beopen

Open framework for boosting EU High Value Datasets from Public Sector

beopen-dep.eu

Project co-funded by the European Union under the Digital Europe Programme. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

beopen

D3.2 BeOpen Pilot services - demonstrator

Document History

Ver.	Date	Description	Author
0.1	04/11/2024	Initial Table of Contents	Giovanni Giacco(LAT40), Juan de Dios Rodríguez (HOPU)
0.2	27/11/2014	First Draft Version	All Partners
0.3	11/12/2024	First version	Giovanni Giacco (LAT40)
0.4	16/12/2024	Improvement with feedback from internal review by DPOW and CERTH	Luca Alessandro Remotti (DPOW), Evgenia Valla (CERTH), Giovanni Giacco (LAT40)
0.5	19/12/2024	Final version	Giovanni Giacco (LAT40)
1.0	20/12/2024	Submission to EC	A. Filograna

Co-financed by the Connecting Europe Facility of the European Union

Publication details
Grant Agreement Number 101069831
Acronym BeOpen

Full Title	BeOpen
Topic	Open framework for boosting EU High Value Datasets from Public Sector
Funding scheme	Horizon Europe
Start Date	01/01/2023
Duration	30 Months
Project URL	https://beopen-dep.eu
Project Coordinator	ENGINEERING
Deliverable	D3.2: BeOpen Pilot services - demonstrator
Work Package	WP3
Delivery Month (DoA)	M24
Version	1.0
Actual Delivery Date	20/12/2024
Nature	Report
Dissemination Level	PU – Public
Lead Beneficiary	LAT40
Contributors	All Partners
Quality Reviewer(s)	DPOW, CERTH
Keywords	Data sharing, trust components, Digital Decade 2030, Open Data, High Value Datasets, public sector, availability, accessibility, quality, reliability, interoperability, licensing and reuse, transparency, accountability

Co-financed by the Connecting Europe Facility of the European Union

D3.2 BeOpen Pilot services - demonstrator

Table of Content

List of Figures

Abbreviations and Acronyms

ACRONYM	Definition
AI	Artificial Intelligence
API	Application Programming Interfaces
EC	European Commission
EAB	Ethics Advisory Board
EU	European Union
DCAT	Data Catalogue Vocabulary
DCAT-AP	DCAT Application Profile for data portals in Europe
DEM	Data and Ethics Manager
DPO	Data Protection Officer
DPIA	Data Protection Impact Assessment
DIGITAL	Digital Europe Programme
CA	Consortium Agreement
CEF	Connecting Europe Facility
GA	Grant Agreement
GDPR	General Data Protection Regulation
HVD	High Value Dataset
IP	Intellectual Property
IPR	Intellectual Property Rights
M	Month
NIS2	The Directive on measures for a high common level of cybersecurity across the Union
ODD	Open Data Directive
WP	Work Package

Executive Summary

This deliverable, **D3.2: BeOpen Pilot Services - Demonstrator**, documents the implementation and validation of the BeOpen Framework across diverse pilot cities in Europe. The BeOpen Framework aims to enhance the usability, accessibility, and interoperability of High-Value Datasets (HVDs) in alignment with the EU's digital transformation goals. These HVDs are leveraged to develop and enhance digital services that improve the management of city and regional life, while also serving as training data for AI-driven services.

This initiative supports data-driven innovation, fosters transparency, and improves public service delivery in areas such as urban resilience, environmental monitoring, and mobility management. The project spans eight cities—Porto, Herne, Vilnius, Athens (Attica Region), Torre Pacheco, Molina de Segura, Cartagena, and Naples—and addresses 15 distinct use cases. Importantly, each pilot demonstrates the implementation of a specific BeOpen Framework application, while the use cases refer to the specific challenges tackled within these pilots, such as climate change mitigation, urban security, mobility integration, and disaster response.

The pilot demonstrations validate the BeOpen Framework through tailored digital services that integrate enriched datasets with innovative tools. Examples include AI-driven flood forecasting in Porto, urban heat island analysis in Cartagena, and advanced road condition monitoring in Herne. Key outcomes showcase the utility of HVDs in addressing complex urban challenges while advancing EU directives for open data.

This report presents the initial outcomes achieved during the pilot execution phase for the development of each digital service.

1 Introduction

1.1 Background on the BeOpen Framework and Pilot Use Cases

The BeOpen Framework is a robust solution designed to enhance the usability, quality, and interoperability of High Value Datasets (HVDs) produced by the European Public Sector. This framework facilitates the implementation of all phases of the data lifecycle—starting from data collection and quality improvement to semantic harmonization, publication, and cross-border accessibility. Its core objective aligns with the broader goals of the EU's digital transformation strategy, aiming to foster data-driven innovations and improve public services across sectors.

The BeOpen Framework integrates a suite of technical tools and standards to address the primary challenges of managing HVDs. It ensures their alignment with European directives and the FAIR (Findable, Accessible, Interoperable, Reusable) principles. Key functionalities include:

- **Data Collection and Quality Improvement**: Tools for aggregating and validating data from heterogeneous sources to ensure its accuracy and reliability. These tools enable automated anomaly detection, error reduction, and metadata enrichment.
- **Semantic Harmonization and Interoperability**: Leveraging established standards like DCAT-AP and INSPIRE, the framework ensures that datasets are interoperable across different platforms, regions, and applications. The use of linked data models enhances the reusability and contextual understanding of data.
- **Data Publication and Accessibility**: The framework supports APIs and web services for seamless data sharing and provides a comprehensive catalogue for easy dataset discovery. Licensing mechanisms are integrated to promote open data usage while respecting intellectual property rights.
- **Privacy and Security**: Ensuring data protection through anonymization, encryption, and robust user authentication protocols. These measures align with GDPR and other European regulations to build trust among users and stakeholders.

The BeOpen Framework serves as the foundation for digital services across the project's diverse pilot sites, addressing challenges in urban mobility, environmental monitoring, and natural disaster management. Sections below outlines the use of the framework for each use case.

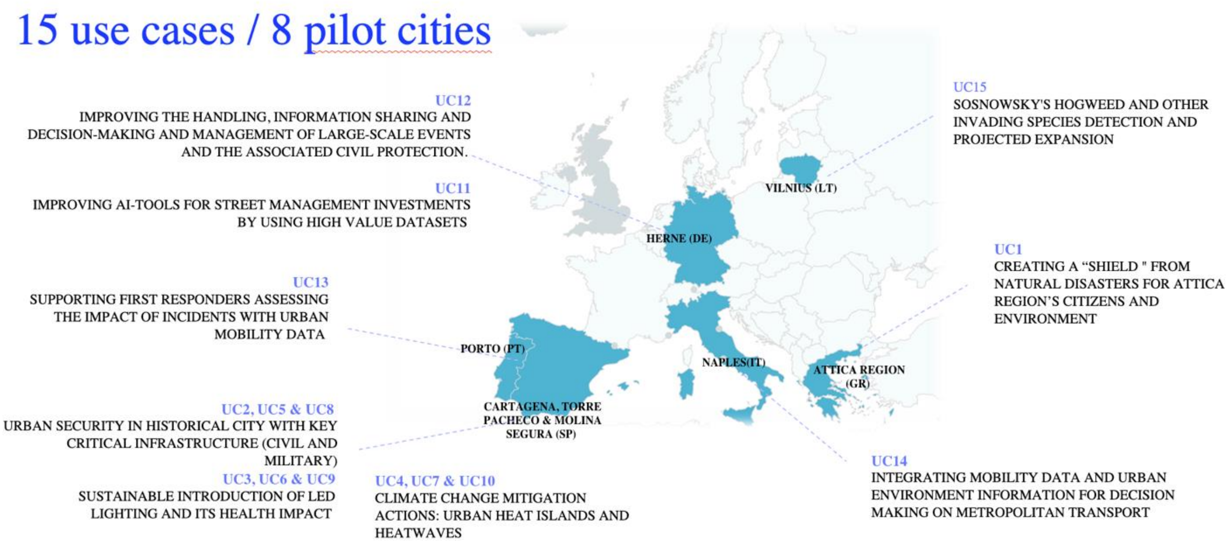


Figure 1. Overview of Use Cases and Pilot cities involved in the BeOpen project.

The project spans eight pilot cities across Europe, addressing 15 diverse use cases aimed at improving urban resilience, sustainability, and safety using HVDs. The cities and regions involved include **Porto (Portugal)**, **Herne (Germany)**, **Vilnius (Lithuania)**, **Athens (Attica Region**) **(Greece)**, **Torre Pacheco**, **Molina de Segura**, and **Cartagena (Spain)**, each contributing with unique challenges and solutions.

The use cases cover a wide array of urban issues, ranging from infrastructure improvements to environmental management:

- Enhancing **civil protection** and the management of large-scale events (UC2, UC13).
- Improving AI tools for **street management investments** (UC11).
- Supporting **first responders** with urban mobility data (UC13).
- Detecting and managing **invasive species**, such as Sosnowsky's hogweed (UC15).
- Strengthening **urban security** in historical city centers and critical infrastructures (UC3, UC8).
- Promoting the **sustainable introduction of LED lighting** and assessing its health impacts (UC4, UC6).
- Developing **climate change mitigation strategies**, including managing urban heat islands and heatwaves (UC4, UC7, UC10).
- Integrating **mobility data** for decision-making in metropolitan transport systems (UC14).
- Building a "shield" to protect the Attica Region's citizens from natural disasters (UC5).

1.2 Objectives of Pilot Execution

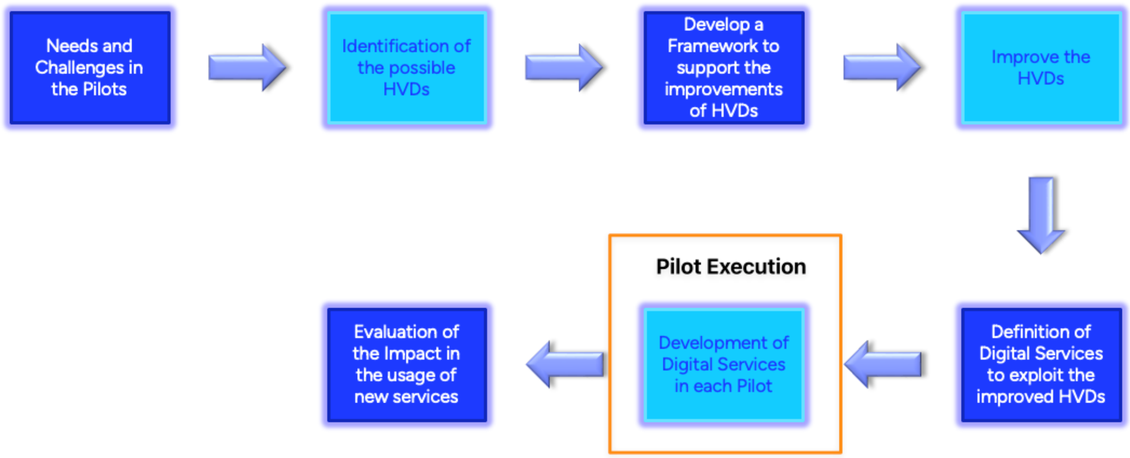


Figure 2. Pilot Execution phase for the BeOpen project.

The pilot execution phase plays a pivotal role in translating the BeOpen Framework's technical advancements into actionable digital services. It focuses on validating the framework by addressing specific challenges identified during the pilot planning phase. The execution aligns with the broader BeOpen methodology and follows these sequential objectives:

- **Improvement of HVDs**: The enhancement of HVDs is a preparatory step essential for the success of pilot execution. An HVD within the BeOpen framework refers to a dataset that has undergone a series of enhancements to maximize its usability, accessibility, and interoperability for specific applications. Improvement involves ensuring that the datasets are of higher quality, more interoperable, accessible, and semantically enriched. These enhancements lay the groundwork for their effective integration into digital tools and AI-driven applications, making them pivotal for addressing the challenges identified in the pilot planning phase.
- **Development of Digital Services**: Tailored digital services are created for each pilot site, leveraging the improved HVDs to address the unique needs and challenges of each location. These services aim to showcase innovative solutions in areas like urban mobility, environmental resilience, and disaster response.
- **Exploitation of Improved HVDs**: The pilots maximize the utility of the enhanced HVDs by incorporating them into digital services that improve decision-making, operational efficiency, and stakeholder engagement. The improved datasets become a foundation for innovation and cross-border applications.
- **Evaluation of Impact**: The pilot execution phase collects data to evaluate Key Performance Indicators (KPIs). These metrics assess the impact of the new services on end-users and stakeholders, focusing on their effectiveness, scalability, and potential for replication.

This phase ensures that the BeOpen methodology—starting with the identification of challenges and with the improvement of HVDs to define and deploy digital services—is tested comprehensively, providing insights for refinement and replication in other regions.

1.3 Scope and Structure of the Document

This document presents the outcomes of the BeOpen pilot services and demonstrates the applicability of the BeOpen Framework in diverse urban contexts. Its primary purpose is to detail the methodologies, tools, and results associated with the pilot implementations, offering insights into how HVDs can be leveraged to address urban challenges effectively. The document also serves as a resource for stakeholders interested in replicating or scaling up these approaches in other regions.

The scope of the document encompasses the entire pilot execution process, from the initial design and development of digital services to their validation and impact assessment. Each pilot is analyzed in detail, highlighting the specific challenges addressed, the integration of the BeOpen Framework, and the outputs achieved.

The document is structured to facilitate a clear understanding of the BeOpen project and its outcomes. Following this section, the document provides a comprehensive background on the BeOpen Framework, its objectives, and its alignment with the EU's digital transformation strategy. Subsequent sections detail the implementation of each pilot, describing their unique challenges, digital service designs, and the results achieved.

2 Use of BeOpen Framework Tools Across Pilots

The BeOpen Framework has been central to the development and enhancement of Digital Services in all pilot cities, providing tools for data management, quality assurance, and interoperability. These tools enable the transformation of diverse datasets into High-Value Datasets (HVDs), ensuring their alignment with EU standards and supporting their integration into local and European-level data ecosystems. The following key functionalities and tools have been applied consistently across pilots:

1. Data Storage and Management:
   - **MinIO**: Used for efficient storage and management of large-scale datasets such as geospatial files, IoT sensor data, and real-time updates. This ensures secure, scalable, and accessible storage for intermediary and processed datasets.
   - **FIWARE Orion Context Broker**: Supports dynamic data management, enabling real-time updates, query functionalities, and the seamless dissemination of dynamic data streams.
2. Data Integration and Harmonization:
   - Data Model Mapper (DMM): Facilitates the semantic alignment of datasets with FIWARE Smart Data Models, ensuring interoperability and compliance with EU data-sharing standards.
   - Metadata Quality Validator: Validates metadata quality to ensure datasets meet high standards of reliability, usability, and compliance with the DCAT-AP standard.
3. Cataloging and Accessibility:
   - **CKAN**: Used to provide a central platform for hosting, accessing, and cataloging datasets, enabling bulk downloads and API-based access for stakeholders.
   - **Idra**: Facilitates the federation of datasets with European Data Portals, enhancing cross-border data sharing and accessibility.
4. Dataset Quality Assessment:
   - **HVD Impact Assessment Tool**: Evaluates datasets for compliance with EU Implementing Regulations for HVDs, ensuring transparency, completeness, and reusability.
5. Enhancement and Visualization:
   - Grafana: Provides advanced visualization capabilities, allowing dynamic and interactive presentation of processed datasets to support decision-making.
   - OpenRefine: Used for cleaning and refining datasets, ensuring high-quality, standardized data ready for integration.

These tools have been applied variably across pilots depending on specific needs.

In the **Attica** use case, BeOpen tools were employed to process geolocated social media posts, predicted fire danger levels, and Sentinel data for wildfire detection. MinIO enabled the efficient storage of large datasets, including JSON files and geospatial shapefiles, while CKAN facilitated API-based dataset access. Metadata quality was enhanced using the Metadata Quality Validator, ensuring compliance with European standards. The ongoing federation of CKAN with Idra is designed to centralize dataset discovery and sharing, allowing broader accessibility and seamless integration.

Cartagena, Torre Pacheco and Molina de Seguro leveraged BeOpen tools to improve IoT datasets from Libelium, covering air quality, environmental data, noise, and crowd monitoring. MinIO provided scalable storage for real-time and historical datasets, while OpenRefine was used to clean and refine data before mapping it to standardized formats via the Data Model Mapper. These datasets were then integrated into CKAN for accessibility and visualized through Grafana to support urban management applications, such as monitoring urban heat islands and pedestrian safety. The HVD Impact Assessment Tool was used to align these datasets with EU HVD regulations.

Herne focused on validating the quality of metadata for time-series data generated by its IoT platform. Tools like the Metadata Quality Validator ensured adherence to high standards, while the HVD Impact Assessment Tool evaluated dataset compliance with EU Implementing Regulations. Although storage tools like MinIO were not used due to city-specific data policies, the pilot emphasized semantic harmonization through the Data Model Mapper, aligning datasets with standardized models before integrating them into the city's IoT platform.

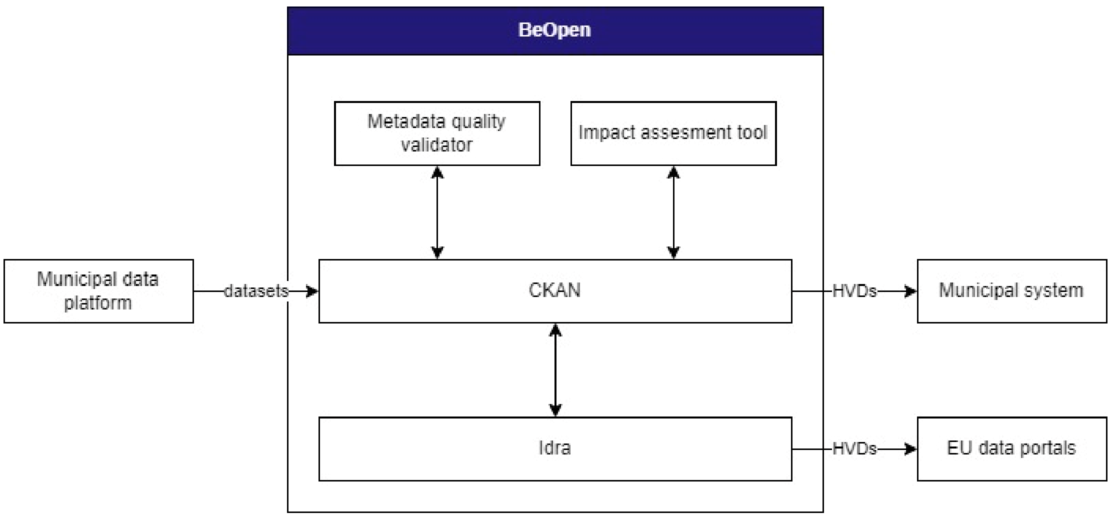


Figure 3. Herne pilot, BeOpen framework usage.

Porto used the BeOpen Framework to manage and validate datasets supporting emergency response services. Sensitive data, such as geospatial information on police and firefighter locations, was securely stored in MinIO, while metadata quality was assessed using the Metadata Quality Validator. The Orion Context Broker facilitated real-time updates for public data, such as weather and drainage system data. Federation efforts linked CKAN with Idra, improving accessibility and integration into European-level data-sharing platforms.

In **Naples**, the Data Model Mapper was pivotal in aligning datasets with FIWARE Smart Data Models, ensuring semantic harmonization and interoperability. MinIO managed the storage of urban mobility and environmental monitoring data, while the HVD Impact Assessment Tool systematically evaluated dataset readiness to meet EU standards. The FIWARE Orion Context Broker enabled dynamic data stream management, supporting real-time updates and query functionalities for digital services focused on urban sustainability.

The **Vilnius** pilot prioritized data accessibility and integration by uploading HVDs to both its local CKAN platform and the BeOpen CKAN platform. This federation facilitated interoperability with Idra and alignment with European Data Portals. The Metadata Quality Validator and HVD Impact Assessment Tool ensured the compliance and quality of datasets, supporting their inclusion in broader European repositories. These efforts underscored Vilnius's commitment to enhancing data visibility and fostering cross-border data sharing.

While the tools and functionalities of the BeOpen Framework were adapted to the unique needs of each pilot, their consistent application ensured high-quality datasets, robust data management, and alignment with EU data-sharing objectives. The BeOpen Framework provided a flexible and comprehensive approach to enhancing HVDs, enabling innovative digital services and fostering data-driven decision-making across all participating cities.

3 Pilot Attica: Natural Disaster Shield for Citizens and Environment

3.1 Digital Service Requirements

The Attica region faces a critical challenge in improving wildfire management, detection, and response mechanisms due to the increasing frequency, severity, and impact of forest fires. Climate change exacerbates this issue by creating favorable conditions for fire ignition and spreading. The proximity of fires to urban areas threatens human life and infrastructure, while the destruction of ecosystems and property results in significant environmental and economic damage. These factors underscore the urgent need for better detection and resource allocation strategies.

The Attica Digital Service aims to address these challenges through more effective wildfire detection, response, and management. The service identifies wildfires events using geolocated social media posts and satellite data to detect unusual patterns. These patterns include clusters of fire-related activity, such as dense posts mentioning smoke, fire, or emergency incidents, and anomalies in environmental parameters, such as elevated NO_2 levels caused by biomass burning. By detecting these indicators, the service enables emergency teams to respond more effectively to fire events and supports the management of natural disasters. The Digital Service will provide actionable insights for emergency responders, enhancing situational awareness and operational efficiency.

The Attica use case involves several key stakeholders critical to effective emergency management and wildfire response. These include:

- **Civil Protection Authority of the Region of Attica:** This authority is responsible for overseeing disaster preparedness, response coordination, and recovery efforts within the region. It serves as the central body for implementing policies, allocating resources, and ensuring that emergency operations are executed efficiently during crises, including wildfires.
- **The Hellenic Fire Brigade:** The Hellenic Fire Brigade plays a frontline role in combating wildfires across the region. It is tasked with fire suppression, evacuation support, and protecting lives and property. The brigade's expertise, equipment, and rapid response capabilities are vital to managing fire incidents and mitigating their impact.
- **The Hellenic Rescue Team:** This group of dedicated volunteers actively participates in emergency response efforts, offering additional support during crises. They contribute to various activities, including search and rescue operations, first aid, and logistical assistance. Their flexibility and local knowledge make them an invaluable part of the region's disaster management framework.

3.2 HVDs supporting the Digital Service

The following improved HVDs are listed with their relevance and enhancements to support the use case. The HVDs include multi-source data critical for wildfire management, covering historical fire events, social media fire-related activity, predicted fire danger levels, and satellite measurements of air pollutants. These datasets provide comprehensive insights for detection, analysis, and response to wildfire events.

Dataset Name	Description
Historical Fires Dataset_NOA	Wildfires occurred between 2019 and 2021 in Attica region, Greece.
Social Media Events related to Fires	Social media posts linked to identified fire events within the Attica region during the period of 2019 to 2021. These events are grouped into clusters based in event anomaly detection analysis of social media posts.
Social media Posts related to fires	Fire-related posts from X social media platform including Tweet ID, Tweet URL, metadata produced from CERTH processing (i.e., timestamp, location extracted from text, links to related images, fire danger level, relevance score) in the Attica region during the period of 2019 to 2021.
Predicted Fire Danger Levels Using EFFIS Data	Predicted fire danger levels for the areas of Athens. Predictions were made using an MLP model trained on embeddings generated by SatCLIP and based on fire danger data from the European Forest Fire Information System (EFFIS)
Copernicus Data Space Ecosystem Satellite Data	Sentinel-5 measurements of atmospheric pollutants from 2019 to 2021 within the Attica region.

3.3 Digital Service Design

The Near Real-Time Fire Event Detection Digital Service focuses on detecting wildfire events characterized by unusual patterns in specific data streams. It integrates Earth Observation (EO) data, such as Sentinel-5 measurements, with non-EO data, such as posts on platform X, using a late fusion technique. Initially, wildfire events are identified through social media by applying a DBSCAN density-based clustering algorithm. These detections are then verified using Sentinel-5 measurements of **atmospheric pollutants**, particularly by analyzing NO2 concentrations and employing additional outlier detection techniques that check values against both temporal and spatial axes. The architecture of the Digital Service is structured to address the critical challenges of wildfire detection and management in a streamlined, data-driven approach. Figure 4 shows an overview of the architecture. The solution integrates social media analysis, Earth Observation (EO) data, and advanced data processing through the BeOpen Framework to deliver actionable insights for emergency response teams. Below is a detailed description of the architecture:

- **Data Sources**:
  a. **Social Media Analysis**: Geolocated social media posts are analyzed in near real-time to detect anomalies and events related to wildfires. This involves processing metadata, extracting fire-related posts, and clustering information to identify potential incidents.
  b. **Earth Observation Data**: Historical EO fire datasets and Copernicus satellite data are used to provide a comprehensive view of fire patterns and environmental conditions. These datasets enhance correlation, classification, and evaluation processes for more accurate fire event detection.
- **Core Digital Service (DS1)**: The central component of the architecture is the **Near Real-Time Fire Event Detection System (DS1)** that integrates social media and EO data for wildfire detection identifying unusual patterns.
- **BeOpen Framework Integration**: The processed data and metadata from DS1 are fed into the BeOpen Framework, which provides essential tools for data management, processing, and visualization. The framework ensures efficient handling of datasets, supports semantic harmonization,

and facilitates interoperability with European data standards, enabling seamless integration and reuse across diverse systems.

