



OneAquaHealth

PROTECTING URBAN AQUATIC ECOSYSTEMS TO PROMOTE ONE HEALTH

Grant Agreement: 101086521

D5.2 Geospatial and satellite information platform vs2



Funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Document Information

Deliverable number:	D5.2
Deliverable title:	Geospatial and satellite information platform vs2
Deliverable version:	Draft version
Work Package number:	WP5
Work Package title:	A.I. supported Environmental surveillance system
Due date of delivery:	M24 (December 2024)
Actual date of delivery:	31.12.2024
Dissemination level:	Public (PU)
Type:	DEC
Author(s):	Georgios Koutalieris, Paraskevi Rapti, Xenofon Karagiannis, Symeon Symeonidis, Anagnostis Delkos, Vasilis Perifanis (ENORA Innovation)
Contributor(s):	Maria João Feio, Sónia Serra, Ana Raquel Calapez, Janine P Silva (UC) Alexander Nikolov (SYNYO)
Reviewer(s):	Dirk Schmeller (INPT)
Project name:	Protecting Urban Aquatic Ecosystems to Promote One Health
Project Acronym:	OneAquaHealth
Project starting date:	01.01.2023
Project duration:	48 months
Rights:	OneAquaHealth Consortium

Document history

Version	Date	Beneficiary	Description
0.1	10.10.2024	ENORA	Set up of the document and Creation of Table of Contents
0.2	17.11.2024	ENORA	Update of chapters about Earth Observation
0.3	02.12.2024	ENORA	Update of content regarding Software implementation updates
0.4	09.12.2024	ENORA	Upload the file for contributions by partners
1.0	23.12.2024	UC	Final version reviewed

Executive Summary

The D5.2 deliverable, related to Task 5.1 and titled "Geospatial and Satellite Information Platform vs2", following the D5.1 report titled "Geospatial and Satellite Information Platform vs1", report detailing the requirements and features of the Geospatial and Satellite Information Platform that is integrated with the Data Backoffice (D5.3) forming an integrated platform named GEOSSIP. This document is based on input gathered by scientists from research sites through workshops and meetings coordinated by WP5.

This deliverable also includes additional justifications and implementation details related to the following points that have been raised during the P1 review discussions. As this process did not include developments up to M24, the following points are further updated in this deliverable:

- **Software Libraries and Deployment details:** the description of the software libraries utilised and the deployment locations of WP5 software components, including the OneAquaHealth (OAH) Geospatial and satellite information services, is updated.
- **Data Management:** This deliverable specifies the types of data to be downloaded, processed, and provided, with an emphasis on time series data over the demonstration sites. Both systematic and on-demand processing approaches are clarified.
- **Application Enhancements and Efficiency of the development approach:** Updates and corrections to the online application are included to meet user expectations and align with similar tools. The benefits of developing tools that process and provide EO Data based on the requirements of OneAquaHealth without depending on existing sources of EO data that would not fully secure the sustainability of the resulting solution is addressed.
- **Demonstration and Deployment:** A live demonstration of the full solution is prepared, and deployment details, including API endpoints, are made accessible via the project website.
- **Implementation details:** The development achievements of WP5, including its integration with open-source tools and services, are highlighted.
- **Software implementation updates:** The software implementation updates include the development of city dashboards for visualising urban metrics, automation of environmental index calculations, and enhancements to the Citizen Science App, such as new wellness-related questions, updated guidelines and instructions with images, user-added site functionality, and improved UI/UX for better engagement and accessibility.
- **Maintenance and Exploitation:** Plans for maintaining these components and contributing to extended strategies for their post-project exploitation are presented.

The software development process continued throughout M24, allowing for the incorporation of end-user feedback and continuous improvements during integration into the OneAquaHealth Information Hub. All updates related to Task 5.2 software development are documented in this deliverable.

This document explores the development aspects of the Geospatial and Satellite Information services, which are accessible through the second iteration of GEOSSIP. The current version is available for testing at <https://app.enora-oah.eu/>. The project follows an agile Scrum methodology, enabling ongoing development and refinement. Until M24, a second version was released, focusing on addressing issues, upgrading functionalities, enhancing user experience, and improving interoperability with the Open Information Hub.

The deliverable also defines the usage requirements for the Geospatial and Satellite Information services provided through GEOSSIP, derived from use cases that outline the system's needs and functionalities. This is a crucial step for ensuring the project's effectiveness. In addition, it details the tools developed for Earth Observation (EO) data access and visualisation, presenting the methodologies, technologies, and tools used in the development process. The entire development is guided by the SCRUM methodology, which connects user requirements (user stories) with software components and applications.

Contents

Executive Summary.....	4
1 Introduction.....	11
1.1 Scope.....	11
1.2 Relation to other tasks and deliverables.....	11
1.3 Structure of the deliverable.....	13
2 System Design and Architecture	13
2.1 GEOSSIP: The platform	13
2.1.1 Conceptual Architecture.....	14
2.2 Geospatial and Satellite services solution architecture.....	16
2.2.1 Components.....	17
3 Summary of Prototyping and selection of EO sources	18
3.1 Development process of the OAH Geospatial and Satellite Information services prototype 18	
3.1.1 EO data sources and results.....	18
3.1.2 Features and functionalities of the initial version.....	19
3.2 Combining EO indices with environmental data	21
4 Technologies for the implementation of GEOSSIP Information services.....	22
4.1 Technologies for the implementation of Geospatial and Satellite data services.....	22
4.1.1 Deployment Environment	22
4.1.2 Security Considerations.....	23
4.1.3 Performance Metrics	23
4.1.4 Libraries and Software Stacks	23
4.2 Processing of EO data on OAH Geospatial and Satellite Information services.....	25
4.3 Description of the tools developed for EO data analytics	25
4.3.1 Data Access and Data Processing	26
4.3.2 On-demand Services via GEOSSIP.....	26
4.4 OAH Geospatial and Satellite services via GEOSSIP.....	27
5 Justification of tools development and the efficiency of the development methodology.....	38
5.1 Relevance of open source and open access tools to OAH research objectives	38
5.2 Cost-benefit justification of using existing commercial solutions.....	38
5.3 Open-source libraries used for the development of Geospatial and Satellite Information Platform	39

5.3.1	Flexibility of the solution and added value for the sustainability of project results	39
6	Data Backoffice Development updates.....	40
6.1	The Data Backoffice architecture	40
6.2	GEOSSIP: The OAH backend platform – development updates.....	41
6.3	City Dashboards.....	45
6.4	Citizen Science App	47
7	Harmonisation of planned development activities in WP2, WP5 and WP6	64
7.1	Integration with WP2 and data collection.....	64
7.1.1	Limitations and Foundations.....	64
7.1.2	Current Implementation.....	65
7.1.3	Proposed Solution.....	65
7.1.4	Post-Project Exploitation	67
7.2	Integration with OAH Open Information Hub (Task 6.1).....	67
7.2.1	Current state	67
7.2.2	Community module	68
7.2.3	Future implementations.....	72
7.3	Collaboration with OAH Decision-Support System Development Initiatives (Task 6.2)	73
7.4	Synchronising planned developments regarding OAH Predictive modelling and scenarios (Task 5.3).....	74
8	Ethical and Safety Considerations	74
8.1	Authentication mechanism.....	74
8.2	Disclaimer and information of users for Citizen Science Application	74
9	Scrum Summary	75
10	Next steps	75
11	Conclusions and Future work.....	76
	References	77
	Appendix.....	77
	Appendix A	77
	Draft Disclaimer – Citizen Science App.....	77

Figures

Figure 1 Earth Observation APIs.....	15
Figure 2 GEOSSIP APIs	16
Figure 3 High-level architecture of EO data in OAH Geospatial and Satellite Information services	17
Figure 4 Figure a) S2 True Colour Composite, b) S2-NDVI, c) PS-NDVI d) PS-True Colour Composite over Coimbra, Portugal.....	19
Figure 5 Temporal variation of EVI values, C1, Coimbra	21
Figure 6 GeoJSON example of Areas of Interest in Coimbra, Portugal	26
Figure 7 EO - Information Page	28
Figure 8 EO - Satellite Imagery	28
Figure 9 EO Satellite Imagery example result	29
Figure 10 Earth Observation Indicators.....	29
Figure 11 Earth Observation Indicators - Area selection.....	30
Figure 12 Earth Observation Indicators - Site selection.....	30
Figure 13 Earth Observation Indicators - Multiple Sites Selection.....	31
Figure 14 Earth Observation Indicators - Indices Selection	31
Figure 15 Earth Observation Indicators - Data Range Selection	32
Figure 16 Earth Observation Indicators - BWDRVI result.....	32
Figure 17 Earth Observation Indicators - EVI result.....	33
Figure 18 Earth Observation Indicators - EVI result.....	33
Figure 19 Earth Observation Indicators - NDVI result.....	34
Figure 20 Earth Observation Indicators - NDWI result.....	34
Figure 21 Earth Observation Indicators - Multiple sites and indices selection.....	35
Figure 22 Earth Observation Indicators - Graph date and historical date selection	36
Figure 23 Earth Observation Indicators - Export data selection.....	36
Figure 24 Earth Observation Indicators - Visualisation of Graph	37
Figure 25 The software architecture of the OneAquaHealth backend	41
Figure 26 Login Screen	42
Figure 27 Research Sites Map	42
Figure 28 List of Research sites and coordinates.....	43
Figure 29 Categories of Data Annotations types.....	43
Figure 30 Preview of Data Submissions	44
Figure 31 Data Submission Form.....	44
Figure 32 City Dashboards - Selection of Site	45
Figure 33 City Dashboards - Selection of Data	46
Figure 34 City Dashboards - Status of nitrates.....	46
Figure 35 City Dashboards - Species richness	47

Figure 36 Citizen Science App - Installation guide.....	48
Figure 37 Citizen Science App Intro Screen	49
Figure 38 First Step before submission	50
Figure 39 Before you submit - About the App	51
Figure 40 Citizen Science App/About the App	52
Figure 41 Selection of OAH urban stream sites.....	53
Figure 42 New Site Addition	54
Figure 43 Site Photos and Video Upload	55
Figure 44 Set of Questions.....	56
Figure 45 Set of Questions and guide to fill.....	57
Figure 46 Questions with images	58
Figure 47 Question and add measure by user	59
Figure 48 Overall Assessment	60
Figure 49 Wellness Questions	61
Figure 50 Submission Summary	62
Figure 51 Submissions for the stream sites generated by the user	63
Figure 52 Field changes with annotation Table	66
Figure 53 Data Upload.....	66
Figure 54 Common data structure creation	66
Figure 55 Data Access.....	66
Figure 56 Snapshot of the current state of the Open Information Hub	68
Figure 57 Snapshot of the Community overview page	69
Figure 58 Snapshot of the "User profile" section with a showcased submission of urban stream observation data.....	70
Figure 59 Snapshot of the "Members" section	71
Figure 60 Snapshot of the "Activity feed" section.....	72
Figure 61 Snapshot of the preliminary design of the City Dashboards.....	73
Figure 62 Jira tickets for the implementation	75

Tables

Table 1 D5.2 Input from other tasks and deliverables	12
Table 2 D5.2 Output for other tasks and deliverables	12
Table 3 Spectral indices	20

Acronyms & Abbreviations

Term	Description
API	Application Programming Interface
EO	Earth Observation
GEOSSIP	Geospatial and Satellite Information Platform
OAHA	OneAquaHealth

1 Introduction

This document is part of the Work Package 5 (WP5) technical deliverables concerning developing the Geospatial and Satellite Information Hub (GEOSSIP) and specific version 2. The platform is designed to provide custom analyses directly on satellite data for the project's five (5) research cities.

The D5.2 provides an overview of the tools developed for accessing and visualising Earth Observation data. These tools, implemented as a web application for the OAH, enable the exploration of ground data and the generation of actionable insights, thereby supporting more effective environmental management and policy formulation. GEOSSIP aims to enhance access to satellite information through these advancements, facilitating various applications such as urban planning and environmental monitoring.

In addition, this deliverable includes updates and progress related to the Data Backoffice and the OAH Citizen Science App, reflecting key enhancements and new functionalities that align with the project objectives. These include the implementation timeline until month 24, with progress documented in this deliverable. Specific advancements include the development of city dashboards for visualising urban stream metrics, automating the calculation of environmental indices to streamline data analysis, and introducing a new set of questions and images in the Citizen Science App to engage users better and enhance their understanding. Furthermore, a dedicated set of wellness-related questions has been added based on the respective guidelines provided by WP4, along with functionality enabling users to contribute by adding new sites. UI/UX enhancements have also been implemented to improve usability and accessibility, ensuring a more seamless and engaging user experience.

1.1 Scope

The deliverable builds upon previous work by refining the system's design and architecture to incorporate solutions and address identified challenges. It further emphasizes the development and implementation phase, ensuring that the technical components are integrated and function as intended. Additionally, it includes an in-depth analysis of gathered requirements to validate that the system aligns with user needs and project specifications, forming a foundation for subsequent phases.

By clarifying GEOSSIP's requirements and functionalities, this deliverable ensures the well-structured design and smooth implementation of the system, fully aligning with the overall goals of the OAH project.

Additionally, the Citizen Science Application has evolved, adding advanced features and enhancing usability to encourage increased user participation and engagement. Its smooth integration with the OneAquaHealth Open Information Hub facilitates data sharing and collaboration, addressing the needs of both contributors and stakeholders.

1.2 Relation to other tasks and deliverables

This deliverable is connected to various tasks and deliverables within the OneAquaHealth project.

The connections are detailed in Table 1, which outlines inputs from related tasks and deliverables, and Table 2, which specifies outputs to other tasks and deliverables.

Receives inputs from:**Table 1 D5.2 Input from other tasks and deliverables**

Deliverable/ Task	Due Date	Input for D5.2
D2.1	M18	This deliverable specifies the in-situ data collected at the OAH research sites that are related to the ecological assessment of stream health.
D7.7	M18	The Data Management Plan has been updated based on the data model utilised in the OAH Backend.
D5.1	M18	Geospatial and Satellite Information Platform vs1
D5.3	M18	Digital app and backoffice

Provides outputs to:**Table 2 D5.2 Output for other tasks and deliverables**

Deliverable/ Task	Due Date	Output from D5.2
T2.4	M26-M38	To be considered in the upcoming iterations of this research documentation.
T3.1	M18	To be considered in the upcoming iterations of this research documentation.
T5.3	M24-42	To be considered in the upcoming development of the A.I. models.
T6.1	M2-24	To be considered in the upcoming development of the OAH Information Hub Services.
T6.2	M24-42	To be considered in the upcoming development of the OAH Information Hub Services.

1.3 Structure of the deliverable

This deliverable provides an overview of GEOSSIP development, implementation, and future direction. It begins with an introduction outlining the scope and relationships to other tasks, followed by a detailed system design and architecture, including the conceptual framework and solution architecture for the Geospatial and Satellite services. The document covers the prototyping process, selection of Earth Observation (EO) data sources, and the integration of EO indices with environmental data. It highlights the technologies and tools used for the platform's implementation, including deployment environments, security measures, performance metrics, and data processing tools. Additionally, the deliverable justifies the development of open-source and commercial tools, and provides an analysis of the efficiency and sustainability of the chosen solutions. Methodology updates discuss harmonisation with OAH research activities and workshops with research sites, while data backoffice updates detail GEOSSIP backend platform and developments in city dashboards and the Citizen Science App. The deliverable also addresses the harmonisation of activities in WP5 and WP6, integration with the OAH Open Information Hub, and the development of the Decision-Support System (DSS) and predictive modeling. Ethical considerations, such as authentication mechanisms and data security, are also covered, along with a summary of the Scrum process, next steps, and future work.

2 System Design and Architecture

2.1 GEOSSIP: The platform

GEOSSIP has been developed to meet the requirements of the backend architecture and serve as a single entry point to demonstrate some of the OAH's data capabilities. The platform provides access to:

- Geospatial and satellite information services that manage Earth observation data are available at OAH research sites. This development is presented in the previous deliverable 5.1 (M18), and the current deliverable D5.2.
- Data Backoffice infrastructure is used to handle in-situ data from research sites. This development has already been presented in D5.3.
- Initially introduced in D5.3 (M18), Citizen Science Application development continues to evolve and improve as part of the project activities. This ongoing effort focused on refining features, enhancing user engagement, and ensuring integration with the broader system architecture and the OneAquaHealth Open Information Hub to serve both contributors and project stakeholders better.

GEOSSIP is designed to provide web-based access to Earth Observation data via APIs, as explained in the sections below. This design solution enables the extension of GEOSSIP to incorporate external systems using standardised integration approaches.

In this deliverable focuses on the short presentation of the design presented in D5.1, development, and deployment of OAH's geospatial and satellite services and improvements until M24. The development framework addresses aspects of API structuring, component integration, and the deployment of tools and functionalities. It also extends this work by presenting the OAH in-situ data management and visualisation capabilities of GEOSSIP.

2.1.1 Conceptual Architecture

As outlined in Deliverables D5.3 Digital App and Backoffice and D6.1 OneAquaHealth Open Information Hub v1, the OAH conceptual architecture is designed to manage and present geospatial and environmental data to support citizen science, experts, and data-driven initiatives. It is structured into several layers, each fulfilling specific functions such as data collection, processing, visualisation, and user interaction. The Data Collection Layer gathers data from in-situ measurements, satellite imagery, user inputs, and health observations. Preprocessing steps are applied to ensure data quality through filtering, validation, and normalisation, providing a solid foundation for analysis.

The Data Processing and Analytics Engine Layer leverages machine learning and A.I. techniques to convert raw data into actionable insights. The processed data is then stored in structured formats, making it easy to use in analytics and APIs. The Visualization Layer produces intuitive visualisations, dashboards, and infographics that facilitate user engagement and decision-making, bridging the gap between complex technical outputs and user-friendly tools. Interactive Endpoints offer API-based access to external applications, enabling seamless data retrieval and integration. These APIs are secure, efficient, and optimised for reliability, with features such as authentication and load balancing.

User-facing applications, including dashboards and the Citizen Science App, directly connect to the platform, providing diverse users access. Security and operational layers ensure the system's reliability while complying with data privacy regulations. Together, these components support OAH's mission to advance environmental and health monitoring.

2.1.1.1 Data Management

GEOSSIP's data management, outlined in Deliverables D7.6 and D7.7, ensures efficient handling, processing, and retrieval of data. It aggregates data from diverse sources—such as in-situ measurements, EO analytics, user content, and health indicators—into a secure, scalable repository. Preprocessing steps like filtering, cleansing, validation, and normalisation ensure data consistency and quality. Processed data is stored in formats like structured databases, JSON files or Apache Parquet data storage format, depending on its intended use. The data is organised in layers (raw, processed, analysis-ready) to maintain integrity, support version control, and ensure accessibility for analytics, visualisations, and compliance with privacy and security regulations.

2.1.1.2 Providing APIs Designing and Implementing Robust API Interfaces

The API layer of GEOSSIP acts as the main interface for accessing, interacting with, and integrating data from external applications. Built using RESTful architecture, the APIs offer standardised endpoints for various functionalities, such as data retrieval, submission, and processing requests. These APIs are developed by defining clear data schemas and models to support a wide range of data types, from spatial datasets to health indicators. Key features include pagination, filtering, and customisable queries to optimise data fetching. Access to the API is secured with API keys and OAuth/JWT tokens, ensuring that only authorised users or applications can interact with the platform.

The API layer also includes an orchestrator component that manages request routing, load balancing, and response caching, improving the overall efficiency and reliability of data exchange. Comprehensive API documentation is provided to assist developers in integrating their applications with the platform.

The developed REST API is fully documented with Swagger and OpenAPI specifications, offering an interactive and user-friendly interface to help users explore and test the available endpoints. These comprehensive resources are designed to simplify the integration process, enabling users to efficiently interact with the API and leverage its full potential for Earth Observation data retrieval and analysis.

