D8.4 –Smart Monitoring System Design Document

Project Information

Grant Agreement Number	958454
Project Full Title	Intelligent Water Treatment for water preservation combined with simultaneous energy production and material recovery in energy intensive industries
Project Acronym	intelWATT
Funding scheme	IA
Start date of the project	1st October 2020
Duration	42 months
Project Coordinator	Andreas Sapalidis (NCSR)
Project Website	https://www.intelwatt.eu

Deliverable Information

Deliverable n°	D8.4
Deliverable title	Smart Monitoring System Design Document
WP no.	8
WP Leader	TECHEDGE
Contributing Partners	FUELICS, NOKIA, NIJHUIS
Nature	Other
Authors	Francisco Javier Corredera Pérez (Techedge)
Contributors	
Reviewers	Kouvelos Evangelos, Andreas Sapalidis (NCSR)
Contractual Deadline	30/09/2021
Delivery date to EC	30/09/2021

Dissemination Level

PU	Public	✓
PP	Restricted to other programme participants (incl. Commission Services)	
RE	Restricted to a group specified by the consortium (incl. Commission Services)	
CO	Confidential, only for the members of the consortium (incl. Commission Services)	

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958454".

Document Log

Version	Date	Authors	Description of Change
V1.0	20/09/2021	Francisco Javier Corredera Pérez	First release
V1.1	23/09/2021	Kouvelos Evangelos	First review
V2.0	26/09/2021	Andreas Sapalidis	Final review

“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958454”.

Table of Contents

List of figures

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958454".

1 Executive Summary

Deliverable 8.4, describes the methodology, focusing on the design of the other capabilities and functionalities that are necessary to support the Smart Monitoring System and have not been described in the previous deliverables.

These new capabilities will cover the administration of the system, definition of notice rules, simulation of operations and what-if scenarios, application of AI recommendations and information about the machine learning models and performance.

“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958454”.

2 Introduction

This report describes the architecture and requirements for a flexible and powerful smart monitoring system for which all sensor data should be collected and, at the time that they are processed, the insights will be used to feed deep learning models, present their AI recommendations and help decision making through monitoring dashboards.

The smart monitoring concept refers to the monitoring and optimization of a process using various software components to build the apps that allow visualizing the current status, defining the rules that will be applied and finally define the changes to optimize the process.

It´s also necessary to provide an admin application that allow controlling the back-end process, the status of all the components and the user's administration.

The solution components will be a common infrastructure to the three case studies and will be hosted on cloud and on the plant facilities (using AI edge capabilities) in order to make the system as robust and reliable as possible and under high security provided by the IMPACT platform that will additionally act as intermediary platform between shop floor devices and other systems that will be hosted in cloud.

Data processing will be carried by elastic infrastructure components based on big data techniques and ensure the exchange of information with PLC systems such as SCADA.

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958454".

3 Theoretical architecture

For the deployment of the proposed solution, a hybrid scheme has been decided to be implemented. This architecture will have:

- On-premise components (installed and deployed inside the plant), which will make it possible to respond to the different plant scenarios with very low latency and in almost real time. This implementation enables the artificial intelligence system to continue operating in the event of communications outage.
- Cloud components (set of solutions and services) that offer all the necessary tools for the implementation of the project.

The choice of this approach takes advantage of the positive aspects of both deployments, while reducing the disadvantages of each of them.

The on-premise components allow high-speed response to the different plant scenarios, as they are close to the machine, and they offer high availability and security, as they are isolated from the rest of the plant's systems.

On the other hand, the cloud components offered by cloud solution providers, such as Microsoft or Google, provide all the necessary components to deploy the solution, without the duty to assume the cost of purchasing a computing platform, nor the costs of maintenance, licensing, etc.

The value points offered by cloud solutions are:

- Capacity

  The computing and storage capacity needed at any time is available, therefore no expensive hardware platforms to support the business is needed.

- Flexibility

  It is not necessary to always have the maximum computing, storage, or connectivity capacity available. Flexible cloud solutions allow adapting to the current situation.

- Low cost

  The cost of this type of solution is much lower than a specific information system and it is spread over time through a monthly fee according to consumption. Some cloud platform providers can even charge only for the minutes the services were used.

- Security

  The level of security is very high and completely configurable, supported by a group of experts who control and limit all access to the platform in a secure way.

- High availability

  Systems will be completely secure, having a copy, replication, and balancing systems. The response to the needs of the system at any time is ensured.

Having identified the different types of components and their interrelationship, a description of each of them will be carried out.

The figure below represents in a schematic and summarised way the complete architecture of the proposed solution, with all the identified data flows. It is the basis for the description of each of the components.