- **Visualization and Dissemination**:
    a. Processed data is made available for visualization through the BeOpen Framework (Figure 7)
    b. The solution also integrates with **EU Portals** through IDRA, ensuring cross-border interoperability and compliance with European standards for open data.

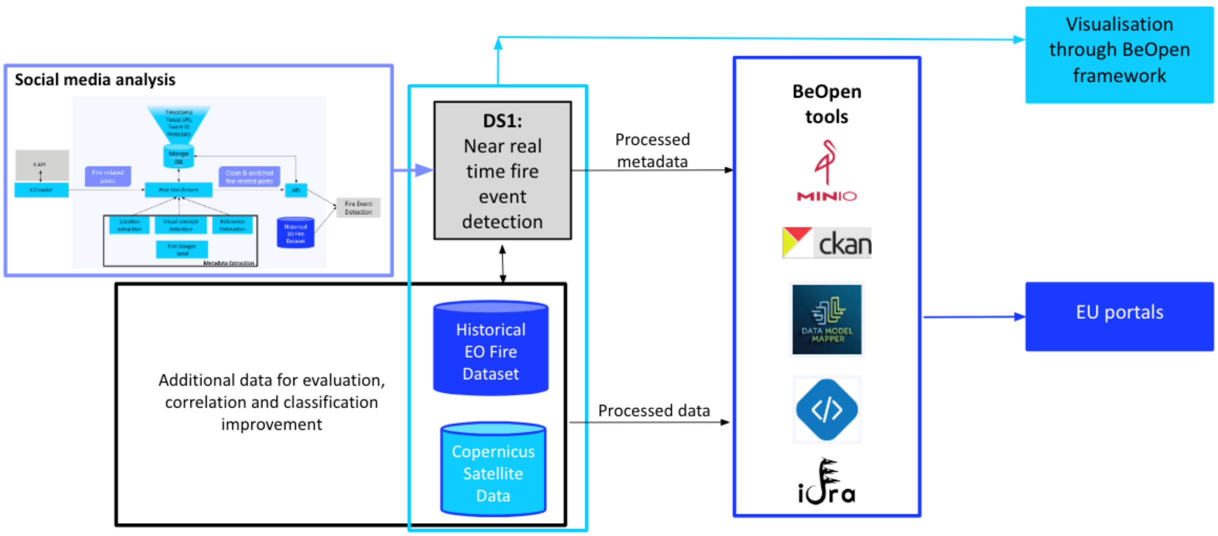


Figure 4. Architecture Attica Digital Service.

The Social Media Analysis (Figure 5) component is a key part of the Digital Service, designed to obtain the enriched and clean fire-related posts. Below is a detailed breakdown of its functionality:

- **Data Ingestion**:
    a. **X API and X Crawler**: Social media posts are collected using an API and a crawler that retrieves relevant data, such as posts related to wildfires.
    b. The extracted data includes metadata such as timestamps, URLs, tweet IDs, and other attributes that are stored in a **MongoDB** database for further processing.
- **Post Enrichment**:
    a. Once ingested, fire-related posts are passed through a **Post Enrichment** module to enhance their quality and relevance.
    b. This module performs various metadata extraction tasks, including:
        i. **Location Extraction**: Identifies the geographical location associated with the post to provide spatial context.
        ii. **Visual Concept Detection**: Analyzes images or videos within posts to detect fire-related visual elements.
        iii. **Relevance Estimation**: Evaluates the post's relevance to wildfire detection based on predefined criteria.
        iv. **Fire Danger Level**: The fire Danger Levels, predicted using an MLP model trained on embeddings generated by SatCLIP and based on fire danger data from the European Forest Fire Information System (EFFIS) are integrated into each post based on its coordinates.

- **Cleaned and Enriched Posts**:
    a. The enriched fire-related posts are cleaned and structured into a standardized format, ready for further analysis or integration with other data sources.
- **Integration with Historical Data**:
    a. The cleaned and enriched posts are combined with **Historical EO Fire Datasets** using late fusion technique. This integration allows for correlation and validation, enhancing the reliability of fire event detection.
- **Output: Fire Event Detection**:
    a. The final output is a comprehensive dataset used for **Fire Event Detection**, enabling emergency responders to identify, assess, and respond to wildfire incidents with greater efficiency.

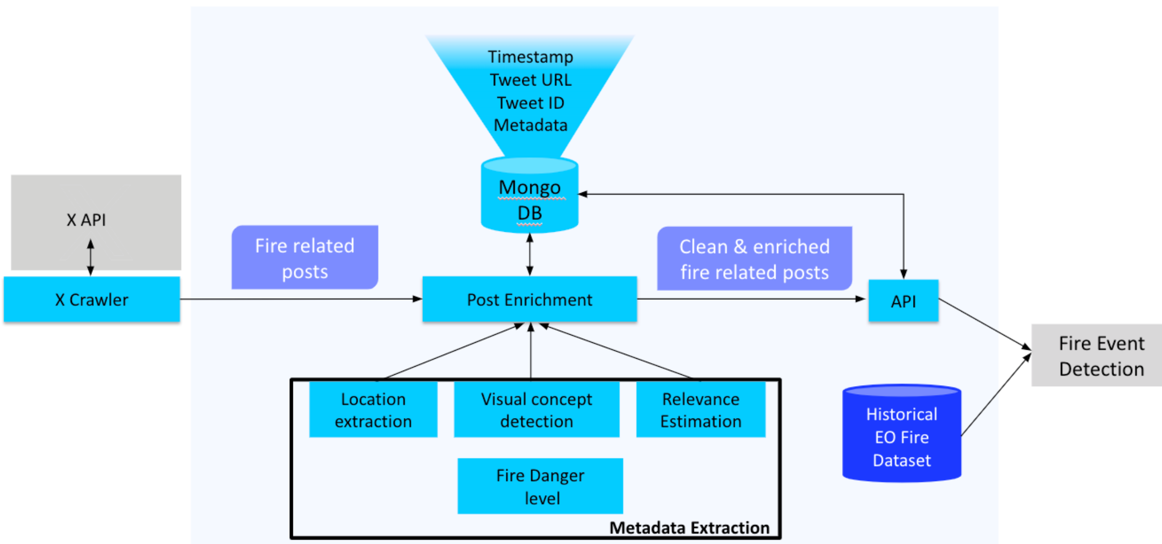


Figure 5. Digital Service Attica: Social media analysis.

3.4 Digital Service Output

The digital service output includes the clustering of social media posts referring to the same wildfire event based on their location and time. The algorithm successfully grouped posts into clusters, as shown in the figure. Each cluster represents a specific fire event, characterized by attributes such as the number of tweets, start and end timestamps, and geographic coordinates. For example, as depicted in Figure 6, Cluster ID "cluster108" includes 92 tweets spanning from July 20, 2019, to July 21, 2019, near coordinates [23.7073102, 38.0840267]. These results demonstrate the system's ability to identify and structure event-specific social media activity for enhanced situational awareness.

```
Cluster ID: cluster108
Usecase: fire_events
Number of Tweets: 92
Start Date: 2019-07-20T13:59:00.138000Z
End Date: 2019-07-21T11:41:12.443000Z
Coordinates: [23.7073102, 38.0840267]
Cluster ID: cluster280
Usecase: fire_events
Number of Tweets: 87
Start Date: 2020-07-16T10:13:36.113000Z
End Date: 2020-07-17T07:31:35.730000Z
Coordinates: [24.0548015, 37.7142551]
Cluster ID: cluster533
Usecase: fire_events
Number of Tweets: 52
Start Date: 2021-08-07T11:29:18.321000Z
End Date: 2021-08-08T09:48:28.161000Z
Coordinates: [23.79940251269328, 37.9946543]
Cluster ID: cluster534
```

Figure 6. Part of the output of the event detection algorithm using social media post from X platform.

These results are visualized on a map within the BeOpen platform. The dashboard screenshot represents all clusters across the entire period (2019 to 2021) in the Attica region.

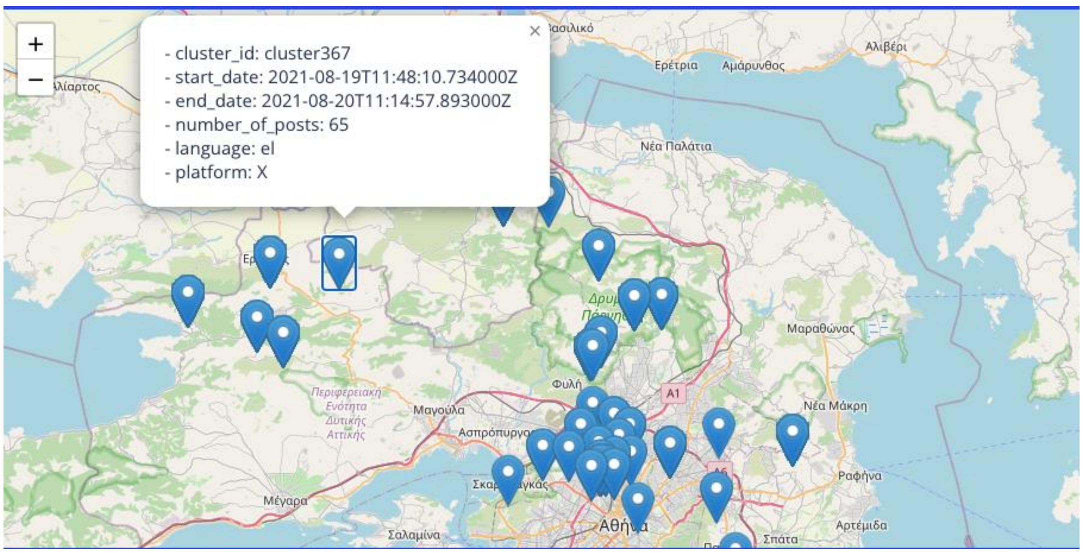


Figure 7. Visualization of Events (clusters) through BeOpen Platform.

The Figure 5 illustrates the wildfire-affected areas in the Attica region during 2021, as visualized on the BeOpen platform. The blue polygons represent the burned scars of historical EO data, enabling detailed analysis of fire-affected regions.

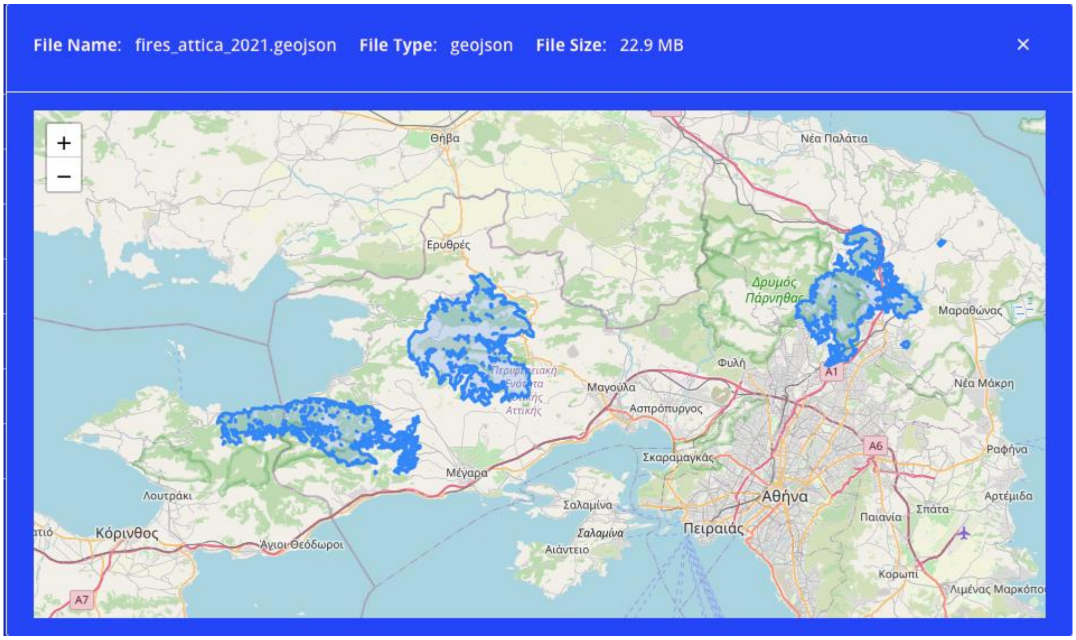


Figure 8. Burned scars of historical EO data in Attica region during 2021.

3.5 Next Steps

Wildfire events are initially detected using a density-based clustering algorithm, which identifies social media posts referring to the same wildfire event based on their location and time. Following this, a late fusion technique is applied to integrate social media data with Earth Observation data. This integration focuses on analyzing NO_2 concentrations by employing anomaly detection techniques across both temporal and spatial dimensions. This approach improves the accuracy of wildfire detections and reduces false positives, ensuring more reliable event identification. Finally, the detected wildfire events are visualized, filtered by date, and displayed alongside the corresponding burned scar areas for the selected time period.

4 Pilot Cartagena: Data Visualisation Platform

The Cartagena pilot is an integral component of the BeOpen initiative, designed to address critical urban challenges through the development and application of innovative digital services. These services are aimed at enhancing urban safety, sustainability, and resilience to climate change by leveraging HVDs and advanced analytical tools. This chapter delves into the methodologies, technologies, and processes that culminate in the deployment of a robust data visualization platform, a central feature of the pilot. While the platform itself represents the final interface for end-users, it is supported by a comprehensive framework that incorporates IoT devices, environmental data, and advanced modeling capabilities.

4.1 Digital Service Requirements

The Cartagena use case presents a comprehensive digital service designed to address three interconnected objectives: (1) Urban Safety in a Historic City with Critical Infrastructure, (2) Sustainable Implementation of LED Lighting and Its Impact on Health, and (3) Climate Change Mitigation with a Focus on Urban Heat Islands and Heat Waves.

- **Urban Safety in a Historic City with Critical Infrastructure:** The primary goal is to enhance the safety of residents and visitors by integrating data on air quality, pollutant levels, meteorological conditions, and road traffic management. This information is used to evaluate pedestrian safety along urban routes, enabling evidence-based measures to mitigate risks. The outcomes of implemented actions are assessed, and iterative improvements are made to ensure optimal safety outcomes. This approach empowers authorities to make data-driven decisions, fostering a safer urban environment.

- **Sustainable Implementation of LED Lighting and Its Impact on Health:** This objective addresses the balance between energy efficiency and adequate lighting levels, which are critical for urban safety. The pilot evaluates the impact of LED lighting, particularly blue-spectrum light, on public health by analysing energy consumption data, lighting characteristics, and population health statistics, including medication usage. The findings inform the development of urban lighting procedures that optimize both energy efficiency and public well-being.

- **Climate Change Mitigation: Urban Heat Islands and Heat Wave:** This component focuses on understanding and mitigating the effects of Cartagena's climatic challenges, particularly heat and excess humidity. By analysing variables such as temperature, humidity, atmospheric conditions, and traffic patterns (pedestrian and vehicular), the service provides insights into the evolution of heat-related phenomena. The results guide public authorities in implementing targeted measures to alleviate urban heat islands and develop policies to protect citizens during extreme heat events.

The use cases are interconnected as they rely on shared datasets and address overlapping urban challenges. For example, urban safety benefits from optimized LED lighting, which also improves energy efficiency and reduces greenhouse gas emissions, contributing to climate change mitigation. Similarly, climate mitigation measures like urban planning and reducing heat islands enhance pedestrian safety and comfort.

The pilot leverages HVDs and digital services to transform decision-making processes and support the development of evidence-based urban policies. The Cartagena use case involves multiple municipal services and stakeholders, each playing a critical role in the effective management of urban policies and emergency responses. The following requirements were established to ensure the proper usability of the digital service:

Functional Requirements

- **Real-Time Data Integration**:
  - Collect and display real-time data on air quality, pollutant levels, meteorological conditions, and road traffic.
  - Support data ingestion from IoT devices, such as air quality sensors, traffic cameras, and weather stations.
- **Risk Visualization**:
  - Provide heat maps and visual analytics to highlight high-risk pedestrian zones.
  - Include overlay capabilities for traffic data and environmental conditions to analyze correlations.
- **Outcome Evaluation**:
  - Enable comparison of historical and real-time data to assess the impact of implemented safety measures.
  - Offer reporting tools to track safety improvements over time.
- **Environmental Data Analysis**:
  - Provide tools to visualize temperature, humidity, and atmospheric data.
  - Overlay traffic patterns to identify areas most affected by urban heat islands.
- **Heat Wave Monitoring**:
  - Display real-time heat maps with alerts for extreme heat events.
  - Enable predictive modelling for heat wave evolution and its impacts on urban areas.
- **Policy Guidance Tools**:
  - Provide actionable insights for authorities to design interventions, such as increasing vegetation or improving urban planning to reduce heat retention.
  - Track the effectiveness of implemented heat mitigation measures.

Technical Requirements

- **Data Format Compatibility**: Support diverse data formats (e.g., JSON, CSV, GeoJSON) and ensure compatibility with existing datasets, including IoT and satellite data.
- **Scalability**: Allow the platform to scale for additional data sources, cities, or use cases in the future.
- Interoperability: Integrate seamlessly with platforms like FIWARE, CKAN, or Open Data APIs for data sharing and visualization.
- **Responsive Design**: Ensure the platform is accessible across devices (desktop, tablet, mobile).

Other Requirements

- **User-Friendly Interface**: Use clear visuals (charts, heat maps, graphs) and intuitive navigation for decision-makers.
- **Data Privacy**:
  - Ensure compliance with GDPR and other local data protection regulations.
  - Anonymize sensitive data.
- **Public Engagement**:
  - Provide an open-access version of the platform for citizens to view non-sensitive data, such as air quality and heat wave alerts.

This comprehensive platform would enable Cartagena authorities to leverage HVDs effectively, providing actionable insights to enhance urban safety, sustainability, and resilience to climate challenges.

4.2 HVDs supporting the Digital Service

The Cartagena Digital Service integrates multiple HVDs to address urban safety, sustainability, and climate challenges. Road traffic data provides insights into vehicular movement, supporting analyses of air quality and urban heat dynamics, while pedestrian count data informs assessments of mobility and safety on pedestrian routes. Luminosity indices are used to evaluate urban lighting levels, contributing to studies on energy efficiency and public safety. Weather and air quality data offer critical information on meteorological conditions and pollutants, essential for addressing climate challenges such as heat islands. Energy consumption data supports the analysis of sustainable energy practices, particularly in relation to LED lighting systems. Together, these datasets enable a comprehensive, evidence-based approach to urban management, equipping Cartagena with the tools to design targeted interventions and policies that enhance urban safety, sustainability, and resilience.

Dataset Name	Description
Road traffic data	Real road traffic information, characterizing type of emissions, speed, vehicle capacity, type. Generates historical data over time
Crowd Monitoring	It characterizes the pedestrian flow through areas of the city. In real time and historically.
Luminosity indices	Night-time lighting measurements of streets with LED lights. Real-time and historical.
Energy consumption	Historical archive of energy consumption by LED street lightning.
Environmental Data	Measures meteorological conditions, temperature, humidity, precipitation, wind. Characterizes environmental conditions of gaseous pollutants and particles. Real time and generates historical data.
Air Quality	Collects measurements of different pollutants measured by local stations deployed in several points of the city.
Noise	Real-time and historical measurement of noise produced by traffic and pedestrians.
Copernicus Satellite Data	Satellite luminosity and heat maps data to validate hyperlocal measurements by IoT devices.

Some of the datasets are collected from various sensors installed throughout the city, providing real-time and spatially distributed data to support the objectives of the Cartagena Digital Service.

4.3 Digital Service Design

The **Cartagena Digital Service** employs a robust architecture built on the **BeOpen framework tools** and the **FIWARE platform** to address three key objectives: Urban Safety, Sustainable LED Lighting Implementation, and Climate Change Mitigation. These tools ensure efficient data management, interoperability, and visualization. Key Components:

FIWARE Platform Integration

- **IoT Devices (Libelium):** These provide real-time data streams for air quality, noise levels, environmental conditions, and crowd monitoring.

- **IoT Agent & Orion Context Broker:** Collect, harmonize, and manage context data from connected IoT devices.
- **Quantum Leap & CrateDB:** Enable efficient historical data storage and time-series analysis, ensuring the system can track changes and trends over time.

Data Processing through BeOpen Tools

- **MinIO:** High-performance storage solution to manage extensive datasets from IoT devices and external sources.
- **OpenRefine:** Cleans and organizes raw datasets for improved data quality.
- **Data Model Mapper (DMM):** Maps data into standardized formats to ensure semantic harmonization with EU standards.
- **Metadata Quality Validator:** Ensures compliance with metadata standards like DCAT-AP before integration into external systems like Idra.

Visualization and Interoperability

- **CKAN:** Centralized data management platform enabling searchable, reusable datasets with API access.
- **Grafana:** Provides dynamic, interactive visualization of processed datasets, enhancing insights for decision-making.
- **Idra Integration:** Federates standardized datasets for cross-platform interoperability and compliance with European standards.

Use Case Applications:

1. **Urban Safety in Historic Areas:**
   1.1. Combines data from IoT devices and external sources, using the FIWARE platform to analyze pedestrian safety and critical infrastructure conditions. Insights are visualized in Grafana to assist authorities in iterative safety improvements.
2. **Sustainable LED Lighting:**
   2.1. Evaluates energy consumption and public health impacts (e.g., blue-spectrum lighting effects). Processed data from IoT and external sources is stored in MinIO, harmonized via DMM, and visualized using Grafana to inform lighting strategies.
3. **Climate Change Mitigation:**
   3.1. Tracks heat and humidity data from IoT and satellite sources to study urban heat islands. Data processed through the BeOpen tools and FIWARE components guides the development of mitigation policies.

Outcome: This integrated solution leverages the strengths of the **BeOpen tools** (MinIO, DMM, OpenRefine, Metadata Validator) and **FIWARE components** to create a comprehensive platform that supports data-driven urban management while ensuring compliance with EU standards for HVDs.

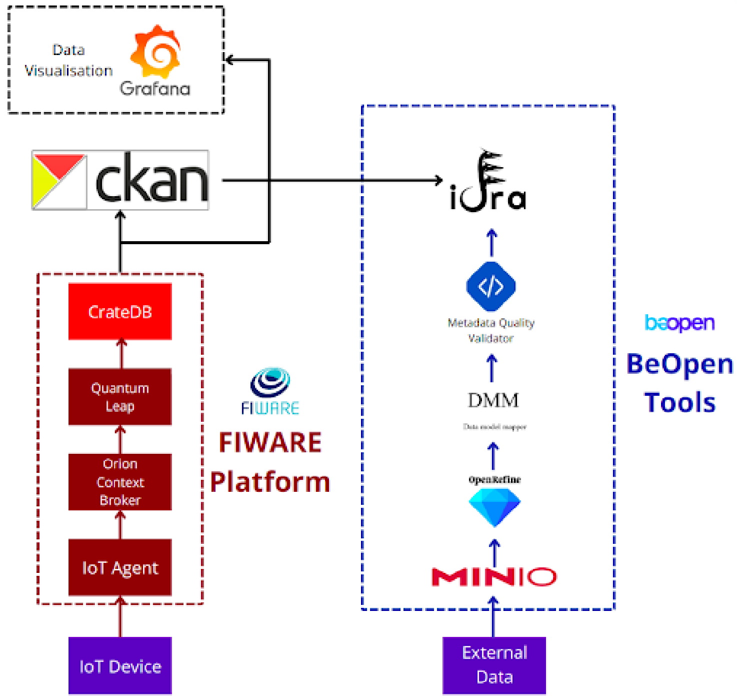


Figure 9. Architecture of the digital service supporting the Cartagena pilot.

4.4 Digital Service Output

The output of the digital service for Cartagena is an **interactive web platform** designed to provide stakeholders with seamless access to HVDs through visually engaging and user-friendly dashboards. These dashboards integrate data from diverse sources such as IoT devices, Earth Observation systems, and external data repositories, presenting insights that address critical urban challenges in safety, sustainability, and climate resilience. The Digital Service counts with the following key features:

1. **Comprehensive Data Visualization** The platform offers interactive charts, maps, and graphs that dynamically display data trends and patterns. These visualizations are tailored to illustrate specific use cases, such as pedestrian safety metrics, LED lighting impacts, and urban heat island analyses.
2. **Real-Time Insights.** Leveraging FIWARE's real-time data processing capabilities, the platform provides up-to-the-minute updates on variables such as air quality, noise levels, and meteorological conditions, enabling stakeholders to monitor urban dynamics as they evolve.
3. **User-Centric Design.** Built with accessibility in mind, the platform ensures intuitive navigation and customization options. Stakeholders can filter, compare, and analyse data specific to their needs, supporting informed decision-making.
4. **Integrated Data Management.** Backed by the BeOpen framework tools (e.g., MinIO for storage, OpenRefine for data cleaning, and CKAN for data cataloguing), the platform ensures that datasets are standardized, interoperable, and enriched with high-quality metadata.
5. **Enhanced Decision Support.** By providing actionable insights through advanced analytics, the platform supports strategic decision-making for urban planners, policymakers, and emergency responders. For example, dashboards related to climate change mitigation highlight high-risk areas during heatwaves and propose potential interventions.