Figures 1 and 2 are showcase the Swagger documentation for the EO (<https://api-enorainnovation.com/docs#/>) and GEOSSIP (<https://api.enora-oah.eu/swagger-ui/index.html>) APIs.

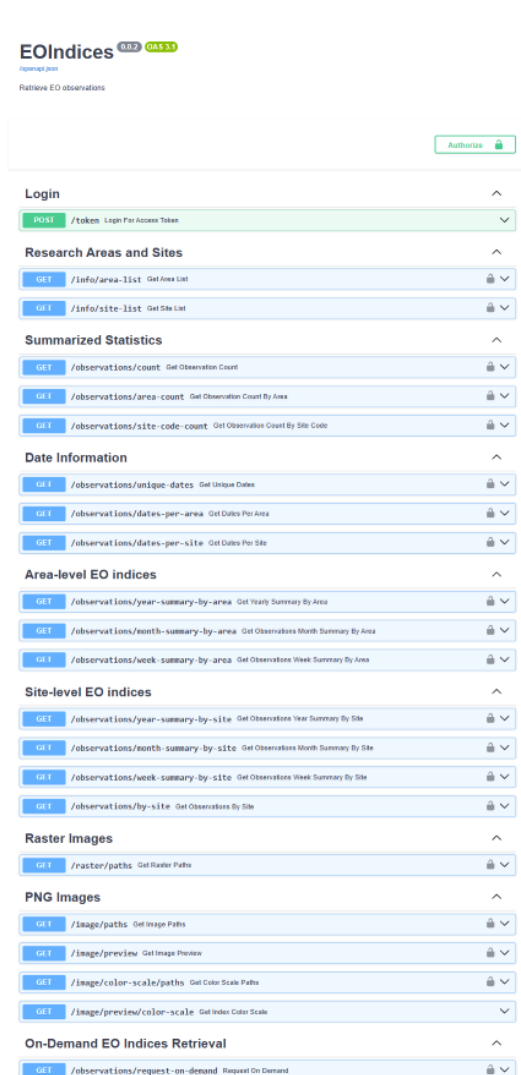


Figure 1 Earth Observation APIs

Citizens Science			^
PUT	/api/citizens/submit	Submit a citizen science observation	▼
GET	/api/citizens/water_flows	Get all water flow types	▼
GET	/api/citizens/water_colors	Get all water color types	▼
GET	/api/citizens/vegetation_types	Get all vegetation types	▼
GET	/api/citizens/submissions_of_external_user/{externalId}	Get all citizen science observations that correspond to an external user	▼
GET	/api/citizens/submissions	Get all citizen science observations	▼
GET	/api/citizens/submission/{id}	Get a single citizen science observation	▼
GET	/api/citizens/stream_assessments	Get all stream assessments	▼
GET	/api/citizens/habitats	Get all habitats	▼
GET	/api/citizens/fallen_biomass	Get all fallen biomass types	▼
GET	/api/citizens/channel_types	Get all channel types	▼
GET	/api/citizens/channel_forms	Get all channel forms	▼
GET	/api/citizens/bank_types	Get all bank types	▼
Users			^
POST	/api/users/edit	Edit an existing user	▼
POST	/api/users/add_external	Add a new user with an external ID	▼
POST	/api/users/add	Add a new user	▼
GET	/api/users/roles	Get all user roles	▼
GET	/api/users/get/{username}	Get an existing user	▼
GET	/api/users/all	Get all users	▼
Dashboards Data Old			^
POST	/api/dashboards_old/timeline/nutrient	Timeline of a nutrient variable	▼
Dashboards Data			^
POST	/api/dashboards/status_of_nitrate/site		▼
POST	/api/dashboards/status_of_nitrate/city		▼
POST	/api/dashboards/species_richness/site		▼
POST	/api/dashboards/species_richness/city		▼
POST	/api/dashboards/simpsons_index/site		▼
POST	/api/dashboards/simpsons_index/city		▼
POST	/api/dashboards/shannon_index/site		▼
POST	/api/dashboards/shannon_index/city		▼
POST	/api/dashboards/pielou_index/site		▼
POST	/api/dashboards/pielou_index/city		▼
POST	/api/dashboards/margalef_index/site		▼
POST	/api/dashboards/margalef_index/city		▼

Figure 2 GEOSSIP APIs

2.2 Geospatial and Satellite services solution architecture

The OAH Geospatial and Satellite services aim to streamline the processing and transformation of satellite imagery into actionable insights. By leveraging advanced algorithms and robust computing resources, these services convert raw satellite data into high-quality, preprocessed imagery and derived products. These outputs can be seamlessly integrated into a common platform, GEOSSIP,

allowing users to relate satellite imagery with in-situ data for enhanced environmental monitoring, and will be investigated in Task 2.4.

The system is designed to handle the complexities of large-scale satellite data processing through a modular and scalable architecture. It encompasses integrated components for data ingestion, storage, processing, and delivery, ensuring efficient and reliable operation. This approach enables the seamless ingestion and preprocessing of diverse satellite data while providing robust storage solutions, such as PostgreSQL, for managing large datasets with flexible querying capabilities, as presented in Figure 3.

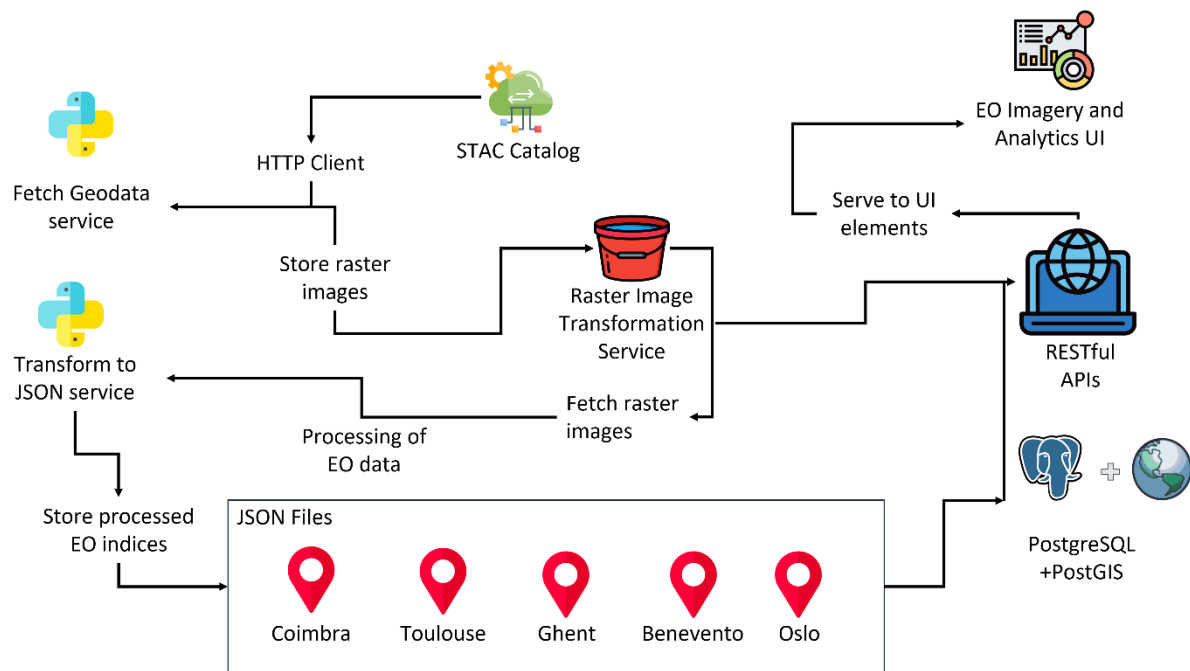


Figure 3 High-level architecture of EO data in OAH Geospatial and Satellite Information services

Central to the system is the data processing component, which uses advanced algorithms to transform raw satellite imagery into usable insights. The user interface, implemented within GEOSSIP platform, offers an accessible, web-based environment for data visualisation and interaction. Built with modern technologies, the UI ensures compatibility across devices and delivers a user-friendly experience for Earth Observation data.

An API gateway complements these features, serving as the integration hub between OAH services and GEOSSIP platform. It ensures secure and efficient communication by handling authentication, request validation, and traffic management. Together, these components create a robust ecosystem for satellite data processing, empowering users with high-quality data and tools for environmental monitoring.

2.2.1 Components

As presented in Figure 3, the high-level architectural design of the OAH Geospatial and Satellite Information services focuses on efficiently processing, managing, and visualising Earth Observation (EO) data. Key components enable the seamless flow of satellite imagery from acquisition to user-friendly visualisation, ensuring the system meets its objectives for geospatial data analysis and accessibility.

The process begins with the Sentinel API Call, which acquires raw satellite imagery from Sentinel satellites. This data is retrieved by the Geodata Fetch Satellite Image Service, which fetches geospatial data based on specific geolocation parameters via HTTP requests and stores the images securely in an S3 Bucket (Raster Images) for further processing. These raw raster images are then converted into the GeoJSON format through the Raster Image Transform to GeoJSON Service, which prepares the data for geospatial analysis and visualisation.

The transformed GeoJSON files are temporarily stored in the Geodata Volume, an intermediary storage solution that ensures the data is structured and ready for persistent storage. The GeoJSON Insert into PostgreSQL function then stores the processed geospatial data in a relational database, facilitating efficient and durable data management. This data is accessed by the GeoJSON Microservice, which provides HTTP endpoints to serve GeoJSON files to the Dashboard UI.

Finally, the Dashboard UI visualises the geospatial data, offering users an intuitive interface to explore and analyse information. By integrating these components, the OAH system efficiently handles the full lifecycle of EO data, from acquisition to real-time visualisation, ensuring reliability and user accessibility.

3 Summary of Prototyping and selection of EO sources

3.1 Development process of the OAH Geospatial and Satellite Information services prototype

The development of the OAH Geospatial and Satellite Information services prototype focused on leveraging Earth Observation (EO) data to monitor and assess the health of urban aquatic ecosystems. The process involves integrating high-resolution satellite imagery from sources such as Sentinel-2 and Planetscope, which provide critical data on vegetation, soil, and water conditions. By using spectral indices like NDVI, EVI, and NDWI or raw bands data, the prototype aims to deliver actionable insights into ecosystem health, water quality, anthropogenic alterations in the channels, and in the riparian vegetation. This section outlines the development process of the prototype, including the selection and analysis of EO data sources, the integration of spectral indices with in-situ environmental parameters, and the approach to refining the prototype to meet the needs of urban stream monitoring.

3.1.1 EO data sources and results

The analysis of Earth Observation (EO) data relies on multiple data sources, including Sentinel-2 and Planetscope, which provide high-resolution imagery for monitoring urban aquatic ecosystems. Sentinel-2 Level 2A data, accessed via Earth Search by Element84, offers multispectral images with a 10-meter resolution, focusing on vegetation, soil, and water information. Planetscope, with a 3-meter spatial resolution, offers daily global coverage and a range of eight spectral bands. The Surface Reflectance - Harmonised product from Planetscope was used in this study, providing data comparable to Sentinel-2. Despite some challenges posed by atmospheric conditions, Planetscope's daily revisits offer more valid images than Sentinel-2 for certain time periods, which is advantageous for continuous monitoring [1].

In comparing the two sources, as Figure 4 presents, Planetscope's higher resolution and daily revisit capability make it more suitable for monitoring narrow urban streams, which are often under vegetation canopies. However, calibration and validation (CalVal) of large sensor fleets can lead to slight discrepancies in measurements. To address this, used median values from the data to minimise these differences. The analysis includes both true-color composites and various spectral indices, such as Normalised Difference Vegetation Index (NDVI) and Normalised Difference Water Index (NDWI), which are used to monitor vegetation health and water content across the studied areas.

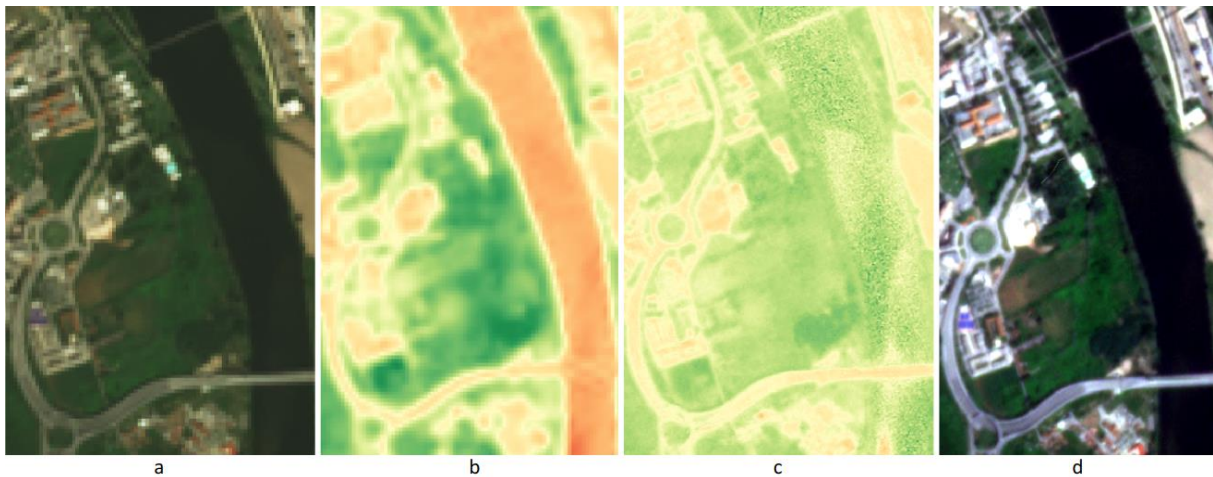


Figure 4 Figure a) S2 True Colour Composite, b) S2-NDVI, c) PS-NDVI d) PS-True Colour Composite over Coimbra, Portugal

3.1.2 Features and functionalities of the initial version

The selection of spectral indices was guided by their relevance in evaluating the health of urban aquatic ecosystems and their ability to be combined with in-situ health indicators. The urban streams monitored in this project are narrow, often just 2-3 meters wide, and typically shaded by vegetation. Therefore, we explored the use of remote sensing data that covers broader areas around these streams, serving as proxy parameters for the surrounding environmental conditions and integrating this data with in-situ environmental and health metrics.

Through an initial analysis using Sentinel-2 imagery, we evaluated several, as Table 3 presents, spectral indices that could provide relevant information on urban streams ecosystem health and degradation. Based on this analysis, we refined our selection to the following key indices:

1. Normalised Difference Vegetation Index (NDVI)
2. Enhanced Vegetation Index (EVI)
3. Enhanced Vegetation Index 2 (EVI2)
4. Blue-Wide Dynamic Range Vegetation Index (BWDRVI)
5. Normalised Difference Moisture Index (NDMI)
6. Normalised Difference Water Index (NDWI)

Table 3 Spectral indices

Indices that we have considered
Vegetation
Normalised Difference Vegetation Index (NDVI)
Modified Normalised Difference Vegetation Index (mNDVI)
Enhanced Vegetation Index (EVI)
Enhanced Vegetation Index (EVI2)
Soil Adjusted Vegetation Index (SAVI) and Transformed Soil Adjusted Vegetation Index (TSAVI)
Blue-wide dynamic range vegetation index (BWDRVI)
Leaf Water Content Index (LWCI)
Global Vegetation Moisture Index (GVMI)
Global Environment Monitoring Index (GEMI)
Normalised Normalised Difference Moisture Index (NDMI)
Leaf Area Index (LAI) (not in the list of Sentinel-2)
Green Leaf Area Index (GLAI) (in Sentinel-2 only the GLI Green Leaf Index)
Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) (not in the list of Sentinel-2)
Water
Land Surface Temperature (LST)
Normalised Difference Water Index (NDWI) (Lu et al., 2020)
Modified Normalised Difference Water Index (MNDWI) (Du et al., 2016)
Augmented Normalised Difference Water Index (ANDWI)
Sentinel-2 Water Index (S2WI)
Land Surface Water Index (LSWI)
NDVIMNDWI NDVI-MNDWI Model
Urbanisation
Built-Up Land Features Extraction Index (BLFEI)
Band Ratio for Built-up Area (BRBA)
Normalised Built-up Area Index (NBAI)
Index-Based Built-Up Index (IBI)
Enhanced Built-Up and Bareness Index (EBBI)
Perpendicular Impervious Surface Index (PISI)
Urban Index (UI)
Vegetation Index Built-up Index (VIBI)

The Normalised Difference Vegetation Index (NDVI) is a common remote sensing metric used to assess vegetation health, with values ranging from -1 to +1, where higher values indicate healthier vegetation. It is calculated from reflectance measurements in the red and near-infrared bands, making it useful for monitoring vegetation dynamics in agriculture, forestry, and ecological research.

The Enhanced Vegetation Index (EVI) improves upon NDVI by offering greater sensitivity in high biomass areas and correcting for atmospheric conditions and canopy background signals. EVI is particularly beneficial in dense vegetation regions where NDVI may saturate and play a key role in climate studies.

Enhanced Vegetation Index 2 (EVI2) simplifies EVI by removing the need for the blue band yet retains its improved sensitivity and accuracy in high biomass areas, making it useful when blue band data is unavailable.

The Blue-Wide Dynamic Range Vegetation Index (BWDRVI) incorporates the blue band to enhance vegetation analysis, providing additional detail and contrast in diverse ecosystems. It is beneficial for tracking subtle changes in vegetation health and detecting plant stress.

The Normalised Difference Moisture Index (NDMI) monitors vegetation moisture content, aiding in detecting drought conditions and plant water stress. It uses the shortwave infrared (SWIR) and near-infrared bands to assess water content, which is essential for water management and drought prediction.

The Normalised Difference Water Index (NDWI) measures moisture content in vegetation and soil, making it an important tool for monitoring drought conditions and assessing water availability, critical for agricultural planning and water resource management.

3.2 Combining EO indices with environmental data

GEOSSIP allows OAH researchers and A.I. prediction services to combine site-specific data with spectral indices like NDVI, EVI, EVI2, BWDRVI, NDMI, and NDWI, creating a robust tool for monitoring urban aquatic ecosystem health. These indices offer insights into vital environmental factors such as vegetation health, moisture levels, and water content, which are essential for ecosystem stability. By analysing these indices, researchers can detect seasonal changes, trends, and shifts in vegetation and water conditions. For example, high NDVI and EVI values indicate healthy vegetation crucial for stabilising riverbanks and filtering pollutants, while NDMI and NDWI highlight areas with sufficient moisture and water for aquatic life. Temporal analysis reveals stress periods and potential environmental threats, enabling timely intervention. Integrating these indices with in-situ data improves understanding of urban ecosystem interactions, supporting effective management strategies, particularly in areas affected by rapid land use changes and pollution. Figure 5 shows the temporal variation of EVI values for site C1 in Coimbra, Portugal, providing key statistics to track vegetation dynamics over time [2].

In addition to these, data analyses performed by the UC team revealed that the original bands obtained from satellite observations could be very relevant to correlate with environmental data [Silva et al. under preparation; 2; 3]. Those bands are: Blue, Coastal, Green, NIR, SWIR16, and SWIR22. These are also described and available through the dashboards.

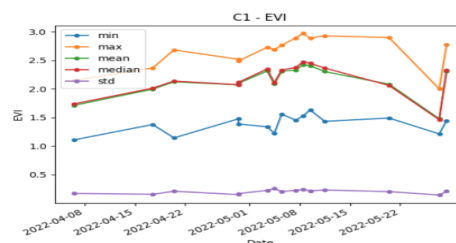


Figure 5 Temporal variation of EVI values, C1, Coimbra

4 Technologies for the implementation of GEOSSIP Information services

GEOSSIP services leverage a carefully selected stack of technologies to deliver a robust, scalable, and efficient platform. Python serves as the primary programming language due to its versatility, simplicity, and extensive ecosystem of libraries, enabling rapid development and seamless integration of data processing and machine learning capabilities.

