



BALTFLOODS, Baltic Flood Resilience and Digital Solutions

Deliverable 1.2

Data Systems Integration Plan

AUGUST 31, 2025

Interreg
Baltic Sea Region



Co-funded by
the European Union

RESPONSIVE PUBLIC SERVICES
BALTFLOODS



This project is co-funded by the European Union through the Interreg Baltic Sea Region Programme.

About BALTFLOODS Project

BALTFLOODS aims to enhance flood preparedness and mitigate runoff pollution in cities across the Baltic Sea region by leveraging digital and technological solutions and engaging citizens as key stakeholders. The project addresses three main challenges aligned with the thematic scope of Priority 1 of the Interreg Baltic Sea Region Programme, particularly Objective 1.2. Firstly, BALTFLOODS will improve disaster preparedness and response to floods by implementing advanced monitoring systems that provide real-time data for timely interventions, benefiting local and national public authorities, infrastructure owners, and service providers. Secondly, the project will decrease the discharge of polluted stormwater, thus enhancing environmental quality and public health. This involves monitoring water quality through innovative approaches that support environmental and public health goals. Thirdly, BALTFLOODS will increase community engagement in flood and water pollution issues through participatory tools, empowering citizens and educational institutions to take an active role in environmental stewardship. By fostering a well-informed and proactive community, the project builds societal resilience to environmental threats. Transnational cooperation will be essential to facilitate knowledge exchange, policy alignment, and resource pooling to enhance the scalability and sustainability of the solutions, ultimately benefiting urban populations and the Baltic Sea Region ecosystem.



Learn more about the project:
www.interreg-baltic.eu/project/baltfloods

Project Name	Baltic Flood Resilience and Digital Solutions
Project No.:	#C063
Submission Date	August 31, 2025
Authors	City of Lappeenranta, Sanna Varis Norwegian University of Science and Technology, Steven Yves Le Moan Norwegian University of Science and Technology, Marcos Xosé Álvarez Cid University of Gothenburg, Sebastian Andreasson



Executive Summary

Data from various sources is needed to solve diverse environmental challenges and support decision-making in municipalities. As the sources of data and the volume of data generated continue to grow, it is important to have data systems that can store, manage, process and present large amounts of data efficiently.

The BALTFLOODS project leverages the existing digital and monitoring infrastructures of two pilot municipalities, enhancing them with new tools and integration layers to overcome current limitations in interoperability, real-time data analytics, and citizen engagement. BALTFLOODS project centers on achieving interoperability across municipal systems, national databases, and third-party applications. Data systems integration plan is modular and standards-based, designed to accommodate varying levels of digital maturity among municipalities.

The integration architecture will adopt the FIWARE system as its backbone. This architecture enables seamless integration of environmental sensors, municipal data systems, satellite and drone feeds, and citizen inputs. This approach supports long-term maintainability and aligns with existing municipal workflows. It allows municipalities to connect their existing data systems rather than replacing them, facilitating replication. The project delivers a reusable methodology for stormwater quality and flood management. Rather than building a centralized system, it establishes a flexible, shared integration plan with transnational applicability.



Table of Contents

Introduction.....	10
1. Overview of Existing Data Systems	11
1.1 Lappeenranta	11
1.2 Gjøvik	11
1.3 Common Challenges Across Both Cities	11
2. The BALTFLOODS Integration Strategy	12
2.1 FIWARE as the Integration Backbone	12
2.1.1 Simplified Description of the FIWARE Architecture and Data Flow	12
2.1.2 Implementation of the FIWARE-Based Data Integration Environment.....	13
2.1.3 Early Deployment and Mock Data Integration.....	13
2.1.4 Containerized and Orchestrated Deployment	13
2.1.5 Deployment Location and Infrastructure Options	13
2.1.6 Use of National e-Infrastructure Services	13
2.2 Standards and Protocols.....	14
2.2.1 Remote Sensing Data (Satellite and Aerial).....	15
2.2.2 In-Situ Monitoring and Terrestrial Platforms	15
2.2.3 Meteorological Data.....	15
2.2.4 Topographical and Geospatial Data	15
2.2.5 Laboratory Analyses	16
2.2.6 Citizen-Generated Data.....	16
2.2.7 Historical and Archival Data	16
2.2.8 Standards and Harmonization.....	16
2.3 Data Security and Privacy Considerations	16
2.3.1 Security of Data Sources and Flows	16
2.3.2 Identity and Access Management	17
2.3.3 Privacy Considerations	17
2.3.4 Security and Compliance Measures	17
2.3.5 Data Ownership and Governance: Balancing Openness and Protection	17
3. Implementation Timeline and Activities.....	18
3.1 Planning (Months 1–5)	18
3.2 Implementation and Procurement (Months 5–20).....	18
3.3 Testing and Adaptation (Months 5–20)	19
3.4 Expansion and Consolidation (Months 13–36)	19
4.1 Transnational Mission	19
4.1.1 Ensuring Compatibility	20
4.1.2 Guaranteeing Future Proofness	20



4.1.3 Enabling Replicability and Baltic-Wide Applicability	20
4.2 Ensuring Interoperability	20
4.2.1 Designing for Future-Proofing	20
4.2.2 Replication in Diverse Municipal Contexts	21
Conclusions	22

Table of Acronyms

Acronym	Full Term	Description
API	Application Programming Interfaces	A set of protocols that enable different software components to communicate and transfer data
AWS	Amazon Web Services	A cloud computing platform by Amazon.
CCO	Creative Commons Zero	A public domain dedication tool
CCTV	Closed Circuit Television	A video surveillance system
CI/CD pipelines	Continuous Integration/Continuous Delivery	A set of automated processes that allow software development teams to build, test, deliver and deploy.
CKAN	Comprehensive Knowledge Archive Network	Open-source data management system to publish, share and manage data through data portals and hubs.
CSV	Comma-Separated Values	Simple file format used to store tabular data such as spreadsheets or databases.
D1.3/ D3.1/ D3.5	Deliverable Numbers	Internal project numbering for deliverables
DMZ	Demilitarised Zone	A physical or logical subnetwork that separates an organization's internal local area network (LAN) from untrusted external networks such as the internet.
DPIA	Data Protection Impact Assessment	A systematic process used to identify, assess and mitigate the privacy risks associated with the collection, use and processing of personal data.
EM-DAT	Emergency Disaster Database	A repository of information on mass disasters that happened since 1900.
ETSI	European Telecommunications Standardization Institute)	Independent, not-for-profit, standardization organization operating in the field of information and communications
EU/EEA	European Union/European Economic Area	Political and Economic union



FIWARE	Future Internet WARE	Framework of Open-source platform components used for the development of smart solutions.
GCP	Google Cloud Platform	A suite of cloud computing services offered by Google.
GDPR	General Data Protection Regulation	A comprehensive data privacy and security law implemented by European Union.
GE	Generic Enabler	A reusable, open-source software component in the FIWARE architecture that provides standard functionality and can be combined with other components to build smart solution platforms
GeoJSON	Geographic JavaScript Object Notation	An open standard format designed for representing simple geographical features, along with their non-spatial attributes.
GeoTIFF	Geographic Tagged Image File Format	A public domain metadata standard which allows georeferencing information to be embedded within a TIFF file
GIS	Geographic Information System	A framework for gathering, Managing, analysing and visualizing spatial and geographic data.
HTTPS	Hyper Text Transfer Protocol Secure	An internet communication protocol that ensures secure data transfer between user's browser and a website.
IoT	Internet of Things	Devices with embedded sensors, processing ability, software and other technologies that allows to connect and exchange data with other devices and systems over the internet or other communications networks
IP	Internet Protocol	It is set of rules that governs how data is sent and received over the internet and other networks.
ISO	International Organization for Standardization.	An independent organization that develops and publishes standards for various industries and services to ensure quality, safety, reliability and efficiency.
LoRa	Long Range	Physical proprietary radio communication technique based on spread spectrum modulation techniques derived from chirp spread spectrum (CSS) technology.
LoRaWAN	Long Range Wide Area Network	Protocol designed to wirelessly connect battery operated devices to the internet in regional, national or global networks.
Modbus-TCP	Modbus Transmission Control Protocol	Version of the Modbus protocol which allows devices to be interconnected directly through Ethernet without any special