“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958454”.

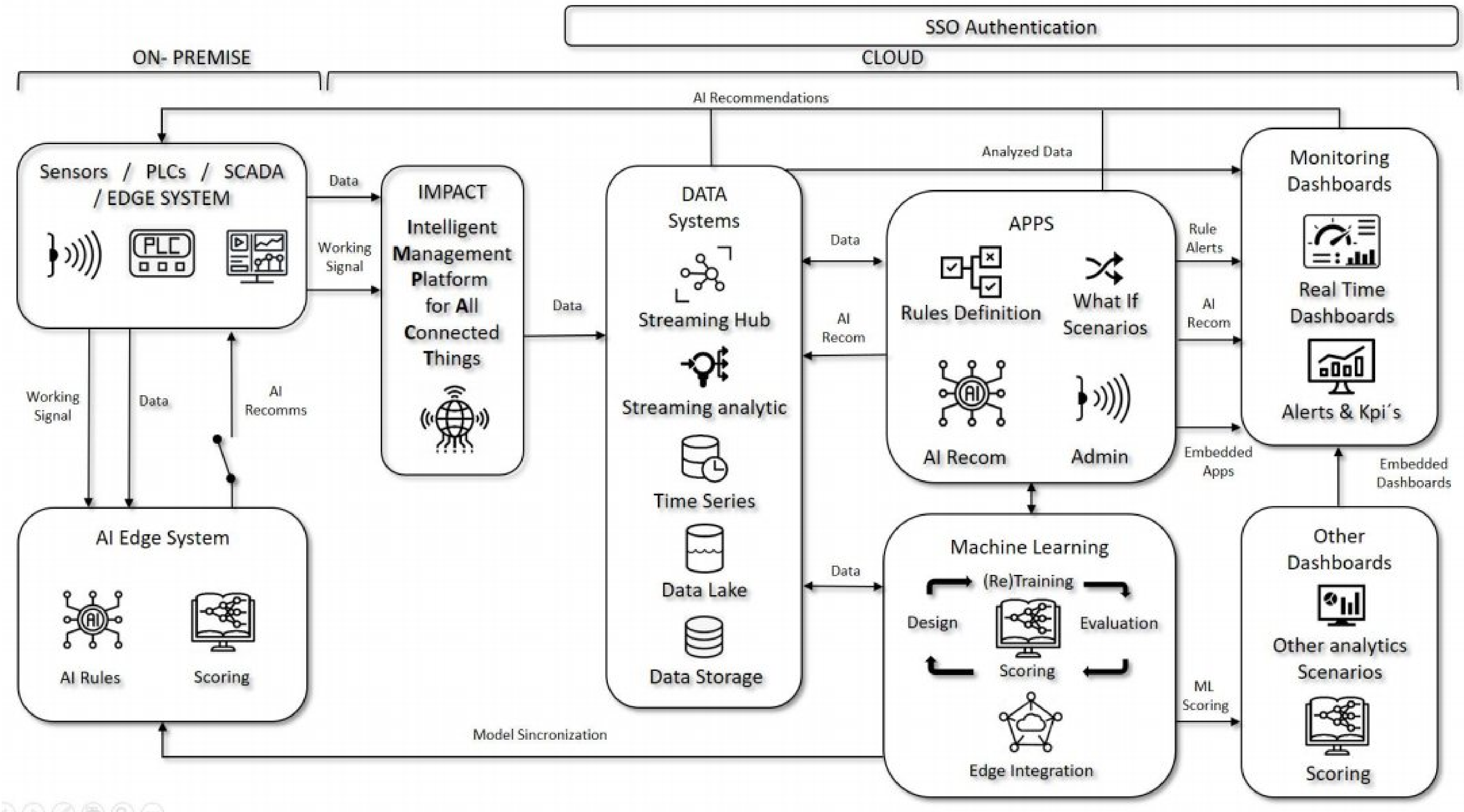

Figure 1 - **Theoretical architecture**

“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958454”.

3.1 Apps

The Smart Monitoring solution requires several applications or points of interaction with the user that give meaning to the functionalities offered. It is very important to bear in mind that this is not a static solution, but that many of the processes to be carried out allow, and even require, interaction with the users of the application.