It is accessible at https://pma.ayto-cartagena.es/visualizador/d/public-smartcity-fi/cuadro-de-mando-tecnico?orgId=1, and the platform provides a technical dashboard that visualizes and analyses key data. This platform provides a dynamic interface for monitoring key urban parameters and supports evidence-based decision-making for municipal authorities.

As demonstrated in the Figure 10, the platform includes a real-time dashboard showcasing air quality metrics, such as nitrogen dioxide (NO_2) levels, along with classifications of air quality. The visualization integrates spatial data, displaying pollutant measurements on a city map, enabling stakeholders to pinpoint specific areas of concern. Additionally, temporal trends are represented through detailed graphs, highlighting variations in air pollutant levels over time, providing insights into patterns and potential triggers. The interactive features allow users to navigate between datasets, filter information by parameters such as pollutant type or time frame, and access additional metadata. This functionality ensures that the platform is both user-friendly and capable of supporting complex analyses required for urban management. The integration of real-time data and historical trends makes the platform a robust tool for addressing challenges related to urban safety, sustainability, and climate resilience.

Figure 10. Main page of the platform. Key features include geolocated air quality indexes, pollutant levels weather station data, and detailed meteorological information.

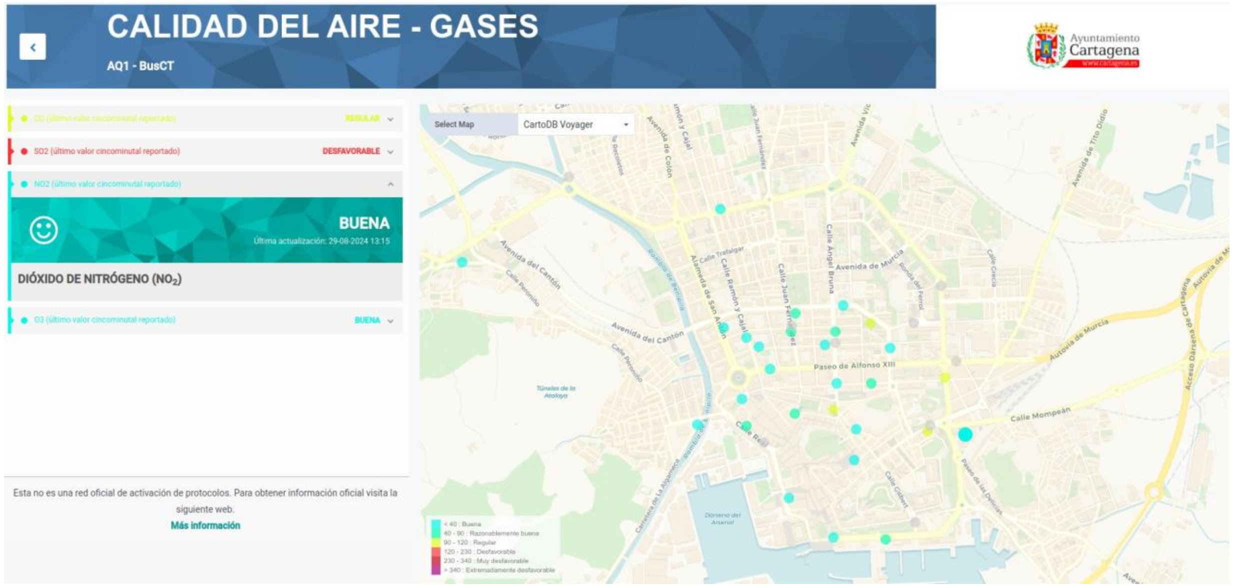


Figure 11. Real-time air quality dashboard from the Cartagena Digital Service platform.

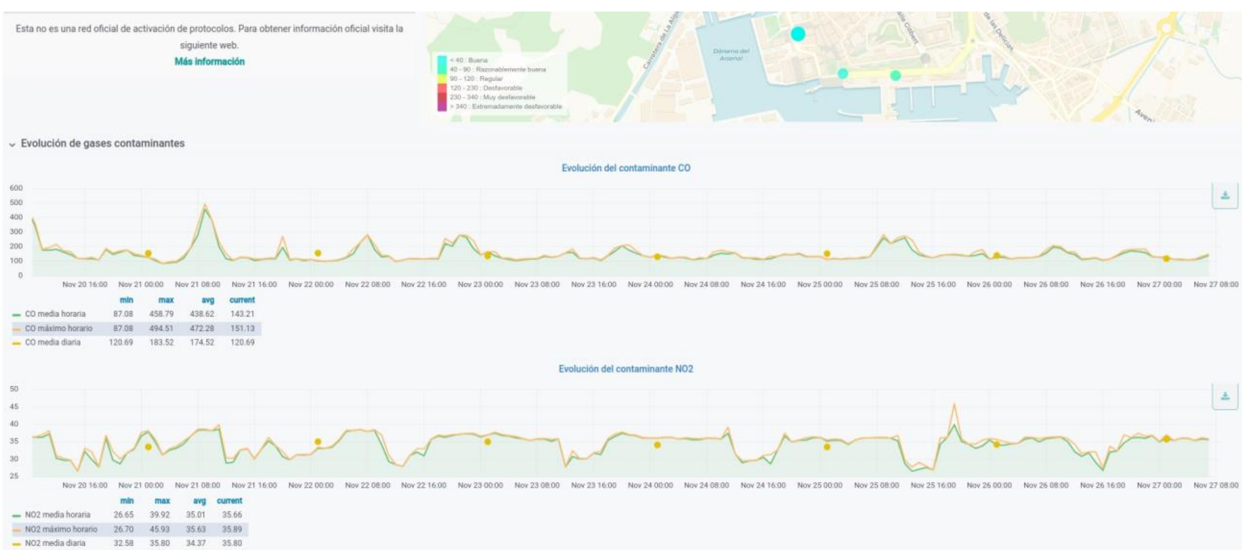


Figure 12. Visualization of real-time air quality monitoring data on the Cartagena Digital Service platform.

4.5 Next Steps

These are the main steps considered to further improve the digital service through the piloting phase to come up with a solid, consistent and robust solution. These steps will be executed through an iterative approach closely collaborating with the stakeholders of each one of the use cases covered.

- **Integration of Remaining Datasets:**
  - **Second Iteration Focus:** The next iteration of the platform will include the integration of two additional datasets:
    - **Urban Noise Levels Dataset:** To expand the scope of urban safety monitoring, noise pollution data will be added, providing insights into areas with excessive noise levels and correlating them with public complaints or safety incidents.
    - **Energy Consumption Data for LED Lighting:** This dataset will further refine the sustainable lighting module, allowing stakeholders to assess the energy efficiency of LED installations in more detail and evaluate their impact on carbon emissions.
- **Stakeholder Feedback and Iterative Improvements:**
  - **User Testing Phase:** As stakeholders interact with the platform, their feedback will be collected systematically through surveys, interviews, and interactive workshops. Key aspects to be addressed include:
    - **User Experience Enhancements:** Improving navigation, interactivity, and customization options in dashboards.
    - **Data Granularity:** Adding options for users to select the level of detail displayed, such as neighborhood-specific analyses or aggregated city-wide trends.
    - **Real-Time Data Integration:** Expanding the capability for near real-time updates, particularly for critical datasets like air quality and weather conditions.
  - **Incorporation of Suggested Features:** Based on feedback, new features (e.g., predictive analytics, alerts for threshold breaches, or mobile accessibility) may be prioritized for inclusion.
- **Performance Optimization and Scalability:**
  - To accommodate increased data volumes and user interactions, performance tuning of the MinIO object storage system and CKAN platform will be conducted. This ensures seamless access to datasets and smooth operation of visualizations.
- **Cross-System Integration:**
  - Exploring opportunities for integration with external municipal platforms or EU-wide systems like the Copernicus Data Services, allowing for enhanced interoperability and broader adoption.
- **Training and Documentation:**
  - Stakeholders will receive targeted training on platform functionality and datasets. Additionally, updated user manuals and documentation will be developed to facilitate ease of use.

By implementing these steps, the digital service will evolve into a more robust, versatile, and stakeholder-centric tool, ensuring its sustainability and long-term value in urban decision-making processes.

5 Pilot Torre Pacheco: Data Visualisation Platform

5.1 Digital Service Requirements

The Torre Pacheco pilot serves as a replica of the Cartagena pilot, implemented in a smaller city to demonstrate the solution's effectiveness and replicability. For this reason, the digital service conceived is very similar to the one for the pilot of Cartagena, sharing its main requirements, design, HVDs and architecture, but being particularly specialized in the necessities of the stakeholders of the Torre Pacheco municipality.

The Torre Pacheco pilot leverages HVDs and digital services to transform decision-making processes and support the development of evidence-based urban policies. The Cartagena use case involves multiple municipal services and stakeholders, each playing a critical role in the effective management of urban policies and emergency responses. The following requirements were established to ensure the proper usability of the digital service:

Functional Requirements

- **Real-Time Data Integration**:
  - Collect and display real-time data on air quality, pollutant levels, meteorological conditions, and road traffic.
  - Support data ingestion from IoT devices, such as air quality sensors, traffic cameras, and weather stations.
- **Risk Visualization**:
  - Provide heat maps and visual analytics to highlight high-risk pedestrian zones.
  - Include overlay capabilities for traffic data and environmental conditions to analyze correlations.
- **Outcome Evaluation**:
  - Enable comparison of historical and real-time data to assess the impact of implemented safety measures.
  - Offer reporting tools to track safety improvements over time.
- **Environmental Data Analysis**:
  - Provide tools to visualize temperature, humidity, and atmospheric data.
  - Overlay traffic patterns to identify areas most affected by urban heat islands.
- **Heat Wave Monitoring**:
  - Display real-time heat maps with alerts for extreme heat events.
  - Enable predictive modelling for heat wave evolution and its impacts on urban areas.
- **Policy Guidance Tools**:
  - Provide actionable insights for authorities to design interventions, such as increasing vegetation or improving urban planning to reduce heat retention.
  - Track the effectiveness of implemented heat mitigation measures.

Technical Requirements

- **Data Format Compatibility**: Support diverse data formats (e.g., JSON, CSV, GeoJSON) and ensure compatibility with existing datasets, including IoT and satellite data.
- **Scalability**: Allow the platform to scale for additional data sources, cities, or use cases in the future.

- **Interoperability**: Integrate seamlessly with platforms like FIWARE, CKAN, or Open Data APIs for data sharing and visualization.
- **Responsive Design**: Ensure the platform is accessible across devices (desktop, tablet, mobile).

Other Requirements

- **User-Friendly Interface**: Use clear visuals (charts, heat maps, graphs) and intuitive navigation for decision-makers.
- **Data Privacy**:
  - Ensure compliance with GDPR and other local data protection regulations.
  - Anonymize sensitive data.
- **Public Engagement**:
  - Provide an open-access version of the platform for citizens to view non-sensitive data, such as air quality and heat wave alerts.

This comprehensive platform would enable Cartagena authorities to leverage HVDs effectively, providing actionable insights to enhance urban safety, sustainability, and resilience to climate challenges.

5.2 HVDs supporting the Digital Service

The Torre Pacheco Digital Service integrates multiple HVDs to address urban safety, sustainability, and climate challenges. Road traffic data provides insights into vehicular movement, supporting analyses of air quality and urban heat dynamics, while pedestrian count data informs assessments of mobility and safety on pedestrian routes. Luminosity indices are used to evaluate urban lighting levels, contributing to studies on energy efficiency and public safety. Weather and air quality data offer critical information on meteorological conditions and pollutants, essential for addressing climate challenges such as heat islands. Energy consumption data supports the analysis of sustainable energy practices, particularly in relation to LED lighting systems. Together, these datasets enable a comprehensive, evidence-based approach to urban management, equipping Cartagena with the tools to design targeted interventions and policies that enhance urban safety, sustainability, and resilience.

Dataset Name	Description
Road traffic data	Real road traffic information, characterizing type of emissions, speed, vehicle capacity, type. Generates historical data over time
Crowd Monitoring	It characterizes the pedestrian flow through areas of the city. In real time and historically.
Luminosity indices	Night-time lighting measurements of streets with LED lights. Real-time and historical.
Energy consumption	Historical archive of energy consumption by LED street lightning.
Environmental Data	Measures meteorological conditions, temperature, humidity, precipitation, wind. Characterizes environmental conditions of gaseous pollutants and particles. Real time and generates historical data.
Air Quality	Collects measurements of different pollutants measured by local stations deployed in several points of the city.
Noise	Real-time and historical measurement of noise produced by traffic and pedestrians.

Copernicus Satellite Data	Satellite luminosity and heat maps data to validate hyperlocal measurements by IoT devices.

Some of the datasets are collected from various sensors installed throughout the city, providing real-time and spatially distributed data to support the objectives of the Cartagena Digital Service.

5.3 Digital Service Design

The **Torre Pacheco Digital Service** employs a robust architecture built on the **BeOpen framework tools** and the **FIWARE platform** to address three key objectives: Urban Safety, Sustainable LED Lighting Implementation, and Climate Change Mitigation. These tools ensure efficient data management, interoperability, and visualization. Key Components:

FIWARE Platform Integration

- **IoT Devices (Libelium):** These provide real-time data streams for air quality, noise levels, environmental conditions, and crowd monitoring.
- **IoT Agent & Orion Context Broker:** Collect, harmonize, and manage context data from connected IoT devices.
- **Quantum Leap & CrateDB:** Enable efficient historical data storage and time-series analysis, ensuring the system can track changes and trends over time.

Data Processing through BeOpen Tools

- **MinIO:** High-performance storage solution to manage extensive datasets from IoT devices and external sources.
- **OpenRefine:** Cleans and organizes raw datasets for improved data quality.
- **Data Model Mapper (DMM):** Maps data into standardized formats to ensure semantic harmonization with EU standards.
- **Metadata Quality Validator:** Ensures compliance with metadata standards like DCAT-AP before integration into external systems like Idra.

Visualization and Interoperability

- **CKAN:** Centralized data management platform enabling searchable, reusable datasets with API access.
- **Grafana:** Provides dynamic, interactive visualization of processed datasets, enhancing insights for decision-making.
- **Idra Integration:** Federates standardized datasets for cross-platform interoperability and compliance with European standards.

Use Case Applications:

1. **Urban Safety in Historic Areas:**

   1.1. Combines data from IoT devices and external sources, using the FIWARE platform to analyze pedestrian safety and critical infrastructure conditions. Insights are visualized in Grafana to assist authorities in iterative safety improvements.
2. **Sustainable LED Lighting:**
   2.1. Evaluates energy consumption and public health impacts (e.g., blue-spectrum lighting effects). Processed data from IoT and external sources is stored in MinIO, harmonized via DMM, and visualized using Grafana to inform lighting strategies.
3. **Climate Change Mitigation:**
   3.1. Tracks heat and humidity data from IoT and satellite sources to study urban heat islands. Data processed through the BeOpen tools and FIWARE components guides the development of mitigation policies.

Outcome: This integrated solution leverages the strengths of the **BeOpen tools** (MinIO, DMM, OpenRefine, Metadata Validator) and **FIWARE components** to create a comprehensive platform that supports data-driven urban management while ensuring compliance with EU standards for HVDs

5.4 Digital Service Output

The output of the digital service for Torre Pacheco is an **interactive web platform** designed to provide stakeholders with seamless access to HVDs through visually engaging and user-friendly dashboards. These dashboards integrate data from diverse sources such as IoT devices, Earth Observation systems, and external data repositories, presenting insights that address critical urban challenges in safety, sustainability, and climate resilience. The Digital Service counts with the following key features:

1. **Comprehensive Data Visualization** The platform offers interactive charts, maps, and graphs that dynamically display data trends and patterns. These visualizations are tailored to illustrate specific use cases, such as pedestrian safety metrics, LED lighting impacts, and urban heat island analyses.
2. **Real-Time Insights.** Leveraging FIWARE's real-time data processing capabilities, the platform provides up-to-the-minute updates on variables such as air quality, noise levels, and meteorological conditions, enabling stakeholders to monitor urban dynamics as they evolve.
3. **User-Centric Design.** Built with accessibility in mind, the platform ensures intuitive navigation and customization options. Stakeholders can filter, compare, and analyse data specific to their needs, supporting informed decision-making.
4. **Integrated Data Management.** Backed by the BeOpen framework tools (e.g., MinIO for storage, OpenRefine for data cleaning, and CKAN for data cataloguing), the platform ensures that datasets are standardized, interoperable, and enriched with high-quality metadata.
5. **Enhanced Decision Support.** By providing actionable insights through advanced analytics, the platform supports strategic decision-making for urban planners, policymakers, and emergency responders. For example, dashboards related to climate change mitigation highlight high-risk areas during heatwaves and propose potential interventions.

It is accessible at https://torrepacheco-opendata.hopu.eu/d/public-smartcity-fi/datos-abiertos?orgId=1&kiosk, and the platform provides a technical dashboard that visualizes and analyses key data.

The platform is accessible via the [Torre Pacheco Open Data Portal](). It features , and the platform provides a technical dashboard that visualizes and analyses key data (Figure 13).

This platform provides a dynamic interface for monitoring key urban parameters and supports evidence-based decision-making for municipal authorities.

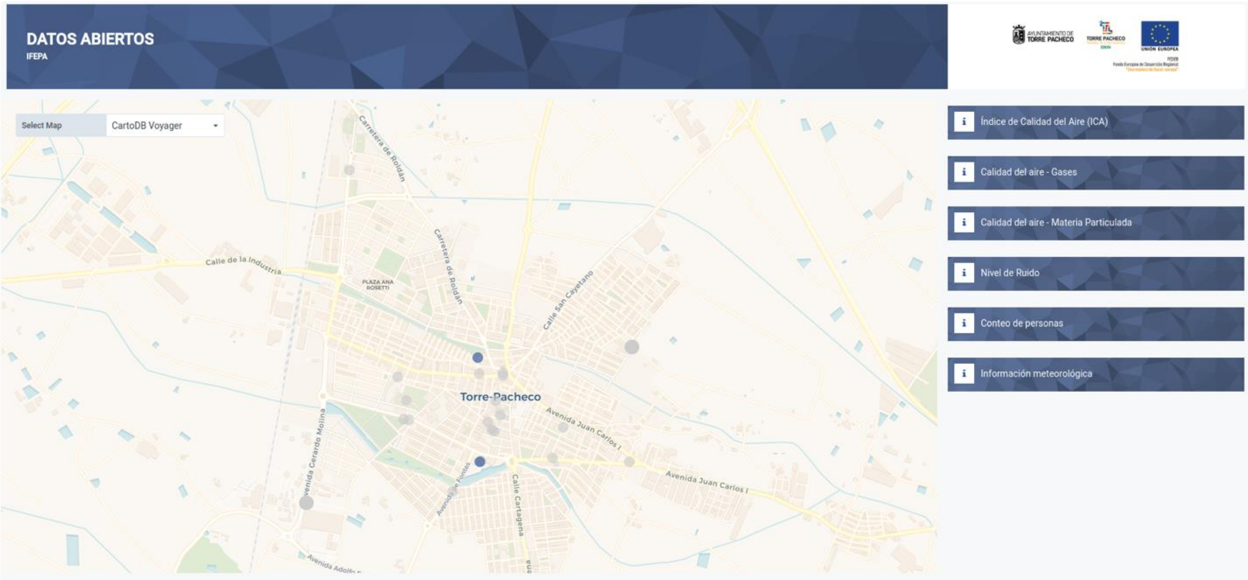


Figure 13. Torre Pacheco, Main page of the platform.

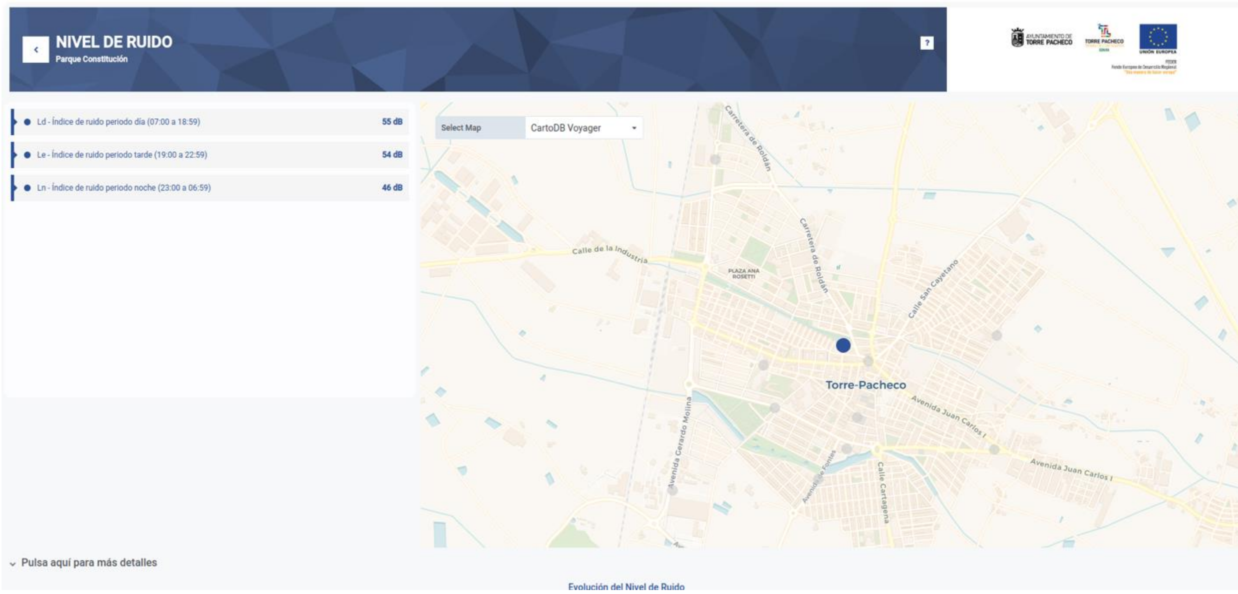


Figure 14. Real-time noise level dashboard from the Torre Pacheco Digital Service platform.

Figure 15. Visualization of real-time noise data on the Torre Pacheco Digital Service platform.

5.5 Next Steps

Similarly to the case of Cartagena, these are the main steps considered to further improve the digital service through the piloting phase to come up with a solid, consistent and robust solution. These steps will be executed through an iterative approach closely collaborating with the stakeholders of each one of the use cases covered.

- **Integration of Remaining Datasets:**

    - **Second Iteration Focus:** The next iteration of the platform will include the integration of two additional datasets:
        - **Energy Consumption Data for LED Lighting:** This dataset will further refine the sustainable lighting module, allowing stakeholders to assess the energy efficiency of LED installations in more detail and evaluate their impact on carbon emissions.
- **Stakeholder Feedback and Iterative Improvements:**
    - **User Testing Phase:** As stakeholders interact with the platform, their feedback will be collected systematically through surveys, interviews, and interactive workshops. Key aspects to be addressed include:
        - **User Experience Enhancements:** Improving navigation, interactivity, and customization options in dashboards.
        - **Data Granularity:** Adding options for users to select the level of detail displayed, such as neighborhood-specific analyses or aggregated city-wide trends.
        - **Real-Time Data Integration:** Expanding the capability for near real-time updates, particularly for critical datasets like air quality and weather conditions.
    - **Incorporation of Suggested Features:** Based on feedback, new features (e.g., predictive analytics, alerts for threshold breaches, or mobile accessibility) may be prioritized for inclusion.
- **Performance Optimization and Scalability:**
    - To accommodate increased data volumes and user interactions, performance tuning of the MinIO object storage system and CKAN platform will be conducted. This ensures seamless access to datasets and smooth operation of visualizations.
- **Cross-System Integration:**
    - Exploring opportunities for integration with external municipal platforms or EU-wide systems like the Copernicus Data Services, allowing for enhanced interoperability and broader adoption.
- **Training and Documentation:**
    - Stakeholders will receive targeted training on platform functionality and datasets. Additionally, updated user manuals and documentation will be developed to facilitate ease of use.

By implementing these steps, the digital service will evolve into a more robust, versatile, and stakeholder-centric tool, ensuring its sustainability and long-term value in urban decision-making processes.

6 Pilot Molina de Segura: Data Visualisation Platform

6.1 Digital Service Requirements

The Molina de Segura pilot is another replica of the Cartagena pilot, implemented in a medium-sized city to extend the demonstration of the solution's effectiveness and replicability. For this reason, the digital service conceived is again very similar to the one for the pilots of Cartagena and Torre Pacheco, sharing its main requirements, design, HVDs and architecture, but being particularly specialized in the necessities of the stakeholders of the Molina de Segura municipality.