		hardware or communication interface.
MQTT	Message Queuing Telemetry Transport	Lightweight messaging protocol for use in cases where clients need a small code footprint and are connected to unreliable networks or networks with limited bandwidth resources.
NGSI-LD	Next Generation Service Interfaces – Linked Data	Information model and API for publishing, querying and subscribing to context information.
NIRD	National Infrastructure for Research Data	Norwegian e-infrastructure dedicated to storing, processing, and archiving scientific research data securely and at scale.
NOAH	Nationwide Operational Assessment of Hazards	Comprehensive platform designed to improve disaster risk reduction and management in the Philippines.
NVE	Norwegian Water Resources and Energy Directorate	Norway's national authority responsible for managing water and energy resources, regulating hydropower, reducing flood and landslide risks, and ensuring a stable and sustainable power supply.
OKD	Origin Community Distribution of Kubernetes	The open-source community distribution of Kubernetes that forms the basis of Red Hat OpenShift, providing a platform to build, deploy, and manage containerised applications.
OPC-UA	Open Platform Communications – Unified Architecture	An industrial communication standard for secure, reliable, and platform-independent data exchange between devices and systems.
OT	Operational Technology	Hardware and software systems used to monitor and control physical devices, processes and infrastructure.
RBAC	Role-based access control	A security model that restricts system access based on a user's role within an organization.
REST API	Representational State Transfer Application Programming Interface	A set of rules and conventions for building and interacting with web services.
SCADA	Supervisory Control and Data Acquisition	An architecture that enables industrial organizations to manage, monitor, and control processes, machines, and plants.
SDGs	Sustainable Development Goals	17 universal goals adopted by all United Nations Member, forming the 2030 Agenda for Sustainable Development.
SSEW	Small Scale Exploratory Work	NIRD service that offers Norwegian researchers limited access to national



		supercomputers and storage resources for a specific period.
SSO	Single Sign-on	An authentication process that allows users to access multiple applications or systems with one set of login credentials.
SYKE	Finnish Environment Institute	Governmental research and expert institute operating under the Ministry of the Environment of Finland.
TLS	Transport Layer Security	Encrypts data sent over the internet to ensure that hackers are unable to see what is transmitted for privacy and secure sensitive information.
URL	Uniform Resource Locator	The address of a web page.
VPN	Virtual Private Network	An encrypted secure tunnel for internet connection that covers real IP address and location to enhance privacy and security online.
WP2/WP3	Work Package	Internal project structure naming



List of Figures

Figure/Image no.	Title	Page
Image 1	Implementation timeline and activities	15



Introduction

The data systems integration plan described in this document is designed to provide a future-proof, standards-based architecture for incorporating flood and storm water data from various sources. The integration plan is built on existing systems used by the two pilot municipalities – Lappeenranta and Gjøvik, but it also provides an alternative for organizations and municipalities that do not currently have a system for combining and presenting environmental data.

Municipalities face similar integration challenges, such as limited interoperability and connectivity, which BALTFLOODS aims to address. To facilitate rapid prototyping, interoperability, and replication, the integration architecture adopts an open-source FIWARE system as its backbone.

Municipalities face similar integration challenges, such as limited interoperability and connectivity, which BALTFLOODS aims to address.

The data system integration plan outlines the key components related to data flow used in the pilots by City of Lappeenranta and Gjøvik municipality along with a brief description of their existing data system. It provides a shared blueprint for integrating flood and storm water quality data during the project timeline between March 2025 until February 2028. It is intended to support real time decision making to improve flood preparedness and reduce polluted runoff in Baltic Sea cities. This document offers guidance for procurement, installation, testing and future expansion of monitoring solutions for cities experiencing similar issues related to flood.

The integration plan also consists of timeline with activities that also included windows for adjusting the solutions underway,

allowing the project team to adapt to changing circumstances and incorporate feedback from stakeholders.

D1.2 is closely linked to D1.4 (Monitoring and Evaluation Approach). Whereas D1.2 outlines the technical environment for data flows and integration, D1.4 specifies how the functioning of this environment will be monitored, how indicators will be collected, and how results will be evaluated during the pilots.



1. Overview of Existing Data Systems

The BALTFLOODS project builds on top of the existing digital and monitoring infrastructures of its two pilot municipalities—**Lappeenranta** in Finland and **Gjøvik** in Norway—while introducing new tools and integration layers to address gaps in interoperability, real-time analytics, and participatory engagement. These existing systems form both the foundation and the integration challenge for the project. They include municipal SCADA and telemetry platforms, national environmental open-data portals, and citizen-science applications from prior projects.

In addition to local systems, the integration plan will also interface with **national environmental data services**, such as the Finnish Environment Institute (SYKE) and the Norwegian Water Resources and Energy Directorate (NVE), to incorporate authoritative hydrological and meteorological datasets. Connections to existing citizen-science platforms from earlier Baltic or European initiatives, such as the **NOAH flood app** or the **EyeonWater Colour app**, are also under consideration as potential sources of crowdsourced flood data or as dissemination channels.

1.1 Lappeenranta

Lappeenranta's primary operational platform is **StreetAI**, a comprehensive urban data management and analysis environment developed through an innovation partnership in 2021–2022. StreetAI provides a **real-time, map-based situational picture** of the city by aggregating data from a wide array of sensors and external data feeds. Its scope extends beyond environmental monitoring to include traffic flows, parking availability, air quality, road surface conditions, and public space usage.

For BALTFLOODS, the most relevant StreetAI capabilities include:

- **Stormwater quality and quantity measurements**, feeding into flood risk assessment.
- **Roadbed moisture monitoring**, which helps assess surface water accumulation.
- **Hyperlocal weather data**, integrating municipal sensor readings with external meteorological feeds.

StreetAI ingests data primarily through **real-time APIs**, with some batch uploads for datasets not yet available via live feeds. This architecture allows the city to respond quickly to environmental events, such as heavy rainfall, by updating situational dashboards in near real-time.

Complementing StreetAI is **CitySen.App**, a progressive web application that serves as Lappeenranta's citizen engagement and citizen-science interface. Accessible via any modern browser—without downloads or installation—CitySen.App allows residents to **view live environmental data** and **submit observations**. Citizens can report water quality parameters such as temperature, clarity, pH, turbidity, presence of blue-green algae, or nitrate levels, and these submissions feed directly into the StreetAI database.

CitySen.App is openly licensed under **Creative Commons Zero (public domain)**, allowing other municipalities, including those in BALTFLOODS, to adapt and deploy it without licensing restrictions. This open approach lowers barriers for adoption and provides a tested vehicle for citizen participation in environmental monitoring.

1.2 Gjøvik

Gjøvik Municipality operates a municipal **SCADA/telemetry system** for managing its sewer and stormwater networks, alongside an integrated IoT data hub that connects multiple sensor vendors and data streams.

The monitoring setup combines:

- **Vendor-supplied sensors** (e.g. Pipelife) for water level, flow, and quality measurements. These sensors are part of a managed service that includes communications (4G, 5G, LoRa) and a basic web dashboard. While the vendor dashboard is functional, it is not part of daily municipal operations and is used mainly for troubleshooting or vendor liaison.
- **Municipally owned sensors**, connected directly to the **IoT Hub API** without passing through vendor-managed systems.
- **Data integration into two main municipal systems:**
 - **SCADA/Operations** (GE iFix or Gemini VA) for real-time operations, alerts, and process control.
 - **Data lake / data warehouse** (Gurusoft Report) for long-term storage, time-series analysis, and open data sharing via APIs.

In this architecture, the IoT Hub functions as a **central intake layer** for all sensor data, regardless of origin. It enables harmonisation and routing to both operational (SCADA) and analytical (data lake) systems, providing Gjøvik with a clear separation between real-time operational monitoring and long-term data management.