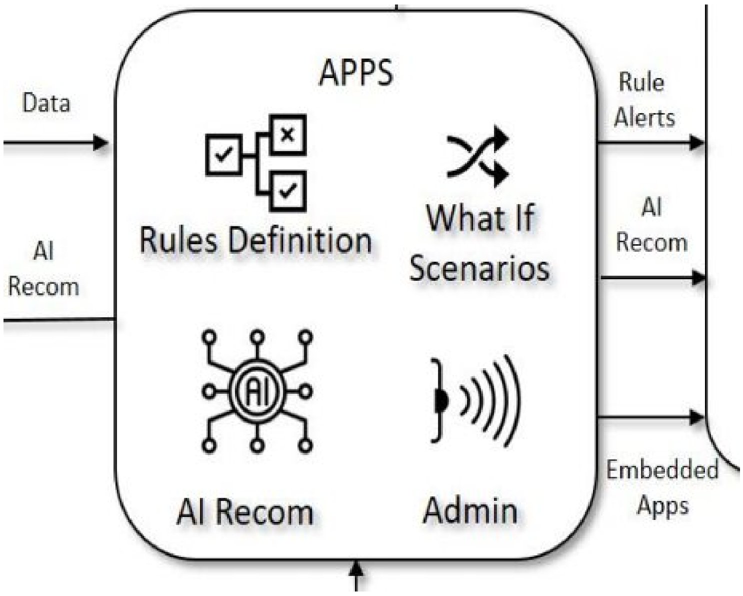

Figure 2 - Apps

Each of the applications will have a specific functionality, focused on covering a requirement of the solution, and will have its own input and output flow. The different applications are described hereafter.

3.1.1 Admin app

This solution will make it possible to register and manage the different devices and signals and their correlation with the IMPACT platform, as well as providing additional information for display on the different dashboards.

The information to be recorded will be stored in the Data storage component.

Through this admin app, we must define all the signals, devices and systems that are going to be monitored. It´s necessary to provide the connectivity information to connect to *Impact system*, and add the metadata info that will be used in the other apps, like name of the component, unit of measure or warning data.

3.1.2 Rules definition

Another functionality to be implemented is the creation of rules and alerts based on the signal from the different devices. The aim is that the values to be considered in these rules can vary according to the different environments of each of the case studies and the platform users themselves can manage them.

As in the previous application, the information is stored in the data storage, as this component is the storage point for all the rules.

“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958454”.

This functionality is proposed to extend the value of the monitoring dashboards, as complex rules, that consider different equipment, signals and results can be created.

The ability to create these rules is provided through a visual tool that allows these processes to be carried out.

3.1.3 What-if scenarios

Another functionality to be deployed is the ability to see what could happen in the system if any of the configuration parameters of a usage scenario are varied. To do this, the capacity of the simulation models created for learning together with the capacity to visualize and vary these parameters will be used.

In this case, the simulation models are the center of the app. The first step is to define the input variables of the simulation, making a snapshot of the actual information, and then the user can change the input parameters. At the time the input parameters are changed, the simulation models will return the output data that the process will return using machine learning algorithms.

The what-if scenarios app allows the user to do these simulations in a visual mode and without any knowledge of the process.

The simulation models must be created to provide data for the recommendation models learning, because there´s not enough data for the recommendation models, but can be also used to the what-if scenarios, and it´s explained in 3.1.4

This application provides support for decision making that can help in the correct configuration of the parameters and avoid damage or errors when manipulating different parameters.

This component allows this type of task to be carried out without risk and on demand. It will require the capabilities of the different predictive maintenance models, as well as data storage tools.

3.1.4 AI Recommendations

When the Smart monitoring solution is in operation, a learning period will be required in which the changes proposed by the artificial intelligence system will need to be validated prior to their application.

This validation will control that the proposal is not applied directly to the system, but it is verified beforehand and applied to the system only if the project manager accepts the proposal.

This procedure will make the solution much more secure by implementing a manual verification layer. This manual verification can be disabled when the artificial intelligence system is fully trusted.

As a subsequent step to this procedure, the ability for the AI Edge System component to apply the proposals directly will be activated, if deemed feasible.

This application will use the data stored in the Data storage, will interact with the machine learning layer, and will be integrated with the rest of the visualization applications.

3.2 Machine learning

This is the main component of the solution, as the design, training, deployment, evaluation and scoring of the artificial intelligence models will take place here.

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958454".

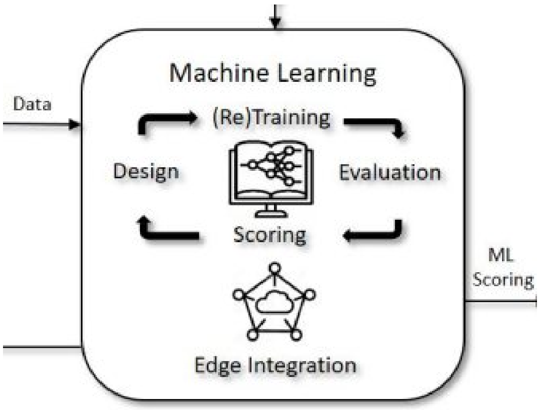

Figure 3 - Machine Learning

Two types of algorithms are proposed within the scope of this project:

- Algorithms for simulating the behavior of each case study.
- Algorithms for optimizing the operation of each case study.