The Molina de Segura pilot leverages HVDs and digital services to transform decision-making processes and support the development of evidence-based urban policies. The Cartagena use case involves multiple municipal services and stakeholders, each playing a critical role in the effective management of urban policies and emergency responses. The following requirements were established to ensure the proper usability of the digital service:

Functional Requirements

- **Real-Time Data Integration**:
  - Collect and display real-time data on air quality, pollutant levels, meteorological conditions, and road traffic.
  - Support data ingestion from IoT devices, such as air quality sensors, traffic cameras, and weather stations.
- **Risk Visualization**:
  - Provide heat maps and visual analytics to highlight high-risk pedestrian zones.
  - Include overlay capabilities for traffic data and environmental conditions to analyze correlations.
- **Outcome Evaluation**:
  - Enable comparison of historical and real-time data to assess the impact of implemented safety measures.
  - Offer reporting tools to track safety improvements over time.
- **Environmental Data Analysis**:
  - Provide tools to visualize temperature, humidity, and atmospheric data.
  - Overlay traffic patterns to identify areas most affected by urban heat islands.
- **Heat Wave Monitoring**:
  - Display real-time heat maps with alerts for extreme heat events.
  - Enable predictive modelling for heat wave evolution and its impacts on urban areas.
- **Policy Guidance Tools**:
  - Provide actionable insights for authorities to design interventions, such as increasing vegetation or improving urban planning to reduce heat retention.
  - Track the effectiveness of implemented heat mitigation measures.

Technical Requirements

- **Data Format Compatibility**: Support diverse data formats (e.g., JSON, CSV, GeoJSON) and ensure compatibility with existing datasets, including IoT and satellite data.
- **Scalability**: Allow the platform to scale for additional data sources, cities, or use cases in the future.
- Interoperability: Integrate seamlessly with platforms like FIWARE, CKAN, or Open Data APIs for data sharing and visualization.
- **Responsive Design**: Ensure the platform is accessible across devices (desktop, tablet, mobile).

Other Requirements

- **User-Friendly Interface**: Use clear visuals (charts, heat maps, graphs) and intuitive navigation for decision-makers.
- **Data Privacy**:
    - Ensure compliance with GDPR and other local data protection regulations.
    - Anonymize sensitive data.
- **Public Engagement**:
    - Provide an open-access version of the platform for citizens to view non-sensitive data, such as air quality and heat wave alerts.

This comprehensive platform would enable Cartagena authorities to leverage HVDs effectively, providing actionable insights to enhance urban safety, sustainability, and resilience to climate challenges.

6.2 HVDs supporting the Digital Service

The Molina de Segura Digital Service integrates multiple HVDs to address urban safety, sustainability, and climate challenges. Road traffic data provides insights into vehicular movement, supporting analyses of air quality and urban heat dynamics, while pedestrian count data informs assessments of mobility and safety on pedestrian routes. Luminosity indices are used to evaluate urban lighting levels, contributing to studies on energy efficiency and public safety. Weather and air quality data offer critical information on meteorological conditions and pollutants, essential for addressing climate challenges such as heat islands. Energy consumption data supports the analysis of sustainable energy practices, particularly in relation to LED lighting systems. Together, these datasets enable a comprehensive, evidence-based approach to urban management, equipping Molina with the tools to design targeted interventions and policies that enhance urban safety, sustainability, and resilience.

Dataset Name	Description
Road traffic data	Real road traffic information, characterizing type of emissions, speed, vehicle capacity, type. Generates historical data over time
Crowd Monitoring	It characterizes the pedestrian flow through areas of the city. In real time and historically.
Luminosity indices	Night-time lighting measurements of streets with LED lights. Real-time and historical.
Energy consumption	Historical archive of energy consumption by LED street lightning.
Environmental Data	Measures meteorological conditions, temperature, humidity, precipitation, wind. Characterizes environmental conditions of gaseous pollutants and particles. Real time and generates historical data.

Air Quality	Collects measurements of different pollutants measured by local stations deployed in several points of the city.
Noise	Real-time and historical measurement of noise produced by traffic and pedestrians.
Copernicus Satellite Data	Satellite luminosity and heat maps data to validate hyperlocal measurements by IoT devices.

Some of the datasets are collected from various sensors installed throughout the city, providing real-time and spatially distributed data to support the objectives of the Cartagena Digital Service.

6.3 Digital Service Design

The **Molina de Segura Digital Service** employs robust architecture built on the **BeOpen framework tools** and the **FIWARE platform** to address three key objectives: Urban Safety, Sustainable LED Lighting Implementation, and Climate Change Mitigation. These tools ensure efficient data management, interoperability, and visualization. Key Components:

FIWARE Platform Integration

- **IoT Devices (Libelium):** These provide real-time data streams for air quality, noise levels, environmental conditions, and crowd monitoring.
- **IoT Agent & Orion Context Broker:** Collect, harmonize, and manage context data from connected IoT devices.
- **Quantum Leap & CrateDB:** Enable efficient historical data storage and time-series analysis, ensuring the system can track changes and trends over time.

Data Processing through BeOpen Tools

- **MinIO:** High-performance storage solution to manage extensive datasets from IoT devices and external sources.
- **OpenRefine:** Cleans and organizes raw datasets for improved data quality.
- **Data Model Mapper (DMM):** Maps data into standardized formats to ensure semantic harmonization with EU standards.
- **Metadata Quality Validator:** Ensures compliance with metadata standards like DCAT-AP before integration into external systems like Idra.

Visualization and Interoperability

- **CKAN:** Centralized data management platform enabling searchable, reusable datasets with API access.
- **Grafana:** Provides dynamic, interactive visualization of processed datasets, enhancing insights for decision-making.
- **Idra Integration:** Federates standardized datasets for cross-platform interoperability and compliance with European standards.

Use Case Applications:

1. **Urban Safety in Historic Areas:**
   1.1. Combines data from IoT devices and external sources, using the FIWARE platform to analyze pedestrian safety and critical infrastructure conditions. Insights are visualized in Grafana to assist authorities in iterative safety improvements.
2. **Sustainable LED Lighting:**
   2.1. Evaluates energy consumption and public health impacts (e.g., blue-spectrum lighting effects). Processed data from IoT and external sources is stored in MinIO, harmonized via DMM, and visualized using Grafana to inform lighting strategies.
3. **Climate Change Mitigation:**
   3.1. Tracks heat and humidity data from IoT and satellite sources to study urban heat islands. Data processed through the BeOpen tools and FIWARE components guides the development of mitigation policies.

Outcome: This integrated solution leverages the strengths of the **BeOpen tools** (MinIO, DMM, OpenRefine, Metadata Validator) and **FIWARE components** to create a comprehensive platform that supports data-driven urban management while ensuring compliance with EU standards for High-Value Datasets.

6.4 Digital Service Results

The output of the digital service for Molina de Segura is an **interactive web platform** designed to provide stakeholders with seamless access to HVDs through visually engaging and user-friendly dashboards. These dashboards integrate data from diverse sources such as IoT devices, Earth Observation systems, and external data repositories, presenting insights that address critical urban challenges in safety, sustainability, and climate resilience. The Digital Service counts with the following key features:

1. **Comprehensive Data Visualization** The platform offers interactive charts, maps, and graphs that dynamically display data trends and patterns. These visualizations are tailored to illustrate specific use cases, such as pedestrian safety metrics, LED lighting impacts, and urban heat island analyses.
2. **Real-Time Insights.** Leveraging FIWARE's real-time data processing capabilities, the platform provides up-to-the-minute updates on variables such as air quality, noise levels, and meteorological conditions, enabling stakeholders to monitor urban dynamics as they evolve.
3. **User-Centric Design.** Built with accessibility in mind, the platform ensures intuitive navigation and customization options. Stakeholders can filter, compare, and analyse data specific to their needs, supporting informed decision-making.
4. **Integrated Data Management.** Backed by the BeOpen framework tools (e.g., MinIO for storage, OpenRefine for data cleaning, and CKAN for data cataloguing), the platform ensures that datasets are standardized, interoperable, and enriched with high-quality metadata.
5. **Enhanced Decision Support.** By providing actionable insights through advanced analytics, the platform supports strategic decision-making for urban planners, policymakers, and emergency responders. For example, dashboards related to climate change mitigation highlight high-risk areas during heatwaves and propose potential interventions.

It is accessible via Molina de Segura Platform, and the platform provides a technical dashboard that visualizes and analyses key data. This platform provides a dynamic interface for monitoring key urban parameters and supports evidence-based decision-making for municipal authorities.

As demonstrated in the Figure 16, the platform includes a real-time dashboard showcasing air quality metrics, such as nitrogen dioxide (NO_2) levels, along with classifications of air quality. The visualization integrates spatial data, displaying pollutant measurements on a city map, enabling stakeholders to pinpoint specific areas of concern. Additionally, temporal trends are represented through detailed graphs, highlighting variations in air pollutant levels over time, providing insights into patterns and potential triggers. The interactive features allow users to navigate between datasets, filter information by parameters such as pollutant type or time frame, and access additional metadata. This functionality ensures that the platform is both user-friendly and capable of supporting complex analyses required for urban management. The integration of real-time data and historical trends makes the platform a robust tool for addressing challenges related to urban safety, sustainability, and climate resilience.

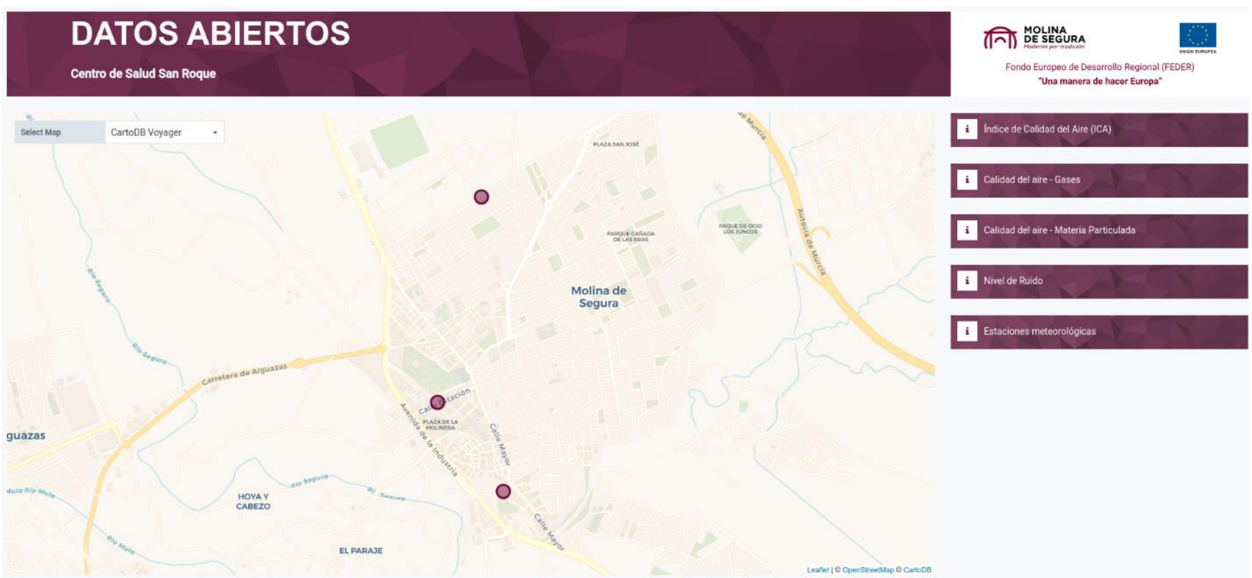


Figure 16. Molina de Segura, main page of the platform.

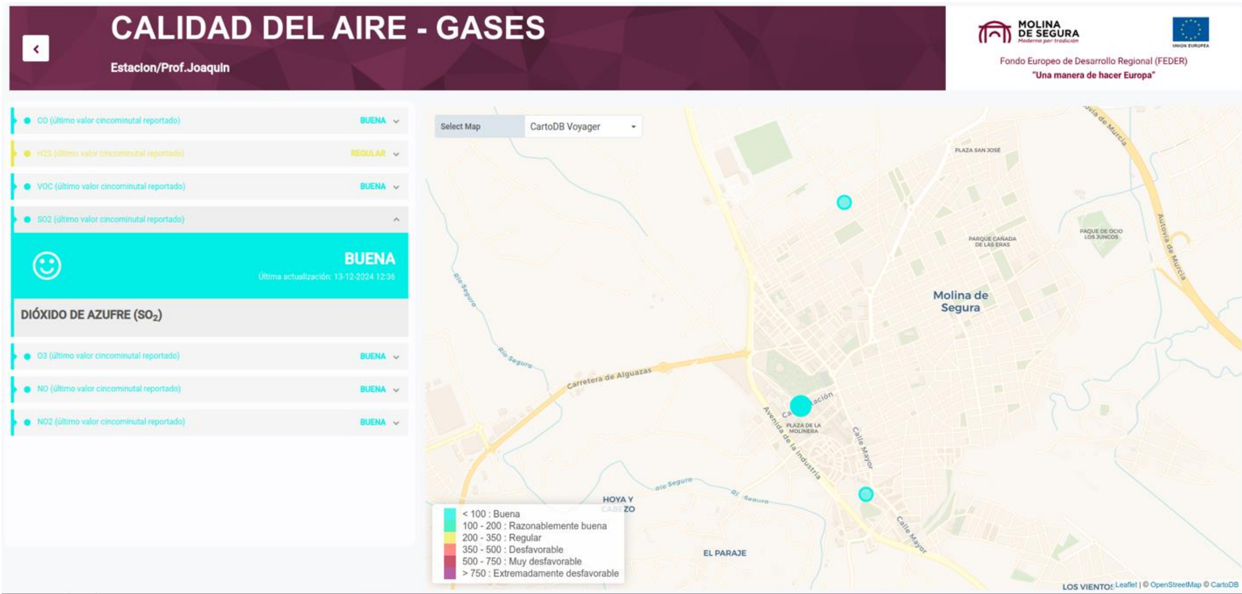


Figure 17. Real-time air quality dashboard from the Molina de Segura Digital Service platform.

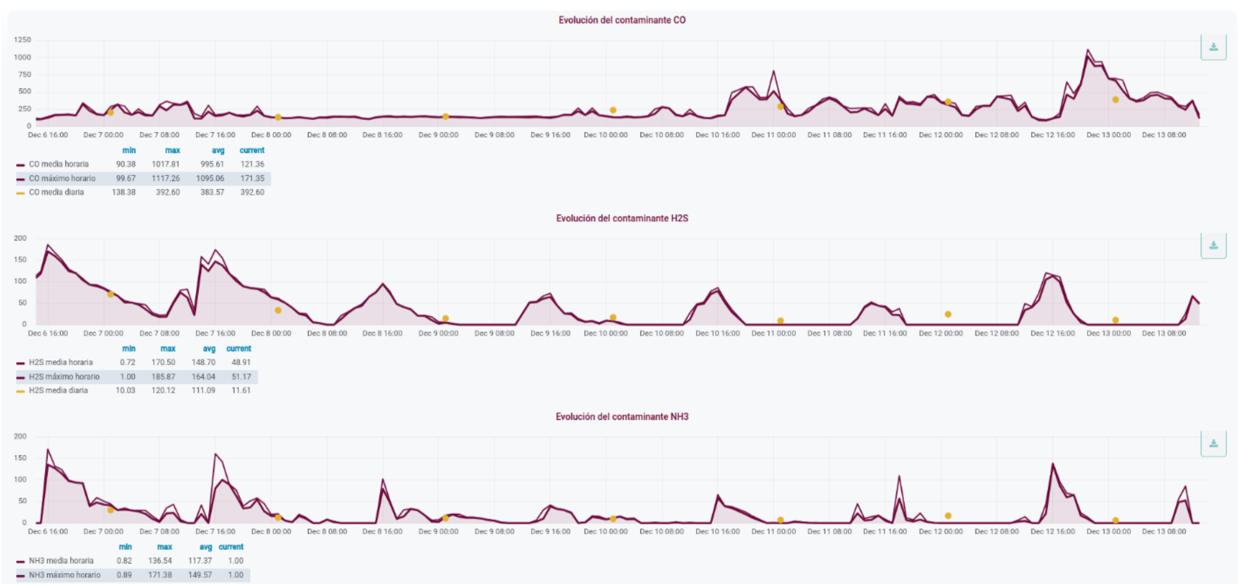


Figure 18. Visualization of real-time air quality monitoring data on the Molina de Segura Digital Service platform.

6.5 Next Steps

There are different main steps considered to further improve the digital service of Molina de Segura through the piloting phase to come up with a solid, consistent and robust solution. These steps will be also executed through an iterative approach closely collaborating with the stakeholders of each one of the use cases covered.

- **Integration of Remaining Datasets:**
    - **Second Iteration Focus:** The next iteration of the platform will include the integration of two additional datasets:
        - **Urban Noise Levels Dataset:** To expand the scope of urban safety monitoring, noise pollution data will be added, providing insights into areas with excessive noise levels and correlating them with public complaints or safety incidents.
        - **Energy Consumption Data for LED Lighting:** This dataset will further refine the sustainable lighting module, allowing stakeholders to assess the energy efficiency of LED installations in more detail and evaluate their impact on carbon emissions.
- **Stakeholder Feedback and Iterative Improvements:**
    - **User Testing Phase:** As stakeholders interact with the platform, their feedback will be collected systematically through surveys, interviews, and interactive workshops. Key aspects to be addressed include:
        - **User Experience Enhancements:** Improving navigation, interactivity, and customization options in dashboards.
        - **Data Granularity:** Adding options for users to select the level of detail displayed, such as neighborhood-specific analyses or aggregated city-wide trends.
        - **Real-Time Data Integration:** Expanding the capability for near real-time updates, particularly for critical datasets like air quality and weather conditions.
    - **Incorporation of Suggested Features:** Based on feedback, new features (e.g., predictive analytics, alerts for threshold breaches, or mobile accessibility) may be prioritized for inclusion.
- **Performance Optimization and Scalability:**
    - To accommodate increased data volumes and user interactions, performance tuning of the MinIO object storage system and CKAN platform will be conducted. This ensures seamless access to datasets and smooth operation of visualizations.
- **Cross-System Integration:**
    - Exploring opportunities for integration with external municipal platforms or EU-wide systems like the Copernicus Data Services, allowing for enhanced interoperability and broader adoption.
- **Training and Documentation:**
    - Stakeholders will receive targeted training on platform functionality and datasets. Additionally, updated user manuals and documentation will be developed to facilitate ease of use.

By implementing these steps, the digital service will evolve into a more robust, versatile, and stakeholder-centric tool, ensuring its sustainability and long-term value in urban decision-making processes.

7 Pilot Herne: AI-Tools for Street Management Investments

7.1 Digital Service Requirements

The road network in the city of Herne covers several hundred kilometers. The road network is usually the largest asset of a municipality. At the same time, the expansion and repair of the road network costs enormous sums. The aim of reducing these costs and making maintenance and repair more sustainable is hampered by the following problems:

- The exact condition of the road network is only partially known and out of date.
- The ratings of road damages are not objective or standardized
- The ratings are not continuous, so the development of the damage cannot be observed
- Existing data is hosted in different systems and is not compatible or interoperable, e.g.:
    - Information regarding additional measures that also require work on the road surface, such as work on the sewer system, water or electricity supply
    - Data on environmental influences that affect the condition of roads

As part of a development partnership with an external technology company, edge devices were developed that help to continuously record the condition of the roads.

The digital service developed as part of the BeOpen project processes this data. It makes it available to urban systems with other relevant, available data sets so that they can be used there for the sustainable planning of construction sites, investments, etc.

Stakeholders in this use case are:

- City Department of Civil Engineering and Transport. Responsible authority for road construction and project owner in the technology partnership with Edgital to develop the sensor system and the data hub for the detection and classification of road damage.
- City Department of Digitalization. Responsible for the city's IT infrastructure, as well as special applications such as the road management system.
- Edgital (technical partner). Technical partner who provides the sensor system and the data hub for the detection and classification of road damage.
- Herne.Digital (technical partner). Technical partner for the implementation of the digital service. Responsible for the implementation of interfaces, as well as implementation and operation of the municipal data platform.

7.2 HVDs supporting the Digital Service

The development of the Digital Service for the city of Herne leverages a comprehensive set of datasets to enhance urban management and planning. The service focuses on integrating diverse data sources related to road conditions, traffic flow, and external factors impacting the urban environment. Key datasets include detailed assessments of road conditions, segmented and rated for standardization, and road traffic data that captures vehicle counts, classifications, and speeds. To provide spatial and contextual insights, external datasets such as city boundaries, construction sites, main traffic routes, and speed zones are incorporated. These factors influence traffic patterns and road conditions, enabling more accurate assessments and predictions. Additionally, the service accounts for climate-related impacts by including data on areas with high heat loads and vulnerabilities, ensuring a holistic approach to urban resilience and infrastructure management. This integration of datasets supports the development of a data-driven platform that informs decision-making, optimizes resource allocation, and enhances the city's capacity to address challenges related to mobility, infrastructure, and climate adaptation.

Dataset	Description
Road condition	Represents a street divided into several segments, with standardized ratings for each segment
Road traffic	Represents the traffic flow on a street based on vehicle count, vehicle classification and vehicle speed
City Boundaries	External factor, providing spatial context for localizing assessment datasets
Construction Sites	External factor, identifying potential disruption affecting road condition.
Main Traffic Routes	External factor, identifying critical roads and their effects on road condition.
Speed Zones	External factor, representing traffic patterns
Areas with High Heat Load and Vulnerability - Climate Change	External factor, considering climate impacts on the road condition

7.3 Digital Service Design

The digital service developed for the city of Herne comprises three interconnected components: the Road Condition Detection Management System, the Databroker, and the BeOpen platform. These components work in tandem to collect, process, harmonize, and disseminate data, supporting informed decision-making for urban management and infrastructure planning.

- **Road Condition Detection Management System**

As part of a collaboration with an external technology company, edge devices have been mounted on garbage trucks to continuously detect road damage using artificial intelligence. The system classifies the detected damage according to standardized procedures and visualizes it on dashboards within a data hub. This enables an objective assessment of the road network's condition, providing a comprehensive view of the infrastructure status across the entire city.

- **Databroker**

The Databroker serves as an intermediary, retrieving road condition data from the data hub via an API. It allows for the extraction of specific datasets based on user-defined parameters, such as particular roads of interest. Once extracted, raw data undergoes model mapping and is transferred to the city's IoT data platform.

Within the IoT data platform, the condition data is persistently stored in a time-series database, enabling temporal analyses of individual road segments. The platform also integrates additional datasets, including external factors such as traffic flow and climate data, ensuring interoperability and standardized access through APIs. Harmonized datasets are then forwarded to the BeOpen platform or specialized municipal applications. These applications include planning tools for construction projects, budget planning systems, and the road management system, which forms part of the city's critical infrastructure. To showcase and test the system's functionality, a demonstrator has been implemented, allowing relationships between road condition changes and external factors like traffic volume or temperature to be analyzed and visualized.

- **BeOpen Platform**

Harmonized datasets are exported from the municipal data platform in JSON format and imported into the CKAN portal within the BeOpen framework. The CKAN portal facilitates metadata management by generating RDF files for individual datasets, which are further linked to Idra for catalog-wide RDF metadata generation. The Metadata Quality Validator evaluates the metadata quality, ensuring compliance with high standards and suggesting improvements if necessary.

Additionally, the Impact Assessment Tool is used to verify the compatibility of datasets with the HVD EU Implementing Regulation. Upon meeting these criteria, the datasets are classified as High-Value Datasets (HVDs) and made available for use in municipal applications or the demonstrator.