1.3 Common Challenges Across Both Cities

Despite the advancement of these municipal systems, both Lappeenranta and Gjøvik face **similar integration challenges** that BALTFLOODS aims to address:

- **Heterogeneous data protocols and interfaces:** From industrial IoT standards (Modbus-TCP, OPC-UA, LoRaWAN) to diverse web APIs, integrating new data sources often requires bespoke adapters, slowing down deployment of innovative solutions.
- **Variable latency requirements:** Flood warnings may require sub-minute updates, while water quality metrics or spatial datasets can be aggregated on hourly or daily intervals. Current systems are not uniformly optimised for both high-frequency and low-frequency data flows.
- **Limited interoperability with external platforms:** While national datasets (e.g., SYKE, NVE) are valuable, the current municipal systems do not have automated, standardised pipelines to ingest or merge them with local data.
- **Regulatory compliance and data sharing constraints:** Data governance requirements differ by jurisdiction, affecting how, when, and with whom environmental data can be shared.



- **Connectivity limitations:** Some sensor locations—particularly in rural catchments—have intermittent network coverage, impacting the timeliness and completeness of real-time data feeds.

These constraints underscore the need for a **robust but adaptable integration environment**—one that can ingest diverse data streams, harmonise them into common formats, and route them efficiently to both operational and analytical systems, while respecting local regulations and technical realities.

2. The BALTFLOODS Integration Strategy

The BALTFLOODS integration strategy is designed to provide a future-proof, standards-based environment for ingesting, harmonising, and disseminating environmental and hydrological data from multiple, heterogeneous sources. It reflects the principles defined in Deliverable D1.1—particularly those in Pillar 3: Data Protocol and Integration Framework—but develops them into an operational plan with a clear technical structure.

The proposed approach enables the seamless integration of:

- **Environmental monitoring data** derived from multi-scale remote sensing platforms, including satellite and aerial systems.
- **Municipal SCADA telemetry** from sewer and stormwater infrastructure (e.g. flow, level, and quality sensors)
- **National open data portals**, such as the Finnish Environment Institute (SYKE) and the Norwegian Water Resources and Energy Directorate (NVE)
- **Citizen-science inputs** from mobile applications, participatory mapping platforms, and structured community reporting

The system must support **real-time data ingestion** where feasible, as well as **batch processing** for periodic datasets. It must also enable interoperability across different municipalities and countries, ensuring that the resulting solutions are replicable across the Baltic Sea Region.

2.1 FIWARE as the Integration Backbone

To facilitate **rapid prototyping**, **interoperability**, and **replication**, the integration architecture adopts the **FIWARE system** as its backbone. FIWARE is an open-source, modular platform widely used for smart city and environmental applications. Its use in BALTFLOODS delivers several strategic advantages:

- **Interoperability through open standards:** Core components use the NGSI-LD standard for context information management, ensuring compatibility across diverse data sources and systems.
- **Modular deployment:** Only the components required for each pilot are installed, reducing complexity and enabling progressive scaling.
- **Rapid integration:** Prebuilt Generic Enablers (GEs) provide immediate capabilities for IoT data ingestion, time-series storage, API management, and identity access management—avoiding the need to develop these from scratch.
- **Replicability:** Configurations and data models can be exported and deployed in new municipalities with minimal customisation.

2.1.1 Simplified Description of the FIWARE Architecture and Data Flow

At the heart of the BALTFLOODS architecture is the principle that all relevant environmental, hydrological, and citizen-generated data should be treated as part of a single, interconnected context. This context is maintained and distributed through the **Orion-LD Context Broker**, a FIWARE Generic Enabler (GE) that acts as the core integration hub of the architecture. The Context Broker continuously ingests updates from multiple, heterogeneous sources, stores their relationships, and makes them available to any authorised system or user that needs them, in real time or on demand.

Data Ingestion Layer

The journey of data into FIWARE begins at the ingestion layer. Here, several different mechanisms operate in parallel:

- **IoT Agents** serve as the translators between physical devices and the NGSI-LD data format used by the Context Broker. In the case of municipal sensor networks for stormwater and sewer monitoring, IoT Agents are configured to understand industrial protocols such as Modbus, OPC-UA, or MQTT. They receive telemetry—such as flow rates, water levels, turbidity readings—and repackage them into structured NGSI-LD entities.
- **API Connectors** link the system to external open data sources. For example, APIs from SYKE and NVE provide hydrological and meteorological measurements, flood warnings, and water quality indicators. These connectors standardise the data so that it can be stored and processed alongside sensor feeds without loss of meaning or precision.
- **Citizen Engagement Interfaces**, including the BALTFLOODS Citizen Engagement App and other participatory monitoring tools, allow residents, community groups, and interest organisations to submit structured observations. Whether this is a photo of a flooded street, a measurement from a turbidity strip, or a location report of a blocked drainage inlet, these submissions are transformed into NGSI-LD entities that have the same standing in the system as those from sensors or satellites.

Data Harmonisation and Storage

Once ingested, data flows into the context management layer, where the Orion-LD Context Broker maintains an always-current representation of the system's state. This representation is semantically rich, meaning that it describes not just raw values but also the relationships between them—for example, linking a turbidity reading to the specific sensor, location, and time it was taken.

For historical analysis and trend detection, the architecture employs **QuantumLeap**, a time-series database component tightly integrated with Orion-LD. QuantumLeap stores historical versions of all context entities, making it possible to query and analyse past events—an essential feature for flood modelling, system performance assessment, and impact evaluation.

Data Access and Use

The processed and harmonised data is then made available to various consumers through standard APIs and subscription mechanisms.

Municipal dashboards can subscribe to receive updates whenever certain thresholds are exceeded—for example, when a flood model predicts



inundation above a critical level or when water quality indicators suggest possible contamination. Researchers can query historical data for in-depth analysis, while public-facing visualisation tools can present selected information to citizens in a clear and accessible way.

Geospatial services can also be integrated, allowing maps to be dynamically updated with overlays showing flood extents, pollutant plumes, or citizen-reported issues. In WP2, these capabilities will underpin both the pilot testing of tools and the participatory evaluation of their usefulness.

Scalable and Modular Design

One of the strengths of this architecture is its **incremental deployability**. Municipalities with mature digital infrastructure can integrate only the components they need to bridge data silos and test new solutions, while municipalities without such systems can deploy the architecture as a complete, standalone platform. Because all components communicate using open standards, new data sources or analytical tools can be added with minimal reconfiguration.

This modularity is crucial for BALTFLOODS' transnational ambitions: it allows a core integration approach to be replicated across different urban contexts, with local adaptations in data sources, tools, and governance.

2.1.2 Implementation of the FIWARE-Based Data Integration Environment

This section describes how the FIWARE-based data integration environment will be **deployed and operated** within BALTFLOODS, moving from the conceptual principles set out in D1.1 to a practical, technical configuration suitable for the WP2 pilots and beyond. The focus here is on the steps, technologies, and infrastructure options that will enable early testing, reliable operation, and future replication of the solution in different municipal contexts.

The implementation approach has been designed to accommodate the diversity of partner infrastructures—ranging from municipalities with mature digital systems to those with minimal existing integration capabilities—while ensuring that the setup remains standards-based, modular, and easy to hand over after the project.

2.1.3 Early Deployment and Mock Data Integration

One of the strengths of the FIWARE approach is that it can be deployed **early in the project timeline**, even before all sensors and external data sources are operational. For example, where certain flood monitoring sensors or water quality probes have yet to be installed, **mock data streams** can be created to emulate their expected output. This allows development and testing of analytical tools, dashboards, and alert systems to begin in parallel with hardware procurement and installation.

By validating data ingestion pipelines and entity models in advance, the project significantly reduces integration risks and avoids costly delays once real data becomes available.

2.1.4 Containerized and Orchestrated Deployment

All FIWARE components—such as the Orion-LD Context Broker, IoT Agents, QuantumLeap, CKAN, and the identity management services—will be deployed as **Docker containers**. Each container will encapsulate a single FIWARE Generic Enabler or supporting service, ensuring modularity, ease of maintenance, and the ability to scale specific components independently.