The platform employs Docker Compose for containerisation, ensuring consistency across development, testing, and production environments. This approach optimises resource management, simplifies scalability, and supports microservices orchestration. Spring Boot is utilised to deploy these services, providing a solid framework for maintaining a scalable microservices architecture while simplifying configuration and deployment processes.

For data storage, PostgreSQL is chosen for its robust querying capabilities, flexible schema design, and support for geospatial data, making it ideal for managing the dynamic metadata associated with satellite imagery. Meanwhile, Apache NiFi facilitates efficient data flow management, enabling the seamless integration of GEOSSIP and other external information consumers through intuitive pipeline design and monitoring.

Finally, GeoJSON is adopted as the standard format for geospatial data, ensuring compatibility with web-based mapping tools and geospatial applications. This facilitates the visualisation and analysis of satellite imagery, enhancing the platform's utility for end-users. Together, these technologies create a cohesive ecosystem tailored to the demands of geospatial and satellite information services.

4.1 Technologies for the implementation of Geospatial and Satellite data services

This section covers the technologies used to implement the OAH Geospatial and Satellite Data Services. The platform is deployed on AWS, ensuring scalability, security, and efficient data processing. Security is maintained through robust authentication and access controls. Performance metrics such as response time and scalability are prioritised, with AWS services like Lambda, EC2, and S3 supporting data processing and storage. Python libraries handle data access, processing, and geospatial analysis. The system supports both systematic and on-demand processing, ensuring effective management and analysis of Earth Observation data for the project.

4.1.1 Deployment Environment

The deployment environment for the OAH Geospatial and Satellite Information services is the Amazon Web Services (AWS) cloud. AWS provides scalable infrastructure and a wide range of services supporting the platform's computing, storage, and data processing requirements. Cloud deployment ensures high availability and flexibility in resource management. AWS was selected to scale effectively, as the geospatial and satellite data managed by the platform can amount to many gigabytes. The elasticity of AWS resources ensures that the platform can handle large volumes of data efficiently while maintaining performance and cost-effectiveness.

4.1.2 Security Considerations

Security is a critical aspect of OAH Geospatial and Satellite Information services and of the access provided via GEOSSIP. Robust mechanisms for authentication and authorisation to protect data and user privacy have been implemented in GEOSSIP and API integration is secured. Access controls and role-based permissions ensure that only authorised users can access sensitive data and functionalities. Network security measures are inherently implemented in the cloud environment (firewalls, encryption, and secure communication protocols), further protecting the development from external threats.

4.1.3 Performance Metrics

Performance metrics are crucial for ensuring the efficiency and reliability of the OAH Geospatial and Satellite Information services. Key performance metrics include response time, throughput, and scalability.

Response Time measures the time to process and deliver satellite imagery to users. This metric is critical for ensuring a responsive and efficient user experience. OAH Geospatial and Satellite Information services are optimised to minimise response time through efficient data processing algorithms and robust infrastructure.

Throughput indicates the volume of data the platform processes within a given timeframe. High throughput is essential for managing large volumes of satellite data and providing timely insights. The OAH Geospatial and Satellite Information services are designed to maximise throughput using scalable data ingestion and processing pipelines.

Scalability refers to the platform's ability to handle increasing data volumes and user demands without compromising performance. OAH Geospatial and Satellite Information services microservices architecture, containerisation with Docker Compose, and deployment in the cloud (AWS) ensure that the platform can scale horizontally. This allows adding more resources as needed, maintaining optimal performance even under varying loads.

4.1.4 Libraries and Software Stacks

Data access, processing, and storage of Earth Observation (EO) data are performed within the AWS cloud environment. A series of AWS Lambda functions has been deployed and is triggered at specific intervals, typically every 30 days. These Lambda functions are integral to the workflow, executing several essential tasks. The primary functionality includes accessing Sentinel-2 Level-2A (L2A) data over designated pilot areas and storing the retrieved data in an Amazon S3 bucket named oah-eo-data-inbound.

4.1.4.1 Accessing EO Data Using STAC

Data access is facilitated through the Earth Search STAC Catalog, a service provided by Element84. The STAC (SpatioTemporal Asset Catalog) specification, detailed at stacspec.org, has emerged as a widely adopted standard for managing spatiotemporal data. Prominent data providers like Sentinel-Hub and Planet have embraced this specification, making it the de facto choice for efficient data management and retrieval. The Sentinel-2 L2A data product is specifically chosen because it has undergone atmospheric correction, a critical prerequisite for analytical applications.

4.1.4.2 Accessing and Processing Sentinel-2 Data

All 12 Sentinel-2 L2A spectral bands are utilised during the data processing workflow. In addition to these bands, the Scene Classification Layer (SCL) data product is also accessed. The SCL data is derived using the Sen2Cor processor and includes 12 classification categories, two of which relate to cloud presence (Cloud Medium Probability and Cloud High Probability). While an initial filtering ensures that only images with less than 70% cloud cover for the entire Sentinel-2 scene are accessed, additional processing is performed to refine the data further. Using the SCL bands, localised cloud masks are created to exclude cloud-covered areas of interest (Aols). These areas are subsequently discarded to ensure high-quality data for downstream applications.

4.1.4.3 Data Storage and Derived Products

Processed data products, such as spectral indices, are stored in another Amazon S3 bucket (oah-eo-data-proc). These spectral indices are used to generate image composites, which are instrumental for visualisation purposes on the application front end. This ensures that the processed data is readily available for further analysis and user interaction.

4.1.4.4 Technical Implementation Using Python

The entire codebase for data access and processing is written in Python. A variety of Python libraries and packages are utilised to manage different aspects of the workflow efficiently. The primary libraries include:

- pystac and pystac-client for managing STAC specifications and data access
- rasterio, rioxarray, and xarray for handling geospatial raster data
- Geopandas, pyproj, shapely, and skimage for geospatial analysis and data manipulation
- boto3 and botocore for interaction with AWS services
- imageio for image processing tasks

These tools form a robust ecosystem for handling complex EO data workflows.

4.1.4.5 Utilised AWS Services

The project extensively leverages AWS services, including Elastic Compute Cloud (EC2) for scalable computation, Simple Storage Service (S3) for data storage, AWS Lambda for serverless processing, and Elastic Container Registry (ECR) for containerised workflows. This integration enables efficient, scalable, and cost-effective processing of EO data.

4.1.4.6 Systematic and On-Demand Processing

Processing is conducted in two modes: systematic and on-demand. Systematic processing occurs at the end of each month, covering all predefined demo sites. On-demand processing, on the other hand, is triggered as needed to validate specific data or perform targeted analyses.

4.1.4.7 Data Validation and Consistency Checks

Sentinel-2 vegetation indices are validated and compared against Landsat data as part of on-demand processing. The methodology ensures the reliability and accuracy of derived data products by assessing the consistency of vegetation indices, such as the Normalised Difference Vegetation Index (NDVI) and

the Enhanced Vegetation Index (EVI), over the same regions and time periods. This validation step provides confidence in using Sentinel-2 indices for various analytical applications.

4.2 Processing of EO data on OAH Geospatial and Satellite Information services

Processing EO data within OAH Geospatial and Satellite Information services involves detailed steps to transform raw satellite imagery into actionable insights. This process includes cropping the data to focus on specific areas, computing spectral indices to highlight environmental characteristics, applying cloud masks to ensure pixel validity, incorporating geo-referenced sampling sites, extracting Areas of Interest (AoI), and performing statistical analysis on spectral information.

The process begins by cropping the data to focus on specific pilot areas, ensuring that all available sampling sites are included. This targeted cropping reduces unnecessary data, improving both processing efficiency and the relevance of the analysis. Next, spectral indices are computed from Sentinel-2 data using specific bands highlighting key environmental characteristics, such as vegetation health, water quality, and soil properties. These indices provide crucial insights for monitoring and assessing environmental conditions.

In addition to these steps, pixel validity is ensured through the application of a cloud mask. This process removes pixels affected by clouds, shadows, or other atmospheric disturbances, ensuring that only clear and unobstructed pixels are considered for analysis. Geo-referenced sampling sites are then integrated into the EO data by overlaying buffered sampling areas, ensuring that the data corresponds to specific locations of interest. Furthermore, Areas of Interest (AoI) are extracted from the EO data based on GeoJSON files, focusing the analysis on the relevant environmental features around these sites.

The final step in the processing involves a detailed analysis of the spectral information. For each AoI, individual bands and derived indices are calculated, followed by statistical analysis to extract key metrics such as minimum, maximum, mean, median, and standard deviation values for each spectral index. This comprehensive analysis delivers valuable insights into the environmental characteristics of the sampled areas, enabling informed decision-making and supporting further research.

4.3 Description of the tools developed for EO data analytics

As described in D5.1, the OAH Geospatial and Satellite Information services are pivotal in supporting the retrieval, analysis, and interpretation of Earth Observation (EO) data. These services are tailored to meet the needs of researchers, urban planners, and environmental scientists, providing access to critical geospatial data relevant to aquatic and urban health. By utilising GeoJSON files to define geographic boundaries and characteristics of interest, the system ensures precise and relevant data retrieval for OneAquaHealth's urban areas and designated research sites.

The implementation of these services is structured around two key components. The data access component focuses on querying satellite images for specific urban areas within defined date ranges, ensuring the timely acquisition of relevant data. The data processing component analyses these images to calculate EO indicators, supporting comprehensive research and planning. These components empower stakeholders to make informed decisions and advance urban and environmental health scientific studies.

4.3.1 Data Access and Data Processing

The OAH Geospatial and Satellite Information services streamline the retrieval and processing of satellite images and geospatial data to support urban analysis and environmental research. The system integrates data access and processing functionalities, ensuring efficiency, accuracy, and relevance.

Using GeoJSON files, as Figure 6 result presents, to define areas of interest (e.g., Coimbra, Portugal), the service retrieves geospatial data from SpatioTemporal Asset Catalogs (STAC), such as the Element84 Earth Search for Copernicus Sentinel-2 imagery. Configuration details guide the search, including urban area names, date ranges, and cloud cover thresholds. The service filters results to exclude high-cloud-cover images, ensuring quality. Retrieved data is masked, cropped, and processed before being uploaded to an S3 bucket for easy access.

Once stored, the data undergoes advanced processing, including rescaling spectral bands to a consistent resolution, generating true-colour composites, and computing indices such as NDVI, EVI, and NDWI. Spatial masking and clipping ensure focus on relevant areas. The results, saved as GeoTIFF files, include processed images, computed indices, and true-color composites. Statistical metrics (e.g., min, max, mean) for each index and band are calculated and stored in JSON format, providing detailed insights for each urban area over specific time intervals.

This integrated approach automates data retrieval, processing, and analysis, enhancing the efficiency and accuracy of geospatial workflows. By supporting the automated management of Sentinel-2 data, the system significantly aids environmental monitoring and urban analysis within the OneAquaHealth ecosystem.

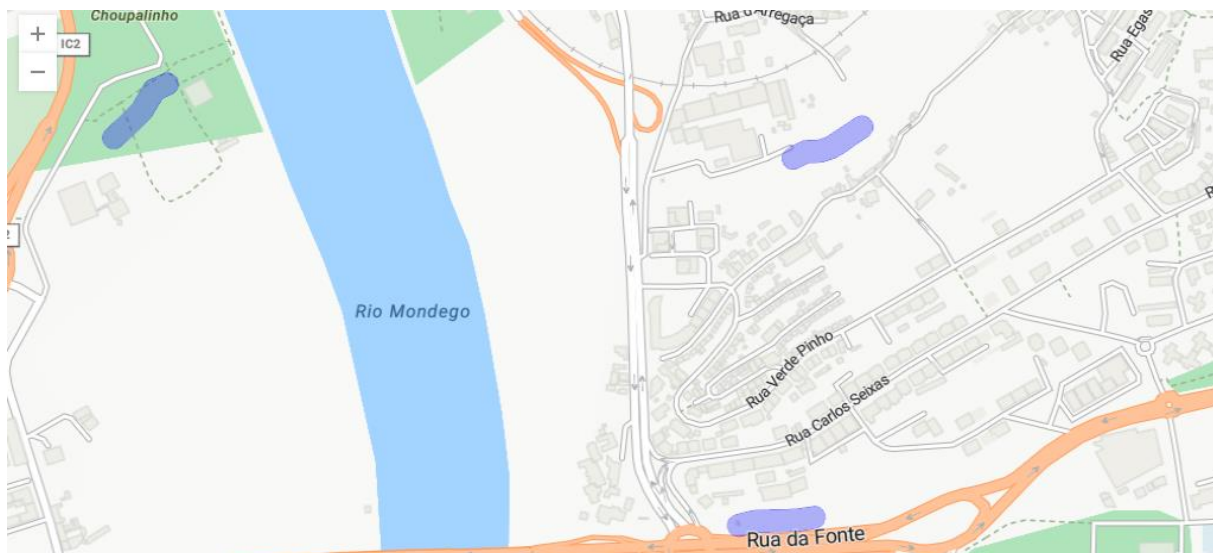


Figure 6 GeoJSON example of Areas of Interest in Coimbra, Portugal

4.3.2 On-demand Services via GEOSSIP

The OAH Geospatial and Satellite Information services query data for urban areas and associated sites cost-efficiently by requesting information at the urban area level rather than for each individual site. GEOSSIP retrieves data stored in a PostgreSQL instance when a user requests Earth Observation information. A dedicated REST API service is responsible for fetching Sentinel statistics for a specific

site within a given date range. The request requires users to specify the area of interest (e.g., Coimbra) and the date of interest. Users must authenticate using JWT token authentication.

Upon receiving a request, the REST API processes the query. If no indices are found for the specified date (e.g., due to the unavailability of Sentinel data), the API responds with an error message indicating that the requested data is unavailable. If data is found for the selected date and area, the system begins processing the request. Users are notified that processing has started, and after a few minutes, users can be presented with the computed spectral indices. These indices include details such as maximum, minimum, mean, median, and standard deviation values for spectral statistics (e.g., BWDRVI, EVI, NDVI).

4.4 OAH Geospatial and Satellite services via GEOSSIP

GEOSSIP offers web-based access to the OAH Geospatial and Satellite services through an open API integration. The results are visualised using the platform's built-in methods. The following Figures showcase the Earth Observation section, displaying satellite imagery and Earth Observation Indicators. Users can select multiple areas of interest, apply multiple indexes such as the Blue Wide Dynamic Range Vegetation Index or Normalised Difference Vegetation Index, and choose a specific date or date range for satellite observations. The data table provides detailed observation information, which is visualised on the map, enabling users to track changes over time.

In the Earth Observation Indicators module, the map shows the selected area and the sidebar lets users choose a region, research site, index type, and date to retrieve analytics. After making these selections, users can click "Get Earth Observation Analytics" to view the data. The Observation Data Table displays key metrics, and an opacity slider allows users to compare different data layers. Moreover, the "Show Earth Observations Graph" option helps users visualise trends, analyse changes over time, and download the data for further analysis.

This work is accessible via <https://app.enora-oah.eu/eo>.

Figure 7 provides an overview of Earth Observation information, including data availability and the current indices. This figure will be updated in the future to incorporate new data and indices as additional information becomes available, ensuring the platform remains current and comprehensive.

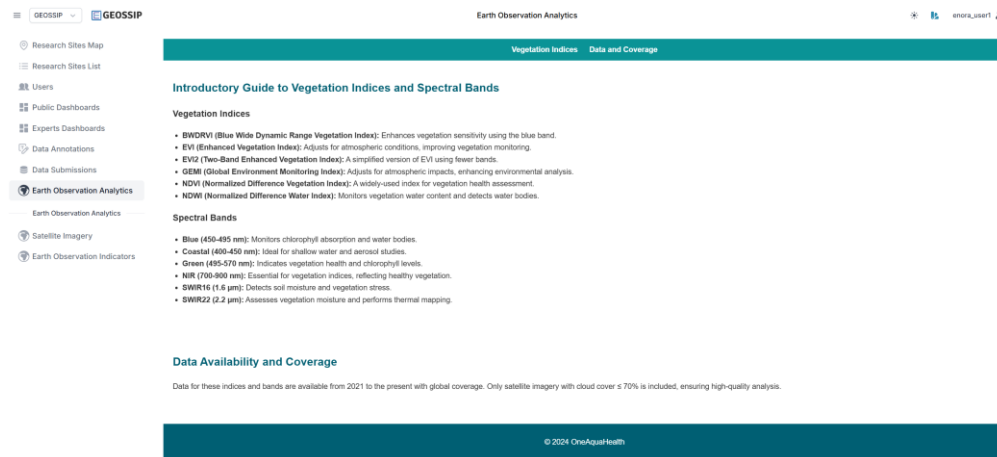


Figure 7 EO - Information Page

The Earth Observation (EO) tools available on the platform enable users to explore and analyse satellite imagery and spectral indices.

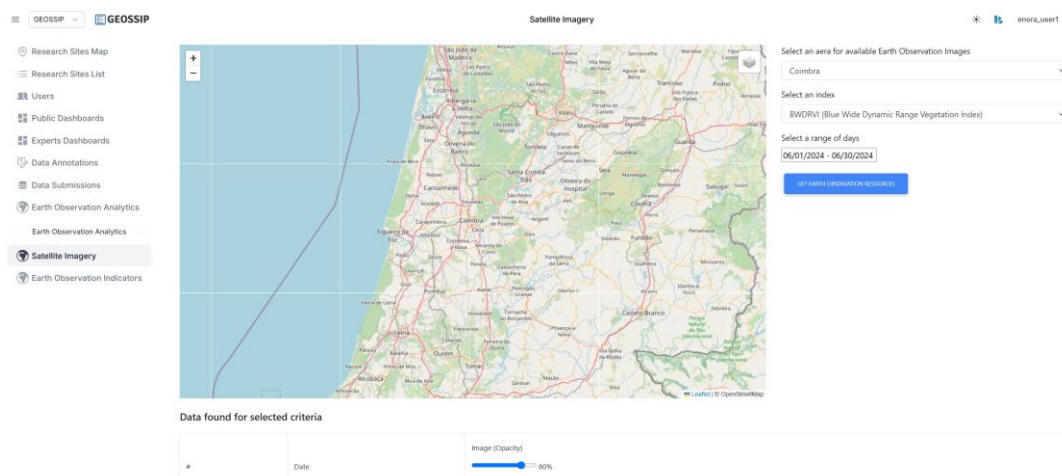


Figure 8 EO - Satellite Imagery

Figures 8 and 9 showcase the satellite imagery and corresponding result. This functionality allows users to select one of five OAH research sites, choose an index from a predefined list, and specify a period. The software then retrieves and displays the available data and images, projecting them onto the map for visualisation.

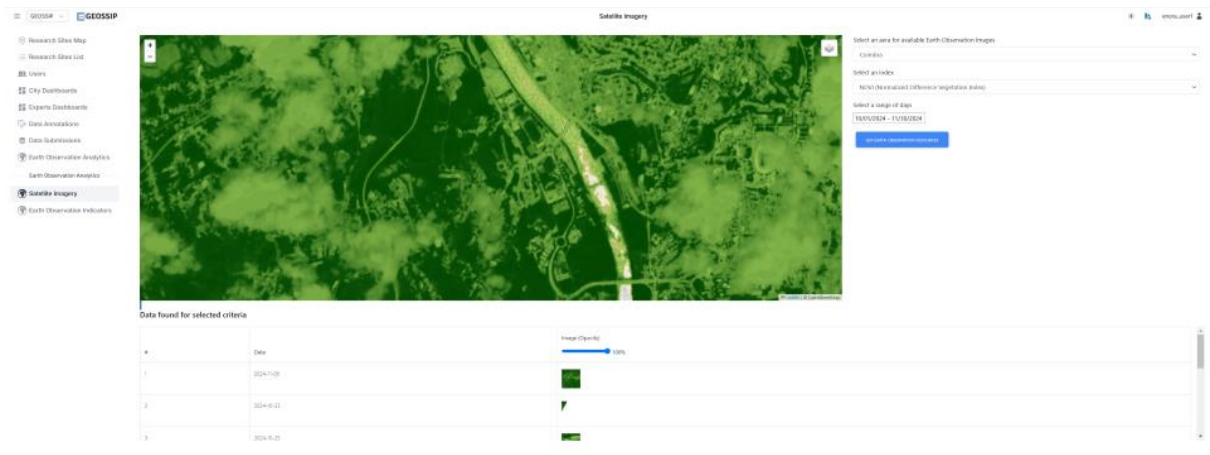


Figure 9 EO Satellite Imagery example result

In the Earth Observation Indicators module, the map dynamically displays the area of interest selected by the user.