One of the main objectives of this project is the automatic optimization of the entire water purification/energy generation process according to the different procedures. For this purpose, it is necessary to know the behavior of each process according to the different input variables. Due to the complexity of the systems and the scenario they represent, the optimal way of imitating this behaviour is through simulation.

In order the process simulation to be as reliable as possible, several algorithms of high accuracy must be created. In this component, this process will be carried out.

These simulation models are of vital importance because they provide a lot of historical information to the optimization models, which will provide the actions to be taken to optimize the process.

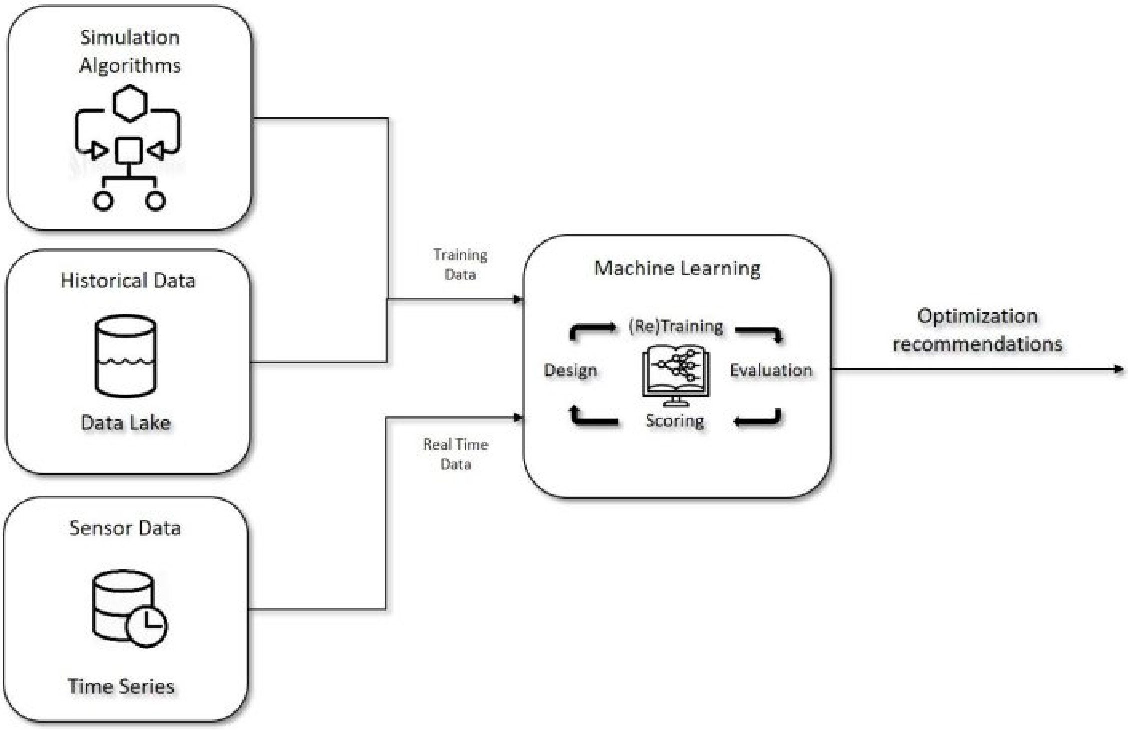

Figure 4 - Machine Learning Data

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958454".

Once the system has a large/sufficient amount of historical information, it moves on to the training phase of the recommendation models. Through various procedures, the correct parameterization of these models is carried out so that they provide the expected results.

As soon as the recommendation models work correctly, the real-time data input is connected, and these models become operational.

To prevent the recommendation models from becoming ineffective, a continuous evaluation, scoring and retraining solution is proposed, whereby the system automatically detects when effectiveness is lowering and retrains itself with historical data.

In a first phase, the optimization recommendations must be validated by the users of the application who also should confirm whether this recommendation is executed. Once it has been validated that the behavior of the solution is correct, the predictive models will be deployed in the AI Edge system component for real-time application in the plant for each case study.

As a summary of this component, the following information sources are identified:

- Training information provided by the simulation algorithms.
- Training information provided by the Data Lake.
- Real-time sensor information provided by the Time series component.

The results of this component are:

- In the first phase, the recommendations to be reviewed by the user of the application.
- In the second phase (automatic recommendations), the predictive models are generated and deployed, which are automatically executed in the AI Edge system component.