These containers will be orchestrated using **OpenShift**, Red Hat's enterprise-grade platform built on Kubernetes. OpenShift offers several advantages over a raw Kubernetes installation:

- **Integrated developer workflows** (source-to-image builds, CI/CD pipelines) to shorten the path from concept to deployment.
- **Enhanced security defaults**, including enforced non-root execution for containers and built-in role-based access control (RBAC).
- **Operational tooling** for logging, monitoring, and automated scaling that reduces system administration overhead.
- **Portability and repeatability**, allowing the same configuration to run in the cloud or on-premises with minimal adaptation.

While OpenShift is a commercial product, it also exists in **free and open-source community editions** (e.g., OKD) that can be deployed at no licensing cost. For the BALTFLOODS pilots, the choice between Red Hat OpenShift (with support contracts) and OKD (community-maintained) will depend on available infrastructure, partner preferences, and operational support requirements. Both options are compatible with the proposed FIWARE deployment model.

2.1.5 Deployment Location and Infrastructure Options

The FIWARE-based environment is flexible in where it can run:

- **On-premises**, within municipal or partner data centres, ensuring full sovereignty over operational and sensitive datasets.
- **Public cloud**, using managed OpenShift/Kubernetes services (AWS, Azure, GCP) for rapid prototyping and elastic scaling.
- **Hybrid**, with sensitive real-time telemetry processed locally, while cloud-hosted analytics and storage handle less sensitive data or larger workloads.

For the WP2 pilots, the most practical approach is to operate a **shared pilot cluster** hosted either in a trusted partner's infrastructure or on a public cloud platform within the EU/EEA. Municipal data sources can connect securely to this cluster via lightweight adapters or IoT Agents installed close to their SCADA systems. This arrangement enables **rapid prototyping and joint testing** without interfering with municipal production environments.

Once the pilots conclude, each partner can choose to:

- Deploy their own standalone FIWARE environment on local infrastructure.
- Continue to connect to the shared environment for certain services.
- Extract and integrate selected solutions directly into their existing municipal platforms, bypassing FIWARE where appropriate.

2.1.6 Use of National e-Infrastructure Services

For partners in Norway—particularly Gjøvik—there is the option to leverage the **Norwegian National e-Infrastructure services** provided by Sigma2. These services offer a cost-effective, secure, and high-performance hosting environment for research and innovation projects. Two services are particularly relevant to BALTFLOODS:



- **Small Scale Exploratory Work (SSEW):** Ideal for initial proof-of-concept deployments, rapid testing of integration pipelines, and running small-scale analytics workloads before full-scale infrastructure is committed.
- **NIRD Service Platform:** Provides robust cloud-based compute and storage resources within Norway’s research infrastructure, suitable for hosting long-running services, large datasets, and more computationally intensive analytics tasks.

By using these national e-infrastructure options, BALTFLOODS can ensure compliance with data residency requirements, reduce hosting costs, and provide a path for Norwegian municipalities to maintain the solution within a trusted, sovereign environment.

2.2 Standards and Protocols

The successful integration of diverse data sources in BALTFLOODS relies on the adoption of **open standards, interoperable protocols, and harmonised data models**. This section describes the concrete technical foundations that will enable pilot sites to share, process, and reuse environmental and hydrological data in a sustainable and future-proof manner.

At the core of the BALTFLOODS integration environment is the **NGSI-LD standard**, developed by ETSI, which provides a common API and semantic model for exchanging context information. NGSI-LD ensures that data from sensors, legacy municipal systems, satellite platforms, and citizen observations can be represented as structured entities with attributes (e.g. location, timestamp, measured value) and relationships (e.g. linking rainfall intensity to a flood event). By applying this standard, heterogeneous sources are brought into a shared “data language,” avoiding one-off connectors and reducing the cost of integration.

Complementing NGSI-LD is the adoption of the **FIWARE Smart Data Models**, a library of domain-specific schemas co-developed with international standardisation bodies and EU-funded projects. Within BALTFLOODS, models such as `WaterQualityObserved`, `FloodingEvent`, and `IncidentReport` will be applied or adapted to ensure semantic consistency across the pilots. Where no suitable model exists, extensions will be developed in line with FIWARE and OGC best practices, keeping them open for reuse by other municipalities.

For **data ingestion and exchange**, the integration plan supports widely used protocols including MQTT and **HTTP/REST APIs**, ensuring lightweight, scalable communication between field devices and the integration layer. Where feasible, sensors will be configured to comply with the **OGC SensorThings API**, providing consistent metadata and facilitating alignment with international Earth observation practices. Satellite data from Sentinel missions will be accessed through Copernicus APIs, while UAV data and citizen-generated inputs will be ingested via secure HTTPS connections to API gateways.

Alongside real-time data integration, the framework also establishes dedicated storage layers to support different types of information:

Time-series data: A time-series database (like CrateDB or Timescale, possibly integrated through FIWARE’s components) will store historical sensor readings, model outputs over time, etc. This allows querying, for instance, “what was the water level at sensor X for the past year” or plotting trends.

Spatial data: A GIS database (PostGIS enabled) will store geospatial datasets (maps of flood zones, infrastructure, land use). This ties into modelling tools and also allows spatial queries (like find all sensors within a flood zone).

Files and Big Data: For larger items like satellite images, the framework will use an object storage or file server where those can be stored and accessed by reference (with metadata in the context broker linking to the file). For example, after processing a Sentinel-2 image, the resulting raster of water quality indices might be saved in cloud storage with a URL, and the context entity would contain that URL and summary stats.

A specific focus is placed on citizen-generated data, which forms an important complement to structured sensor and Earth observation streams. This includes both citizen data (e.g. flood photographs, incident reports, simple measurements submitted via the Citizen Engagement App) and citizen science data collected through organised campaigns (e.g. schools using test kits to measure water temperature, turbidity, pH, or nutrients). These contributions will be stored in a structured way: textual and numeric observations will be represented as context entities (e.g. `IncidentReport`, `WaterQualityObserved`), while images and other media will be stored in object storage, linked to metadata in the integration layer. To maintain data quality and trustworthiness, validation protocols will be applied, ranging from automated checks (e.g. sensor auto-validation) to manual moderation and photographic verification. All citizen-generated inputs will be handled under strict privacy and data protection protocols. Personal identifiers will be minimised, images anonymised or blurred where necessary, and contributions stored in compliance with GDPR.

Furthermore, both pilot municipalities have some existing data flows (e.g., Gjøvik’s water utility SCADA, Lappeenranta’s StreetAI). The framework plans for connectors or adaptors to integrate such systems. For example, a CSV export from an existing database can be parsed by a small script and pushed into the context broker regularly. Or if they have an IoT hub already, we might bridge it with our system by subscribing to its data and converting format. Because we prioritize standard APIs, connecting to our platform is also easier: any system that can call a REST API can essentially push or pull data from the FIWARE broker. Furthermore, the protocol ensures that after project conclusion, the data integration continues with minimal fuss. By implementing maintainable interfaces (like documented APIs and not overly custom code), the cities’ IT staff can manage it. We deliberately avoid proprietary protocols that would lock the system – everything is open or widely used. This aligns with the objective to enable maintainability after the project ends and adapt to existing workflows in the municipalities. If a city wants to swap a sensor vendor or incorporate a new data type (say water temperature sensors in the future), they can do so by adhering to the defined API/data model without rebuilding the whole system.

A data management plan –included in preliminary version in D1.1 as Annex B— will outline responsibilities (who can access what data), update frequencies, and data lifecycle (how long data is kept, archival). The integration framework includes aspects of data governance – e.g., making sure sensitive data (like any personal info from app users) is segregated and access-controlled. The FIWARE context broker will likely hold mostly non-sensitive environmental data that can be open, whereas any personal user data is kept in a secure user database not openly accessible.

The standards and protocols described here not only enable technical interoperability but also define the basis on which performance indicators can later be measured. For example, ensuring consistent timestamping, geolocation, and metadata standards is essential for the monitoring indicators pre-identified in D1.4, which depend on comparable and harmonised data streams.