The sidebar provides options to select one of five OAH, then multiple research sites, multiple index types, and date range, allowing for tailored data retrieval, as presented in Figures 10-15.

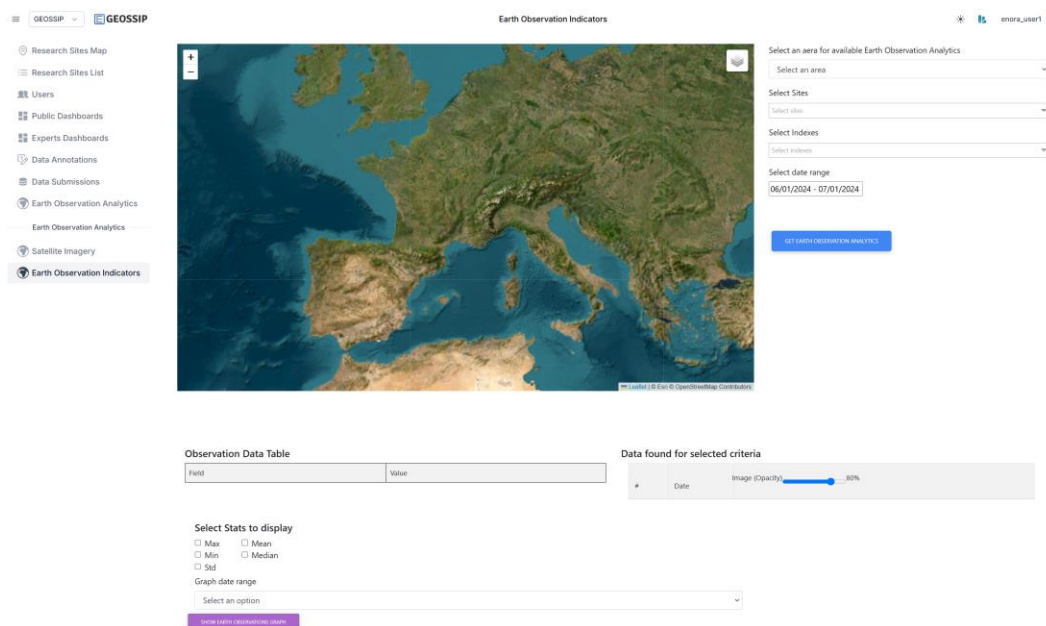


Figure 10 Earth Observation Indicators

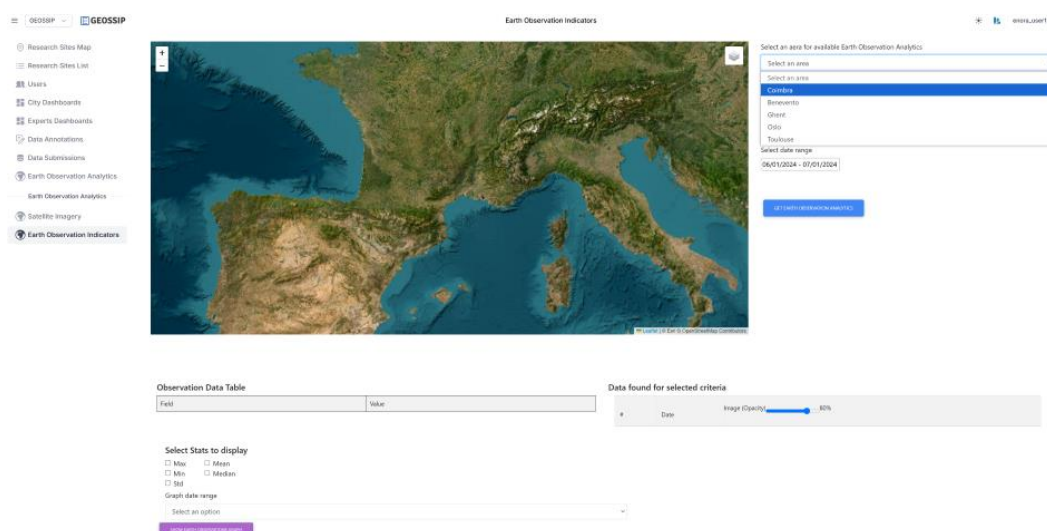


Figure 11 Earth Observation Indicators - Area selection

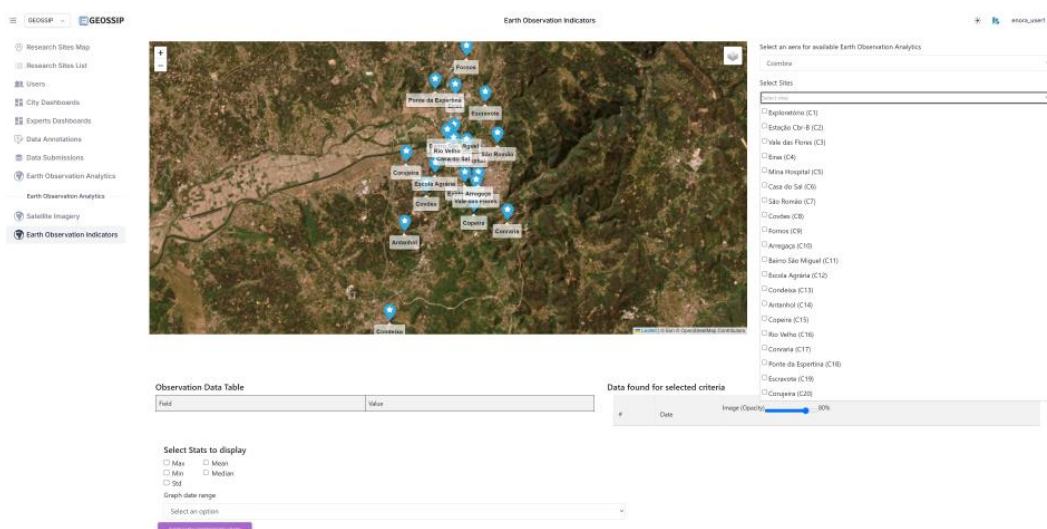


Figure 12 Earth Observation Indicators - Site selection

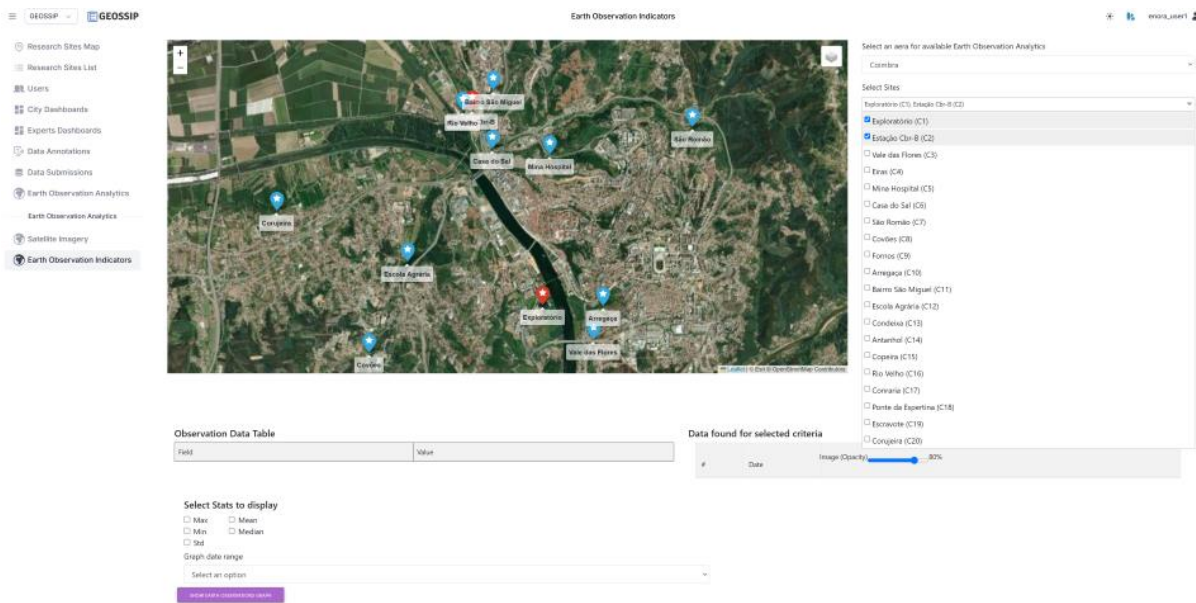


Figure 13 Earth Observation Indicators - Multiple Sites Selection

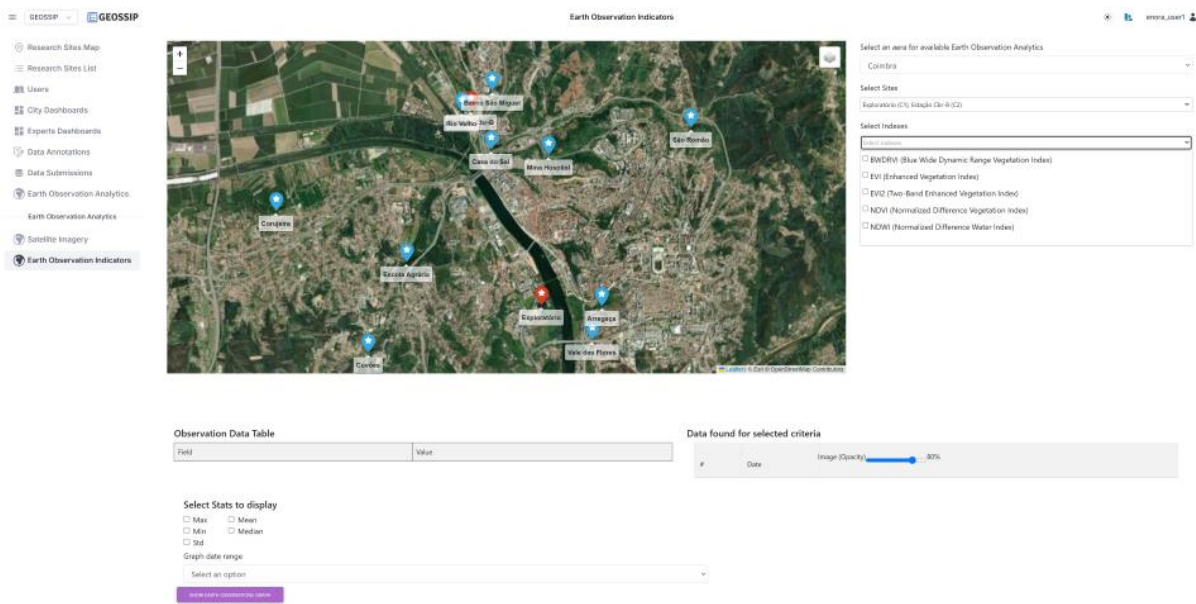


Figure 14 Earth Observation Indicators - Indices Selection

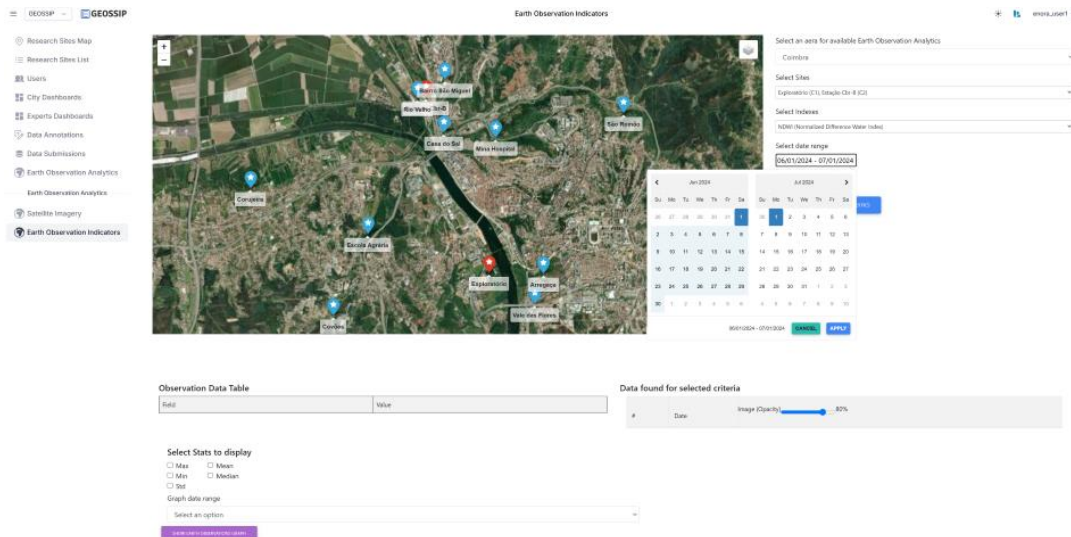


Figure 15 Earth Observation Indicators - Data Range Selection

Upon making these selections, the user can initiate the data retrieval process by clicking the 'Get Earth Observation Analytics' button.

The Observation Data Table then presents the relevant key metrics, as Figures 16 to 21 display, while an opacity slider enables the comparison of multiple data layers.