This integrated design not only enables efficient road condition monitoring and management but also ensures data standardization and accessibility

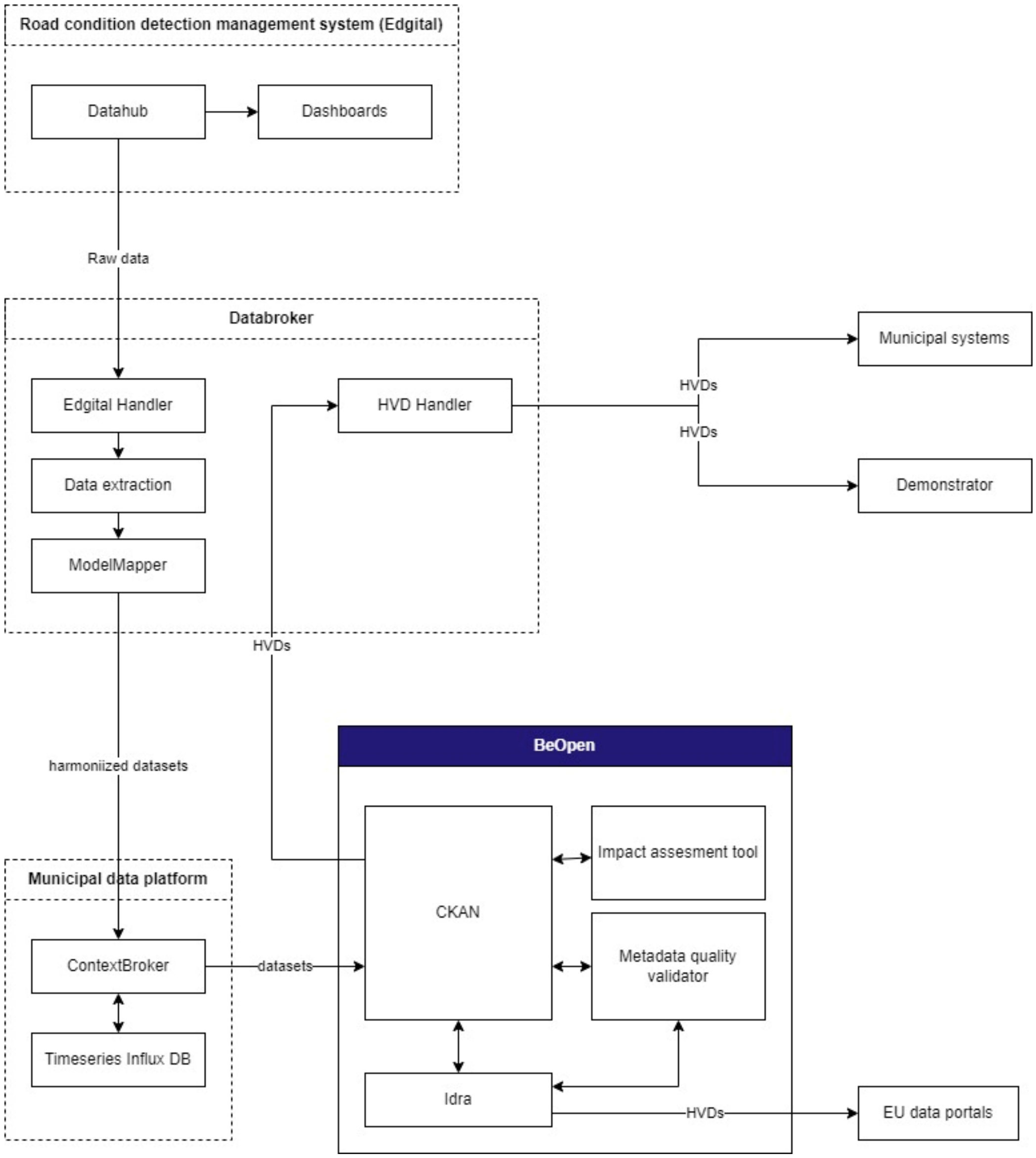


Figure 19. Architecture diagram for the Digital Service for the Pilot Herne.

7.4 Digital Service Output

The primary output of the digital service is a collection of harmonized time series-based datasets that can be seamlessly integrated with other urban data in specialized applications. These datasets enable the city of Herne to analyze and manage road infrastructure more effectively, offering insights into road conditions and their relationship with external factors such as traffic flow and climate data.

To demonstrate the potential of these datasets, a dedicated **demonstrator** was developed. This tool combines and evaluates road condition data from selected roads with additional datasets, such as traffic data, to provide a comprehensive analysis. For example, condition data from a specific road segment is extracted, harmonized, and visualized to ensure relevance for targeted applications. The harmonization process filters an extensive original dataset—containing over 400,000 rows—into a refined dataset tailored to the application's needs. An extracted harmonized dataset looks like showed in Figure 20.

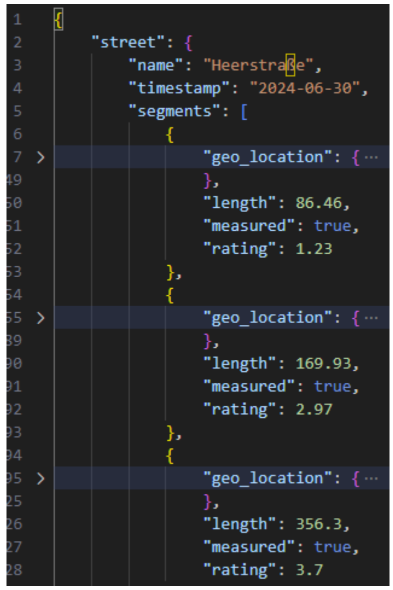


Figure 20. Harmonized Dataset

The demonstrator allows for a detailed analysis of the extracted dataset by integrating it with supplementary data.

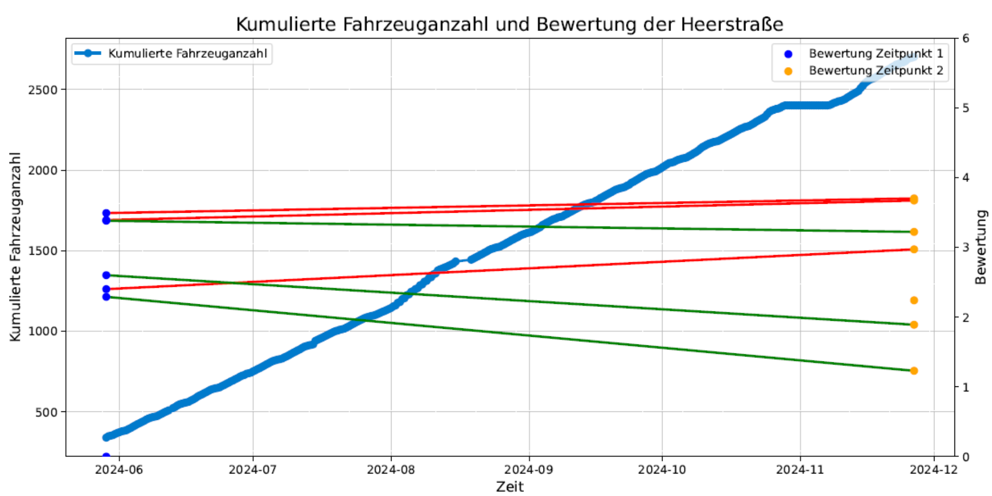
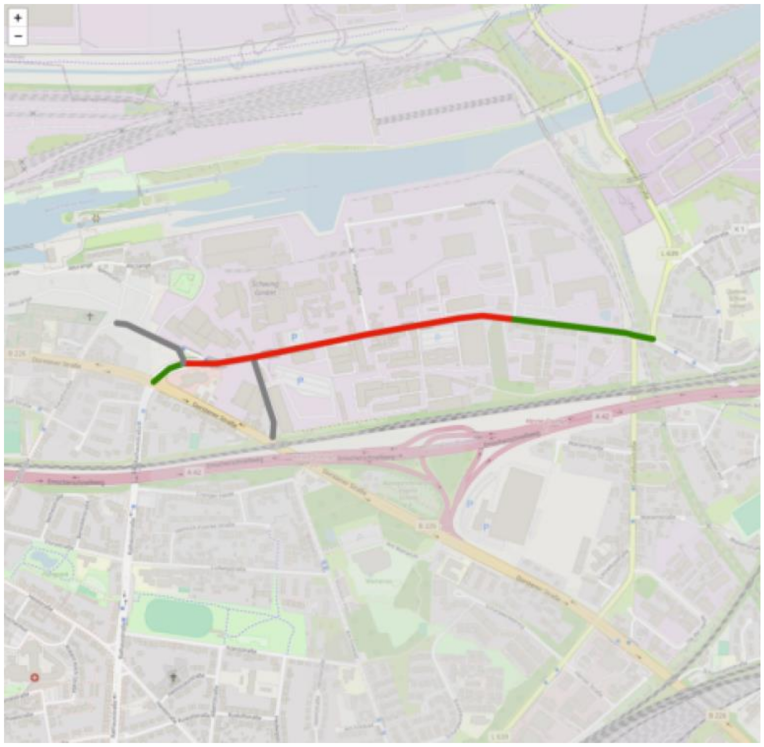


Figure 21. Detailed Time-series Analysis

Outputs include a graphical representation of the road segment under investigation, highlighting its spatial location and current condition (Figure 22), and a detailed analysis of the segment, providing actionable insights for urban planners and decision-makers.



Segment	Bewertung 2024-03-31	Bewertung 2024-06-30	Änderung
1: Segment ((7.170699222, 51.545026783) - (7.175776864, 51.545348887))	3.49	3.70	0.21
2: Segment ((7.16421789, 51.544387068) - (7.166647808, 51.544515888))	2.40	2.97	0.57
3: Segment ((7.170699222, 51.545026783) - (7.166647808, 51.544515888))	3.39	3.67	0.28
4: Segment ((7.163169908, 51.543992985) - (7.16421789, 51.544387068))	2.29	1.23	-1.06
5: Segment ((7.180642091, 51.544913379) - (7.180003845, 51.545013166))	2.60	1.89	-0.71
6: Segment ((7.175776864, 51.545348887) - (7.180003845, 51.545013166))	3.38	3.22	-0.16

Figure 22. Road segment visualization and the corresponding current conditions.

7.5 Next Steps

The future development of the digital service will focus on enabling the utilization of harmonized datasets in additional specialized urban applications. This will enhance the scalability and versatility of the service, allowing it to address a broader range of urban challenges. In addition, currently, many data transfer processes, such as retrieving raw data from the data hub and importing it into the data broker, are conducted manually. To improve efficiency and reduce potential errors, some of these processes will be automated. Automation will streamline data flows, ensure real-time availability of datasets, and facilitate their integration into various urban management systems. These advancements will further optimize the digital service's functionality.

8 Pilot Herne: Management of Large-Scale Events and Civil Protection

8.1 Digital Service Requirements

This use case focuses on improving safety and security during large, crowded events, specifically the "Cranger Kirmes," an annual fair held in August. Attracting approximately 4 million visitors over 10 days, the event presents unique logistical and safety challenges due to its scale and urban location.

The fair takes place in the heart of the city, utilizing public streets and squares. This setting creates narrow pathways and areas with no designated entrances or exits to control visitor flow. Additionally, there is a single primary access road to the fair's parking areas, which serves all visitors arriving by car or public transport. This road is a dead-end, further complicating traffic management.

These circumstances create two primary risks:

1. Severe traffic congestion spreading through the city and potentially reaching the highway.
2. Mass panic on the fairgrounds due to overcrowding.

The objective of this use case is to enable authorities to detect these critical situations promptly and implement appropriate countermeasures in a timely manner. Ideally, the system would allow preemptive action before conditions reach critical thresholds. This will be achieved by collecting and processing relevant real-time data, improving data sharing among responsible authorities, and providing a comprehensive view of the current situation.

Furthermore, the authorities will benefit from an AI-based forecasting system that utilizes real-time data and historical HVDs from previous fairs. This predictive capability will enhance their ability to anticipate potential developments and take proactive measures to ensure the safety and smooth operation of the event.

8.2 HVDs supporting the Digital Service

The development of the digital service for enhancing safety at large events relies on the integration and analysis of HVDs specifically tailored to address the challenges associated with crowd and traffic management. These datasets provide critical insights into key factors influencing the safety and efficiency of the event. The table below highlights the HVDs adopted for this purpose. They include data on parking space occupancy, offering real-time information on parking availability; road traffic, capturing vehicle flow, classification, and speed on key access routes; and people density, monitoring crowd concentration in high-traffic areas of the fairgrounds. Together, these datasets form the foundation for a data-driven approach to ensuring the safety and smooth operation of the event.

Dataset	Description

Parking space occupancy	Represents the percentage occupancy of all parking spaces at the fair.
Road traffic	Represents the traffic flow on a street based on vehicle count, vehicle classification and vehicle speed.
People density	Represents the density (People count / square meter) on well-known hot spots on the fairgrounds.

8.3 Digital Service Development

The digital Service consists of 4 components:

1. **Realtime data collection**
    a. Parking space occupancy. Using an app, parking lot operators can report the occupancy of the parking spaces in percentage. This data is transferred directly to the city's data platform and converted into absolute numbers.
    b. Road traffic. Traffic on the main access road is monitored using solar-powered radar sensors. The number of vehicles at defined time intervals and the average speed are transmitted. This allows conclusions to be drawn about the traffic flow. The data is transmitted directly to the city's data platform via LoRaWAN.
    c. People density. This is the most important data set. Edge cameras were installed at hotspots where it is particularly crowded. These devices record the number of people in a virtually gated area. The data is transmitted via MQTT over a directional WiFi to the city's data platform and converted into the population density (people/square meter).
2. **Realtime visualization**
    The data mentioned under 1. is visualized in real time in a special Grafana dashboard for the authorities. In addition, limiting values are drawn that were defined in advance with the stakeholders.
3. **BeOpen framework integration**
    The corresponding historical data from previous years are exported as JSON files and loaded into the CKAN portal. The data catalog is then linked to Idra. RDF files with the metadata for the individual datasets are generated via CKAN and RDf files with the metadata of the entire catalog are generated via Idra and evaluated in the Metadata Quality Validator so that improvements to the metadata can be made if necessary until a sufficient score is achieved.
    The Impact Assessment Tool is used to check the compatibility of the datasets with the HVD-EU Implementing Regulation. If this is also the case, these datasets can be transferred as high value datasets to the digital service as training data for time series forecasting.
4. **Time series forecasting**
    This component essentially consists of a machine learning component for time series forecasting. The HVDs serve as the basis for training the forecast model. After training, forecasting is carried out on the real-time data and also integrated into the grafan dashboard.

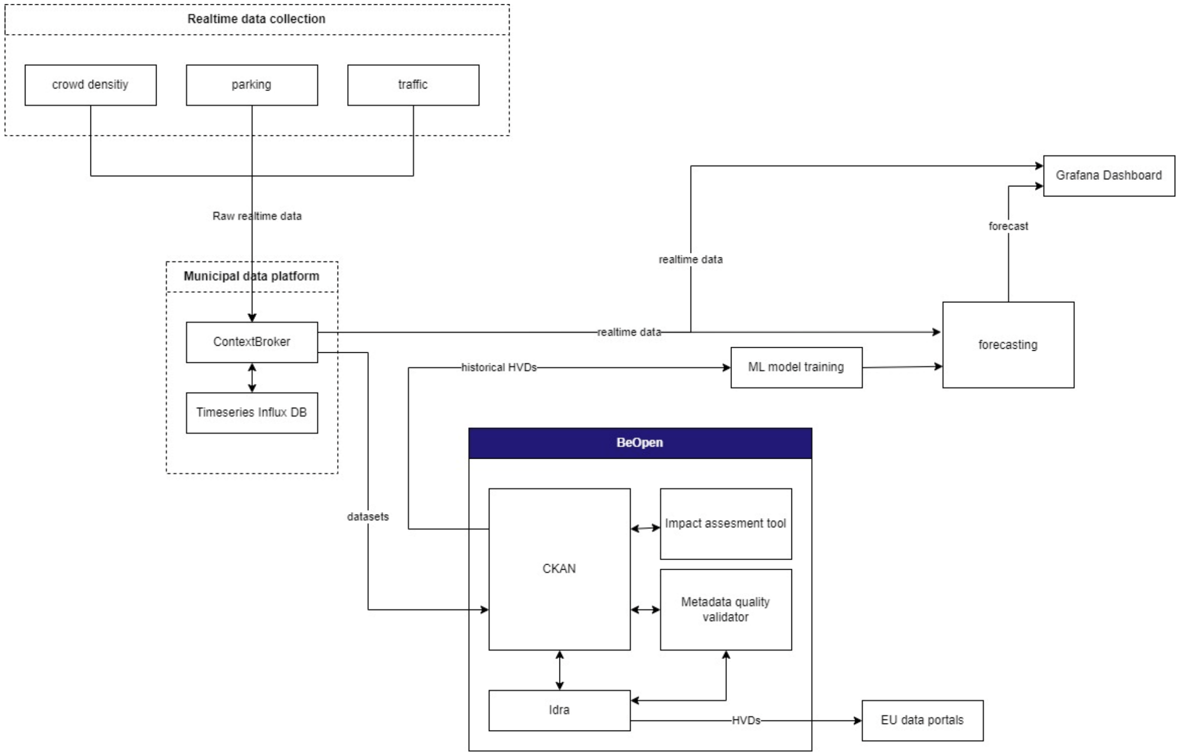


Figure 23. Architecture diagram for the Digital Service for the Pilot Herne.

8.4 Digital Service Output

The result of the digital service is a complex dashboard, which provides real-time data for the road traffic, parking space occupancy and people density.

Due to the fact that the parking spaces are owned by companies in the immediate vicinity of the fair, no permanent sensors can be installed to monitor parking space occupancy. Therefore, an app was developed that the parking lot operators can use to regularly transmit the occupancy level. The app is kept as simple as possible and looks like showed in Figure 24. Axis Edge cameras were used to monitor the crowd density. These cameras perform AI-supported person recognition in the device, count the people and only transmit numerical values.

Figure 25 and Figure 26 show images from the 2024 fair, as well as the corresponding graph (blue = real-time, red = forecasting) in the dashboard as showed in .

Figure 24. Mobile App for parking lot operators

Figure 25. People identification with AI algorithm with graph of number of people during the fair.

Figure 26. The parking space occupancy from 2024 in the dashboard.

8.5 Next Steps

In Future iterations of the digital service, planned for 2025 and beyond, aim to enhance the forecasting capabilities in two key areas.

First, while the current forecasting model is trained exclusively on historical data from the previous year, the next phase will incorporate real-time data into the training process. This will require automating the provisioning of new HVDs to facilitate regular retraining of the model, potentially on a 24-hour cycle. This improvement will ensure the forecasting model remains adaptive and reflective of the most recent conditions.

Second, the forecasting model will be expanded to include additional contextual data, such as weather information. Integrating such external variables will enhance the model's accuracy and enable a more comprehensive understanding of the factors influencing events like traffic flow and crowd density. These advancements will further improve the service's ability to anticipate critical situations and support proactive decision-making.

9 Pilot Porto: Machine Learning for Urban Flood forecast

9.1 Digital Service Requirements

Urban flooding is a significant, recurring challenge in Porto, particularly during the winter months when the city experiences heavy rainfall. The flooding disrupts daily life, damages infrastructure, and puts communities at risk. Although this problem is well-known, the entities responsible for responding to these events—such as firefighters, civil protection teams need better tools to provide timely information to act swiftly and more efficiently.

The BeOpen Porto Pilot aims to address this challenge by developing data-driven solutions to improve flood prediction, response, and recovery. By integrating historical flood data, weather forecasts, and real-time sensor information, the digital services will provide accurate, localized predictions of when flooding is most likely to occur in Porto. This allows emergency teams to act proactively by preparing resources and issuing warnings ahead of time.

Stakeholders play a crucial role in the development and implementation of the digital service, contributing both practical expertise and essential resources to the project. These entities operate on the front lines, directly managing urban flood events, and their experience and insights are invaluable in ensuring the service is both effective and relevant.

To ensure alignment between the project and pilot goals and the actual needs of those directly impacted by urban floods—particularly the teams within the Porto Integrated Management Centre—the Porto Steering Committee for the BeOpen project was established. This informal committee was designed to foster a sense of ownership and engagement across various dimensions of the initiative. It also serves as a key resource, supplying nearly all the datasets used in the development of the Digital Service while addressing trust issues related to data-sharing processes.

To maintain strong collaboration, the Steering Committee is regularly updated on the project's progress, and their feedback is actively sought. This continuous interaction ensures the solutions developed are practical, user-focused, and closely aligned with operational realities.

The Porto Steering Committee is composed of:

- Porto Digital, a partner in the BeOpen consortium and the technical lead for Porto city in the project.
- Stakeholders/beneficiaries who will benefit from the pilot results, including Águas do Porto (municipal water utility), Regimento de Sapadores Bombeiros (municipal firefighters), Centro de Gestão Integrada (Integrated Management Center), the Municipal Police, and Civil Protection.
- Municipal data policymakers from the Data Protection Municipal Department.

9.2 HVDs supporting the Digital Service

The development of the digital service for the city of Porto leverages a variety of HVDs that provide critical insights into urban infrastructure, environmental conditions, and emergency response. These datasets form the foundation for data-driven decision-making, enabling enhanced management of urban systems and improved response to critical events.

The HVDs include data on firefighter occurrences, which record events requiring emergency response across the city, and observed meteorological parameters captured by Porto Digital's sensory network. Additionally, real-time data from the drainage network sensing system (LACROIX) provides insights into water quality and flow conditions across Porto's sanitation, rainwater, and stream systems. Flood monitoring is supported by the Porto Water and Energy Flooding system (TAGO), which uses sensors to detect real-time water levels and identify potential flood risks in key infrastructures such as rivers, rainwater systems, and the Douro River. Furthermore, the H2PORTO platform records all interventions in the public drainage and water supply networks, offering georeferenced data on locations and types of actions taken.

Together, these datasets provide a comprehensive understanding of Porto's urban environment, supporting the digital service's goal.

Dataset	Description
Firefighters Occurrencies	Represents events in Porto where the firefighters were called.
Observed Meteorology – Porto Digital sensory network	Represents the observations of observed meteorological parameters by the sensors of Porto Digital sensory network
Drainage network sensing - LACROIX	Sensorization system of Porto's drainage network (sanitation + rainwater + streams), which allows, in real time, get information related to water quality (e.g. pH) and flow conditions (e.g. speed).
Porto Water and Energy Floodings - TAGO	Sensorization system of water level to detect flood situations, in real time, in riverside, rainwater and Douro River infrastructures.
H2PORTO - NAVIA	Platform for recording all requests and interventions, carried out by the company, in the public drainage and water supply networks in the city of Porto, with georeferencing service of the intervened locations by the teams and type of action carried out.

Some of the datasets used in Digital Service 1 are being opened to Open Data and then improved to HVD; we expect this to be concluded at the end of the project.

9.3 Digital Service Development

The architecture has been designed as a scalable system to integrate diverse data sources, process them efficiently, and deliver actionable insights for flood management.

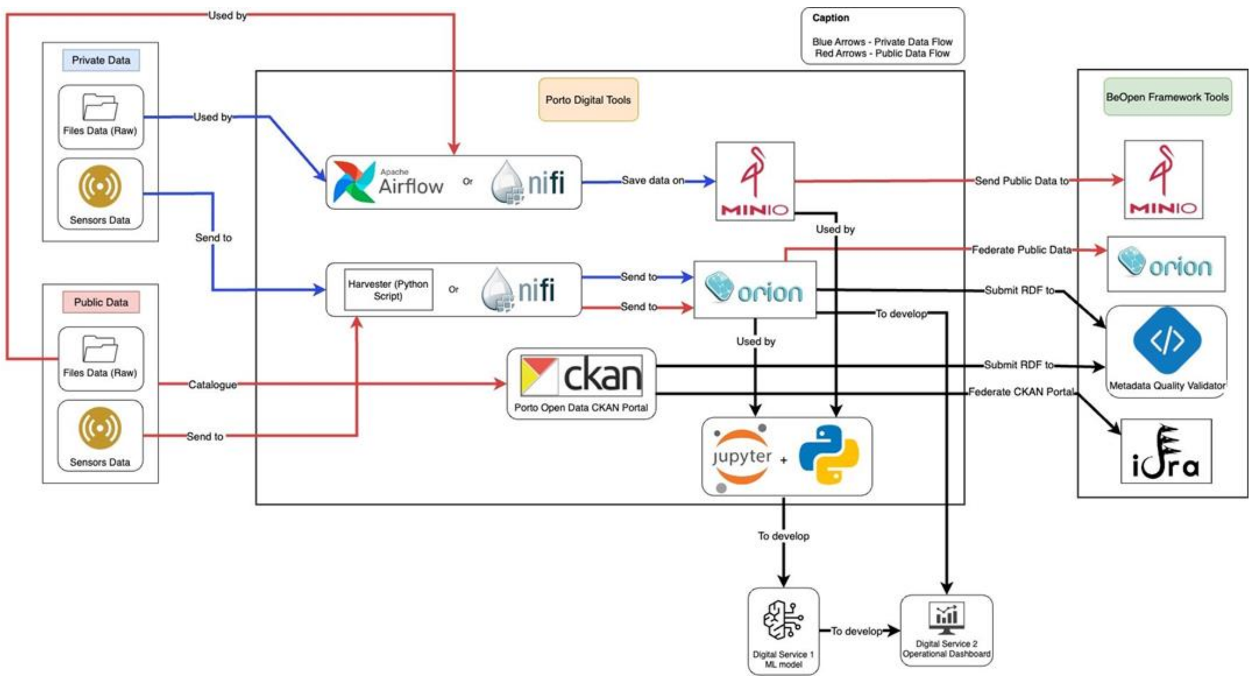


Figure 27 – Porto Digital Service 1 Architecture

Figure 27 shows in detail the whole architecture. Following details about each section of the solution are outlined.

Data Ingestion and Processing

The system collects **private data**, such as raw files and sensor data from the drainage network, and **public data**, including meteorological observations, watershed boundaries, and traffic system data. These datasets are processed through tools like **Apache Airflow** or **Apache NiFi**, which handle the data ingestion pipelines for both real-time and historical datasets.

- **Private Data Handling**: Sensitive information, such as sensor data, is stored securely in an internal **MinIO object storage** system to ensure data security and compliance with privacy requirements.
- **Public Data Handling**: Public datasets are cataloged and processed using NiFi pipelines and made accessible via a **CKAN portal**, allowing integration with the BeOpen Framework.

Data Storage and Management

- **MinIO** is used to store and manage private data securely, ensuring controlled access.
- **Orion Context Broker** integrates real-time public data and federates it with the BeOpen Framework's Orion instance, enabling interoperability across platforms.

Machine Learning Component

The processed data is fed into a **Machine Learning (ML) pipeline** designed to forecast urban floods. The ML model combines data from various HVDs. The model utilizes advanced time-series algorithms to predict flood likelihoods, leveraging the data's temporal and spatial patterns. It outputs forecasted flood events with geospatial precision, aiding proactive decision-making.