By grounding the BALTFLOODS system in open standards, modular databases, and clear governance rules, the project ensures not only interoperability between Lappeenranta and Gjøvik, but also replicability for other cities across the Baltic Sea Region. This standards-based



foundation is what allows the integration environment to remain flexible, scalable, and future-proof, capable of evolving as new data sources and technologies emerge.

BALTFLOODS integrates a broad spectrum of data types to capture the complexity of stormwater management, flood preparedness, and runoff pollution monitoring. These data sources span from spaceborne observations to ground-level sensors, laboratory analyses, citizen contributions, and historical records. Mapping them early is essential to ensure access agreements, technical compatibility, and harmonisation of formats. The following subsections describe the main categories, their sources, and their roles within the project's integration framework.

2.2.1 Remote Sensing Data (Satellite and Aerial)

Earth observation provides a synoptic view of flood dynamics and water quality indicators across the Baltic Sea Region. Data from the **Copernicus Sentinel constellation** is central:

- **Sentinel-1 radar (SAR)** delivers flood inundation mapping regardless of cloud cover, crucial during storm events when optical sensors are obstructed.
- **Sentinel-2 multispectral** imagery provides proxies for water quality (e.g., turbidity, chlorophyll-a) and land cover information to assess runoff potential.
- **Sentinel-3** can contribute broader water monitoring (temperature, colour indices).
- **Sentinel-5P** provides atmospheric data (e.g., aerosols, NO₂) relevant for rainfall and air–water quality interactions.

These data are accessed via Copernicus Open Access Hubs (free and open), typically provided in **GeoTIFF raster format** with standard metadata.

To complement satellites, **UAV (drone) surveys** equipped with RGB, multispectral, and thermal cameras will capture **centimetre-level detail** in targeted hotspots: retention ponds, urban drainage zones, or river stretches prone to overflow. UAV data will be particularly valuable where Sentinel-2 is limited by cloud cover or where SAR is hampered by urban occlusion. UAV orthomosaics and 3D point clouds (produced via photogrammetry or LiDAR where available) will provide fine-scale validation of remote sensing products.

Municipal **aerial orthophotos** (updated annually in many cities for planning) may also be integrated, providing historical context and consistent basemaps.

2.2.2 In-Situ Monitoring and Terrestrial Platforms

Ground-based monitoring provides continuous, high-frequency observations of stormwater and water quality dynamics. BALTFLOODS combines **newly installed sensors** with **existing municipal and national monitoring infrastructure**, ensuring both local innovation and alignment with established systems.

- **Hydrological sensors:** water level loggers in rivers and lakes; flow meters in pipes and open channels; rain gauges measuring precipitation intensity.
- **Water quality sensors:** probes for turbidity, conductivity, temperature, pH, dissolved oxygen, nitrates, phosphates, deployed at stormwater outlets, retention ponds, and receiving waters.
- **Overflow detection devices:** including the **V-overflow modules** in combined sewer systems, detecting overflow events and estimating volumes.
- **Multi-sensor manhole units:** measuring runoff water quality and quantity simultaneously, delivering real-time data streams from within drainage networks.
- **CCTV/smart cameras:** visual evidence of flooding and water levels, producing video/image data streams.

These sensors will stream data in real time, typically via **4G/5G or LoRaWAN** managed services. Vendor dashboards (e.g., Pipelife) may serve as intermediate platforms, but integration is achieved through municipal **IoT hubs** (as in Gjøvik) or citywide systems such as **StreetAI** (in Lappeenranta). From there, data is standardised through adapters into **NGSI-LD JSON for streaming** or **CSV for bulk export**.

Existing **terrestrial monitoring platforms** from national institutes (e.g., SYKE gauging stations in Finland, NVE hydrological stations in Norway) will complement municipal deployments, ensuring consistency with national datasets and regulatory standards.

2.2.3 Meteorological Data

Meteorological inputs are indispensable for both short-term flood forecasting and long-term climate resilience assessment. BALTFLOODS will rely on:

- **Real-time weather station data** (precipitation, wind, temperature, humidity).
- **High-resolution weather forecasts** (ultra-local where available).
- **Air quality data** (e.g., particulate matter, ozone) relevant for stormwater interactions.

These are sourced from **national meteorological services** and accessed via API (with API keys where needed). Data formats include **NetCDF** and **CSV/JSON** outputs, which will be parsed into harmonised context entities for integration.

2.2.4 Topographical and Geospatial Data

Geospatial reference data provides the **spatial backbone** of the BALTFLOODS system, contextualising all dynamic measurements. This includes:

- **Digital Elevation Models (DEMs)** for catchment delineation and flood modelling.
- **Land cover and land use maps**, essential for runoff estimation and pollution modelling.
- **Drainage network maps** of sewers, culverts, and open channels, to connect overflow events with downstream impacts.
- **Infrastructure and cadastral layers** (roads, parcels, buildings) to identify vulnerable areas and exposure.
- **Zoning and urban planning data**, integrating stormwater requirements into city development.



These layers, typically available through municipal GIS services or national mapping agencies, are delivered as **GeoTIFF rasters** (e.g., DEMs, orthophotos) and **GeoJSON/shapefiles** (vector data). They are foundational for decision-support tools: hydrological/hydraulic modelling, flood susceptibility mapping, and risk communication dashboards.

2.2.5 Laboratory Analyses

Where sensors cannot capture certain pollutants or require validation, BALTFLOODS employs laboratory-based water analyses. Samples collected during rainfall and overflow events will be analysed for:

- Nutrients (nitrogen, phosphorus).
- Heavy metals.
- Microbiological parameters (e.g., *E. coli*).
- Emerging pollutants such as microplastics.

These results provide **ground-truth validation** of in-situ and remote sensing data and extend the monitoring scope to parameters not measurable in real time. Laboratory outputs are stored in structured databases with metadata linking them to sampling events, locations, and chain-of-custody records.

2.2.6 Citizen-Generated Data

Citizen science and stakeholder engagement add an important human dimension to the data ecosystem. Through the **Citizen Engagement App** and school/community networks, BALTFLOODS will collect:

- Flood incident reports (with structured fields: who, what, where, when).
- Geotagged photos and videos of floodwaters or pollution.
- Simple manual measurements (e.g., street water depth with rulers).
- Citizen science results from distributed test kits (pH, turbidity, nutrients).
- Survey responses and lived experiences of flood impacts.

Contributions are submitted via **secure HTTPS APIs**, and processed into structured formats (e.g., NGS-LD entities like `IncidentReport` or `ImageObservation`). Media files are stored in object storage, with metadata linking them to structured context entities. Privacy and GDPR compliance are ensured through anonymisation and clear consent procedures.

2.2.7 Historical and Archival Data

Past events provide critical context for present and future resilience. BALTFLOODS will incorporate:

- Municipal records of flood incidents and infrastructure damage.
- Emergency services reports and response logs.
- National flood and hydrology databases (e.g., SYKE in Finland, NVE in Norway).
- International sources such as EM-DAT for major disaster events.

These datasets will support baseline assessments, calibration of models, and retrospective evaluation of project impact.

2.2.8 Standards and Harmonization

To manage this diversity, BALTFLOODS applies a **standards-based integration approach**:

- **Remote sensing and geospatial data:** GeoTIFF (rasters), GeoJSON/shapefiles (vectors).
- **Time-series data:** NGS-LD JSON for real-time streaming, CSV for bulk archival.
- **Qualitative data:** Structured fields (who, what, where, when) for machine readability.
- **Citizen contributions:** Metadata-driven storage with anonymisation protocols.

By aligning each data type with well-documented standards, BALTFLOODS ensures that data flows from satellites, sensors, laboratories, municipal systems, and citizens can be harmonised within one coherent integration framework.

2.3 Data Security and Privacy Considerations

While the BALTFLOODS data integration environment is being designed to maximise openness and reusability of environmental data, the project must also respect **technical, regulatory, and operational constraints** around sensitive information. Many municipal systems—particularly SCADA and other operational technology (OT) platforms—contain data that cannot be publicly released due to security, safety, or privacy considerations. Therefore, the integration approach is governed by the principle “**as open as possible, as closed as necessary**”, ensuring that the system enables broad data sharing without exposing sensitive assets or compromising operational integrity.