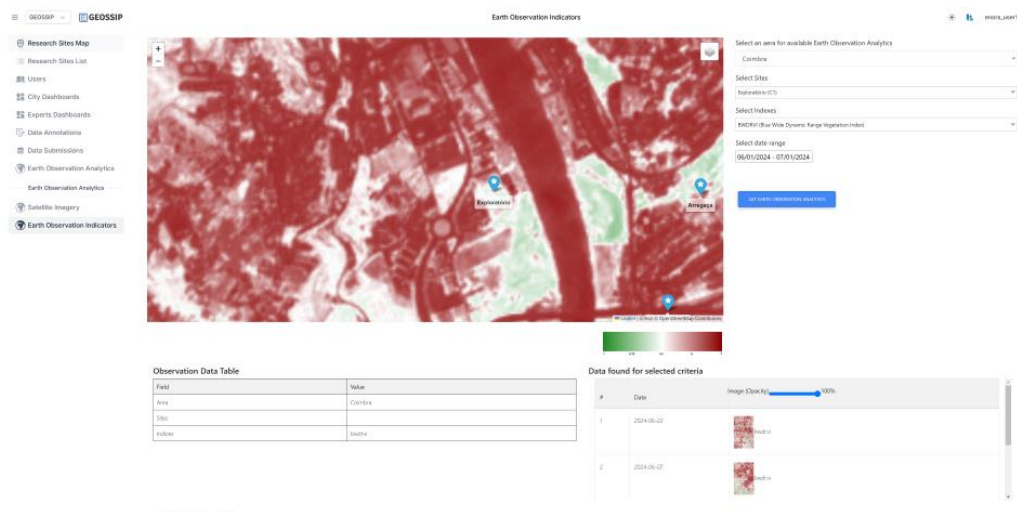


Figure 16 Earth Observation Indicators - BWDRVI result

D5.2 Geospatial and satellite information platform vs2

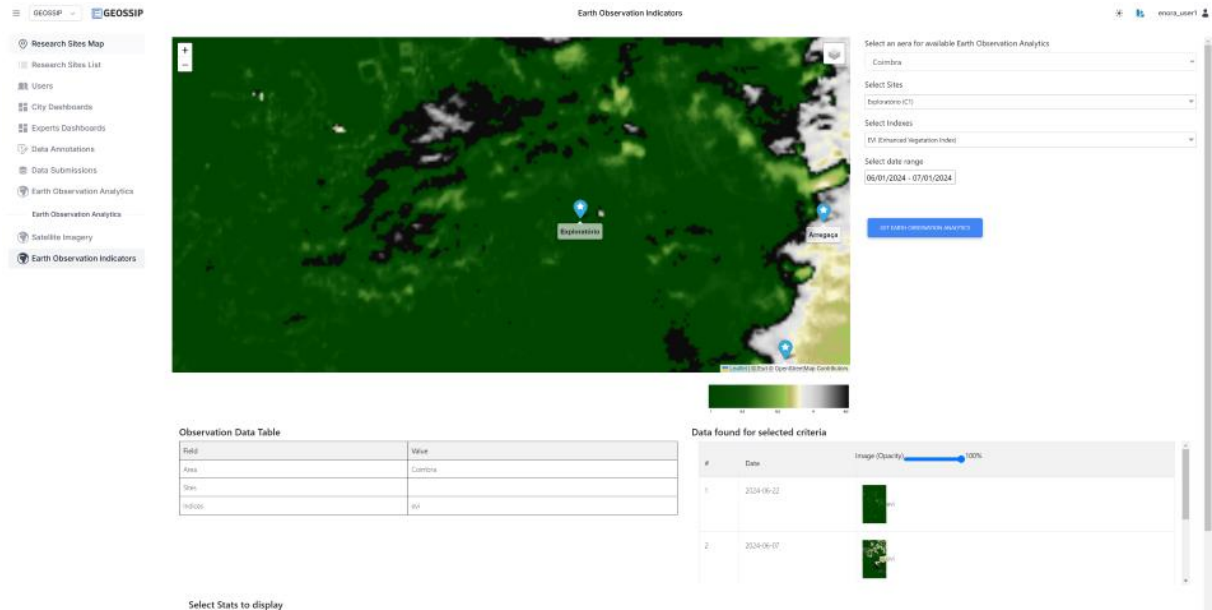


Figure 17 Earth Observation Indicators - EVI result

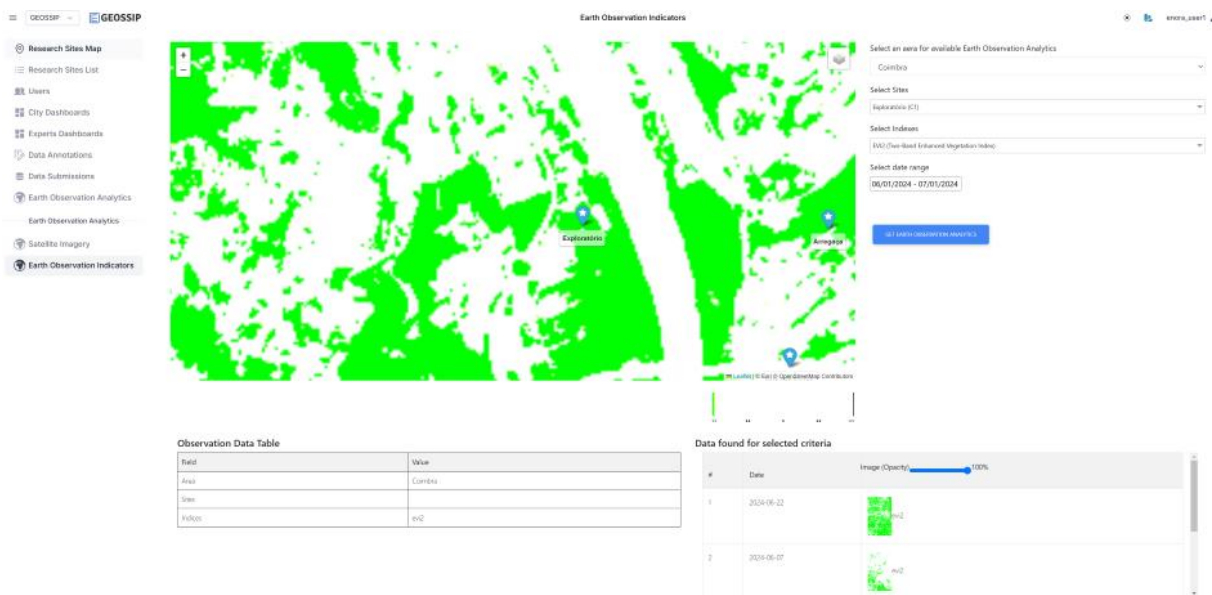


Figure 18 Earth Observation Indicators - EVI result

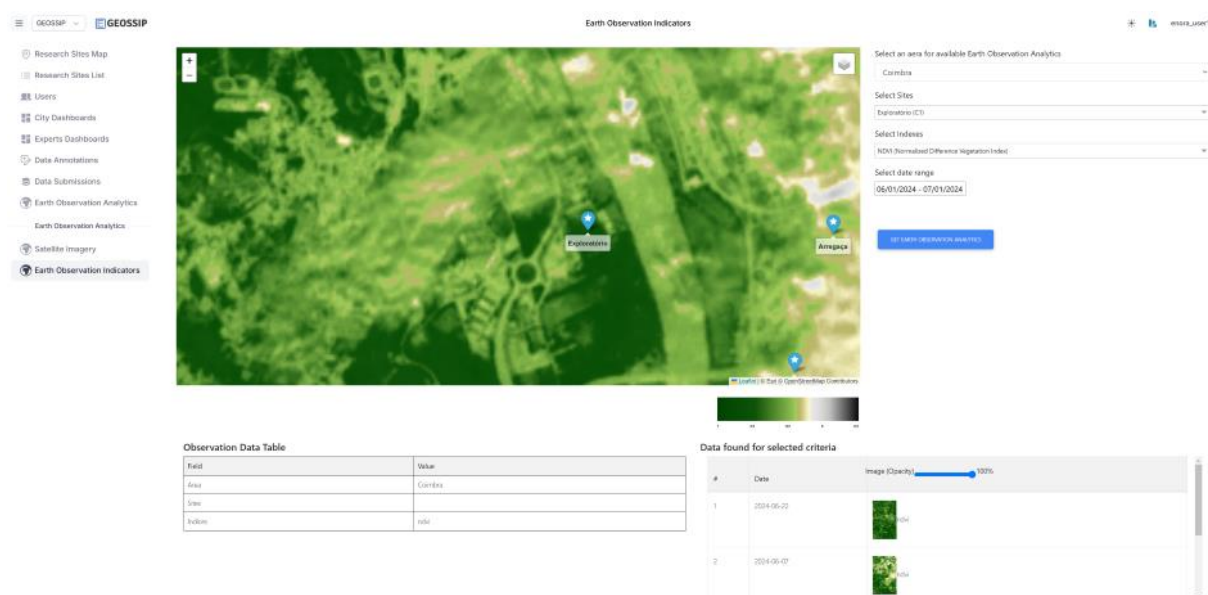


Figure 19 Earth Observation Indicators - NDVI result

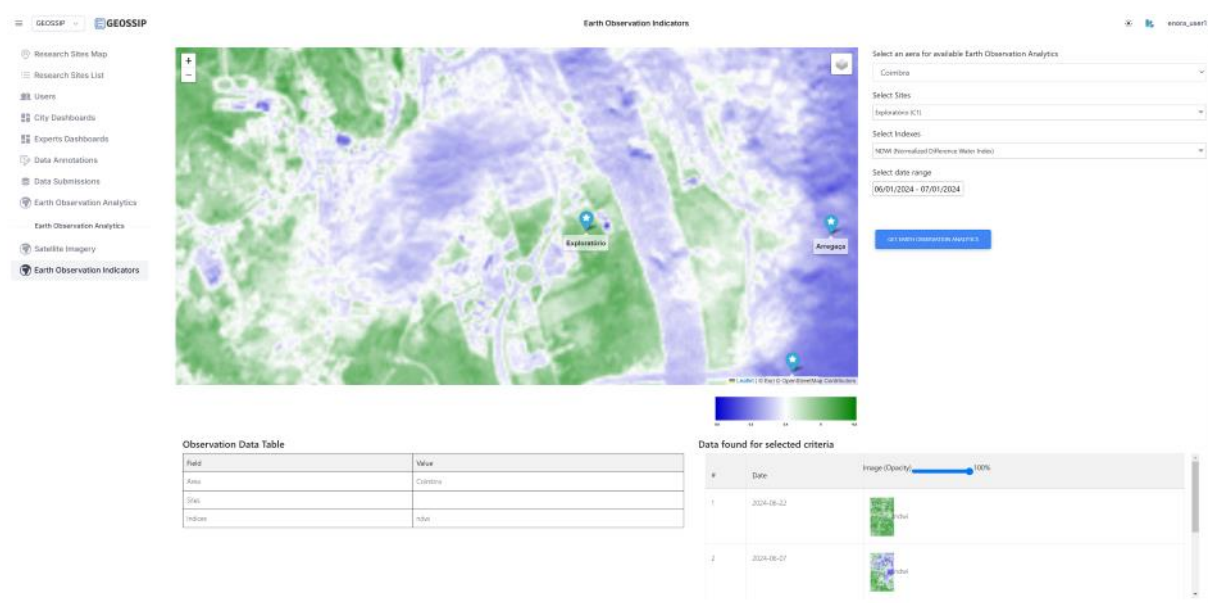


Figure 20 Earth Observation Indicators - NDWI result



Figure 21 Earth Observation Indicators - Multiple sites and indices selection

Additionally, the 'Show Earth Observations Graph' option allows the user to visualise trends, analyse temporal changes, review the analytics for the selected index, and export the data to a file. Figures 22, 23 and 24 show these functionalities.

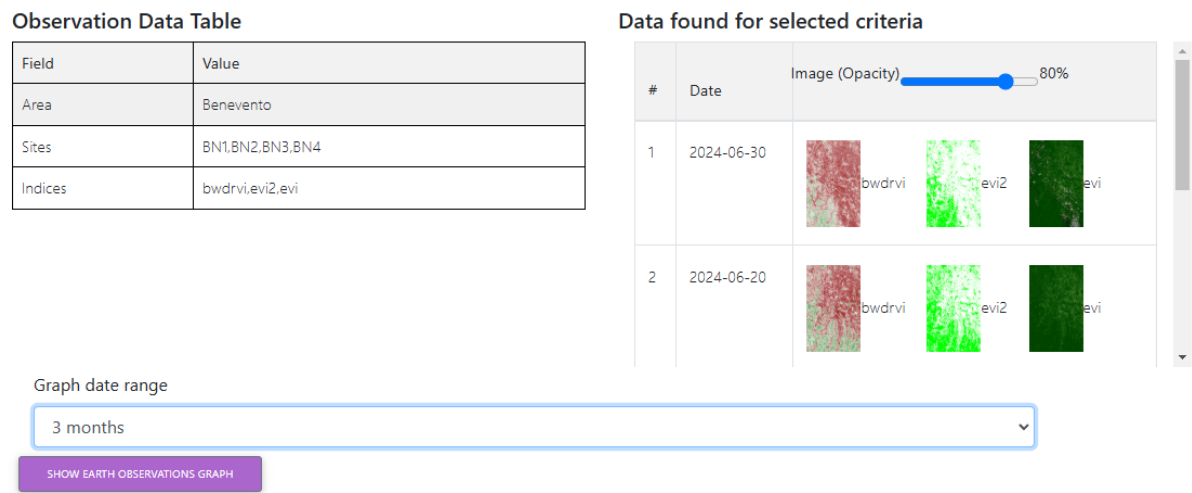


Figure 22 Earth Observation Indicators - Graph date and historical date selection

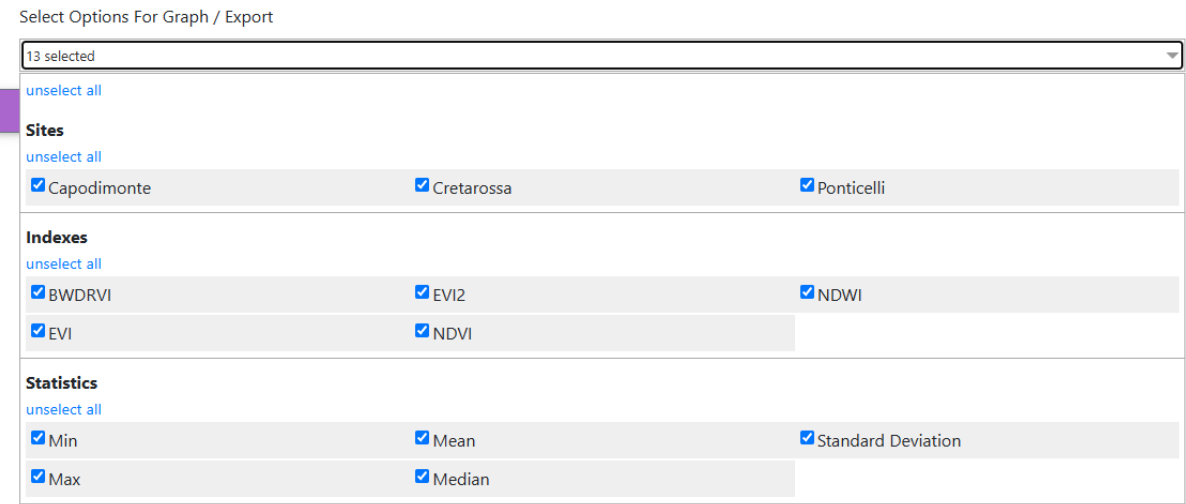


Figure 23 Earth Observation Indicators - Export data selection

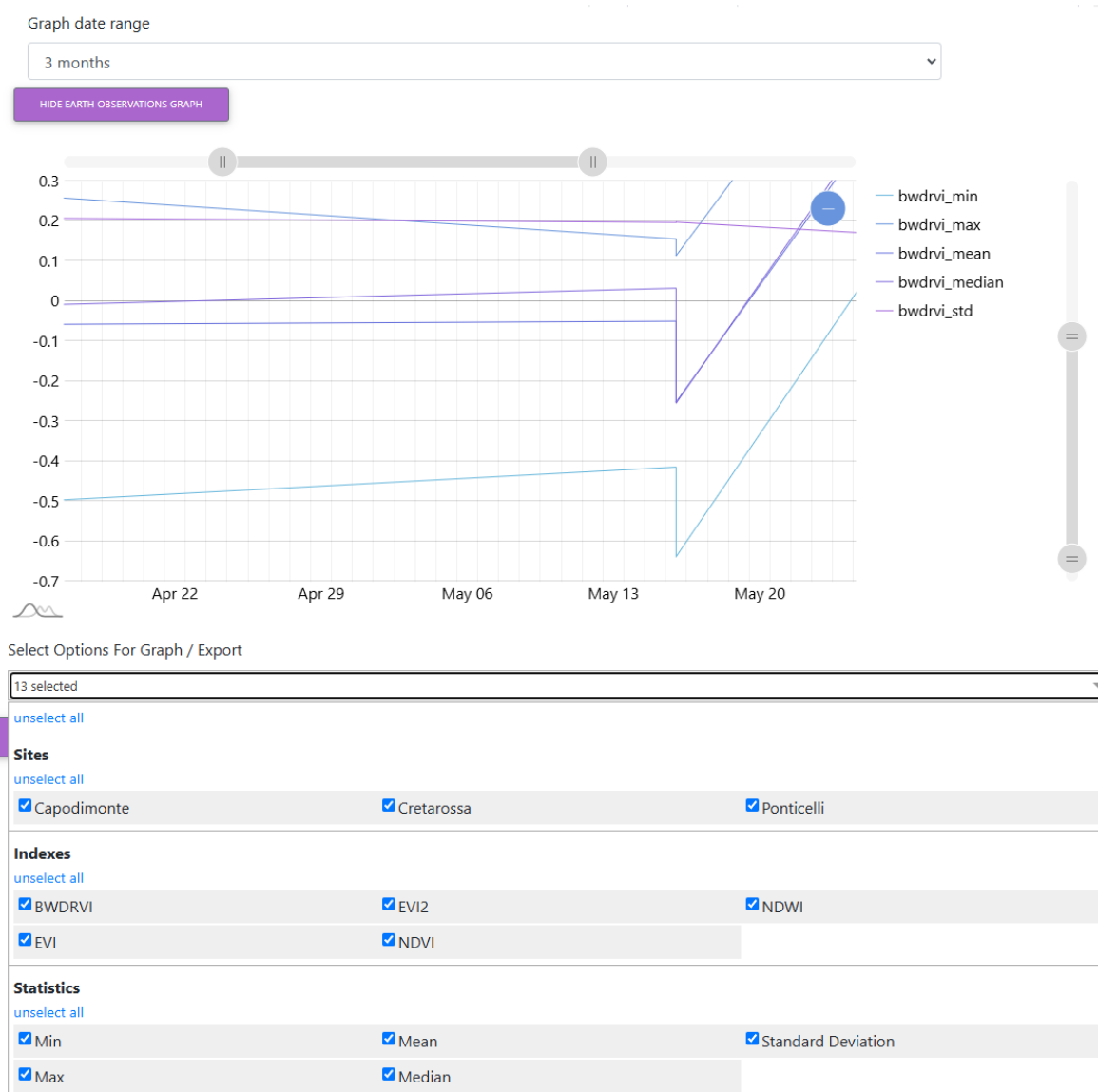


Figure 24 Earth Observation Indicators - Visualisation of Graph

5 Justification of tools development and the efficiency of the development methodology

The development of GEOSSIP and Geospatial and Satellite Information services is guided by the strategic selection of tools and methodologies designed to enhance efficiency, scalability, and maintainability. By adopting versatile programming frameworks, robust containerisation solutions, and streamlined data management practices, the platform ensures adaptability to evolving project needs. This section provides an overview of the rationale behind these choices and highlights how the development approach supports delivering high-quality, reliable services effectively.

5.1 Relevance of open source and open access tools to OAH research objectives

Adopting open-source and open-access tools aligns seamlessly with the research objectives of OAH, promoting accessibility, reproducibility, and innovation. Open-source libraries provide essential functionalities for geospatial and satellite data processing while ensuring that workflows remain transparent and adaptable. Using the SpatioTemporal Asset Catalog (STAC) specification, a community-driven standard, exemplifies the commitment to leveraging widely recognised, open-access solutions. This approach enhances collaboration with stakeholders, researchers, and organisations by using tools that are both accessible and well-supported by the global community.

Open-access tools also reduce dependency on proprietary systems, mitigating risks associated with vendor lock-in and enabling the long-term sustainability of project outcomes. This openness fosters innovation, allowing researchers to customise tools to meet specific analytical needs while benefiting from active community support. Furthermore, such tools empower stakeholders in regions with limited resources to replicate and extend the methodologies developed, thereby amplifying the project's impact.

5.2 Cost-benefit justification of using existing commercial solutions

Leveraging existing commercial solutions for developing and operating a Geospatial and Satellite Information Platform offers significant cost and operational benefits. First and foremost, commercial platforms like AWS provide a well-established, scalable infrastructure that eliminates the need for upfront investments in hardware and its ongoing maintenance. For example, the AWS Lambda services used for periodic and on-demand processing ensure that computational resources are allocated dynamically, enabling cost-effective scaling without requiring over-provisioning.

Moreover, commercial solutions come with robust security, compliance, and reliability measures built in. As a cloud service provider, AWS adheres to stringent security standards and offers tools like identity and access management (IAM) to secure sensitive Earth Observation (EO) data. This reduces the need for in-house expertise in security architecture, thereby saving time and resources. Integrating commercial services with established data standards, such as the Earth Search STAC Catalog, further ensures interoperability and access to a wide range of data providers, enhancing data usability without additional development overhead.

In addition to cost efficiency, commercial solutions often include advanced tools and pre-built functionalities. AWS services, such as S3 for data storage and EC2 for computation, enable the

seamless handling of massive EO datasets. Furthermore, the ability to integrate advanced analytics, machine learning capabilities, and visualisation tools accelerates the deployment of applications, reducing time-to-market. For example, the ability to generate image composites directly within the AWS ecosystem provides immediate visualisation support, which is vital for end-user engagement.

Lastly, relying on commercial solutions ensures ongoing support and updates. Providers like AWS continuously evolve their services, ensuring compatibility with the latest technologies and standards. This approach reduces the technical debt often associated with maintaining custom-built infrastructure and ensures that the project can adapt to changing requirements and technologies, supporting long-term sustainability.

5.3 Open-source libraries used for the development of Geospatial and Satellite Information Platform

The Geospatial and Satellite Information Platform development relies heavily on open-source libraries, which offer cost-effective and highly customisable solutions for handling EO data. By using Python-based libraries such as pystac, pystac-client, and rasterio, the platform benefits from robust tools tailored for managing geospatial data workflows. These libraries are designed to handle large datasets efficiently and integrate seamlessly with commercial services like AWS to optimise data access and processing.

Open-source tools like rioxarray and xarray allow developers to work with multi-dimensional raster data, enabling sophisticated analysis and visualisation. This flexibility is critical for processing Sentinel-2 L2A bands and generating spectral indices, such as NDVI and EVI. Libraries like Geopandas and shapely provide powerful geospatial analysis capabilities, allowing the creation of custom cloud masks to refine data quality. The modularity of these tools facilitates rapid development and ensures that the platform can adapt to evolving requirements without incurring additional licensing costs.

A significant advantage of open-source libraries is the active community support and frequent updates. These communities often provide extensive documentation and examples, reducing the learning curve for developers and ensuring that the platform remains aligned with best practices. Additionally, open-source tools ensure transparency, as developers can audit and modify the code to meet specific project needs, fostering innovation and collaboration.

5.3.1 Flexibility of the solution and added value for the sustainability of project results

Integrating open-source libraries with AWS services ensures a flexible and sustainable solution for managing EO data. This hybrid approach leverages commercial platforms' scalability and reliability while maintaining open-source tools' adaptability. For instance, AWS Lambda functions can be tailored to execute custom workflows written in Python, combining the benefits of serverless architecture with the precision of community-developed algorithms.

This flexibility extends to the ability to handle diverse datasets. The platform can process Sentinel-2 data systematically while accommodating on-demand processing for data validation or new analyses. This dual-mode functionality ensures the platform remains relevant and can address current and emerging user needs. The project ensures data reliability by validating Sentinel-2 indices against Landsat data, enhancing user confidence and fostering long-term adoption.

Sustainability is further supported by the modular design of the platform. Standard protocols, such as the STAC specification, ensure interoperability with other systems and datasets, reducing vendor lock-in and supporting integration with future technologies. Additionally, by storing processed data in widely accessible formats on S3, the platform provides a reusable repository of high-quality data products, amplifying its value for subsequent research and applications.

Finally, the reliance on open-source libraries contributes to sustainability by minimising costs and promoting collaboration. Organisations can build on the platform's foundation without being constrained by proprietary restrictions, enabling further innovation. This approach aligns with the project's goals of creating a cost-effective and impactful solution in the long term, ensuring that the results remain accessible and valuable to stakeholders.

6 Data Backoffice Development updates

Recent updates to the software ecosystem have significantly expanded its functionality and usability. Highlights include:

- **City Dashboards:** Developed to visualise urban metrics, these dashboards provide stakeholders with a comprehensive overview of key environmental and health indicators, enabling data-driven decision-making.
- **Citizen Science App Enhancements:** The Citizen Science App has been upgraded with new wellness-related questions, updated guidelines and instructions with images, user functionality to add sampling sites, and improved UI/UX for greater accessibility and engagement.

The following sections provide a detailed overview of all the updates and enhancements made to the system.

