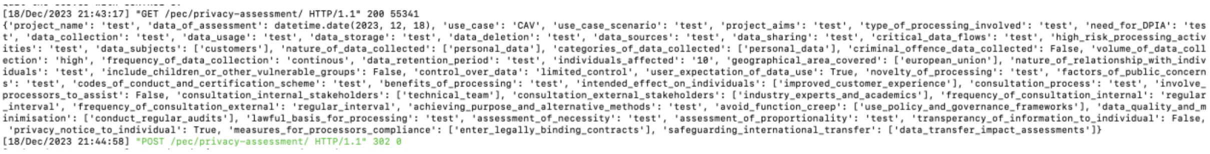
Categories of data collected (multiple selection possible):

Necessity and Proportionality

Does the processing actually achieve your purpose? Is there another way to achieve the same outcome?

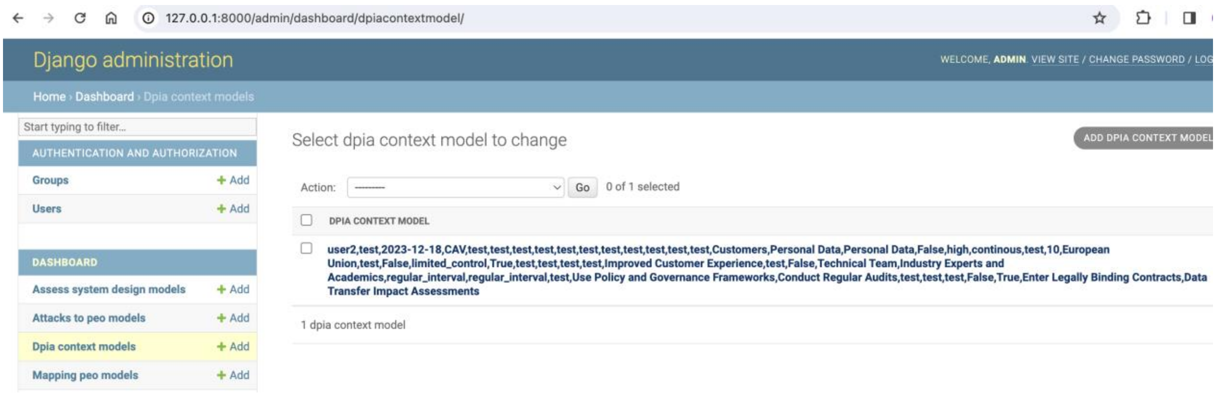
Figure 44. PEC UI Screenshots

Figure 45 presents a response from the backend that the input provided using the privacy risk assessment form has been save to the database. Figure 46 presents a snapshot of the respective database showing the saved response.



```
[18/Dec/2023 21:43:17] "GET /pec/privacy-assessment/ HTTP/1.1" 200 55341
{'project_name': 'test', 'data_of_assessment': datetime.date(2023, 12, 18), 'use_case': 'CAV', 'use_case_scenario': 'test', 'project_aims': 'test', 'type_of_processing_involved': 'test', 'need_for_DPIA': 'tes
t', 'data_collection': 'test', 'data_usage': 'test', 'data_storage': 'test', 'data_deletion': 'test', 'data_sources': 'test', 'data_sharing': 'test', 'critical_data_flows': 'test', 'high_risk_processing_activ
ities': 'test', 'data_subjects': ['customers'], 'nature_of_data_collected': ['personal_data'], 'categories_of_data_collected': ['personal_data'], 'criminal_offence_data_collected': False, 'volume_of_data_coll
ection': 'high', 'frequency_of_data_collection': 'continous', 'data_retention_period': 'test', 'individuals_affected': '10', 'geographical_area_covered': ['european_union'], 'nature_of_relationship_with_indiv
iduals': 'test', 'include_children_or_other_vulnerable_groups': False, 'control_over_data': 'limited_control', 'user_expectation_of_data_use': True, 'novelty_of_processing': 'test', 'factors_of_public_concern
s': 'test', 'codes_of_conduct_and_certification_scheme': 'test', 'benefits_of_processing': 'test', 'intended_effect_on_individuals': ['improved_customer_experience'], 'consultation_process': 'test', 'involve_
processors_to_assist': False, 'consultation_internal_stakeholders': ['technical_team'], 'consultation_external_stakeholders': ['industry_experts_and_academics'], 'frequency_of_consultation_internal': 'regular
_interval', 'frequency_of_consultation_external': 'regular_interval', 'achieving_purpose_and_alternative_methods': 'test', 'avoid_function_creep': ['use_policy_and_governance_frameworks'], 'data_quality_and_m
inimisation': ['conduct_regular_audits'], 'lawful_basis_for_processing': 'test', 'assessment_of_necessity': 'test', 'assessment_of_proportionality': 'test', 'transperancy_of_information_to_individual': False,
 'privacy_notice_to_individual': True, 'measures_for_processors_compliance': ['enter_legally_binding_contracts'], 'safeguarding_international_transfer': ['data_transfer_impact_assessments']}
[18/Dec/2023 21:44:58] "POST /pec/privacy-assessment/ HTTP/1.1" 302 0
```