2.3.1 Security of Data Sources and Flows

Municipal SCADA/OT integration will follow a strict security boundary. Operational systems will never be exposed directly to the internet or to the FIWARE integration layer. Instead, an **edge gateway**—preferably in a **demilitarised zone (DMZ)**—will provide read-only access where possible, translating industrial protocols such as **Modbus, OPC-UA, or MQTT** into **NGS-LD** entities. This gateway will forward data securely to the platform via **mutual TLS connections over VPN**, ensuring both encryption and authentication at the transport layer.

National open data portals (e.g., SYKE, NVE) will be accessed over the public internet using secure API endpoints. API keys or IP whitelisting will be applied where supported, and responses will be cached to minimise repeated external calls, improving resilience and reducing dependency on external service availability.

Citizen-science contributions—including photos, free-text reports, and sensor readings from the engagement application—will be transmitted over **HTTPS** to an **API gateway**. Here, **rate-limiting** and **moderation pipelines** will be applied before data is persisted to FIWARE components



such as **Orion-LD** or **QuantumLeap**. Moderation may include automated validation (e.g., sensor plausibility checks) and manual review, particularly for unstructured inputs like photographs or text.

2.3.2 Identity and Access Management

The system will employ a **centralised identity and access management (IAM)** solution, using FIWARE's **Keyrock** or an equivalent (e.g., Keycloak) to provide **single sign-on (SSO)** and **role-based access control (RBAC)**. Different roles will be defined for municipal operators, research partners, external stakeholders, and public users. Access to sensitive datasets will be strictly limited to authorised personnel, while open datasets and aggregated results will be available via public endpoints.

2.3.3 Privacy Considerations

Certain data streams—especially those from citizen engagement tools—may contain **personal or potentially identifiable information**. Examples include photographs, descriptive location details, or metadata embedded in user submissions. To address this:

- Users will be **clearly informed** about how their data is stored and used, including any anonymisation or aggregation applied.
- Data retention policies will be enforced, with the ability to **delete unrelated or inappropriate data** promptly via an administrator dashboard.
- Where feasible, **derived datasets** (e.g., aggregated pollution indicators) will be published as open data, even when raw data cannot be released.

All citizen-contributed data flows will be subject to a **Data Protection Impact Assessment (DPIA)** to ensure compliance with the **General Data Protection Regulation (GDPR)**, particularly around data minimisation, purpose limitation, and user consent.

2.3.4 Security and Compliance Measures

The BALTFLLOODS integration environment will adopt a **zero-trust security model** from the outset, ensuring that every interaction between components, whether internal or external, is explicitly authenticated and authorised. All communications between services will be encrypted using **mutual TLS**, preventing unauthorised interception or tampering with data in transit. Within the containerised deployment, **network policies** will control the flow of traffic between services, ensuring that only those components with a defined operational relationship can communicate.

Sensitive credentials and configuration parameters—such as API keys, encryption keys, and database passwords—will not be stored in plain configuration files but securely managed in a **dedicated secrets vault**, with access strictly controlled and audited. **Audit logging** will be an integral feature of the environment, capturing all data access and modification events for later review, ensuring accountability and traceability in line with regulatory requirements.

Operational resilience is also central to the compliance strategy. Regular **backups** will be performed to protect against data loss, and **restore procedures will be tested** to confirm that recovery objectives can be met. Critical dashboards—such as those displaying live flood alerts or key infrastructure status—will have a **minimal offline mode**, allowing core situational awareness to be maintained even during temporary network outages or service disruptions.

By combining these measures—proactive encryption, strict network segmentation, secure secret management, continuous auditing, and tested continuity planning—the BALTFLLOODS platform will meet both the security expectations of municipal partners and the compliance obligations of operating within the European regulatory environment.

2.3.5 Data Ownership and Governance: Balancing Openness and Protection

Clear rules on data ownership are essential to ensure trust, transparency, and long-term sustainability of the BALTFLLOODS integration environment. The project follows the principle that data remains under the control of those who generate it, while also promoting openness and reuse wherever possible.

- **Municipal and utility data** (e.g., SCADA telemetry, infrastructure records, operational logs) remains the property of the respective municipalities or service providers. BALTFLLOODS will not alter these ownership rights; instead, the integration layer provides mechanisms for controlled access and interoperability. Municipalities retain full authority to decide which datasets are shared openly, which are restricted, and under what conditions they can be reused.
- **Citizen-contributed data** (e.g., flood photos, observation reports, citizen science measurements) belongs to the contributors themselves. By participating in BALTFLLOODS, citizens license their contributions for use in research, pilot evaluation, and municipal decision support, under terms that are transparent and compliant with GDPR. Contributions are anonymised or aggregated when published to protect individual privacy.
- **Project-generated datasets** (e.g., harmonised data products, derived indicators, analytical outputs) will, wherever possible, be shared as open data under permissive licenses. This approach maximises the project's impact, enabling replication and broader uptake in other Baltic municipalities. Where restrictions apply—for example, due to embedded sensitive information—access will follow the principle of being “as open as possible, as closed as necessary.”

In many cases, municipal partners will be able to provide processed or aggregated outputs—such as statistical summaries, heatmaps, or performance indicators—that can be released as open data, even if the underlying operational datasets remain closed. This ensures that the BALTFLLOODS platform delivers maximum value to stakeholders while respecting the integrity of critical infrastructure and complying with applicable regulations.

By combining robust security controls with a selective openness policy, the BALTFLLOODS data integration environment will protect sensitive assets, enable trusted collaboration between partners, and contribute valuable open datasets to the wider Baltic Sea region community.



These ownership principles will be embedded in the Data Management Plan (D1.1, Annex B), which specifies responsibilities for access rights, update cycles, retention periods, and licensing conditions. Together, they ensure that BALTFLOODS respects existing rights, safeguards citizen trust, and promotes transnational knowledge transfer.

3. Implementation Timeline and Activities

The implementation of the BALTFLOODS integration environment has been designed as a gradual and layered process that balances structure with flexibility. The work is organized into four main steps—**Planning, Implementation including Procurement, Testing, and Expansion**—which together chart the pathway from early preparation to long-term consolidation (see figure below).

Although these steps follow a logical order, they are not strictly sequential. In practice, activities overlap: testing begins as soon as the first data streams become available; procurement can run in parallel with the preparation of mock data flows; expansion may begin in certain pilot sites while refinement continues in others. This overlapping rhythm avoids unnecessary downtime and ensures that feedback from users and stakeholders can be incorporated without halting progress. Each step is marked by milestones that provide a clear sense of achievement and allow the project partners to verify that the system is moving in the right direction.

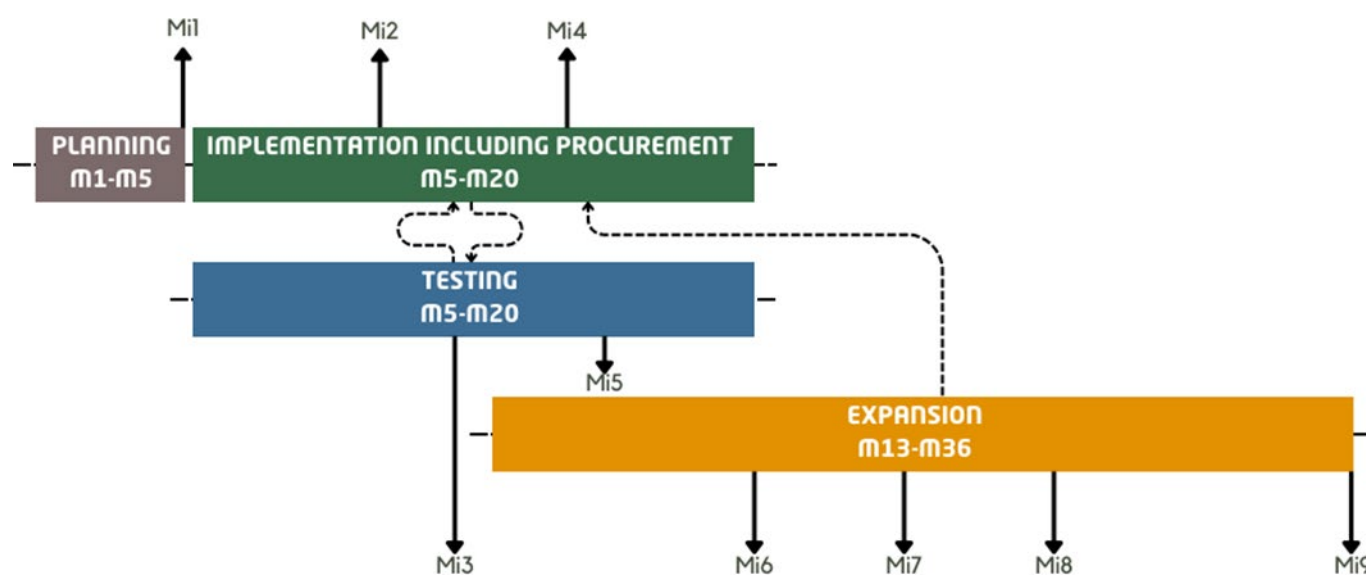


Figure 1. Implementation timeline and activities.