Integration with BeOpen Framework

The system federates public data through the **BeOpen CKAN Portal** and evaluates metadata quality using the **Metadata Quality Validator**. RDF files describing the data are submitted to Idra, further enhancing interoperability.

Visualization and User Interaction

The visualization and user interaction component of the Urban Floods Predictive Model focuses on enabling stakeholders to interpret and act on flood predictions effectively. While initial outputs are visualized through **Jupyter notebooks** and Python-based tools for analytical purposes, the broader objective is to integrate these insights into a dashboard visualization on a web map. This aligns with the goals of the other use case within the Porto pilot that are specifically focused on dashboard development. Detailed plans and implementation strategies for these dashboards are outlined in the dedicated section below addressing the dashboard development.

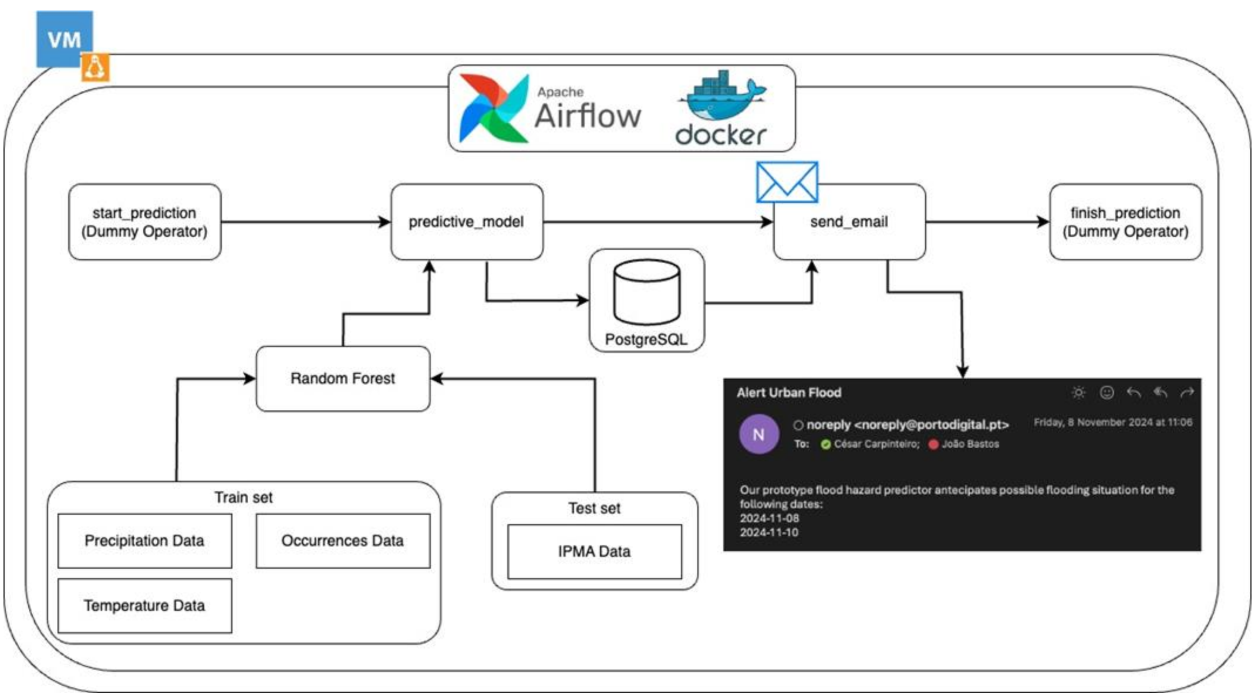


Figure 28 – Predictive Model Technical Architecture (UC 13)

For the deployment of the predictive model and its outputs, we are using **Apache Airflow** to manage the pipeline. The pipeline consists of several tasks, with the process concluding with an email notification.

- The **Start Prediction** and **Finish Prediction** tasks are dummy operators, used to mark the beginning and end of the pipeline execution.
- The core task in the pipeline, predictive model, triggers a series of Python scripts that make API calls to the WeatherObserved Data (for precipitation and temperature) and fetch historical firefighter occurrence data from MinIO. These datasets, which provide historical data on past incidents, are then used to train and fit a Random Forest model. After training, the model is tested using IPMA forecast data to predict the likelihood of urban flood events. The prediction results are saved in a Postgres database.
- The **Send Email** task queries the results from the database, iterating through each row to identify the predicted dates of potential urban flood events. Once these dates are collected, the task triggers a warning email, notifying relevant stakeholders as outlined in the technical setup architecture.

The development of the Urban Floods Predictive Model relies on a combination of datasets. These datasets encompass observed data, enabling the machine learning model to identify patterns and make accurate predictions. The table below outlines the key datasets utilized, along with their features, highlighting the type of information each dataset contributes to the modeling process. These data sources form the foundation for training and validating the predictive model.

Dataset	Features
Firefighters' Occurrences (csv sample)	• Urban Floods (1 or 0) - binary target variable • Timestamp (data and hour) when happened
WeatherObserved Data (Orion API/Mintaka)	• Precipitation (mm) • Minimum Temperature (°C) • Maximum Temperature (°C) • Timestamp (hourly)
IPMA Forecast API*	• Precipitation Probability (mm) • Minimum Temperature Probability (°C) • Maximum Temperature Probability (°C) • Timestamp (daily)

* Dataset provided by an external supplier, not included in BeOpen HVD listing.

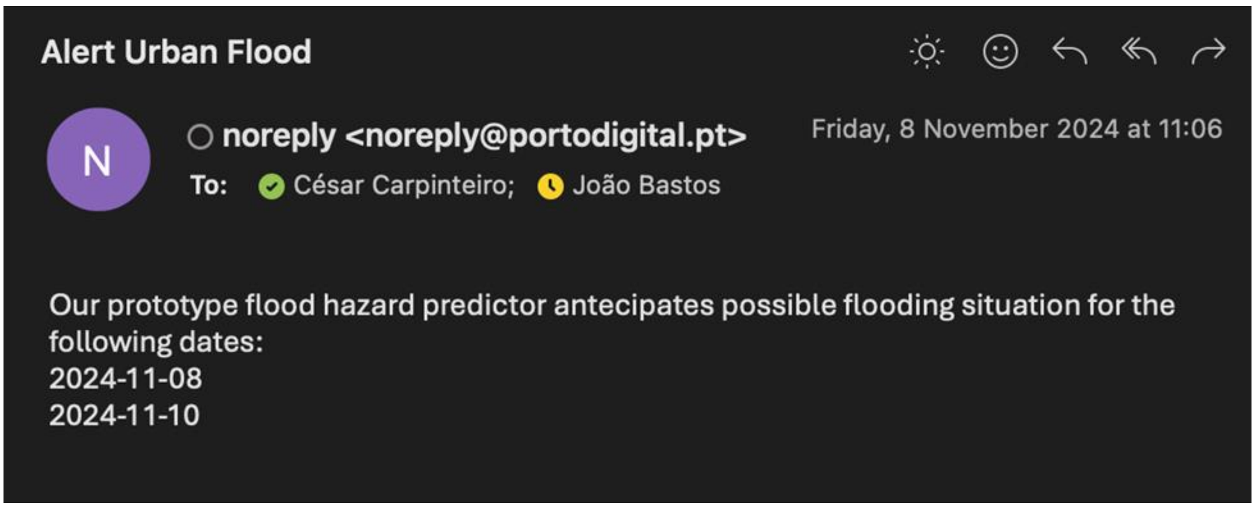
9.4 Digital Service Output

As outlined in the technical setup, the output of our digital service is an email. Currently, these emails are sent internally to Porto Digital, but in the future, once we achieve more accurate results, we plan to send them to our stakeholders to notify them of urban flood event probabilities.

The process begins with the predictive model inserting into the database the predictions for the next 5 days, showing the date and whether the model predicts an urban flood (1) or not (0).

```
id | forecastdate | predicted_value
----+-------------+----------------
  1 | 2024-11-21  |               1
  2 | 2024-11-22  |               1
  3 | 2024-11-23  |               0
  4 | 2024-11-24  |               1
  5 | 2024-11-25  |               1
```

The Email Task then reads the values and sends an email containing the predicted urban flood events, listing the dates when the model anticipates a flood, such as:


Alert Urban Flood

noreply <noreply@portodigital.pt> Friday, 8 November 2024 at 11:06

To: César Carpinteiro; João Bastos

Our prototype flood hazard predictor antecipates possible flooding situation for the following dates:
2024-11-08
2024-11-10

9.5 Next Steps

For upcoming versions, the following improvements are expected:

V2:

- Gain a deeper understanding of Neural Prophet, a time-series model that leverages deep learning for more accurate predictions and integrate this model into our digital service.
- Shift from predicting whether an urban flood will happen on a specific day to forecasting the number of urban flood events.
- Expand data sources and incorporate real-time data to feed the model daily, ensuring it is updated with the latest inputs.

V3:

- Begin making predictions at a more granular level, focusing on specific locations within the city. This will allow us to reduce the action area and concentrate efforts on zones with a higher likelihood of flooding.

10 Pilot Porto: Dashboard for real time visualization of occurrences and emergency teams

10.1 Digital Service Requirements

The primary objective of this Digital Service is to develop a real-time dashboard that facilitates the visualisation of occurrences and emergency response teams. The dashboard will serve as a centralised platform to aggregate, process, and display information, with a particular focus on enhancing emergency response times and optimising route planning. The key objectives and features of the Digital Service are as follows:

1. Centralised Information Aggregation

   The dashboard consolidates data from various reporting platforms into a single platform. This will be accomplished by leveraging insights from the Urban Platform, developed by Ubiwhere and already deployed in Porto Municipality, as well as other municipalities. The Urban Platform acts as a digital twin of the city, providing municipalities with a comprehensive, real-time, and integrated view of the city. It aggregates data across multiple verticals, including mobility and environmental data, enabling a more informed decision-making in emergency situations. By using this digital twin, the dashboard will provide emergency teams with a complete and up-to-date picture of the urban environment, optimising response strategies and resource deployment across different urban settings.

2. **Route Optimisation**

   A key feature of the dashboard is the optimisation of emergency vehicles routing, ensuring the most efficient travel from the emergency response base (e.g., fire stations and police stations) to the incident location. The routing algorithm will consider both the characteristics of the emergency vehicles (e.g., size and maneuverability) and the unique characteristics of the street network, including road width and traffic conditions. This feature is particularly important in Porto's historical centre, where narrow streets and other infrastructure constraints may limit the access of larger vehicles. By addressing these challenges, the dashboard aims to significantly enhance the emergency responses in urban areas with complex layouts.

3. **Real-time Data Integration**

   The dashboard will incorporate real-time data from emergency services, allowing for immediate adjustments to the optimised route as the situation evolves. This information includes key information such as the location and type of occurrence, the current status of resolution, and details on the number of personnel and vehicles involved. To further enhance the system's responsiveness, sensors will be installed in two emergency vehicles - one of smaller size and one with larger dimension - ensuring the algorithm can adapt to varying vehicle types in real-time conditions.

4. **Integration of High-Value Datasets**

   In addition to real-time data, the dashboard will integrate existing high-value datasets available through Porto's Data Portal, such as geospatial information. The project aims to encourage the continuous enhancement of these datasets to ensure they meet the quality standards required for effective emergency management. This process will also promote stakeholder awareness of the importance of real-time and high-quality data, while ensuring data privacy and compliance with relevant regulations. As future iterations of the dashboard are developed, additional datasets - such as environmental data from meteorological stations, weather forecasts, tree inventory, and water lines - will be integrated, providing a more comprehensive view of the urban environment.

5. **Stakeholder Engagement and Use Case Alignment**

   Throughout the development of the dashboard, design decisions have been informed by ongoing engagement with municipal stakeholders in Porto, particularly those involved in the Porto Integrated Management Center (CGI). This includes key representatives from emergency services, such as firefighters, police, and civil protection. While the dashboard is built on the fundamental knowledge of the Urban Platform, its design is being fully customised based on insights gathered through meetings with Porto's municipal stakeholders. These meetings have been - and will continue to be - crucial for identifying critical information required for real-time decision-making and ensuring that the dashboard meets the specific needs of emergency services.

By addressing the unique challenges faced by emergency services - such as navigating narrow streets, optimising vehicle routing, and accessing real-time incident data - this Digital Service aims to enhance the effectiveness of emergency response operations within Porto municipality. Continuous stakeholder input ensures that the dashboard is customised to meet the real-world necessities of emergency responders, ultimately enabling a faster, safer, and more coordinated response.

10.2 HVDs supporting the Digital Service

Dataset	Description
Firefighters Occurrences	Historical data on fire-related incidents for emergency response analysis.
Observed Meteorology – Porto Digital sensory network	Real-time meteorological data from local IoT sensors.
Drainage network sensing - LACROIX	Monitoring of drainage systems to assess flood risks.
Porto Water and Energy Floodings - TAGO	Data on water and energy infrastructure impacts.
H2PORTO - NAVIA	Integrated water resource management data for urban planning.
Police and firefighters geolocation buildings	Locations of key emergency service facilities for incident response.
Watersheds boundaries	Spatial data defining watershed areas for hydrological analysis.

Roadway and street dimensions	Detailed measurements of road infrastructure for traffic and safety planning.
Fire hydrants	Geolocated data of hydrants to support firefighting operations.
Trees	Urban tree inventory data
Traffic Management System - CCTV	Real-time traffic data from surveillance cameras for mobility analysis.

10.3 Digital Service Development

The first iteration of the Digital Service focuses on the visualisation of occurrences and the optimisation of emergency vehicle routes. In this iteration, several HVD are incorporated, which are being improved through the BeOpen Framework. The BeOpen Framework is essential in ensuring that the datasets meet the quality standards required for effective decision-making in the context of emergency response. This iteration includes Geospatial Data from Porto Municipality, particularly datasets related to Security, Police and Firefighters. This dataset provides key information on the locations of police and firefighters buildings, which are directly linked to the Digital Service for identifying the start point for the calculation of optimised routes.

The second iteration expands the Digital Service by integrating real-time occurrence data and additional road-related information, such as road width and street dimensions, which supports the real-time optimisation of emergency vehicle routes. This is in response to the specific use case need of optimising response time in areas with complex street configurations. This iteration also integrates real-time firefighters occurrence data, allowing for more detailed tracking of incidents to which firefighters are dispatched. Additionally, the real-time location of two emergency vehicles will be integrated into this iteration, which will further improve the ability to adapt routes based on the actual vehicle locations. While this data is not classified as an HVD, it plays a crucial role in responding to the evolving needs of emergency services.

The third iteration of the Digital Service focuses on integrating additional verticals. This iteration includes HVD on Observed Meteorology from Porto's Digital Sensory Network, which provides meteorological data essential for emergency response, especially during weather-related incidents. Furthermore, the integration of an HVD related to Porto's drainage network sensors - LACROIX - will enable the Digital Service to track and manage flood risks. Additionally, a water level sensor dataset will be included (Porto Water and Energy Floodings - TAGO dataset) to help manage flood events, an essential part of the use case aimed at mitigating the impact of weather-related disasters. The spatial information defining the geographical boundaries of watersheds in Porto is also planned to be included in the Digital Service. Furthermore, the geolocation of trees, CCTVs, and fire hydrants is also expected to be integrated in this iteration of the Digital Service.

These HVDs are evaluated through the BeOpen Framework, which ensures that each dataset meets the necessary standards for integration and effective use in the Digital Service. The connection between the BeOpen Framework, the Digital Service, and the HVDs is critical to the successful development and deployment of the dashboard. The BeOpen Framework plays a key role in identifying and validating high-value data, ensuring it is aligned with specific use cases related to emergency management.

The dashboard for visualisation and emergency teams is being developed based on the knowledge from the Urban Platform, developed by Ubiwhere and already deployed in Porto Municipality, as well as other municipalities. However, as previously mentioned, this platform is being fully customised based on insights from Porto's municipal stakeholders. As a result, Figure 29 depicts the architecture overview of the dashboard developed within the scope of the BeOpen project.

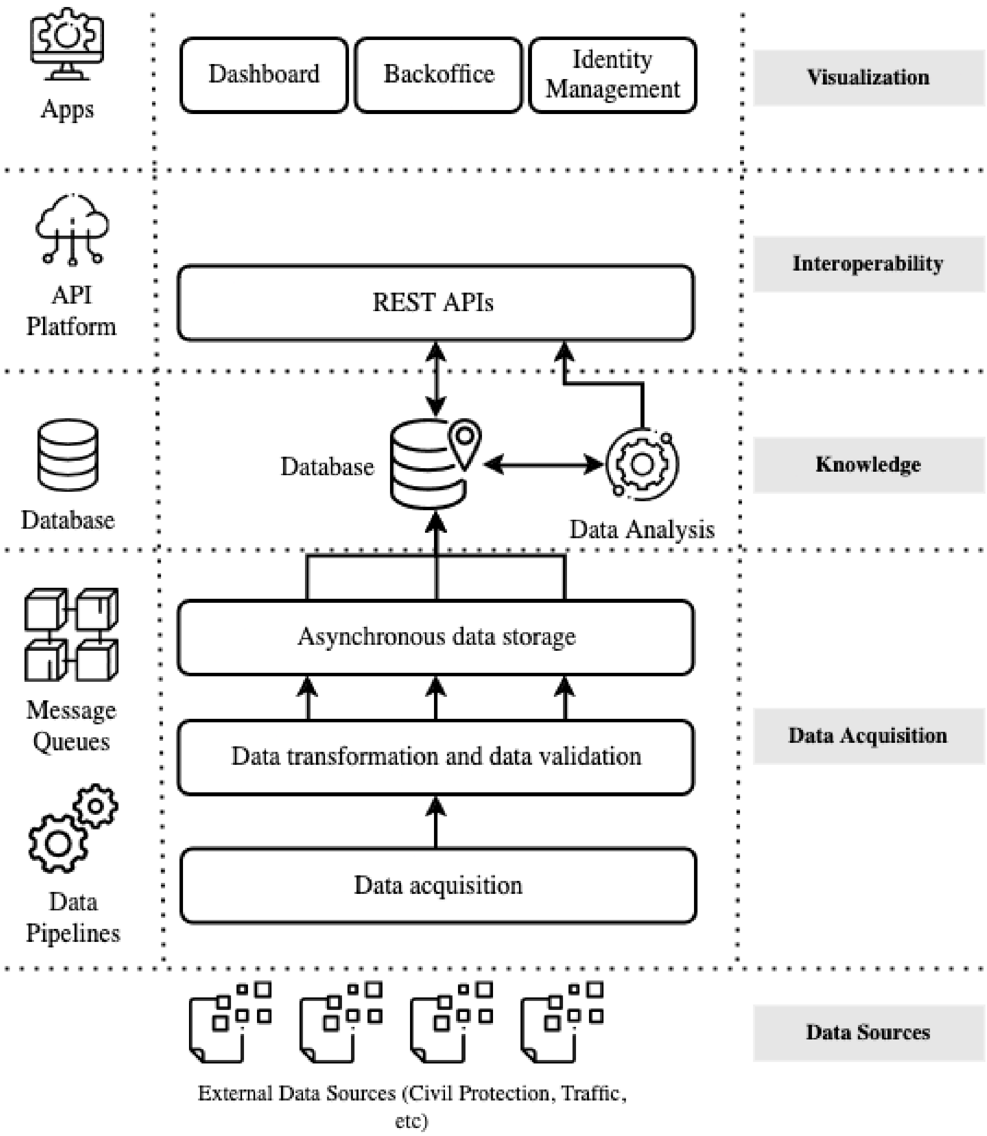


Figure 29. Architecture overview of the dashboard for real time visualisation of occurrences and emergency teams.

1. **Data Sources Layer**

The dashboard is being developed using various datasets provided by stakeholders, which were identified as relevant for its objectives. A detailed description of the data sources used in the dashboard can be found in Sections 9.2 and 19.3. In summary, these datasets include information from meteorological sources, water lines, roadway and street dimensions, geolocation data for police and firefighters, fire hydrants, trees, and CCTVs.

2. **Data Acquisition Layer**

This layer is responsible for integrating all the data into the platform. It is developed to perform it in an incremental and controlled way and ensure that the domains are standardised and harmonised . This layer interacts in a controlled manner with external systems, other platforms and devices of interest and can normalise the data from different protocols and standards into common data models.

3. **Knowledge Layer**

This layer is where all the platform's data are stored, already harmonised, in a central and scalable form. This data is the support of all the analytical applications, such as KPIs to support decision-making.

The database used in this layer is PostgreSQL, one of the most supported open-source databases currently available. PostgreSQL typically works seamlessly with Django (the web framework used to develop the platform). PostgreSQL is designed to handle large datasets and high concurrency since it can handle millions of spatial features and scale horizontally across multiple servers.

4. **Interoperability Layer**

The interoperability layer contains the software components that facilitate communication and data exchange between different systems, applications, or components, acting as a bridge between different technologies, protocols, and formats, enabling them to work together seamlessly.

The interoperability layer consists of open APIs (Application Programming Interfaces) and protocols that provide a standardised interface for communicating with other systems. This allows the integration of different components and services with the dashboard, even if they are built with different technologies or use different data formats.

The interoperability layer ensures that data is consistent and accurate, even if it is coming from different sources, by providing translation and mapping services to normalised ones (using SmartDataModels). Overall, it is a crucial component, as it provides a standard interface for communication and data exchange and helps to ensure that data is accurate, consistent, and secure.

5. **Visualisation Layer**

The User Interface/Usability Layer (UI/UX) is being fully adapted to the BeOpen project, integrating recommendations from Porto's municipal stakeholders. This adaptation is an ongoing process, with three iterations planned, during which stakeholders will provide feedback on the dashboard at each stage. The main screens from the first iteration of the dashboard, focused on the visualisation and emergency teams, are presented with further detail in the next Section (Section 9.5).

10.4 Digital Service Output

The results from the first iteration of the Digital Service include the identification of occurrences and the calculation of response routes.

1. **Visualisation of Occurrences**

   Occurrences are visualised on a map (Figure 30) and listed in a side panel, which provides details on the severity level and resolution status. Users can filter occurrences by severity. The dataset used to identify these occurrences will be updated in a second iteration to include real-time data from emergency services.

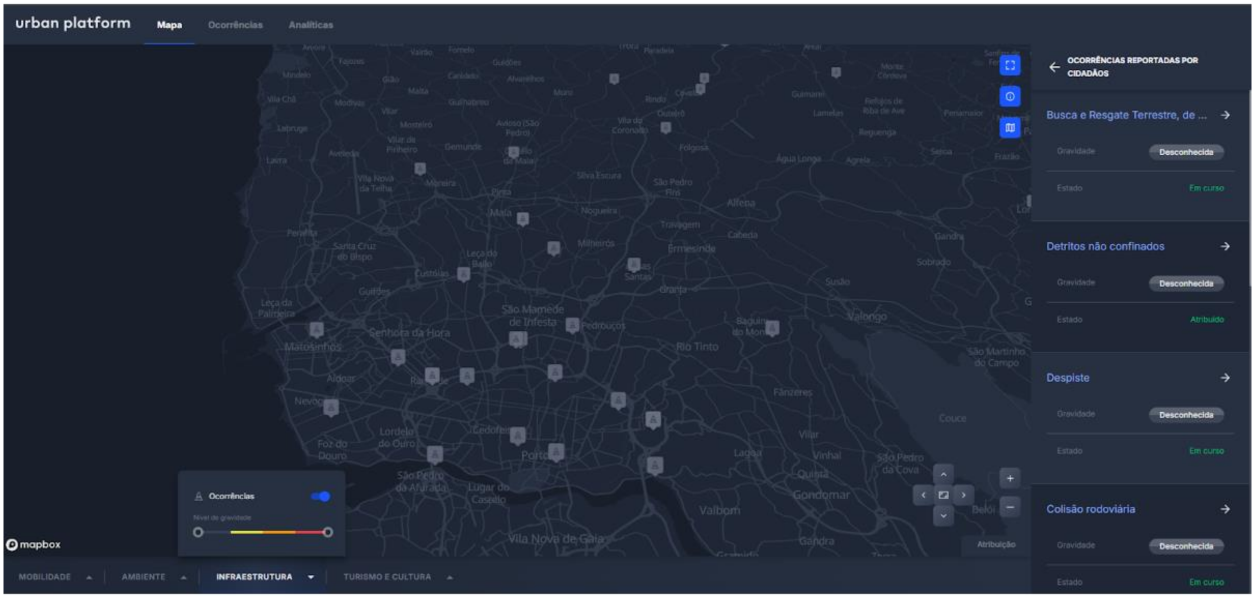


Figure 30. Dashboard showing the map of occurrences.

2. **Detailed Occurrence Information**

By selecting a specific occurrence on the map, the view zooms in, and the side panel displays detailed information (Figure 31), such as:

- ID: A four-digit code assigned to each occurrence for easy identification. This feature will allow different departments to coordinate on the same occurrence using the same identification, which is not possible at this moment.
- Location: The complete address and geographic coordinates of the incident.
- Description: A brief description of the incident, such as, road collision, animal rescue, tree fall, pedestrian accident, or structural collapse.
- Resolution status: The current status of the occurrence (colour coded):
    - In dispatch: Resources are in transit to the incident
    - In progress: The occurrence is ongoing, with no area limitations
    - In resolution: The incident is contained, with no danger of spread
    - In conclusion: The occurrence is resolved, but still under observation
- Number of involved resources: The total number of human resources (emergency responders) and vehicles (terrestrial or aerial) assigned to the incident.