Figure 45. Backend response on successfully saving user input to database



127.0.0.1:8000/admin/dashboard/dpiacontextmodel/

Django administration

WELCOME, ADMIN. VIEW SITE / CHANGE PASSWORD / LOG

Home › Dashboard › Dpia context models

Start typing to filter...

AUTHENTICATION AND AUTHORIZATION

- Groups + Add
- Users + Add

DASHBOARD

- Assess system design models + Add
- Attacks to peo models + Add
- Dpia context models + Add
- Mapping peo models + Add

Select dpia context model to change

ADD DPIA CONTEXT MODEL

Action: --------- Go 0 of 1 selected

□ DPIA CONTEXT MODEL

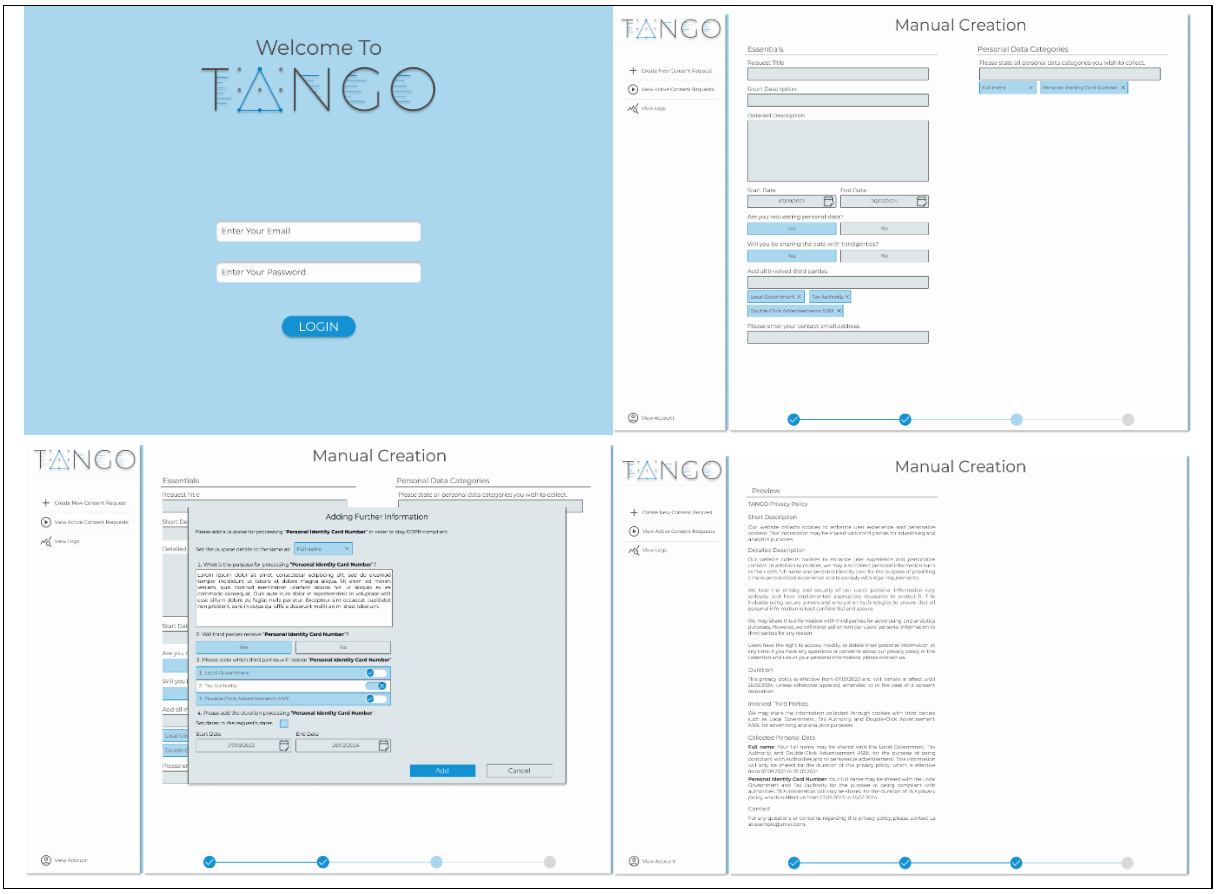
□ user2,test,2023-12-18,CAV,test,test,test,test,test,test,test,test,test,test,test,Customers,Personal Data,Personal Data,False,high,continous,test,10,European Union,test,False,limited_control,True,test,test,test,test,Improved Customer Experience,test,False,Technical Team,Industry Experts and Academics,regular_interval,regular_interval,test,Use Policy and Governance Frameworks,Conduct Regular Audits,test,test,test,False,True,Enter Legally Binding Contracts,Data Transfer Impact Assessments