3.1 Planning (Months 1–5)

The first months of the project are devoted to careful preparation. This is the stage where the ambitions of BALTFLOODS are translated into detailed technical and operational requirements. Municipalities and research partners jointly identify priority sites for pilot deployments, such as sewer overflow points or flood-prone catchments, and clarify what types of data will be most useful for both operational monitoring and long-term planning.

At the same time, a thorough mapping of existing systems is carried out. In Lappeenranta, this means documenting how StreetAI currently ingests and processes stormwater data; in Gjøvik, it involves analyzing the flow of information between the SCADA/IoT Hub and Gurusoft's reporting warehouse. Understanding these existing data pathways is essential to designing integration points that are both technically feasible and secure.

Planning also includes drafting the specifications for new sensors and other equipment. These specifications are shaped by local conditions—such as the need for water quality probes capable of withstanding seasonal ice cover or rain gauges robust enough for high-intensity storms. Where equipment is not yet available, the project sets up mock data feeds using simulated values, allowing developers to begin working on data integration and visualization tools without waiting for hardware deliveries.

By the end of this phase, the pilots will have a clear blueprint: requirements are agreed, sites are selected, procurement documents are ready, and preliminary data flows are already being tested in a simulated environment.

Milestones:

Mi1: Completion of detailed requirements, site selection, procurement specifications, and initial mock data flows. (Month 5)

3.2 Implementation and Procurement (Months 5–20)

Once the groundwork is complete, the project moves into full implementation. This step is marked by the arrival and installation of new equipment, the first real data connections with municipal legacy systems, and the deployment of the technical backbone of the integration environment.

Procurement follows public procedures to ensure transparency and fairness, while also allowing municipalities to source specialized expertise where necessary. Deliveries include water quality sensors for parameters such as turbidity or nutrient levels, overflow monitoring devices, rainfall gauges, and weather stations. Once installed in the pilot locations, each device is calibrated, connected to appropriate networks (ranging from 4G/5G to LoRa, depending on location), and tested for reliability.

In parallel, municipal systems begin to interface with the new environment. In Lappeenranta, the StreetAI platform is linked to the shared integration layer, while in Gjøvik, the IoT Hub starts forwarding data to the central environment alongside the existing SCADA and reporting structures. These connections are not trivial: each requires the creation of adapters or API links that translate between system-specific formats and the shared standards of the integration layer.



The deployment of the core platform infrastructure also takes place during this period. Services are containerized to ensure flexibility and portability, and they are deployed either in municipal IT environments or on national e-infrastructure platforms, depending on each city's needs. By the end of this step, most of the hardware is in place, data is flowing from both legacy systems and new sensors, and the foundation for testing is established.

Milestones:

Mi2: First batch of sensors operational (Month 10)

Mi4: Initial data integration from municipal legacy systems confirmed (Month 15)

Mi6a: Full suite of planned sensors deployed and baseline data collection established (Month 20)

3.3 Testing and Adaptation (Months 5–20)

Testing does not wait until all devices are installed; it begins as soon as the first data becomes available. During this step, the system is observed under real conditions and gradually fine-tuned to improve its accuracy, reliability, and usability.

Sensor outputs are carefully validated. For instance, turbidity readings are checked against manual water samples, rainfall data is compared to official meteorological observations, and overflow sensors are monitored during actual storm events. Equally important is the verification of integration pathways: data flowing from StreetAI or Gjøvik's SCADA system must appear correctly in the shared dashboards, with consistent timestamps and harmonized formats.

At the same time, users and stakeholders are brought into the process. Municipal staff test dashboards and alert functions, while citizens interact with the engagement app to access live data or contribute their own observations. Structured workshops and review sessions provide spaces for feedback, and this feedback is directly incorporated into iterative improvements. Adjustments may involve technical fixes, such as correcting an API mapping, but they can also be practical, like refining how alerts are phrased or how dashboards present information during high-stress flood events.

This iterative cycle ensures that the system evolves in response to actual needs. By the end of this phase, the integration environment is no longer a prototype but a functioning system that has proven its stability under field conditions.

Milestones:

Mi3: First round of stakeholder feedback integrated into the system (Month 12)

Mi5: Revised dashboards and alert mechanisms deployed (Month 16)

Mi6b: System validated as stable under real-world operating conditions (Month 20)

3.4 Expansion and Consolidation (Months 13–36)

Once the pilots are stable, attention turns to expansion. This expansion happens in two senses: geographically, by extending sensor setups and monitoring practices to additional sites within the pilot cities; and functionally, by adding new data sources or engagement modules. For example, an initial focus on sewer overflows may be expanded to include nutrient monitoring in runoff water or additional citizen-science protocols for local schools.

Equally important is the consolidation of knowledge and practices. Municipal staff are trained to operate and maintain the systems independently, while project partners compile detailed documentation of configurations, lessons learned, and troubleshooting procedures. These materials will form the basis of the Sustainability Handbook and Best Practice Guide produced later in the project.

Finally, this step connects directly to the project's transnational mission. The integration environment developed in Lappeenranta and Gjøvik is documented in a way that other cities can replicate. WP3 activities ensure that best practices, methodological frameworks, and technical guidelines are shared across the Baltic region. By the end of the project, both pilot cities will have mature, locally embedded solutions, and the wider region will have access to a proven and adaptable blueprint.

Milestones:

M7: Expansion to additional pilot sites completed (Month 24)

M8: Core pilot systems stabilized and documented (Month 28)

M9: Integration and knowledge transfer completed, ensuring replicability (Month 36)

4. Transnational Value and Replicability

4.1 Transnational Mission

The BALTFLOODS Data Integration Environment has been developed with a clear transnational mission:

1. **To ensure the compatibility and interoperability of measurement data across different data systems,**
2. **To guarantee the future-proofness of the implemented solutions, and**
3. **To secure the replicability and wider applicability of the data measurement, analysis, and integration approaches across the Baltic Sea Region.**

Importantly, in BALTFLOODS, **transnational applicability does not come from real-time cross-border data sharing**. Flooding and stormwater events are localised phenomena, and the operational value of live data is inherently tied to the municipality or catchment in which it is collected. Instead, the real cross-border value comes from **codifying, sharing, and transferring** the methods, technical designs, governance models, and evaluation procedures that have been tested in the pilot cities.



4.1.1 Ensuring Compatibility

A key challenge in municipal stormwater management is the heterogeneity of systems—different sensor vendors, communication protocols, data formats, and database structures. BALTFLOODS addresses this by harmonising these inputs through **open standards** and **semantic data models**.

By using formats such as **ETSI NGSI-LD** and domain-specific Smart Data Models, measurements from municipal SCADA/telemetry systems, national environmental portals (e.g., SYKE, NVE), and citizen engagement platforms can be described in a consistent way.

In practice, this means that a turbidity reading from a Finnish sewer outfall and one from a Norwegian stormwater pond can be described with the same attributes, units, and metadata. Even if the data never needs to cross national borders in real time, the **standardised modelling ensures that methods, analytical tools, and workflows can be shared without major adaptation**.