6.1 The Data Backoffice architecture

The Data Backoffice architecture, as presented in the D5.3 Digital app and backoffice, is designed to be robust, scalable, and flexible, supporting the platform's objectives to monitor and improve aquatic ecosystems and human health. It integrates diverse data sources and systems, ensuring seamless data persistence, processing, and visualisation through a structured framework consisting of three main layers: Persistence, Controller, and View. These layers collaborate to manage data storage, transformation, and user interaction effectively. The Persistence Layer serves as the foundation for storing and managing data. This layer is available to integrate external data sources, including Open Data APIs, EU Data Spaces, and publicly available scientific datasets, ensuring data coverage to meet the platform's goals. The Controller Layer orchestrates data flow across the system, with its Integration Backend ensuring reliable data integration and transformation. This includes managing APIs, authentication, and data consistency.

Meanwhile, the Machine Learning Backend processes the ingested data using algorithms to generate actionable insights, essential for informed decision-making and analysis. The View Layer focuses on providing an intuitive user experience. It includes a front-end interface for data visualisation and interaction and headless routes for API-driven data access. This setup enables users to explore and

manipulate data through interactive dashboards and contribute through the Citizen Science App, fostering active engagement and participation. External systems are seamlessly integrated into the architecture. For example, the Earth Observation platform supplies critical geospatial and environmental data via APIs, enriching the platform's dataset. The OAH Information Hub is a central repository, supporting applications like the Citizen Science App and City dashboards that enhance data dissemination and user engagement.

The software architecture of the OneAquaHealth backend is shown in Figure 25.

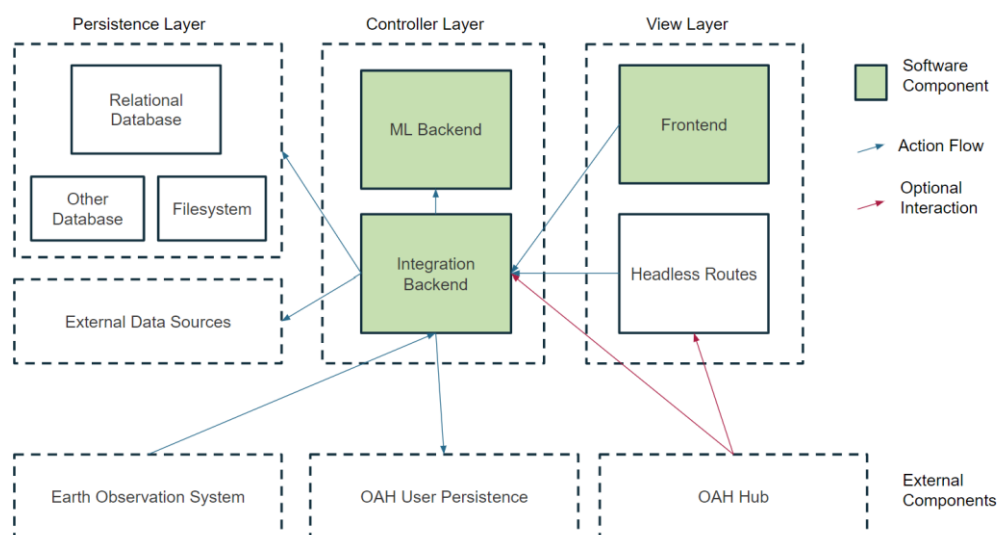


Figure 25 The software architecture of the OneAquaHealth backend

6.2 GEOSSIP: The OAH backend platform – development updates

The following Figures 26-31 provide an overview of the GEOSSIP platform's functionalities, designed to meet the backend architecture requirements of the OneAquaHealth (OAH) initiative. Serving as a centralised access point, GEOSSIP integrates a comprehensive suite of tools to support OAH's research and decision-making. The platform enables seamless interaction with diverse data sources, including Earth Observation (EO) tools, in-situ measurements, and citizen science data, all through a unified interface.

GEOSSIP streamlines data analysis by visualising and processing environmental, biological, and health-related data. It consolidates information from various domains, ensuring users have access to the latest and most accurate datasets for informed decision-making. Additionally, the platform's flexible architecture allows for easy integration of new data sources and research needs, supporting continuous improvements in environmental monitoring and management strategies.

GEOSSIP provides a user-friendly interface for accessing and managing data related to the OneAquaHealth project. Upon navigating to the "Research Sites" section, users are directed to a login screen where they can authenticate, granting them access to the system. After logging in, users are presented with a map displaying the five primary cities involved in the project. The map allows users to zoom in and view the available research sites, and a tabular list provides the exact coordinates for each location. Users can also interact with a dashboard to select a city and research site, and specify the date and quantity of interest, triggering both tabular and graphical representations of the data. The platform organises data into specific according to different environmental categories:

Pharmaceuticals, Nutrients, Macroinvertebrates, Microbiomes, Fish, Diptera, Diatoms, Birds, Hydromorphology, all of which are available for exploration. In the "Data Submissions" section, users can view previously submitted data, add new observations for any research city or site, and complete specific categories related to environmental types. Successful submissions are stored in the backend database for future access and analysis.



Figure 26 Login Screen

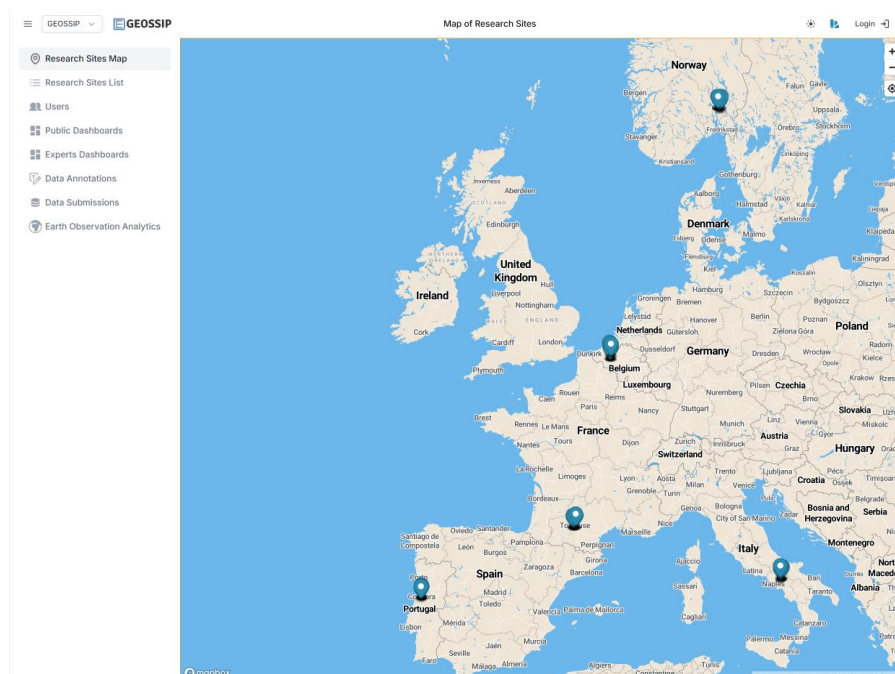


Figure 27 Research Sites Map

D5.2 Geospatial and satellite information platform vs2

GEOSIP GEOSSIP

List of Research Sites

Research Sites Map
Research Sites List
Users
Public Dashboards
Experts Dashboards
Data Annotations
Data Submissions
Earth Observation Analytics

Code	Name	City	Latitude	Longitude	Altitude	
C1	Exploratório	Coimbra	40.1500	-8.3500	15	
C2	Estação Cbr-8	Coimbra	40.1500	-8.3500	15	
C3	Vale das Flores	Coimbra	40.1500	-8.3500	15	
C4	Eiras	Coimbra	40.1500	-8.3500	15	
C5	Minha Hospital	Coimbra	40.1500	-8.3500	15	
C6	Casa do Sal	Coimbra	40.1500	-8.3500	15	
C7	São Romão	Coimbra	40.1500	-8.3500	15	
C8	Covões	Coimbra	40.1500	-8.3500	15	
C9	Fornos	Coimbra	40.1500	-8.3500	15	
C10	Arreagaça	Coimbra	40.1500	-8.3500	15	
C11	Bairro São Miguel	Coimbra	40.1500	-8.3500	15	
C12	Escola Agrária	Coimbra	40.1500	-8.3500	15	
C13	Condeixa	Coimbra	40.1500	-8.3500	15	
C14	Antanho	Coimbra	40.1500	-8.3500	15	
C15	Copela	Coimbra	40.1500	-8.3500	15	
C16	Rio Velho	Coimbra	40.1500	-8.3500	15	
C17	Conraria	Coimbra	40.1500	-8.3500	15	

Figure 28 List of Research sites and coordinates

GEOSIP GEOSSIP

List of Biotic Species

Research Sites Map
Research Sites List
Users
Public Dashboards
Experts Dashboards
Data Annotations
Data Submissions
Earth Observation Analytics
Data Annotations
Biotic Species List
Nutrient Types List
Pharmaceutical Types List
Hydromorphology List

Code T ₁	Category T ₁	Name T ₁
AC001	Algae	Acetabularia medusa-like - constrictum (Greville) Andersen, Stegman & Wirth
AC002	Algae	Acetabularia medusa-like (Greville) Andersen
AC003	Algae	Acetabularia medusa-like (Greville) Andersen
AC004	Algae	Acetabularia medusa-like (Greville) Andersen
AC005	Algae	Acetabularia medusa-like (Greville) Andersen
AC006	Algae	Acetabularia medusa-like (Greville) Andersen
AC007	Algae	Acetabularia medusa-like (Greville) Andersen
AC008	Algae	Acetabularia medusa-like (Greville) Andersen
AC009	Algae	Acetabularia medusa-like (Greville) Andersen
AC010	Algae	Acetabularia medusa-like (Greville) Andersen
AC011	Algae	Acetabularia medusa-like (Greville) Andersen
AC012	Algae	Acetabularia medusa-like (Greville) Andersen
AC013	Algae	Acetabularia medusa-like (Greville) Andersen
AC014	Algae	Acetabularia medusa-like (Greville) Andersen
AC015	Algae	Acetabularia medusa-like (Greville) Andersen
AC016	Algae	Acetabularia medusa-like (Greville) Andersen
AC017	Algae	Acetabularia medusa-like (Greville) Andersen
AC018	Algae	Acetabularia medusa-like (Greville) Andersen
AC019	Algae	Acetabularia medusa-like (Greville) Andersen
AC020	Algae	Acetabularia medusa-like (Greville) Andersen

Figure 29 Categories of Data Annotations types



6.3 City Dashboards

The City dashboards provide a tool for visualising urban environmental metrics, offering insights into the health of urban aquatic ecosystems to the public that were selected by the UC team. These dashboards display various indices important for monitoring the ecosystem quality and biological diversity in urban areas, as Figures 32 and 33 presents. Among the key indices presented is for example the Status of Nitrate, which tracks nitrate levels in water, an important indicator of water quality, which is the focus of the Nitrates directive (integrated in the Water Framework Directive of 2000). Elevated nitrate levels can indicate pollution, often from agricultural runoff or wastewater discharge, and monitoring these levels helps identify areas at risk of eutrophication, which can harm aquatic life.

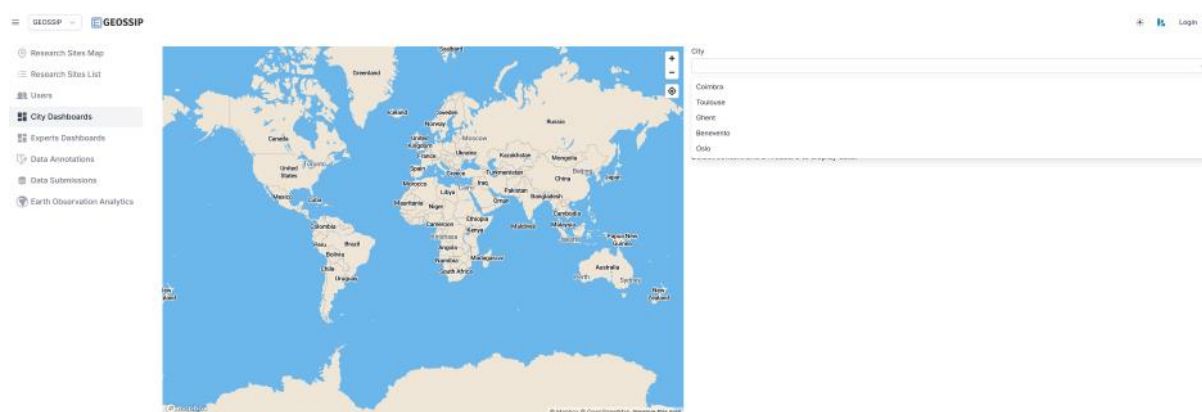


Figure 32 City Dashboards - Selection of Site

In addition to water quality metrics, the City dashboards include various biological quality indices. These include the Species Richness, which indicates the total number of species in a given area, and several diversity indices that provide more nuanced insights into ecosystem balance. The Simpson's Index measures the probability that two randomly selected individuals from a sample will belong to the same species, helping assess species dominance and biodiversity. Similarly, the Shannon-Weiner Index quantifies biodiversity by considering both species richness and the distribution of individuals among species, offering a more detailed view of ecological balance. Pielou's Evenness Index assesses the evenness of species distribution, while Margalef's Index estimates species richness relative to the number of individuals sampled. Additional indices are biological quality indices applied to fish,

macroinvertebrate and diatoms' data collected at the research sites and streams, which are under the scope of ecological monitoring of rivers in the Water Framework Directive (2000).

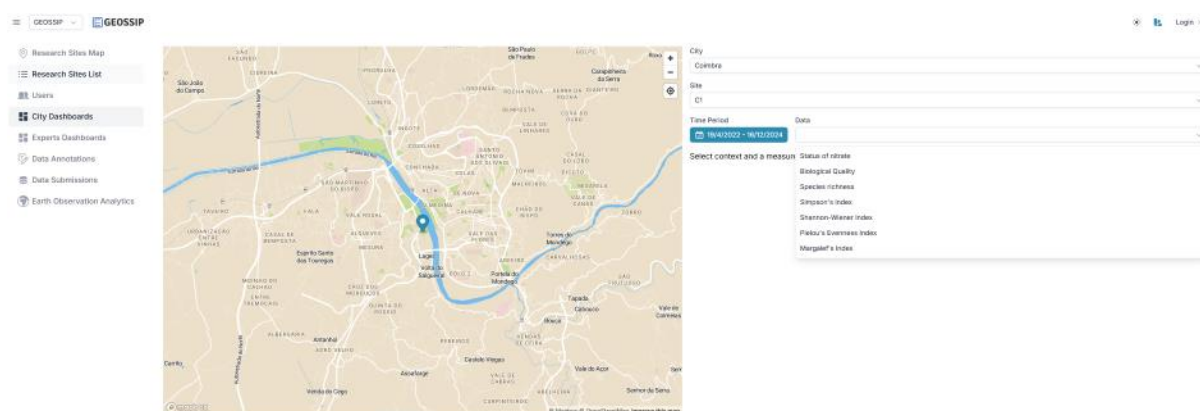


Figure 33 City Dashboards - Selection of Data

An example of these Dashboards are presented in the following Figures 34 and 35.

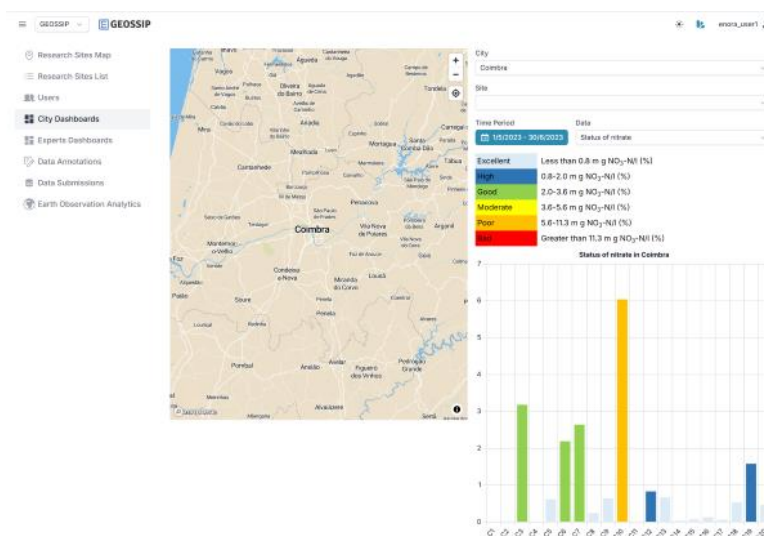


Figure 34 City Dashboards - Status of nitrates

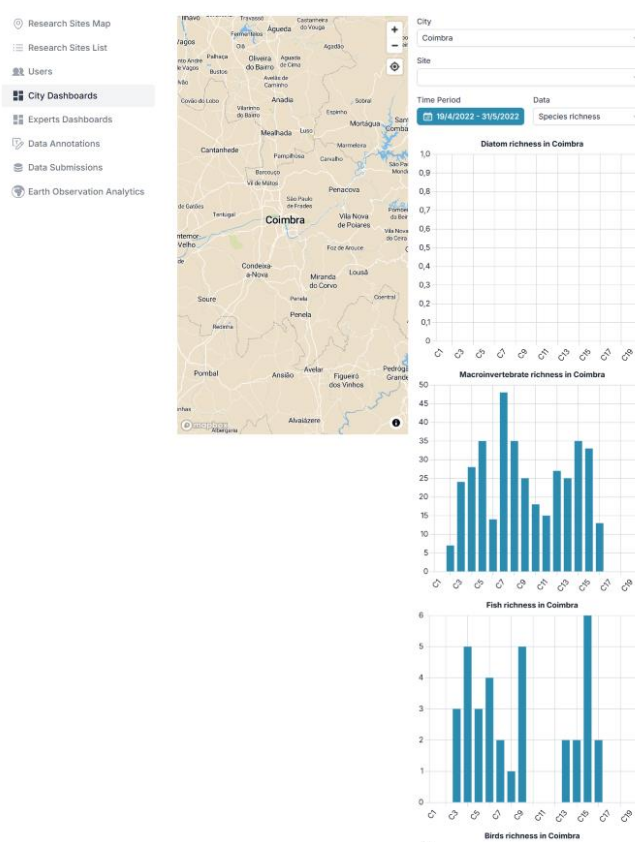


Figure 35 City Dashboards - Species richness

These indices are visualised on the city dashboards and made accessible via API to the Open Information Hub, ensuring that stakeholders and decision-makers can easily retrieve and utilise the data. By integrating these indices into the Open Information Hub, OAH enhances the ability of researchers, policymakers, and the public to access real-time data on urban ecosystems, facilitating informed decisions and timely interventions to preserve biodiversity and improve water quality.

6.4 Citizen Science App

The software implementation updates encompass a series of enhancements to the Citizen Science App, aimed at improving both functionality and user experience. Key updates include the addition of new wellness-related questions, designed to expand the scope of data collection and enhance the app's utility for health and environmental research. Additionally, the app's guidelines and instructions have been updated, incorporating visual aids to improve clarity and ease of use for participants. Another significant enhancement is the introduction of functionality that allows users to add new research sites, further expanding the app's flexibility in gathering diverse data across different locations. To complement these features, the user interface (UI) and user experience (UX) have been optimised to promote better engagement and accessibility, ensuring that users can interact with the app more intuitively and efficiently. These improvements collectively enhance the app's overall performance, providing users with a more seamless and effective tool for contributing to citizen science initiatives.

The following Figures below provide a comprehensive walkthrough of the Citizen Science App, from installation to submission. They detail each stage of the data entry process, including selecting research sites, adding new locations, uploading multimedia, and answering various environmental and wellness-related questions. The app facilitates both standard and user-generated research site submissions, ensuring flexibility in data collection. Additionally, users are guided through each step with instructional screens and visual aids, promoting accurate and effective participation in the OneAquaHealth initiative. These figures demonstrate the app's commitment to user engagement, ease of use, and contribution to valuable scientific research.