As an additional functionality of this component, the evaluation, scoring and retraining processes of the models are managed, and the evaluation data is sent to the visualization dashboards.

3.2.1 Dashboards

The Smart monitoring solution should display information on the current status of the plant at all times, along with information on alerts and KPI's as deemed necessary.

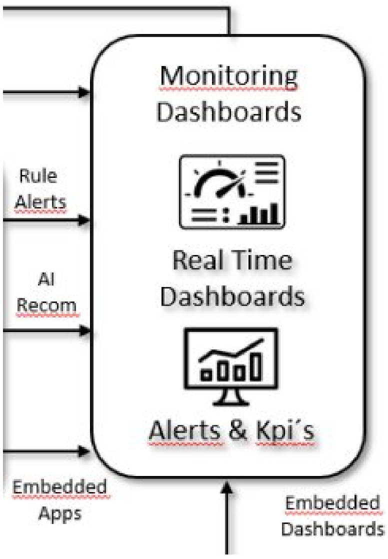

Figure 5 - Dashboards

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958454".

The dashboards will be defined in deliverable D8.14, but the information flows are identified at this point.

The main data input streams are the Streaming Analytics and Time Series components, which provide information on the status of the case studies in near real time.

As a complement to this information, the business data or metadata are employed to identify the names of devices, information zones, name of the plant and any other information about the structure of the case studies. This information is provided by the Data storage component.

The analytical information (KPI's) and alerts are also provided to the monitoring dashboards component. This information is retrieved from all the data sources of the data system, such as Streaming analytic or the Data Lake.

To complete all the expected functionality of the solution, other types of dashboards must be added to allow the visualization of other types of information.

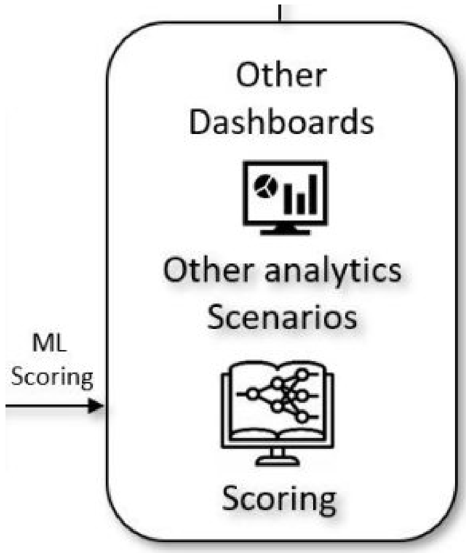

Figure 6 - Dashboards

The first of these is the model tracking or scoring dashboard. This dashboard allows users to visualize the effectiveness of each model, the number of optimization recommendations it has proposed and other historical information about its operation, such as the number of re-trainings it has undergone and whether it is running in the cloud (phase I of the recommendation system) or on-premise (phase II).

An analytical dashboard will also be deployed, where users can visualize the evolution of the recommendation system over time, trends, and comparisons between different time periods.

The purpose of this dashboard is to present a more strategic view of the Smart monitoring system and to allow different users to discover what is happening and how to solve it.

The information from the system will be made available to the users and can be consumed directly by them according to their needs and permissions.

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958454".

4 Platform and architecture components

The previous section described the functional flow of data and explained each of the components identified. This section describes the actual components to be used, their functionality and capabilities.

The first step is to identify the choice of the platform where the solution will be deployed and why it is chosen. The second step describes the way to select the different components to be used, their operation as well as the interaction between them.

It should be noted that, as described above, the Smart monitoring solution have components installed in the plant (on-premise) and on cloud platforms (cloud), with the advantages that this type of architecture brings to the solution.

The architecture diagram of the proposed solution, with its components and data flows, is presented below:

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958454".