This interoperability is key for replication: a city adopting the BALTFLOODS approach does not need to replace its systems, only to connect them via well-documented adapters into the integration environment.

4.1.2 Guaranteeing Future Proofness

To ensure longevity, the integration plan embraces:

- **Open-source, modular components** – such as the CCO-licensed CitySen.App and FIWARE middleware services – that can be customised without licensing restrictions.
- **Cloud-native deployment** (Docker/OpenShift) for portability and scalability, including the possibility to run on national research infrastructures like **Sigma2's NIRD Service Platform**.
- **Extensibility** through configuration rather than redesign—new sensor types or data sources can be added by creating new adapters or context entities.

By avoiding vendor lock-in and aligning with international best practices, the system remains adaptable to emerging needs and technological changes, ensuring that the structured approach remains relevant across the region for years to come.

4.1.3 Enabling Replicability and Baltic-Wide Applicability

The BALTFLOODS methodology is transferable because it is **structured, documented, and adaptable**. The pilots in Lappeenranta and Gjøvik follow the same high-level process but tailor it to their local contexts—producing two different but comparable examples of implementation.

The transnational value lies in the **ability to package these experiences into D3.2 Transfer sessions and best practice handbook**, which includes:

- Detailed system integration designs based on open standards,
- Clear governance and data management guidelines,
- Engagement methods for involving citizens, stakeholders, and interest groups,
- Performance evaluation indicators aligned with international frameworks (e.g., SDGs).

This handbook along with open-source tools and permissive licensing, allows any municipality in the Baltic Sea Region to adopt the BALTFLOODS approach without starting from scratch. The result is a **shared, proven methodology for stormwater and runoff water quality management**, underpinned by open standards but refined through practical experience in two real-world contexts.

In short, **BALTFLOODS builds a common language and method, not a single cross-border data system**—and it is this structured, adaptable, and open approach that makes it genuinely transnational in value.

4.2 Ensuring Interoperability

At the heart of BALTFLOODS' integration effort lies the commitment to **interoperability**—the capacity for municipal systems, national databases, and third-party applications to exchange information without technical or proprietary barriers. This is achieved not through a single technology, but through a carefully designed **ecosystem of interconnected components** that includes existing municipal SCADA and telemetry platforms, national environmental data portals (e.g., SYKE in Finland, NVE in Norway), open citizen-science applications, and new analytical and visualization tools.

To make this possible, all participating systems and services are connected through **open standards and clearly defined interfaces**. Where possible, data is harmonised using **ETSI NGSI-LD** semantic formats and aligned with the **Smart Data Models** initiative, enabling consistent descriptions of entities such as flood events, water quality observations, and rainfall measurements. This common “data language” ensures that a stormwater sensor in Lappeenranta, a flood model in Gjøvik, and a national open-data feed can all be ingested, processed, and compared within the same integration environment.

Importantly, this standards-based approach means that **legacy municipal systems do not need to be replaced**. Instead, lightweight adapters or gateways can translate data from formats like Modbus, OPC-UA, or vendor-specific APIs into the shared NGSI-LD model. This avoids costly, bespoke, one-off integrations and instead creates a **plug-and-play environment** where any city—whether in Finland, Norway, or elsewhere—can connect its existing assets to the BALTFLOODS environment.

Because the ecosystem is loosely coupled, each data source or consumer connects via a clear, well-documented interface. This design not only facilitates local integration but also enables **cross-border data compatibility and aggregation**—for example, visualising comparable water level data from different countries on a single dashboard, or producing region-wide hydrological indicators from harmonised local inputs.

4.2.1 Designing for Future-Proofing

Long-term value requires an environment that is both **technically robust** and **adaptable to change**. The BALTFLOODS integration approach emphasises future-proofing at multiple levels:



- **Technology Choice** – Wherever possible, open-source and modular technologies are used. The **CitySen.App** citizen engagement platform is openly licensed (CC0) and built with modular components, enabling municipalities to adopt, extend, or rebrand it without licensing costs. FIWARE components, deployed as Docker containers in an OpenShift cluster, provide the underlying middleware for real-time data integration, but any service can be swapped or upgraded without re-engineering the entire system.
- **Cloud-Native Deployment** – Running services in containerised, cloud-native environments ensures scalability and portability. Partners can host their own instances or use shared deployments during the project. This flexibility is critical for cities with varying levels of IT infrastructure: those with in-house capacity can manage their own stack, while others can connect to a shared regional instance. Norwegian partners may, for example, deploy on **Sigma2 national e-infrastructure services** such as the NIRD Service Platform, while Finnish partners could use local or EU cloud resources.
- **Evolving Standards Alignment** – By aligning with international standards such as NGSI-LD and ISO metadata norms, the environment can integrate future sensor types, datasets, or analytics tools with minimal rework. Data models and APIs can be extended incrementally, preserving compatibility while embracing innovation.

This approach ensures that the BALTFLOODS integration environment is not only relevant for the pilot phase but remains a sustainable, upgradeable asset in the face of evolving climate resilience needs and digital transformation trends.

4.2.2 Replication in Diverse Municipal Contexts

The BALTFLOODS methodology is deliberately designed to be **replicable in cities with widely differing digital maturity**. This is achieved by combining:

- A **core, standards-based integration layer** that defines how systems interconnect.
- A set of **open, reusable tools** (citizen engagement apps, data visualisation dashboards, analytics pipelines).
- A **documented implementation pathway** that guides partners from needs assessment to live deployment.

In the pilots, Lappeenranta and Gjøvik will follow the same high-level integration process, ensuring that the outputs—data flows, APIs, governance procedures—share a common logic. This consistency allows other municipalities to treat the BALTFLOODS setup as a **blueprint** while still leaving room for local adaptation. Modules can be swapped, new data sources can be added, and engagement approaches can be tailored to cultural or institutional contexts without breaking the overall system design.

Replicability is further reinforced through the **openness of the technical and procedural outputs**. Sensor configurations, API specifications, and data processing scripts will be documented and shared under permissive licenses. Citizen engagement protocols will be detailed in Deliverables D1.3 and D1.4, and evaluation metrics will be aligned with **globally recognised frameworks** (e.g., Sustainable Development Goals indicators) to maximise international relevance.

Finally, replication is not left to passive uptake: the integration work in WP2 feeds directly into the **knowledge transfer activities of WP3**. Peer-learning workshops, transfer clinics, and the Best Practice Handbook (D3.2) will distil the lessons from the pilots into actionable guidance for follower cities. The result will be not only a technically proven integration environment but a **regionally validated, community-driven model** for digital flood resilience—ready to be adopted, adapted, and scaled across the Baltic Sea region and beyond.



Conclusions

Deliverable 1.2 outlines a comprehensive data integration plan that builds upon the current landscape of existing data systems of City of Lappeenranta and Gjovik municipality. The plan is drawn from FIWARE architecture as inspiration and proposes a scalable technical framework that supports interoperability and long-term adaptability. Beyond immediate integration, the simplified FIWARE architecture positions the platform for future scalability.

The key principle of the plan is compatibility which ensures the new integrations can be implemented without major changes or complex modifications to the current data system. This approach minimized disruption and implementation cost.

The plan also defines clear standards and protocols that allow different systems and components to interpret and share data seamlessly. The data in this project are published as open data, the plan ensures the application of principle of data minimization which includes only essential and non-sensitive information is stored. In case of citizen engagement application that allows users to upload files, pictures, additional privacy considerations are planned in order to obtain explicit users consent, clear guidelines on acceptable content and ensuing the uploaded data is processed in compliance with data protection regulations such as GDPR.

Furthermore, the plan includes detailed implementation timeline and activities covering phases such as procurement, adaptation, installation and feedback loops. While the primary focus is on public infrastructure service providers, the plan is designed to be replicated by similar organizations in diverse contexts.





*Towards flood resilient cities
powered by data and people.*



Co-funded by
the European Union

RESPONSIVE PUBLIC SERVICES
BALTFLOODS



This project is co-funded by the European Union through the Interreg Baltic Sea Region Programme.