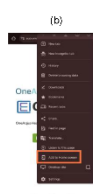
Figure 36 illustrates the step-by-step installation process of the Citizen Science App on users' devices, ensuring easy access for participants.

Installation of the OAH Citizen Science App

Step 1. Open your phone's browser and visit: <https://app.enora-oah.eu/>

Step 2. Look for the app installation prompt:

- If a pop-up window appears suggesting app installation, follow the onscreen instructions to proceed
- If no prompt appears, continue to the next step



Step 3. Add to Home Screen:

- For IOS (a): Tap the share icon (box with arrow) at the bottom of the screen, then select "Add to Home Screen"
- For Android (b): Tap the menu icon (three dots) and choose "Add to Home Screen" or "Install App"

Step 4. Confirm the installation by tapping "Add" or "Install"



Step 5. Find the newly installed app icon on your home screen and tap to launch

OneAquaHealth

OneAquaHealth
GEOSSIP
Open your phone's browser and visit: <https://app.enora-oah.eu/>

OneAquaHealth
GEOSSIP
Open your phone's browser and visit: <https://app.enora-oah.eu/>



Figure 36 Citizen Science App - Installation guide

The introductory screen of the app, in Figure 37, providing a welcome message and a brief overview of its functionalities for new users.



Figure 37 Citizen Science App Intro Screen

This step in Figure 38 explains the initial process users must complete before submitting any data, setting the stage for a structured submission process.

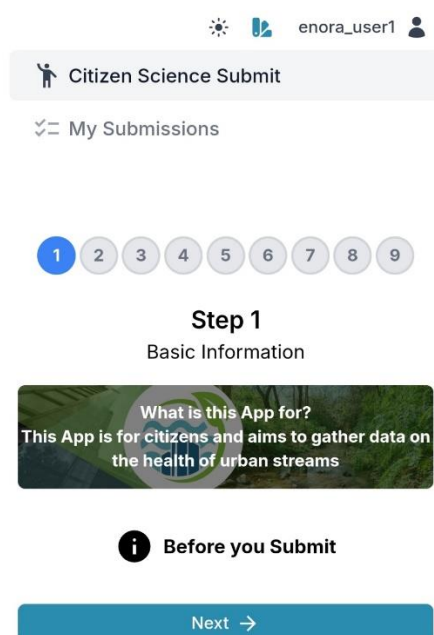


Figure 38 First Step before submission

A brief overview of the app's purpose and its role in citizen science, is presented in Figure 39 preparing users for the data collection process. This introductory text with the respective pictures was elaborated by the freshwater researchers of the UC team.

What is an Urban Stream? – Streams in all landscapes!

When asked to visualize a running stream, you won't envision skyscrapers, office buildings, city buses, or stores. Nonetheless, there are urban streams coexisting with urban development, typically greatly influenced by the humans residing in their watershed. Streams are vital components of all landscapes, including forests, cities, and suburbs. Streams, stream banks, and the lowlands around them provide important habitat for animals, plants and many other organisms that share landscape with us. They are also part of a drainage network for rain and melting snow from our streets, parks, and yards.

Healthy Stream Ecosystems

Healthy streams must be surrounded by specific vegetation – riparian vegetation – capable of living with their roots in saturated or high humidity soil without rotting. Forested streams are full of life (from micro to macroscopic, terrestrial to aquatic) including mammals, birds, fish, amphibians, reptiles, terrestrial and aquatic invertebrates, fungi and high biomass of microorganisms

~~constituting intricate food networks. Some of~~

Figure 39 Before you submit - About the App

Further details on the app's goals, how it contributes to scientific research, and how users can benefit from participating, are shown in Figure 40.

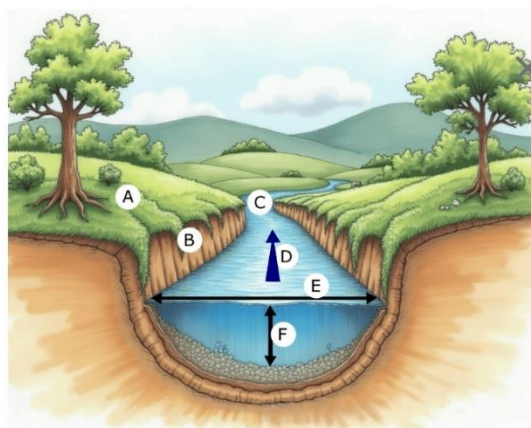


Figure 2. Stream cross section illustration where :
 A) bank top; B) margin; C) main wet channel; D)
 flow direction (downstream view); E) width; and F)
 depth. Illustration of a stream cross-section,
 highlighting key features: A) bank top; B) margin; C)
 main wet channel; D) flow direction (looking
 downstream, with the flow moving away from the
 viewer); E) width; and F) depth.



Figure 3. A simple ruler in many cases is enough to

Figure 40 Citizen Science App/About the App

Figure 41 shows the process of selecting the appropriate research site within the OneAquaHealth framework for data submission.

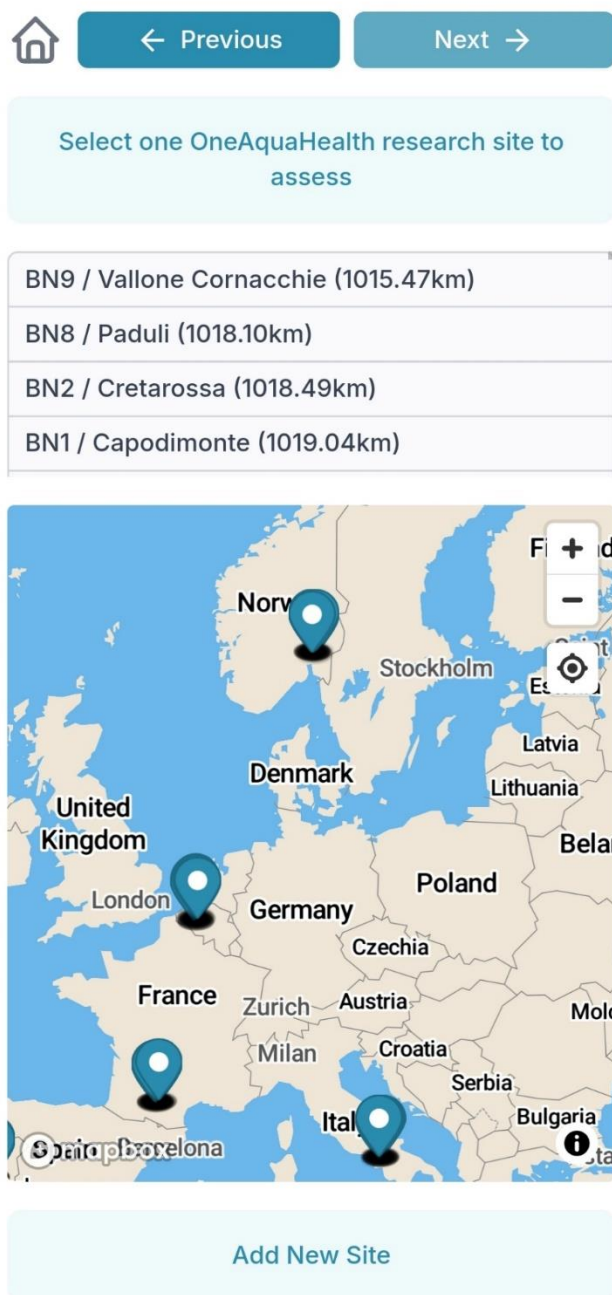
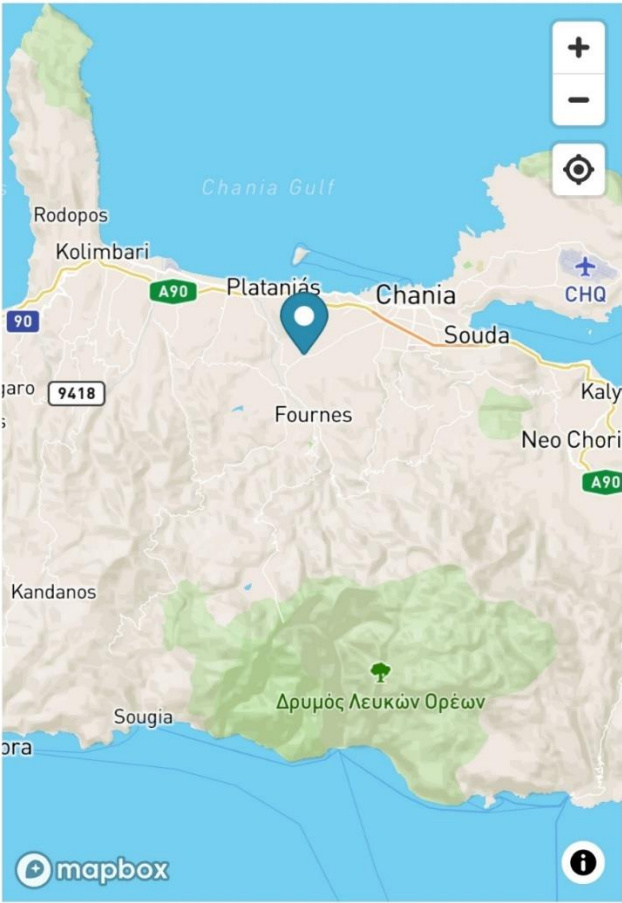


Figure 41 Selection of OAH urban stream sites

Figure 42 illustrates how users can add a new site to the system, expanding the scope of the citizen science data collection.

UGSX1WpWx / Agia



Longitude

23,933392

Latitude

35,475978

Site Name

Agia

Figure 42 New Site Addition

Figure 43 shows the functionality for uploading photos and videos of the selected research site, enhancing the data submission process.

1 2 3 4 5 6 7 8 9

Step 3

Media Upload

← Previous Next →

Select your images

Upstream photo
Upstream photo
Select File

Downstream photo
Downstream photo
Select File

Photo with the surrounding context
Photo with the surrounding context (houses, roads...)
Select File

Photo with an interesting element of biodiversity
Photo with an interesting element of biodiversity (plant, animal, funga, other)
Select File

Make a short video (5 or 10 seconds)
Make a short video (5 or 10 seconds)
Select File

Figure 43 Site Photos and Video Upload

From this step on, a set of questions aiming to characterize several hydromorphological features of the urban streams appear to the user. These questions are relevant information for research and freshwater ecosystems management, and address several aspects related to the health of these ecosystems but that usually are not monitored by the responsible environmental or municipal agencies. By responding to them when visiting an urban stream or by just passing in their way to home or work, the citizens can provide valuable scientific information in a quick way.

All the questions and respective pictures illustrating the various response options were elaborated by the freshwater ecologists of the UC team. Figure 44 illustrates the first question, related to Channel Form.

Step 4

Questions (1/3)

[< Previous](#)[Next >](#)

What do you see from where you stand (in ca. 100m)?

Channel Form

The channel form is...

- ☐ Flat (A)
- ☐ U Shape (B)
- ☐ V Shape (C)
- ☐ Don't know

**Bottom Type**

The bottom of the wet channel is...

- ☐ Natural (A)
- ☐ Artificial (concrete or stones with concrete) (B)
- ☐ Don't know

Figure 44 Set of Questions

Figure 45 presents (in the pop-up box) a guide to help users understand how to properly fill out the set of questions, ensuring accurate data collection.

Habitats
The habitats present are...

Are there any habitats present? ☐

i You may select several types of habitats. To do so just first tick the circle behind the alternatives given, and afterwards select all you identify.



Natural Debris
There are...

Are there any natural debris present? ☐

i



Water Flow
How is the water flowing?

☐ Fast (with waves or high velocity) (A)

☐ Slow (B)

☐ Stagnant/intermittent (C)

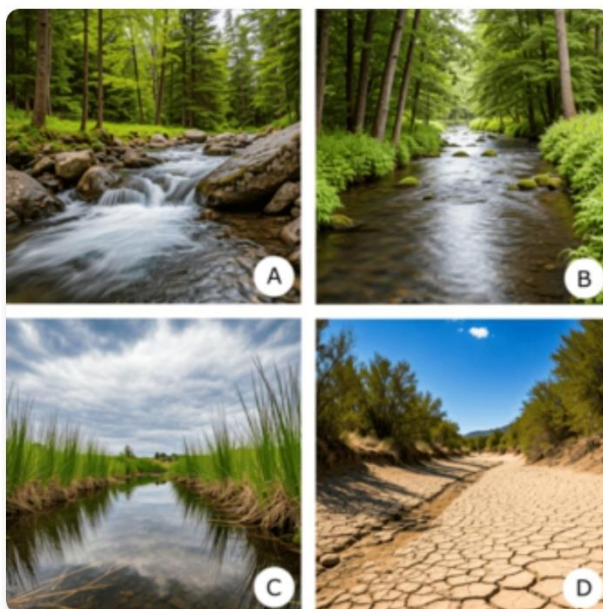
Figure 45 Set of Questions and guide to fill

Figure 46 highlights the integration of images with questions to guide users in providing contextually accurate responses.

Water Flow

How is the water flowing?

- ☐ Fast (with waves or high velocity) (A)
- ☐ Slow (B)
- ☐ Stagnant/intermittent (C)
- ☐ Dry (D)
- ☐ Don't know



← Previous

Next →

Figure 46 Questions with images

The ability for users to add specific measurements or observations as part of their responses to questions is shown in Figure 47.

Water Height

What is the water height?



Please measure in metres – m; and use a "." as a decimal separator

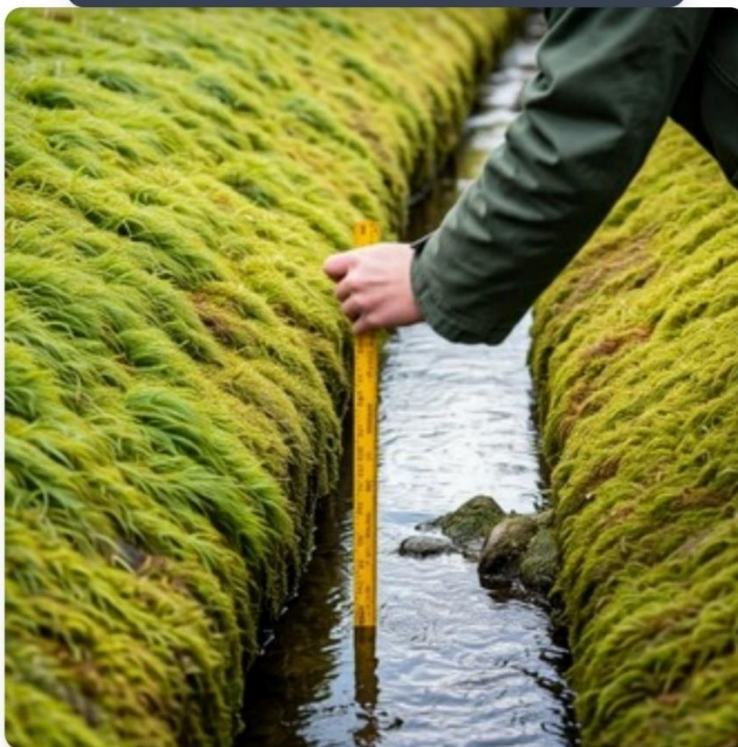
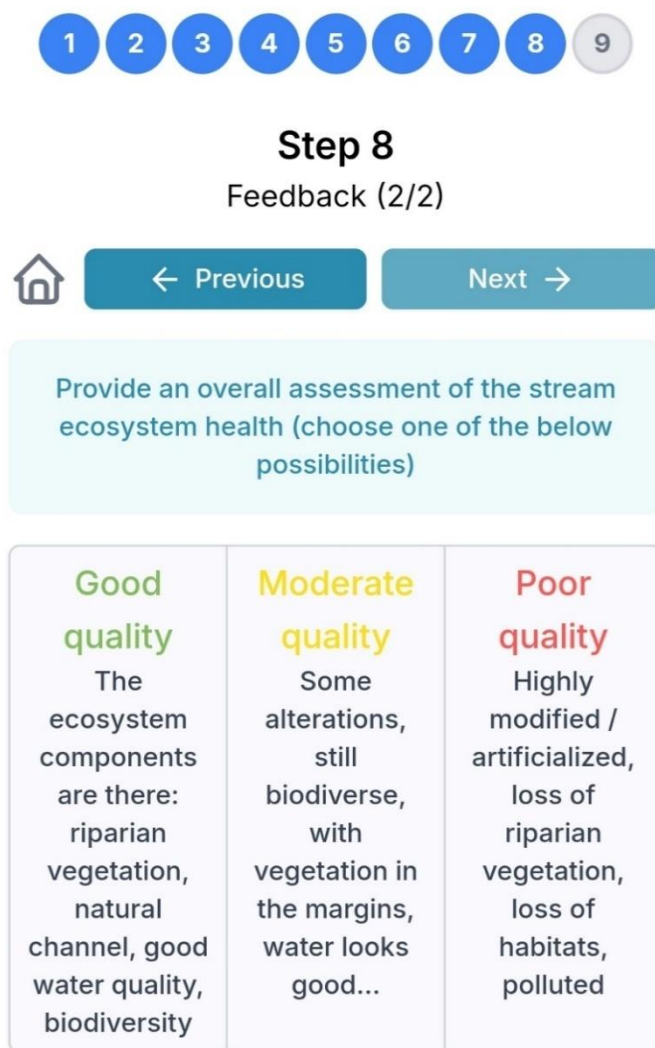
[< Previous](#)[Next >](#)

Figure 47 Question and add measure by user

The last step is a summary screen that allows users to review their responses and assess the overall quality of their data submission, is presented in Figure 48.



1 2 3 4 5 6 7 8 9

Step 8

Feedback (2/2)

← Previous Next →

Provide an overall assessment of the stream ecosystem health (choose one of the below possibilities)

Good quality	Moderate quality	Poor quality
The ecosystem components are there: riparian vegetation, natural channel, good water quality, biodiversity	Some alterations, still biodiverse, with vegetation in the margins, water looks good...	Highly modified / artificialized, loss of riparian vegetation, loss of habitats, polluted

Figure 48 Overall Assessment

Finally, Figure 49 shows a last set of questions that appears to the user that are related to the human wellness that can be provided by the urban stream ecosystems. The user has to answer to four questions about their feelings when visiting the stream and after filling the remaining questions related to the stream characteristics.

Drag your fingers on the slider to rate the intensity of what you are feeling. A higher rating means more intense feeling.

Joy

3

☐ Not Applicable

Serenity

3

☐ Not Applicable

Anger

3

☐ Not Applicable

Fear

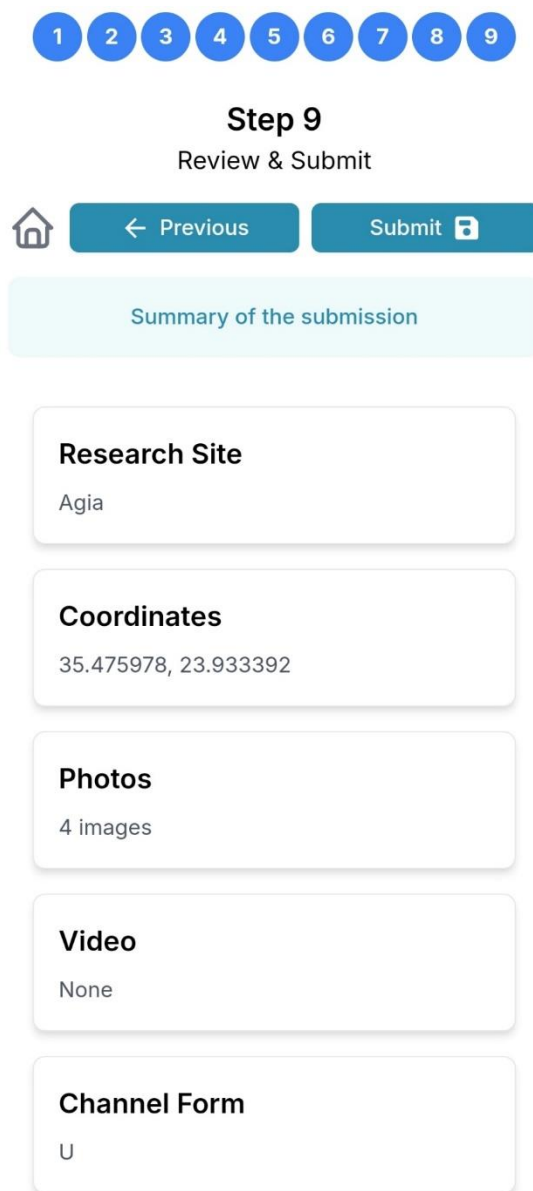
3

☐ Not Applicable

 [< Previous](#) [Next >](#)

Figure 49 Wellness Questions

Figure 50 displays the final summary of the data submission, ensuring that users can review and confirm all entries before finalising their submission.