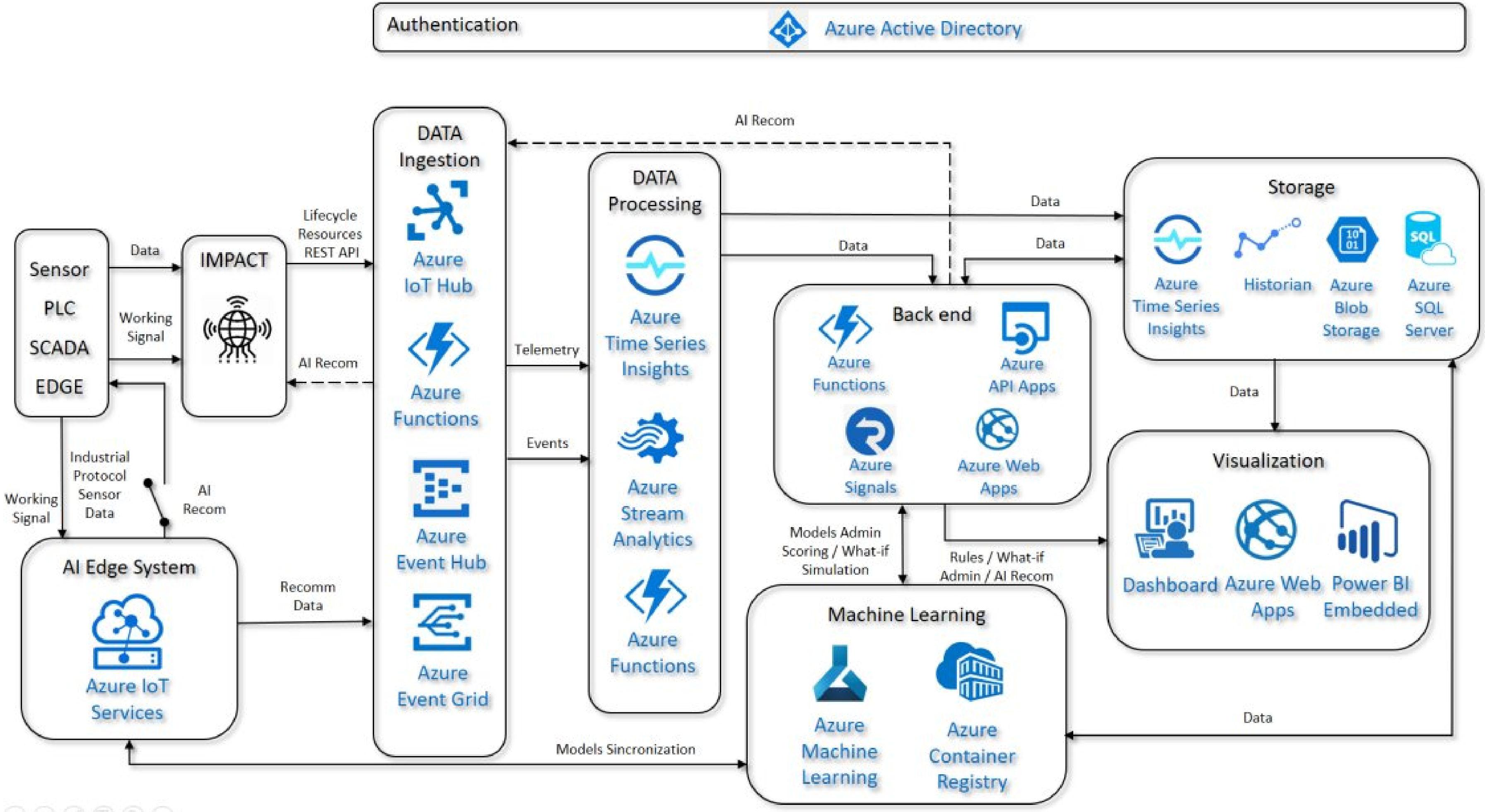

Figure 7 - Architecture components

“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958454”.

4.1 Backend

From this set of services, the applications that manage the functionality of the system will be executed. In a first step, the different functionalities to be deployed will be identified. Subsequently, the functionality offered by each service will be described.

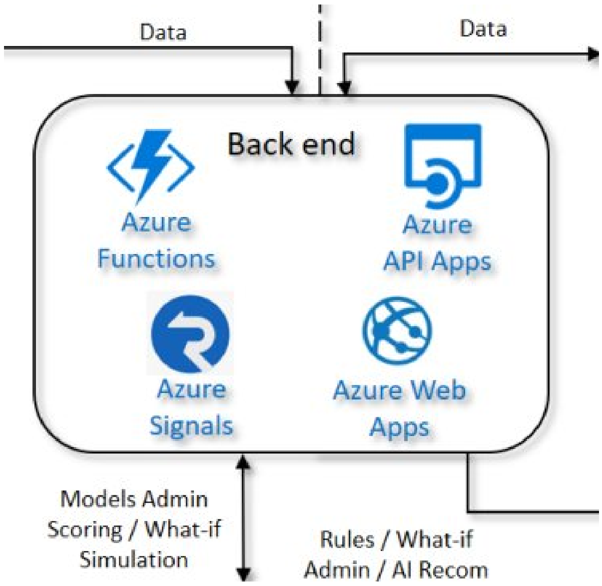

Figure 8 - Backend

The applications to be managed have been already described. The sources and destinations of the information and the components to be used are discussed in this section.