1 dpia context model

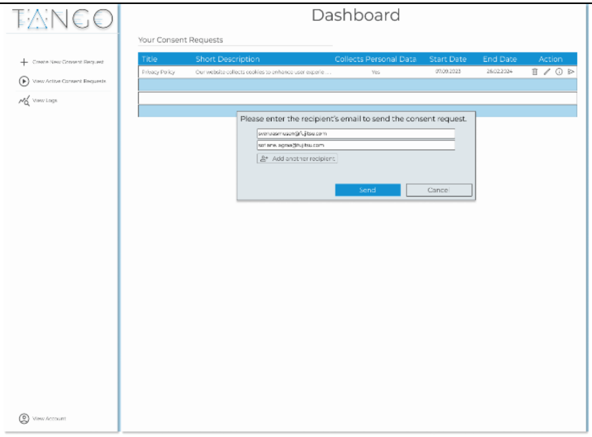
Figure 46. User input recorded in the respective database table

Privacy Assurance Tool (PAT):

Figure 47 and Figure 48 show mock-ups using Figma for the data requestor (organization side) and data subject's (data subject side), respectively. Figure 47 outlines the steps involved in creating a consent by the data requestor for the data subject. It provides a guide on the essential information that must be included in the consent, the crucial information that the data requester imposes on the data subject, as well as the requested data (mandatory or not). Figure 48 outlines the steps involved in accepting a consent by the data subject. The data subject may accept or decline each requested non-mandatory data by the data requestor. The data subject can also choose the duration of the stored data in the data requestor database.



Welcome To
TANGO
Enter Your Email
Enter Your Password
LOGIN
Manual Creation
Essentials
Personal Data Categories
Request Title
Short Description
Detailed Description
Start Date
End Date
Are you requesting personal data?
Will you be sharing the data with third parties?
Add all involved third parties
Local Government
Tax Authority
Double-Click Advertisements ASBL
Please enter your contact email address
Create New Consent Request
View Active Consent Requests
View Logs
View Account
Adding Further Information
Add
Cancel
Preview
TANGO Privacy Policy