This detailed information supports the responders and operators in quickly assessing the status, severity, and resources allocated to each occurrence.

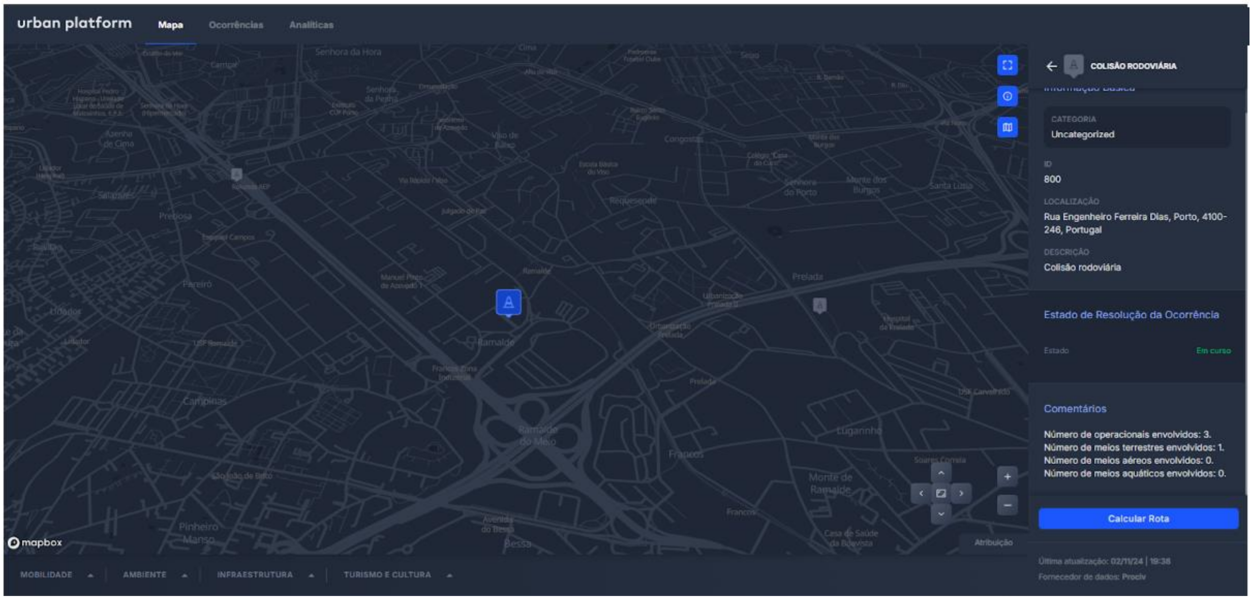


Figure 31. Dashboard showing detailed information for an individual occurrence.

3. Route calculation

The side panel for each individual occurrence features a button to calculate the route, which opens a popup window (Figure 32). This window allows the user to select the relevant department from which the route will be calculated to the incident location. The available departments are derived from the Geospatial Data provided by Porto Municipality's HVD, as described earlier. This dataset currently includes a total of 7 emergency stations (Figure 33).

After selecting the relevant department, the route is calculated from the selected department to the incident location and displayed on the map (Figure 34).

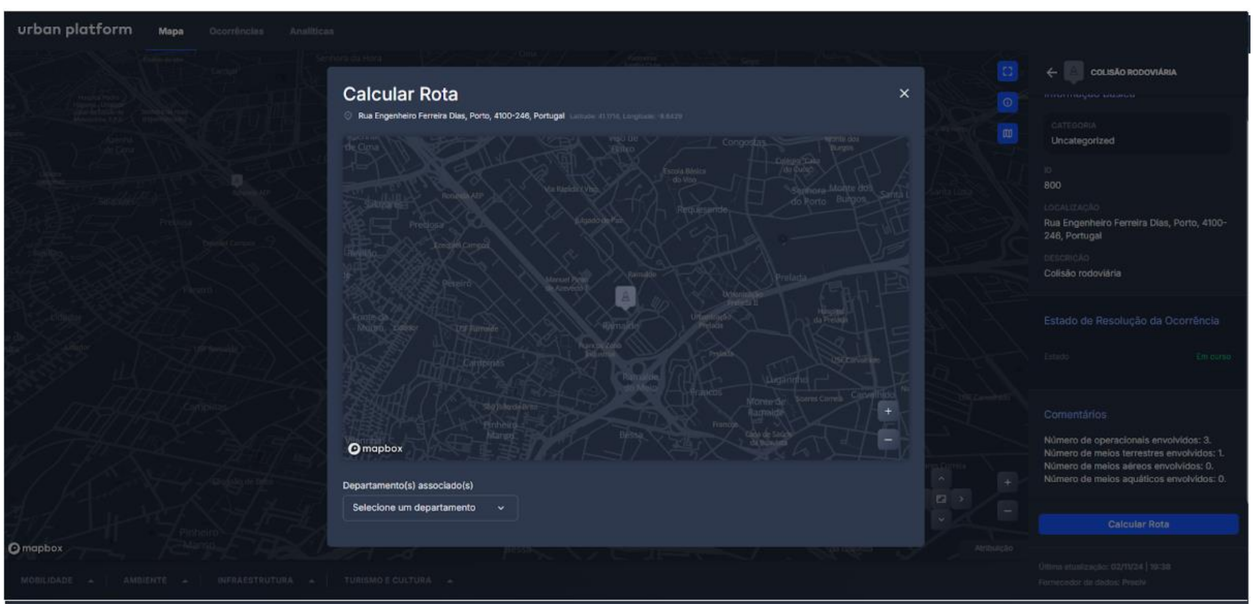


Figure 32. Popup window for route calculation from an emergency station to an incident location.

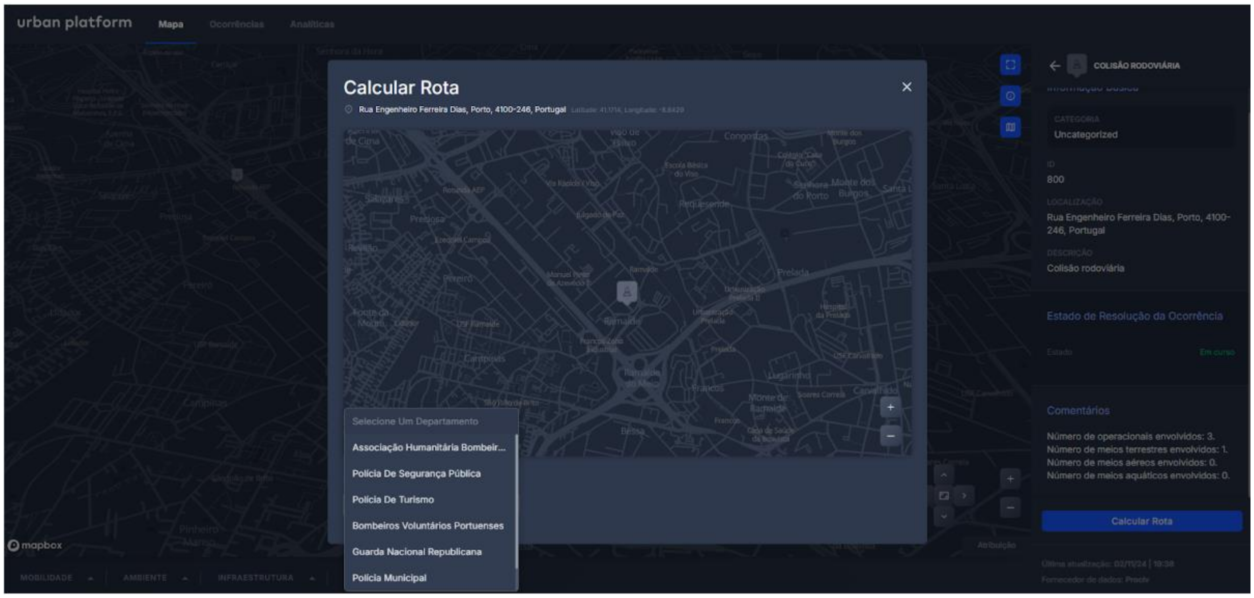


Figure 33. Dashboard displaying the available emergency departments for route calculation.

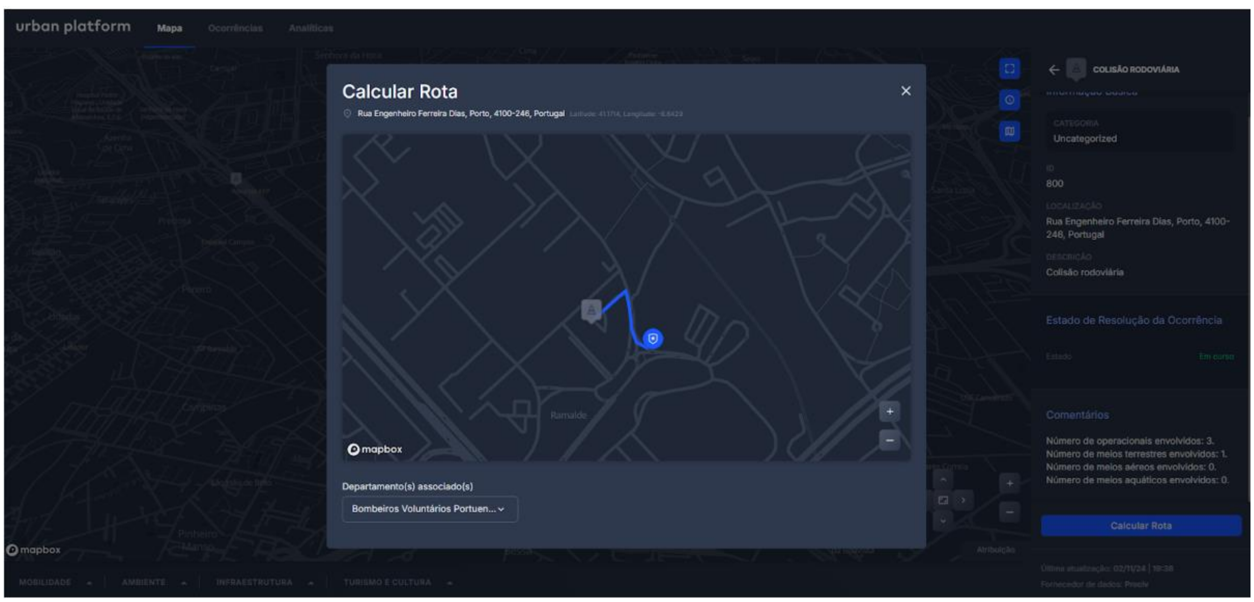


Figure 34. Dashboard showing the route from the selected emergency department to the incident location.

The first iteration of the Digital Service provides valuable insights into the real-time visualisation of occurrences and the calculation of response routes. By displaying detailed information on each incident along with the calculation of routes, this Digital Service helps emergency teams make quick and informed decision-making. Future iterations will incorporate additional HVDs, further enhancing the Digital Service capabilities and improving the efficiency and effectiveness of emergency management.

10.5 Next Steps

The next steps of the Digital Service will promote an iterative approach.

1. **Second Iteration: Enhancement of Occurrence Visualisation and Route Optimisation**

   The objective of the second iteration is to improve the visualisation of occurrences and the optimisation of route calculation. The main updates for this iteration include:

- Real-time occurrences: The Digital Service will be updated to display real-time occurrence data, particularly for relevant data from municipal stakeholders, such as firefighters. This enhancement will provide a more accurate overview of ongoing incidents, improving the response coordination.
- Road characteristics: The Digital Service will integrate additional road characteristics, such as road width, to enhance route optimisation. This will enable more accurate route calculation that considers not only traffic conditions and distance, but also the specific characteristics of emergency vehicles, ensuring they have sufficient space to pass the roads and perform manoeuvres when necessary. It is important to note that this dataset is not yet classified as an HVD.
- Emergency vehicles tracking: This iteration will also integrate real-time tracking of two emergency vehicles with different characteristics, enhancing route planning and emergency response coordination in real-time.

2. **Third Iteration: Integration of New Verticals with a Focus on Environmental Data**

   The objective of the third iteration is to integrate additional data verticals, providing a more comprehensive Digital Service, with a main focus on environmental data. Key features to be integrated in the Digital Service include:

- Meteorological information: This iteration will incorporate data on environmental variables, including meteorological observations from Porto Digital sensory network. This dataset will cover key parameters such as minimum and maximum temperature, humidity, UV index, precipitation, and wind speed. Additionally, historical weather data for the previous 7 days will be included to support historical analysis.
- Weather forecasting: Daily weather forecasts for the Porto municipality will be integrated, focusing on temperature (minimum and maximum), precipitation probability and intensity, as well as wind speed and direction for the following 5 days.
- CCTV location: A new vertical will include the locations of CCTV cameras installed across the municipality.
- Fire hydrants location: A new vertical will include the locations of fire hydrants installed across the municipality.
- Water basins and caudal monitoring: This iteration will include the integration of a water-related vertical, which provides data on water basins, including their boundaries and identification.

Furthermore, information from caudal sensors, detailing current water levels as well as historical data over recent days will also be included.

- Tree inventory and mapping: This iteration will also introduce a tree inventory, which includes mapping the location of trees, identifying species, and displaying their age and unique ID.

These improvements and new verticals will be integrated using the datasets and HVD explained in more detail in Section 9.3 (Use of BeOpen Framework Tools). Furthermore, these iterations will include the feedback from municipal stakeholders to better align the Digital Service with their actual needs. Therefore, while the next steps are outlined here, they may require further adjustments as the process develops and are dependent on data availability.

11 Pilot Naples: Integrating Mobility and Environment Data for Metropolitan Transport and Public Space decision making

11.1 Digital Service Requirements

The first digital service developed for Naples focuses on addressing key challenges in urban mobility, environmental monitoring, and public space management through an integrated, data-driven approach. Its main objective is to enhance the accessibility, usability, and interoperability of urban datasets by enabling data sharing, interactive visualization, and detailed querying via a geographic map. This service empowers municipal authorities, urban planners, and citizens with intuitive tools to access high-quality spatial and non-spatial data. Through its user-friendly mapping interface, the service facilitates the exploration and analysis of datasets related to urban infrastructure, transportation networks, and environmental conditions, supporting informed and evidence-based decision-making.

The second digital service is specifically designed to monitor mobility parameters, as mandated by the Sustainable Urban Mobility Plan (PUMS). With the periodic update of the PUMS initiated in November 2024, this service plays a pivotal role in integrating and assessing new datasets alongside baseline information. Initially tailored for use by municipal officers managing mobility, the service will later support public access, fostering citizen participation in urban planning through adapted versions of the tool.

This digital service is structured around two interactive maps. The first map focuses on sustainable mobility datasets enhanced through the BeOpen project, providing stakeholders with streamlined access to critical mobility metrics. The second map highlights urban public spaces, presenting insights into environmental factors, including urban green areas and climate-related risks. These tools enable decision-makers to pinpoint areas or buildings that require adaptation measures to mitigate climate risks such as urban heat islands or flood-prone zones.

These services also facilitate the redesign of public spaces using adaptation strategies and guidelines derived from prior European research projects, such as the Clarity EU initiative (2020). Proposed interventions include urban afforestation, regeneration of recreational parks, repaving, and de-sealing solutions to combat extreme climate events. Furthermore, contextual data on points of interest supports the evaluation of existing public transport options and informs strategies to promote soft mobility.

The primary stakeholders for these digital services include the citizens of Naples, with plans underway to launch a CKAN portal for Open Data to broaden accessibility. These tools will complement ongoing efforts to foster transparency and participatory urban governance. The BeOpen results will enhance these strategies by improving the flow of information and strengthening citizen engagement in urban planning. Additionally, making open datasets available in machine-readable formats and accessible through APIs will facilitate integration with European and national open data platforms.

Another anticipated aspect of the digital services involves integrating data from the NA2.2.1.B project, *"Intelligent Infrastructures and Technologies for the Management of Traffic Flows – Sustainable Mobility, Strategies for Smart Mobility and MAAS Data."* However, access to this data will require collaborative engagement with stakeholders to address tailored data-sharing protocols. The BeOpen framework is well-positioned to guide this process, providing a clear pathway toward advanced data-sharing practices.

Emerging from BeOpen is a new focus on mobility opportunities driven by private companies, addressing data that is currently outside municipal collection efforts. Insights from other BeOpen pilots and collaborative exchanges with city officers have provided innovative strategies to tackle this challenge. Future initiatives are expected to involve IoT data collection and AI-based use cases, reflecting Naples' commitment to innovative solutions for mobility management.

Ethical and legal considerations surrounding information management have been central to the development of the BeOpen framework. By addressing uncertainties and providing actionable guidelines, the framework ensures compliance with legal interoperability standards while maximizing the dissemination and reuse of public open data. Links to information repositories and relevant guidelines will be integrated into the digital service to support this effort.

Ultimately, these digital services aim to deliver user-friendly, HVD-tailored tools to civil servants involved in urban planning and public administration. By streamlining decision-making processes and enhancing the management of mobility and environmental data, these services will support a more transparent, inclusive, and participatory approach to urban governance.

11.2 Digital Service Design

The services are implemented in the city GIS sharing environment, ESRI Portal technologies. Improved data from BeOpen Dataspace are downloaded on local machines and uploaded on the Esri Portal. The Portal accepts geoJson formats for point features. A problem has arisen with polygon data, and geojson from BeOpen data need a passage in Esri shapefile format (processed in Qgis software, open-source GIS software) to be uploaded and published. The problem is to be investigated for an easier workflow.

The services host additional data from different domains. Data updating is possible with a simple upload of a new file, provided that attribute structure and filename is the same. Otherwise uploading and thematization has to be made again. Datasets belong to Mobility –Environment –Administrative- Statistics domains.

The Digital Services for the city of Naples are implemented within the municipal Geographic Information System (GIS) environment, utilizing ESRI Portal technologies. This robust and scalable infrastructure facilitates the integration, management, and dissemination of geospatial data, serving as a centralized platform for diverse stakeholders, including municipal authorities, urban planners, and citizens.

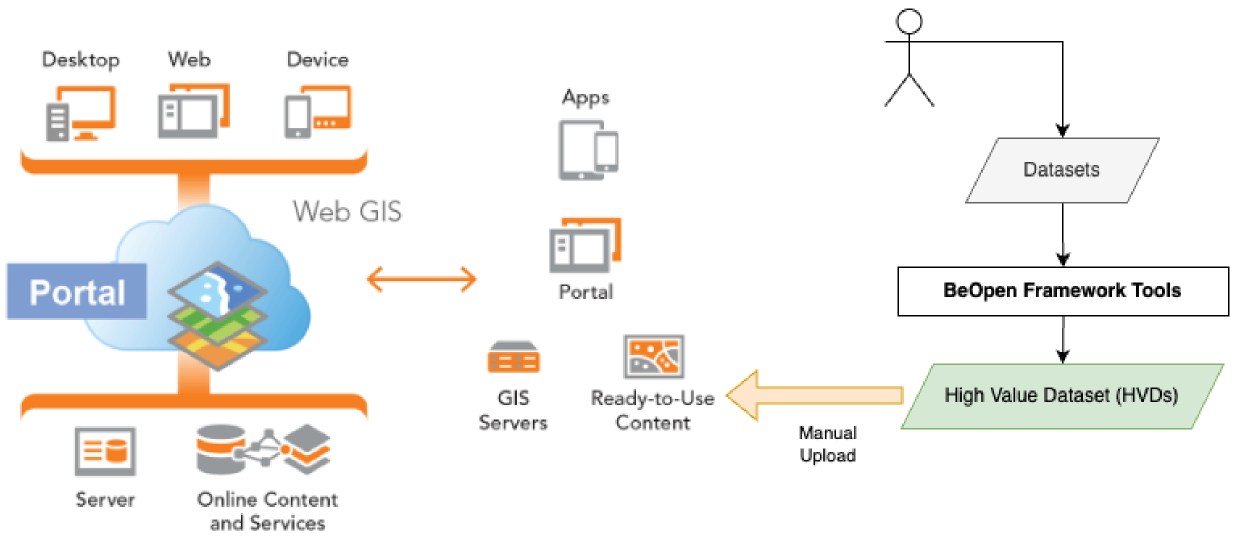


Figure 35. Digital Service Naples architecture.

The architecture, as illustrated in the Figure 35, demonstrates a seamless flow of data between the Portal, GIS servers, and users. BeOpen-enhanced datasets are first processed and improved within the BeOpen Dataspace, after which they are manually downloaded to local machines and subsequently uploaded to the ESRI Portal. While point features are readily integrated using GeoJSON formats, challenges persist with polygon data, which require conversion to ESRI shapefile format via QGIS, an open-source GIS platform, before uploading and publishing. This additional step adds complexity to the workflow, and ongoing efforts are aimed at streamlining the process to support direct GeoJSON uploads for all data types.

The Portal supports a variety of additional datasets spanning mobility, environment, administrative, and statistical domains. These datasets can be updated efficiently through the upload of new files, provided the attribute structure and filename remain consistent. If these parameters are altered, datasets must be re-uploaded and re-thematized to ensure accurate visualization and functionality within the system.

The integration of datasets processed using the BeOpen Framework tools with locally sourced geospatial data enables a comprehensive and cohesive analysis of urban challenges. The framework enhances the datasets to HVD status, ensuring compliance with European standards and fostering interoperability. Through this architecture, users can interact with GIS-ready content across devices, including desktops, web platforms, and mobile devices, enabling decision-makers to derive actionable insights.

Additionally, the services incorporate real-time updates and multi-domain datasets, offering a holistic perspective on urban challenges. This design empowers stakeholders to address critical issues such as mobility management, environmental risks, and public space planning with precision and efficiency. The architecture promotes scalability, facilitating future integration of advanced data types and supporting the city's evolving needs in urban planning and management.

11.3 Digital Service Output

The Digital Services implemented in the city of Naples have resulted in the creation of intuitive and functional dashboards designed to support decision-making in urban mobility, environmental monitoring, and public space management. These dashboards provide stakeholders with detailed insights derived from harmonized HVDs processed through the BeOpen Framework.

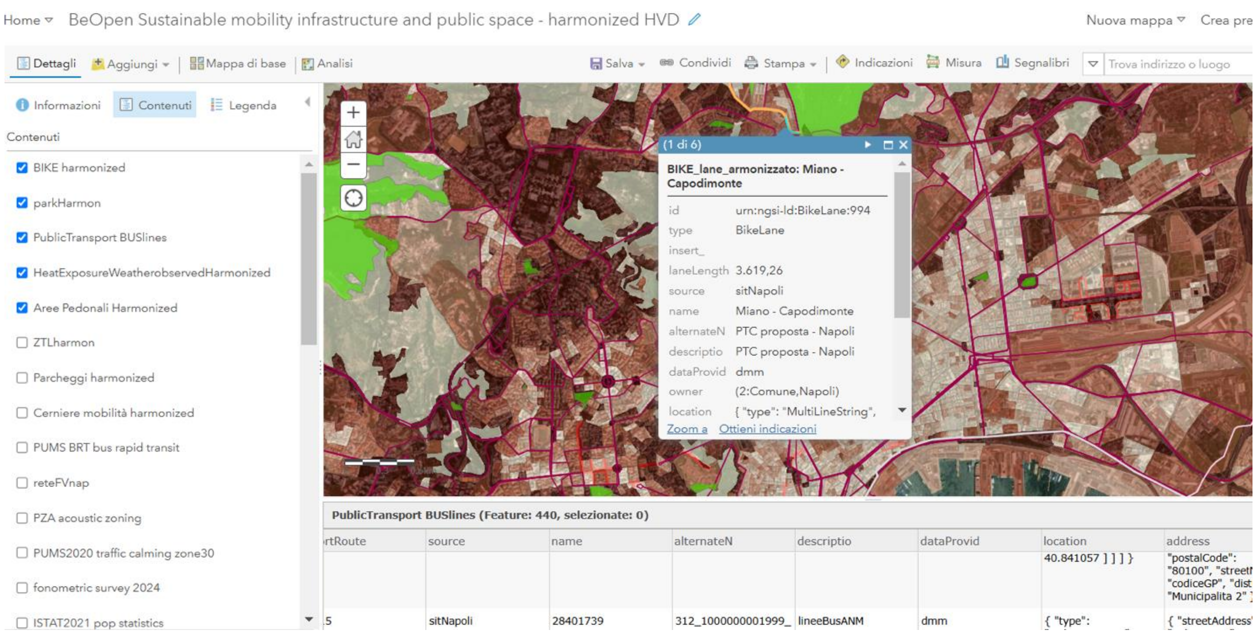


Figure 36. Naples Digital Service, data visualization.

As depicted in the Figure 36, the first Digital Service provides a highly customizable and user-friendly interface. Users can interact with individual map elements through detailed pop-up windows (Figure 36), which display attributes such as bike lane lengths, source data, and geographic locations. The web map allows users to personalize their experience by filtering content—for example, isolating public parks smaller than 2,000 square meters (Figure 37). This functionality enhances the ability to focus on specific datasets relevant to particular planning or operational needs.