The set of data processing services provides the alert information to the back-end system, so that it can provide the necessary warnings to both the users involved and the visualization tools.

The Azure SQL Server service reads and writes the management data of the devices registered in the system, their connection and subscription strings, as well as all the data to identify all the signals and devices involved in each case study.

The machine learning system provides information on the what-if scenarios and the results of the inference of the recommendation models for the generation of the different recommendations to be applied.

Finally, the Data ingestion module receives the orders to apply the recommendations authorized by the system users on the plant.

4.1.1 Azure Functions

As described in other sections, this component allows the creation of small logical applications to carry out certain processes.

Azure Functions is a serverless solution that allows writing less code without the necessity of a physical infrastructure or resources. Instead of worrying about deploying and maintaining servers, the cloud infrastructure provides all the up-to-date resources needed to keep the applications running when it´s needed.

These services allow forgetting the infrastructure management tasks and only deploy the pieces necessaries to orchestrate and control all the data ingestion components and flows.

“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958454”.

The principal scenarios where this component must be applied will be:

- Run scheduled tasks as keep-alive signal review.
- Analyze some data streams to create an alert.
- Transform data streams from one protocol to other.

Technically, Azure Functions allow using all the industry standard programming languages and the developers can automate de deployment easily.

This service covers the functionality of the APPS theoretical component.

4.1.2 Azure Web Apps

This service allows the rapid deployment of applications in a web environment, from which the user can configure the plant and review the plant's optimization recommendations.

Provide a whole platform to create web apps using .NET, Java, Node.js, PHP, and Python on Windows or .NET Core, Node.js, PHP or Ruby on Linux. Use a fully-managed platform to perform OS patching, capacity provisioning, servers and load balancing.

Azure Monitor provides detailed views of resource usage, while Application Insights provides deeper insights into the app’s throughput, response times, memory and CPU utilization, and error trends.

The key capabilities are:

- Continuous deployment with Git, Team Foundation Server, GitHub, and DevOps
- High availability with auto-patching
- Built-in auto-scale and load balancing

4.1.3 Azure API Apps

This service supports backend applications by enabling their integration with each other as well as additional components through API services.

Azure App Service enables you to build and host web apps, mobile back ends, and RESTful APIs in the programming language of your choice without managing infrastructure. It offers auto-scaling and high availability, supports both Windows and Linux, and enables automated deployments from GitHub, Azure DevOps, or any Git repo. Learn how to use Azure App Service with our quick-starts, tutorials, and samples.

Allows building and consuming cloud APIs easily, adding features like security and backup, and can be integrated with many other Azure services.

4.2 Visualization

Due to the special characteristics of the application, three layers of data visualization are proposed:

- Plant monitoring at various levels.
- Plant analytics at strategic level.
- Performance analytics of recommendation models and their implementation.

The information received by this system comes from the following sets of applications:

- Storage, with the signal information (at real-time and historical level) and metadata.
- Backend, with information on alerts and what-if scenarios.
- Machine learning, with the state of the models and their performance.

“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958454”.

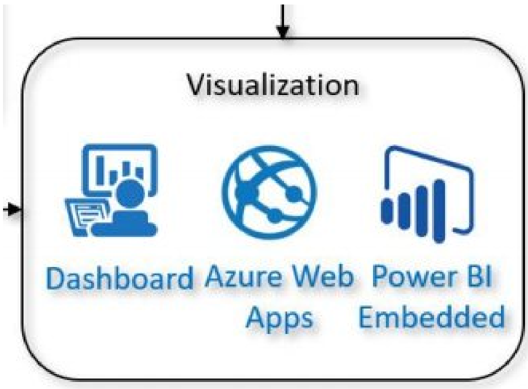

Figure 9 - Visualization

The proposal is the integration of all types of visualisation, which facilitates their consultation and avoids the need to display several types of visualisation windows.

4.2.1 Power BI Embedded

This solution allows the generation of analytical reports in a fast, intuitive, and efficient way and will be used for the analytical dashboards of the different signals. Additionally, this also makes accessible to users the available data sets for exploitation by those involved, according to their needs.

Furthermore, it shows the scoring, status and evolution of the recommendation algorithms, their valuation, and other characteristics, as well as historical information on their recommendations.

The information to be consulted will come from the different storage services, which will be responsible for providing all the necessary information.

As an additional comment on this solution, the tool will be accessible to users for the exploitation of all available information and the auto discovery. Each user will be able to create its own analytical dashboards with the plant's information.

The Embedded capability, which is associated with the Power BI solution, allows any analytical table to be embedded in an external website, so that the data shown in the different dashboards can be extracted to other platforms if necessary.