Dashboard

Your Consent Requests

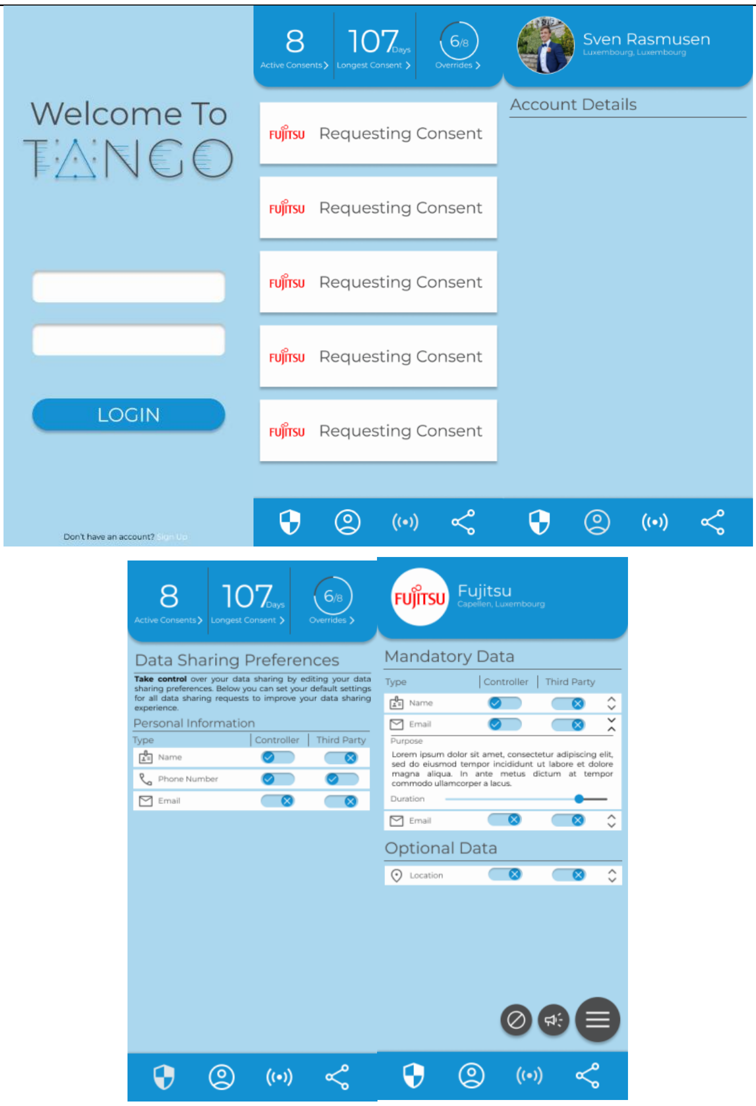
Title	Short Description	Collects Personal Data	Start Date	End Date	Action
Privacy Policy	Our website collects cookies to enhance user experie…	Yes	07.09.2023	26.02.2024	

Please enter the recipient's email to send the consent request.

svenrasmussen@fujitsu.com

[illegible]@fujitsu.com

Figure 47. Data requestor User Interface mock-up. Consent creation and request



Welcome To TANGO

Don't have an account? Sign Up

Sven Rasmusen
Luxembourg, Luxembourg

Account Details

Fujitsu
Capellen, Luxembourg

Data Sharing Preferences

Take control over your data sharing by editing your data sharing preferences. Below you can set your default settings for all data sharing requests to improve your data sharing experience.

Personal Information

Type	Controller	Third Party
Name	✓	☒
Phone Number	✓	✓
Email	☒	☒

Mandatory Data

Type	Controller	Third Party
Name	✓	☒
Email	✓	☒

Purpose

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. In ante metus dictum at tempor commodo ullamcorper a lacus.

Duration

Type	Controller	Third Party
Email	☒	☒

Optional Data

Type	Controller	Third Party
Location	☒	☒

Figure 48. Data subject's User Interface mock-up

The following tables represent a list of endpoints and their allowed operations for users' organisations (organisation side), clients (user side), and consents.

Table 18: List of endpoints for PAT users' organisations

Endpoint	Operation	Parameters	Description
/api/users	GET	No parameter	Retrieves all users

/api/users/save	POST	Request body :
		{
		name: String,
		email: String,
		purposes: String[],
		actions: String[],
		role: String,
		location: String,
		duration: String
		}
			Creates or updates a user

/auth/login	POST	Request body :
		{
		email: String
		}
			Logs in an existing user and returns a JWT

Table 19: List of endpoints for PAT consents

Endpoint	Operation	Parameters	Description
/api/consents	GET	No parameter	Retrieves all consents
/api/consents/{id}	GET	Path variable : id	Retrieves a specific consent

/api/consents/save	POST	Request body:
		{
		consent_ID : Long,
		client_ID : Long,
		consentDetails: [
		{
		role : String[],
		action: String[],
		intendedPurposes:
		String[],
		location: String,
		duration: String,
		dataType: String[]
		}
		]
		}
			Creates or updates a consent

/api/consents/delete/{id}	DELETE	Path variable : id	Deletes consent by ID

Table 20: List of endpoints for PAT Clients

Endpoint	Operation	Parameters	Description
/api/clients	GET	No parameter	Retrieves all client data that a user can access
/api/clients/{id}	GET	Path variable : id	Retrieves specific client data

/api/clients/save	POST	Request body:
		{
		id : Long,
		mandatoryData: {
		firstname: String,
		lastname: String,
		passportID: String,
		nationality: String,
		credit_card: String,
		arrivalDate: String,
		departureDate: String,
		email: String,
		phone: String
		},
		preferencesData: {
		room: String[],
		food: String[],
		breakfastTime: String,
		checkoutTime: String
		}
			creates or updates client data

/api/clients/delete/{id}	DELETE	Path variable : id	deletes specific client data