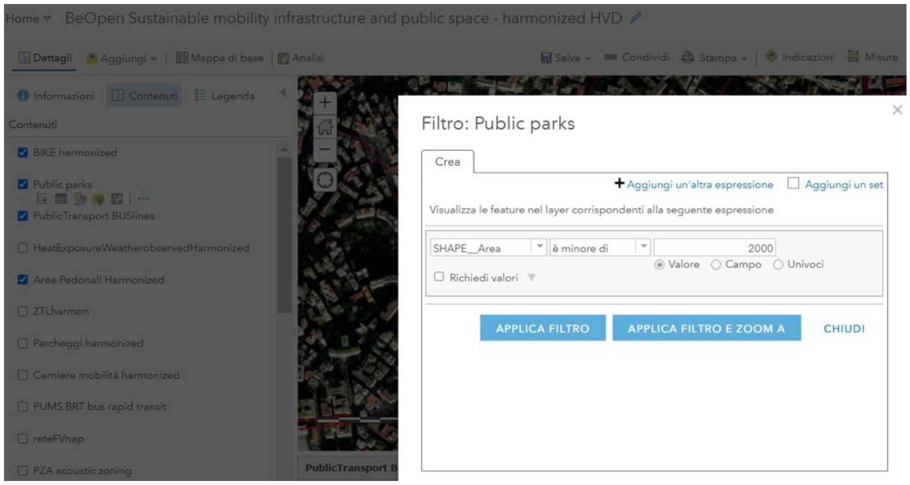


Figure 37. Naples Digital Service. Filtering feature.

The maps can also be overlaid on various base maps, including orthophotos, OpenStreetMap (OSM), and custom imagery, allowing flexibility in visualization. Users can toggle layers on and off to refine their view and create a streamlined analysis environment. Furthermore, the map content can be incorporated into simplified web applications (WEBAPPs), offering intuitive access to essential tools such as geocoded street searches. This streamlined approach makes the digital services particularly effective for urban and environmental data sharing in city workflows.

Figure 38. Digital Service Naples, Monitoring and analysis dashboard.

The Figure 38 illustrates a dynamic monitoring and analysis dashboard, which aggregates real-time and historical data for visualizing key urban indicators. This dashboard provides a holistic view of urban sustainability, integrating data on heat exposure, green infrastructure, and public transport usage. The heat exposure layer, for example, identifies schools at risk due to high temperatures, while the green index layer assesses vegetation coverage across urban areas. The dashboard also incorporates infrastructure data such as bus stop locations and public park counts, offering a comprehensive perspective on urban resource distribution. The user interface is designed to enhance usability, featuring visual elements such as heatmaps, bar charts, and interactive panels. Indicators such as the number of public parks and bus lines are dynamically updated, providing real-time metrics for stakeholders. This level of granularity supports precise and data-driven decision-making, particularly in identifying areas for environmental or mobility interventions.

These dashboards represent a significant advancement in urban data accessibility and usability for the city of Naples. By consolidating harmonized datasets into intuitive tools, they empower municipal authorities, urban planners, and citizens to make informed decisions. The combination of spatial visualization and geoanalytics enhances the city's capacity to address critical challenges such as climate adaptation, mobility optimization, and resource allocation.

11.4 Next Steps

The city of Naples is set to release the second version of its digital service in early January 2025, expanding the audience to a broader group of potential users. This next iteration will serve as an opportunity to showcase the capabilities of the service and demonstrate the potential workflows powered by the BeOpen Framework. The primary goal is to enhance the user experience by incorporating feedback from initial stakeholders, addressing their needs for more advanced data collection, and refining the service based on their recommendations.

Several key actions are planned for the next steps:

1. **Expansion to a Wider User Base**: The second version of the digital service will be shared with a larger group of users across various municipal departments and external stakeholders. This will include representatives from urban planning, mobility management, environmental monitoring, and other municipal services, facilitating a more comprehensive use of the platform and ensuring broader adoption.
2. **Tailored Data Collection and Improvement**: Based on the feedback gathered from the initial version, the next iteration will include tailored sessions to improve data collection processes. These sessions will focus on enhancing the quality and granularity of existing datasets, addressing specific gaps identified by users, and ensuring that new datasets are aligned with user needs and city objectives.
3. **Extension to Other Municipal Areas**: The digital service will be extended to additional areas of the municipality, promoting a citywide approach to data sharing and management. This expansion will encourage cross-departmental collaboration and facilitate a more holistic approach to urban management, supporting more integrated decision-making across various city domains.
4. **Improvement of High-Value Datasets (HVDs)**: One of the key objectives for the next version will be the improvement and expansion of datasets shared as HVDs. This will include enhancing existing datasets by linking them with additional relevant information, improving their usability for various urban planning tasks.
5. **Enhancement of Monitoring and Statistical Tools**: The monitoring service will be further refined to improve the accuracy and relevance of the data it provides. New statistical tools and measurement features will be integrated into the system, allowing users to track trends over time, visualize patterns, and make more informed decisions based on comprehensive data analysis.

12 Pilot Vilnius: Detection and Projection of Invasive Species Spread

12.1 Digital Service Requirement

The digital service for the Vilnius pilot aims to tackle the ecological threat posed by the invasive species *Heracleum sosnowskyi*. By leveraging machine learning techniques, the service focuses on identifying the current locations of this invasive plant and projecting its potential spread if left untreated. The primary objective is to create a digital map of detected locations and predict future spread scenarios, enabling the city to devise informed strategies for limiting and eradicating the plant. This service supports ecological preservation by providing actionable insights to improve resource allocation, enhance planning effectiveness, and monitor critical sites impacted by invasive species.

The pilot leverages strong support from local authorities, digital service providers, and research institutions to ensure alignment with ecological and urban management goals. Key stakeholders include:

1. **Municipality of Vilnius**: Acts as an innovation hub, providing institutional support and media visibility for the pilot.
2. **ID Vilnius**: a separate working unit, with the data center providing digital services to the city and closely engaging with the municipality in managing, conceptualizing and creating various services.

12.2 HVDs supporting the Digital Service

Dataset	Description
Natural framework	Geospatial data on natural landscapes and ecosystems within urban and rural areas.
Extensively and intensively used green spaces	Categorized information on green spaces.
Green area accessibility	Data on the accessibility of green spaces to the public.
Invasive species	Records of invasive plant species distributions. Labels collected into a comfortable geospatial format for visualization to stakeholders.

The HVDs produced by the Vilnius pilot adhere to open data principles, ensuring accessibility, interoperability, and reusability. These datasets are:

1. **Dublin Core Metadata Standard**: All datasets are described using the Dublin Core standard, ensuring rich and structured metadata for better discoverability and interoperability.

2. **CC BY-SA 4.0 License**: The datasets are made available under the Creative Commons Attribution 4.0 license, enabling free use, sharing, and adaptation as long as appropriate credit is given. No sensitive data is included, ensuring compliance with privacy regulations.
3. **Free Accessibility**: The datasets are provided at no cost, promoting transparency and encouraging widespread use by various stakeholders, including researchers, public authorities, and the general public.
4. **CKAN API Integration**: The datasets are accessible through the CKAN API on the Vilnius open data portal, enabling seamless programmatic access for developers and data users.

By following these standards, the Vilnius pilot ensures that its datasets are not only useful for addressing the spread of *Heracleum sosnowskyi* but also contribute to broader data sharing and open innovation efforts.

12.3 Digital Service Design

The architecture of the Vilnius pilot's digital service is a comprehensive system designed to detect and predict the spread of *sosnowskyi*, combining machine learning, environmental data, and web map visualization. It integrates multiple components into a seamless pipeline that processes raw data, generates predictions, and provides actionable insights through an interactive mapping platform. Figure 39 outlines the whole pipeline.

The pipeline begins with GPS coordinates collected from locations where *Heracleum sosnowskyi* has been detected (labels). These data points serve as the foundation for the entire system. The collected labels are processed and stored in a central database, where it is used to train an AI algorithm capable of recognizing patterns and predicting the potential spread of the invasive species.

The AI model analyzes environmental variables, including wind direction and wind displacement, which are critical in determining how seeds or particles from *Heracleum sosnowskyi* may travel over time. By integrating these factors, the algorithm generates a dataset of shifted GPS coordinates, simulating how the species is likely to spread under specific conditions. This predictive capability is planned to be further enhanced by incorporating HVDs, the first being the output of the DS – invasive species HVD. This dataset is given more context through other HVDs such as natural framework, green spaces accessibility data in a user-friendly digital map. The contextual HVDs also have the potential to serve as input for prioritizing and planning drone flights to cover sensitive areas and collect data.

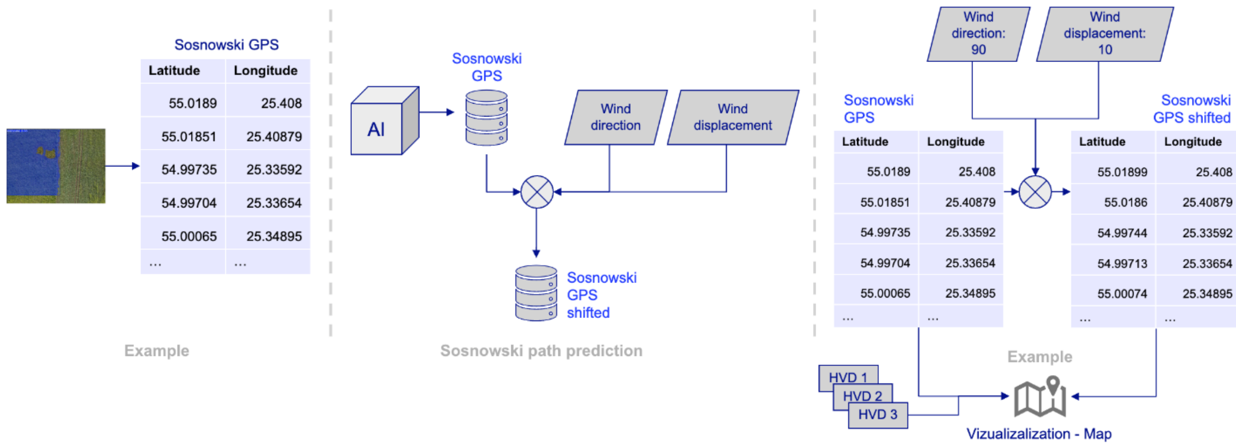


Figure 39. The Sosnowsky's hogweed path prediction process implemented.

The output of this analysis is a refined map of both current locations and predicted future spread of *Heracleum sosnowskyi*. This information is visualized through an intuitive mapping interface, enabling stakeholders to easily identify affected areas and prioritize eradication efforts. The map not only highlights existing infestations but also provides forward-looking insights that help planners and environmental managers allocate resources effectively and mitigate the ecological impact of the invasive species. The supporting HVDs show sensitive areas, where the reachability of greenspaces by citizens is high or showcase natural areas that need protection. This information could be helpful setting priorities, planning data collection since it's not possible to cover the whole city with drones. The meteorological wind data is utilized to predict the possible spread of the invasive species identified in collected RGB images.

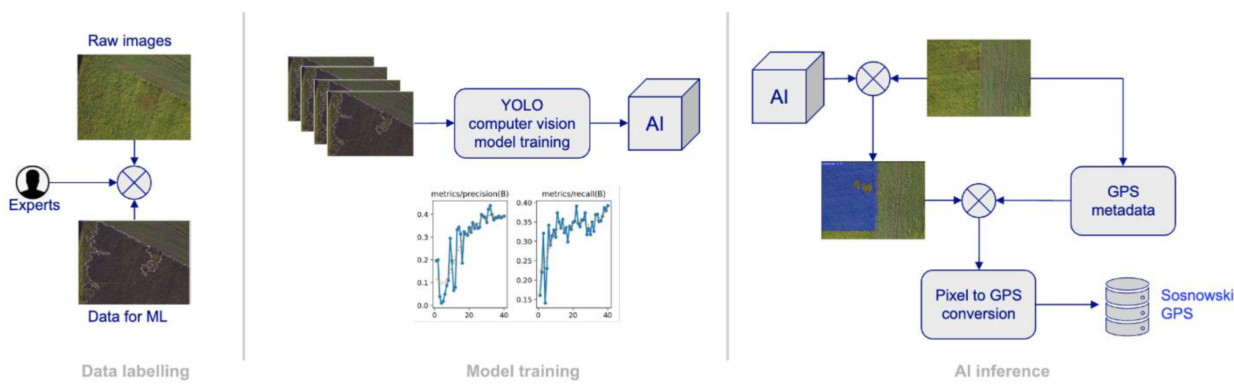


Figure 40. The implemented approach for the digital service developed for the city of Vilnius, showcasing the process of detecting Sosnowsky's hogweed using AI.

Figure 40 outlines the machine learning process implemented to identify the location of the Sosnowsky's hogweed.

Starting from the data labelling, the pipeline starts with the acquisition of **raw imagery**, which is processed to identify regions potentially affected by *Heracleum sosnowskyi*. These raw images are labeled under the guidance of experts who manually annotate the invasive species' locations. The labeled data serves as the ground truth for the machine learning model. This step is crucial for creating a high-quality training dataset that accurately reflects the characteristics of the target species.

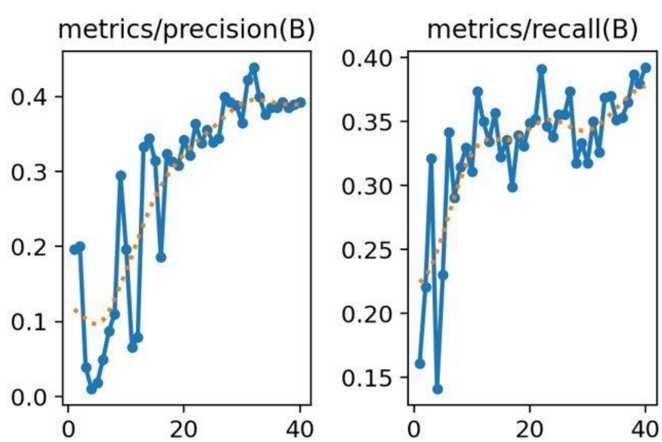
Annotated datasets are prepared with pixel-level precision, ensuring that the labels capture even subtle variations in vegetation that indicate the presence of *Heracleum sosnowskyi*. This data is then preprocessed to ensure consistency in resolution, format, and structure before being passed into the model training phase.

The annotated images are fed into a **YOLO (You Only Look Once)** computer vision model, a state-of-the-art neural network architecture optimized for object detection tasks. The YOLO model is trained to recognize *Heracleum sosnowskyi* by learning patterns, textures, and spatial characteristics specific to the species. During training, the model is exposed to multiple image samples, and its parameters are iteratively adjusted to improve performance. The model training process includes evaluating precision and recall metrics, ensuring the model is capable of detecting the species accurately while minimizing false positives and false negatives. These metrics are monitored and optimized during training to achieve a balance between high detection accuracy and computational efficiency.

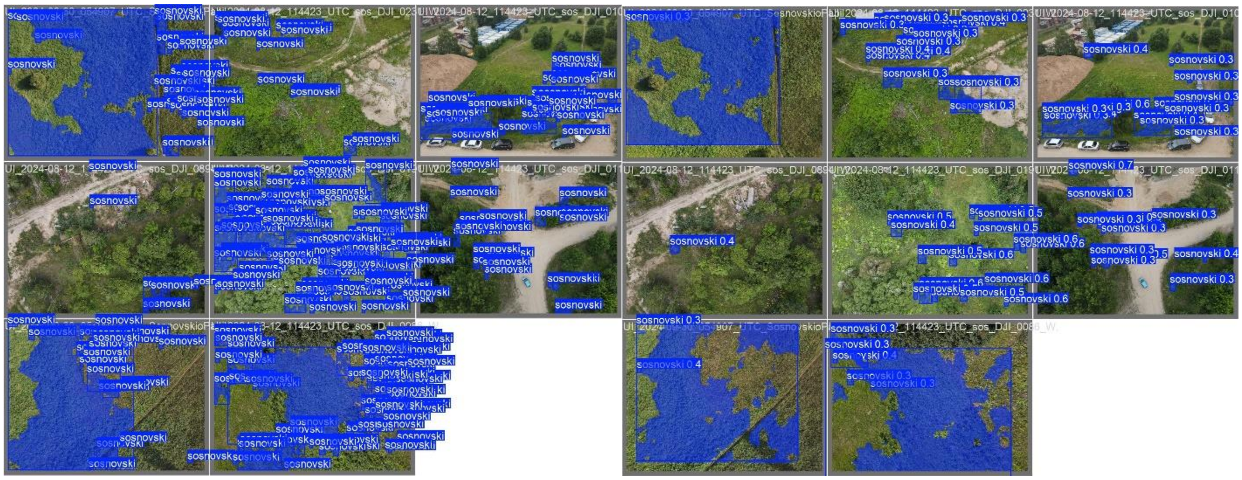
Once trained, the YOLO model is deployed for **AI inference**. In this phase, the model is used to analyze new images, identifying *sosnowskyi* in previously unexamined areas. Detected regions are processed to convert pixel-level information into **georeferenced GPS metadata**, providing precise location coordinates for the identified invasive species. The final step involves converting pixel-based predictions into GPS coordinates to create a geospatial dataset of *Heracleum sosnowskyi* locations. This dataset is stored in a database, where it can be accessed for visualization and further analysis. The georeferenced predictions enable stakeholders to monitor current infestations and plan interventions effectively.

12.4 Digital Service Output

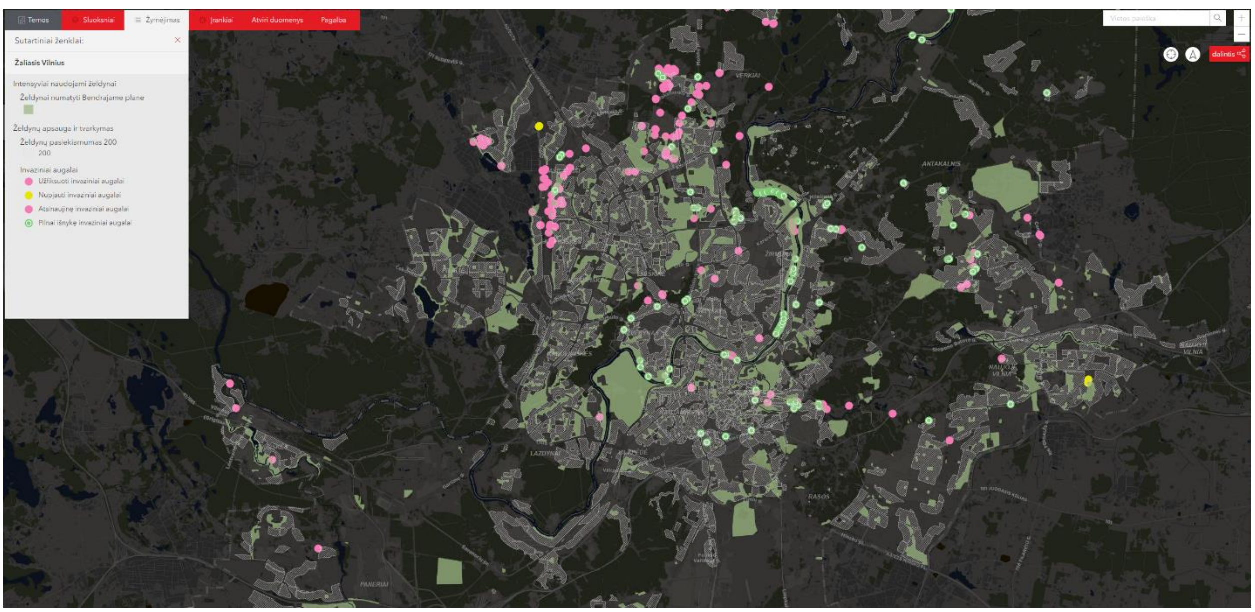
Currently, the YOLO model is capable of predicting the presence of Sosnowski hogweed with roughly 40% accuracy and 40% recall. Judging by the curve of the graph, the model should increase in accuracy as more data will be provided with the next growth season:



An example of the labelled plants and model predictions for the same areas:



A test output was created for shifted Sosnowski coordinates, where wind data was used for modelling. Also, see below the first version of the visualization map with previously collected various invasive species, green space reachability and extensively and intensively used green spaces (natural framework layer is not enabled):



12.5 Next Steps

Moving forward, the next steps will focus on improving the invasive species dataset as well as the shifted coordinates version, because it lacks the closing coordinate of the polygon and currently requires more steps to process the data. The map will encompass all data at the end with additional attributes for the invasive species. The pilot will also be enhancing the service's accuracy, scalability, and usability while incorporating innovative data sources and methodologies to improve its performance. With a new plant growth season approaching, ML data collection will continue.

One key area of improvement is the integration of **multispectral satellite imagery** into the current system. By leveraging satellite data, the service can enhance its ability to detect *Heracleum sosnowskyi* in areas that are difficult to access or not covered by ground surveys. Multispectral imagery provides valuable insights into vegetation health and type, enabling the algorithm to differentiate between invasive species and native flora more effectively. This addition is expected to significantly improve the accuracy and reliability of the predictive model, particularly in regions with diverse ecological conditions.

13 Conclusion

13.1 Key Challenges and Resolutions

The execution of the BeOpen pilot digital services across multiple cities presented a diverse array of challenges. These issues spanned technical, operational, and collaborative domains. Below, we outline the key challenges encountered during the development and implementation phases, along with the resolutions that were instrumental in overcoming them.

- **Data Availability and Quality**

Challenge: Inconsistent availability and varying quality of HVDs posed significant barriers. Many datasets lacked the necessary metadata, were incomplete, or did not comply with interoperability standards.
Resolution: The integration of tools such as the Metadata Quality Validator within the BeOpen Framework addressed these issues by ensuring metadata completeness and adherence to standards. Collaborative efforts with local stakeholders also enhanced data collection and validation processes.

- **Semantic Harmonization**

Challenge: The diverse nature of datasets across pilot cities complicated the semantic harmonization required for interoperability. Variations in data formats and lack of standardization led to inefficiencies in integration.
Resolution: The implementation of BeOpen's semantic harmonization tools, including the DCAT-AP standards, streamlined the process. The use of model mapping and linked data approaches significantly improved the datasets' reusability and cross-border compatibility.

- **Integration with Existing Systems**

Challenge: Integration with legacy systems in municipalities proved challenging due to differing technical architectures and limitations in data-sharing capabilities.
Resolution: Customized adapters and APIs were developed to ensure compatibility with existing systems. This iterative approach allowed the seamless integration, minimizing disruptions to current operations.

- **Stakeholder Engagement**

Challenge: Encouraging active participation from diverse stakeholders, including municipal authorities, technical partners, and end-users, was a complex task.
Resolution: A structured engagement strategy was adopted together with an iterative approach.

The lessons learned and best practices identified will guide future implementations, ensuring smoother deployments and more impactful outcomes.

13.2 Summary of Achievements

The BeOpen pilot services demonstrator is successfully validating the potential of the BeOpen Framework to enhance the usability, interoperability, and accessibility of HVDs across diverse urban contexts. Through the implementation of digital services in eight European cities, the project has demonstrated its ability to address a wide range of urban challenges, including environmental monitoring, mobility management, disaster response, and infrastructure optimization.

The pilot deployments showcased the capability of the framework to transform data into actionable insights, enabling innovative digital services tailored to the specific needs of each city. For example, real-time data processing and AI-driven analytics were leveraged to improve urban resilience, optimize resource allocation, and enhance public safety. The integration of predictive models, combined with advanced visualization tools, provided stakeholders with actionable intelligence to address complex urban issues.

Another significant achievement was the successful engagement of a diverse range of stakeholders, including municipal authorities, technical partners, and end-users. Through a structured and iterative approach, the project fostered collaboration, built local capacities, and ensured that the developed solutions were both relevant and impactful.

Finally, the digital service development outlines the contribution to the advancement of open data principles by ensuring that datasets and tools were accessible and transparent.

13.3 Next Steps for BeOpen Pilot Services

The BeOpen pilot services have established a robust foundation for advancing the integration and application of HVDs across diverse urban contexts. Building on the successes achieved during the pilot phase, the next steps will focus on enhancing scalability, deepening the integration of advanced tools, and fostering greater collaboration across regions and sectors.

Future efforts will prioritize the refinement and scaling of digital services developed during the pilots. This includes addressing the unique needs of additional municipalities and regions, ensuring that the framework is adaptable to a wider range of contexts. Emphasis will be placed on expanding the interoperability and accessibility of datasets, enabling cross-border and cross-sectoral applications that maximize the impact of shared data resources.

Advanced analytics and real-time capabilities will be further developed to meet the growing demand for data-driven decision-making. Incorporating machine learning models, predictive analytics, and dynamic visualization tools will enhance the framework's ability to address complex challenges in urban management, environmental monitoring, and public safety. Iterative improvements to these tools, guided by feedback from stakeholders, will ensure their continued relevance and effectiveness.

The project will further invest in evaluating the long-term impact of the pilot services. By tracking key performance indicators and conducting comprehensive assessments, insights will be gained to refine methodologies and inform future iterations.