“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 958454”.

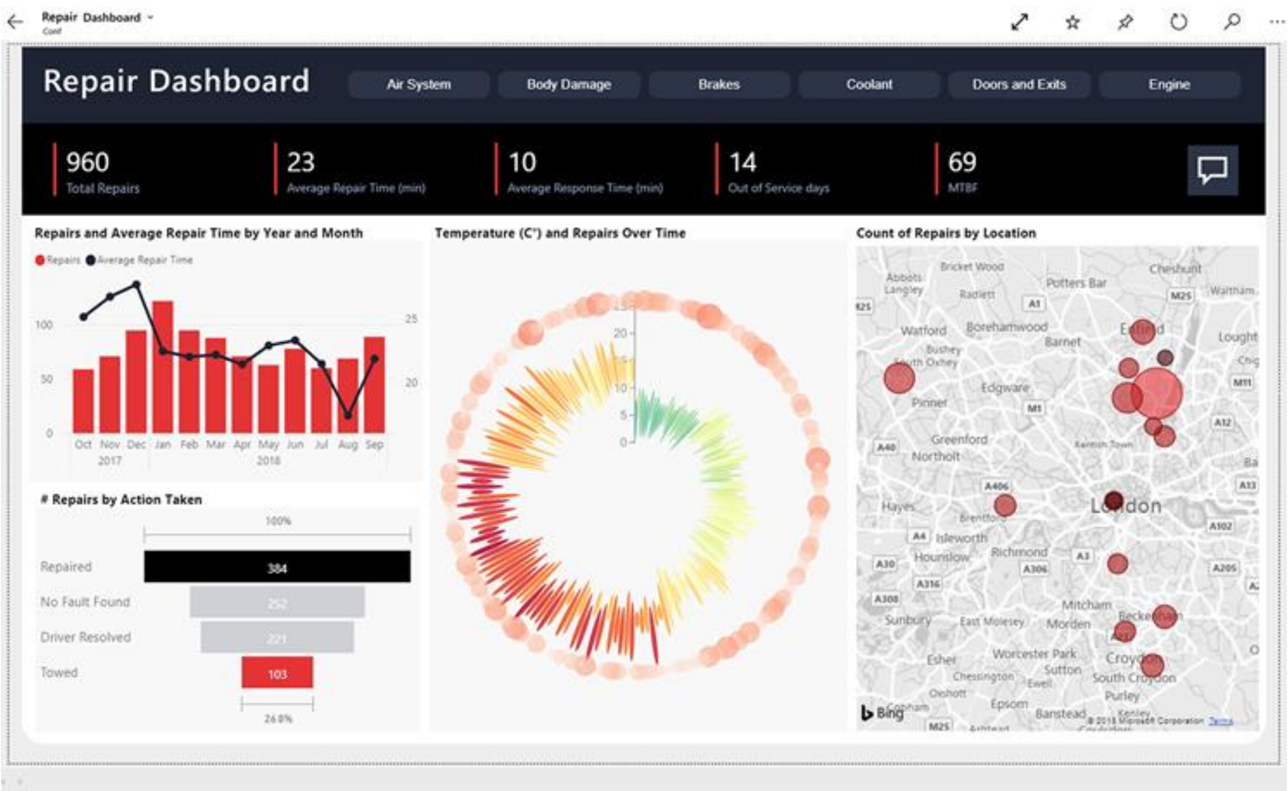

Figure 10 - Dashboard Example

This service covers the functionality of the Monitoring Dashboards theoretical component.

4.2.2 Azure Web Apps

To visualize the status of the AI Edge Service component, whether it is in operation and what type of recommendations it is automatically applying, this service will be used. The integration of this service with the previous ones makes it possible to consult from a single platform.

In this way, it is possible to visualize the real status of the Smart monitoring solution in real time, extracting the information directly from the plant.

This service allows the rapid deployment of applications in a web environment, from which the user can configure the plant and review the plant's optimization recommendations.

It is also provides a whole platform to create web apps using .NET, Java, Node.js, PHP, and Python on Windows or .NET Core, Node.js, PHP or Ruby on Linux based on a fully-managed platform in order to perform OS patching, capacity provisioning, servers, and load balancing.

Azure Monitor provides detailed views of resource usage, while Application Insights provides deeper insights into the app's throughput, response times, memory and CPU utilization, and error trends.

The key capabilities are:

- Continuous deployment with Git, Team Foundation Server, GitHub, and DevOps
- High availability with auto-patching
- Built-in auto-scale and load balancing

This service covers the functionality of the Monitoring Dashboards theoretical component.

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958454".