1 2 3 4 5 6 7 8 9

Step 9

Review & Submit

 [← Previous](#) [Submit !\[\]\(27fbe1a8bc60e9a387bc7db08282bfc2_img.jpg\)](#)

Summary of the submission

Research Site
Agia

Coordinates
35.475978, 23.933392

Photos
4 images

Video
None

Channel Form
U

Figure 50 Submission Summary

Finally, Figure 51 shows the submission process for stream sites that users have generated, called **“Your urban stream sites”**, providing additional flexibility in the data collection process.

Submissions for Generated Sites		
Longitude ↑↓	Latitude ↑↓	Research Site ↑↓
23.933392	35.475978	UGSX1WpWx
23.933392	35.475978	UGSX1WpWx

Figure 51 Submissions for the stream sites generated by the user

7 Harmonisation of planned development activities in WP2, WP5 and WP6

7.1 Integration with WP2 and data collection

The following sub-section introduces a proposed semi-structured data model to enhance flexibility, scalability, and efficiency in handling diverse data submissions while maintaining data integrity and consistency.

The current period is primarily focused on meeting the demands of WP5 and scientific partners, specifically including the following:

- **Administration of Environmental Entities:** The system will include functionalities to manage environmental entities, including the administration of fields and annotations, allowing for efficient categorisation and tracking of data.
- **Automatic Data Transfer:** All data added to pilot sites will be automatically transferred, ensuring integration with the system and preventing data entry discrepancies.
- **Scientific Terminology Management:** The system will be capable of managing and standardising scientific terminology and term codes, enabling consistent data classification and cross-referencing across different datasets.
- **Inclusion of Historical Data:** The platform will support the integration of historical data by allowing the use of a common data format, specifically standardised Excel or CSV files, to upload multiple entries simultaneously. This functionality ensures compatibility with existing datasets and facilitates large-scale data entry.

7.1.1 Limitations and Foundations

The development and implementation of the system are based on several key foundations. First, a common data format for the transfer of multiple entries will be established through agreement among the scientific teams. Once established, any changes to the underlying databases will be reflected in this shared format, ensuring compatibility and enabling seamless updates to the schema without disrupting the ability to upload multiple entries.

Second, the focus of the data development will be on targeted information specifically tailored to support the dashboard and Decision-Support System (DSS). The system is not intended to manage generic scientific information about urban resource streams. Instead, priority will be given to data that directly aligns with and supports these critical project components to ensure their effectiveness.

Third, functional specifications will be clearly defined and approved by relevant stakeholders before the implementation phase begins. This step is essential to ensure that system development adheres to agreed-upon requirements and effectively supports its intended applications.

Finally, the core implementation of the system will be subject to continuous improvement throughout the project lifecycle. Using an iterative approach, the system will be refined and enhanced based on emerging needs and feedback, increasing its overall effectiveness and ensuring it meets evolving project objectives.

7.1.2 Current Implementation

The current system employs a structured database, specifically designed to manage scientific data and information, comprising nine distinct tables, each representing a specific data category, such as pharmaceuticals or nutrients.

Users can input data through a form corresponding to the schema of each table. This approach guarantees that all entries adhere to the defined data structure. Additionally, the system allows users to bulk upload data via CSV or Excel files. However, the uploaded files must align precisely with the structure of the corresponding database table, meaning that any file with additional columns or differing data types will be rejected. This strict adherence to schema is essential for maintaining the integrity of the structured database.

The use of structured databases presents challenges when modifications are required, particularly when adding or altering fields in a data category or table. Schema updates must be applied across all associated interfaces, including forms and batch upload templates. Furthermore, if new or unexpected fields are introduced in future data submissions without an updated schema, these submissions will fail. Any such modifications require backend engineers to validate and update the schema to ensure compatibility with the database.

7.1.3 Proposed Solution

To overcome the limitations of the current structured database and introduce greater flexibility for accommodating future changes, we propose implementing a semi-structured data model as an abstraction layer atop the existing structured database. The proposed approach, as presented in Figures 52-55 retains the structured database for critical functions, such as user authentication, which necessitate a specific schema for security and reliability. Utilising the semi-structured model, the system can adapt to evolving data requirements without compromising the foundational structure required for authentication and other core processes.

The introduced abstraction level would allow:

- Dynamic schema changes, i.e., users could submit data with variable fields without strictly adhering to a predefined schema.
- CSV or Excel files could contain additional fields that are stored without errors or the need for immediate schema updates.
- The system could evolve to include new data categories or fields without significant re-engineering.

The key components of the architecture are as follows:

1. Annotation Table in the Structured Database
 - A dedicated annotations table is included in the structured database.
 - This table contains codes and names for each data category, enabling the management of data categories.
 - Admin users can manage the annotations (e.g., add, edit, or delete them) through a dedicated interface.

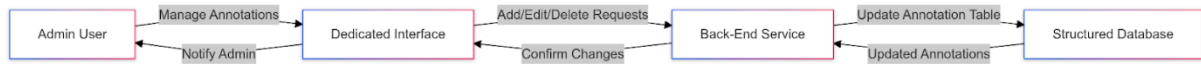


Figure 52 Field changes with annotation Table

2. File Upload from the Front-End

- Users upload the file via the front-end interface after filling out the CSV or Excel template.
- The system allows uploading one file at a time, ensuring manageable and trackable submissions.

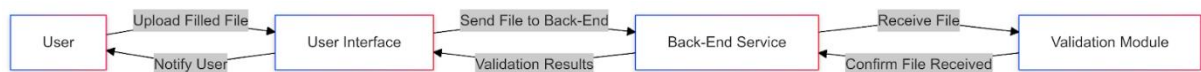


Figure 53 Data Upload

3. Storage and Master File Creation

- If the file passes validation:
 - The data is stored in a semi-structured database.
 - A master file is updated, merging the new submission with all prior submissions.
 - During the merge, the system checks for and removes duplicate entries.

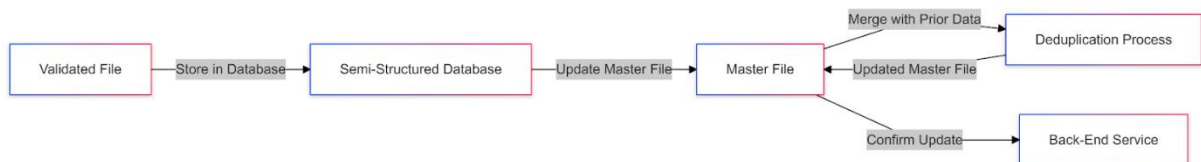


Figure 54 Common data structure creation

4. Data Presentation

- Users can query the back-end to retrieve data from the master file.
- The back-end responds with the requested contents, ensuring users always access the latest and consolidated data.

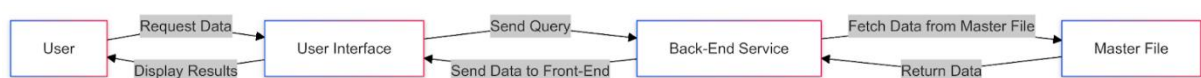


Figure 55 Data Access

Overall, the annotation (structured) table and the abstraction layer enable dynamic schema changes and easier management of data categories, making the system flexible. Users must strictly adhere to the evolving structured templates to ensure high-quality and data consistency. The master file consolidation approach allows the integration of new data without disrupting existing records, which offers scalability. Metadata tracking provides a clear audit trail of data submissions and modifications and finally the whole system can be adapted to include new data categories.

7.1.4 Post-Project Exploitation

For broader exploitation of the system, additional implementations could be required after the official project timeline. These enhancements would ensure the system's scalability and adaptability for future applications beyond the project's initial scope.

To ensure the long-term sustainability of GEOSSIP and the Citizen Science App beyond the project's completion, both tools have been designed with scalability and adaptability in mind. Developed by ENORA, the GEOSSIP platform will support analysis, and visualization, with the flexibility and functionalities as needed. Similarly, ENORA and UC, will support as for the Citizen Science App remain a valuable tool for community-driven environmental monitoring, with ongoing updates and support to engage new users and enhance its features.

7.2 Integration with OAH Open Information Hub (Task 6.1)

The project has developed an Open Information Hub based on ecosystem assessment findings and ongoing stakeholder insights. The Hub has a user-friendly information architecture that features clear and intuitive menus and sub-menus. The design prioritises easy navigation, enabling users to access relevant content seamlessly. Each section is clearly labelled to distinguish between the future project's tools to be integrated and other project knowledge resources, enhancing the overall user experience and accessibility. The OneAquaHealth Open Information Hub is a dynamic platform for stakeholders to comprehensively view the ecosystem and share stories by users, analysis results, and other knowledge materials. As the main communication channel, the Hub targets a diverse audience, including citizens, policy makers, authorities, researchers, civil society organisations, standardisation bodies, investors, industry representatives, business partners, and related projects and initiatives. The following part of the document outlines briefly the current status of the Hub and includes projections for future developments based on feedback from consortium organisations and external stakeholders.

7.2.1 Current state

The Hub was constantly updated due to the consortium's extensive communication and dissemination work in the past months but also based on the design and development of the project's tools. Subject to particular extension of the Hub were the sections "News", "Events", "Research Cities", "Knowledge", "Community", and "Project Info".

Sharing news from the field keeps the Hub current and engages stakeholders, building trust. The "News" section drives more visitors to the Hub and improves search engine rankings by regularly producing content. The Hub ranks for more search terms as contributions increase, boosting organic traffic. This section's clean design ensures accessibility across devices, and new items are automatically shared on social media, reaching a broader audience. Currently, the consortium partners have created more than 55 published news articles.

The "Events" section is a marketing tool that enhances the project's credibility and visibility by announcing important dates and raising awareness of activities. It has two sub-sections: Community Events, which highlights events and conferences for various stakeholders, providing industry updates, guidance, and networking opportunities; and Project Events, which features webinars (currently 4

events) for ecosystem stakeholders to discuss management efficiency, health status assessments, and health risk analysis. This section effectively communicates and disseminates project outcomes.

As noted in the previous deliverables (D7.1 and D7.2), the consortium created a section ("Research Cities") detailing research activities in five urban locations across Europe. These actions engage the public in science, evaluate the project's tools, and disseminate findings. These sub-sections have grown and will continue to expand with planned actions.

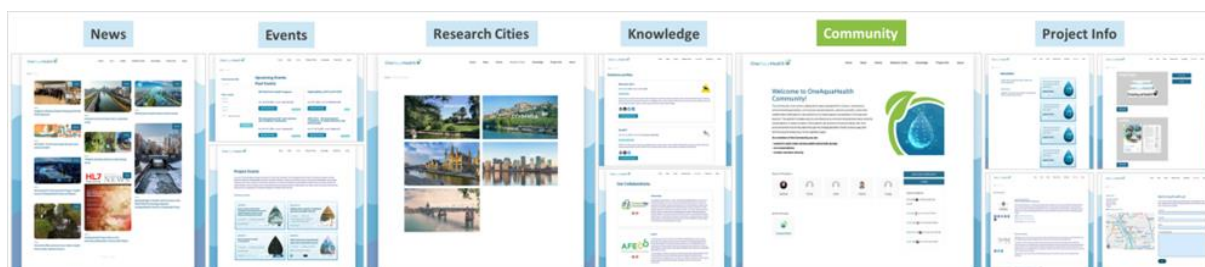


Figure 56 Snapshot of the current state of the Open Information Hub

The "Knowledge" section under "Solutions" includes citizen science projects and tools for public engagement in environmental monitoring and conservation. It centralises practical solutions, highlights the impact of public participation, and may attract more funding and support. It also showcases "Collaborations" with related initiatives, emphasising engagement, collaboration, and knowledge exchange, and may list specific partnerships or joint activities, as shown in Figure 56.

The "Project Info" section serves as a virtual space for all stakeholders to learn about and connect with the project, fostering a sense of community and shared purpose around the project's environmental conservation efforts.

7.2.2 Community module

The project developed the OneAquaHealth Community module within the Open Information Hub to support sustainable growth of the knowledge base on One health. This module facilitates networking and knowledge exchange among experts, providing access to up-to-date information, resources, and best practices. It enhances the visibility and impact of the project's research by showcasing digital tools and encouraging feedback. Connected via API with the CitizenApp, it allows members to submit and share local urban stream data, promoting EU-wide cooperation and knowledge exchange. The established cooperation among various stakeholders can be leveraged for ongoing needs analysis and monitoring of innovation outputs beyond the project's funding period. This collaboration also contributes to shaping policy initiatives and guidelines that extend beyond the project's duration.

In the following section of the deliverable, the different sub-modules of the "Community" will be shortly described.

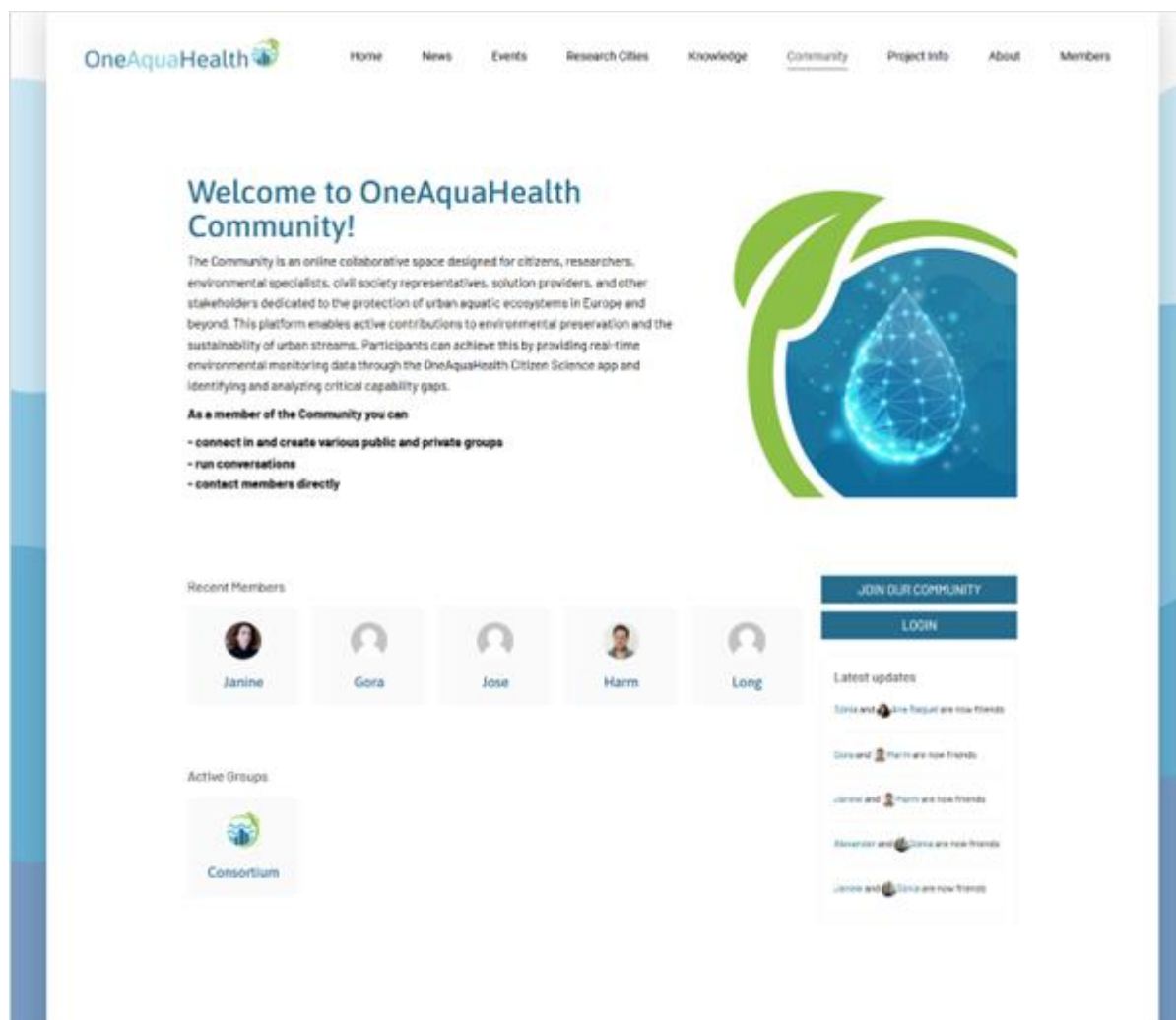


Figure 57 Snapshot of the Community overview page

The overview page (Figure 57) offers details about the "Community," which is designed for citizens, researchers, environmental specialists, civil society representatives, solution providers, and other stakeholders committed to protecting urban aquatic ecosystems in Europe and beyond. The page also provides information on recently registered users and created groups and showcases current activities. This encourages visitors to register, engage with the "Community" members, and support the success of the OneAquaHealth project.

Each member of the OneAquaHealth Open Information Hub Community has a "User profile" (Figure 58). This profile allows them to manage connected features according to their needs and facilitates communication with other registered members of the Community. Additionally, the user profile enables members to customise their experience, track their contributions, and access personalised content and updates. Importantly, each member's submission, which includes local urban stream observation data, is showcased and shared via the "Activity feed" with the rest of the Community. This feature not only highlights individual contributions but also fosters a collaborative environment where members can engage with and learn from each other's data and insights, enhancing the overall impact of the OneAquaHealth initiative.

The screenshot displays a user profile for '@jose' with a circular profile picture and buttons for 'Add Friend', 'Public Message', and 'Private Message'. Below the profile picture, it indicates the user is 'Active 1 day, 3 hours ago'. A sidebar on the left lists navigation options: Profile, Activity (highlighted), Notifications, Messages, Friends, Settings, Groups, and Invitations. The main content area is titled 'Member Activities' and shows a 'Submissions (1)' section. This section contains a table with various data points for a submission.

ID:	803
Longitude:	3.71431
Latitude:	50.88344
Research Site:	G10
Created At:	2024-12-04
Channel Form:	FLAT
Bottom Channel Type:	NAT
Banks Channel Type:	ART
Habitats:	RF, SB
Fallen Biomass Types:	FB, FT
Water Flow:	STA
Water Color:	MU
Water Abstraction:	Yes
Has Dams:	Yes
Number of Dams:	
Pipes:	No
Water Discharge:	Yes
Construction:	Yes
Water Height:	3.5
Overall Assessment:	MODERATE
Joy:	1
Serenity:	3
Anger:	5
Fear:	0
UpstreamPhoto:	
DownstreamPhoto:	No image available
SurroundingPhoto:	No image available
InterestingPhoto:	No image available

Figure 58 Snapshot of the "User profile" section with a showcased submission of urban stream observation data

The OneAquaHealth Community fosters a supportive environment for members to learn and contribute to urban aquatic and human health issues. Registered users can access the "Members" area (Figure 59) to view profiles, including names and user IDs, and establish connections. By adding friends, they can see each other's posts, comments, activities, and send private messages.

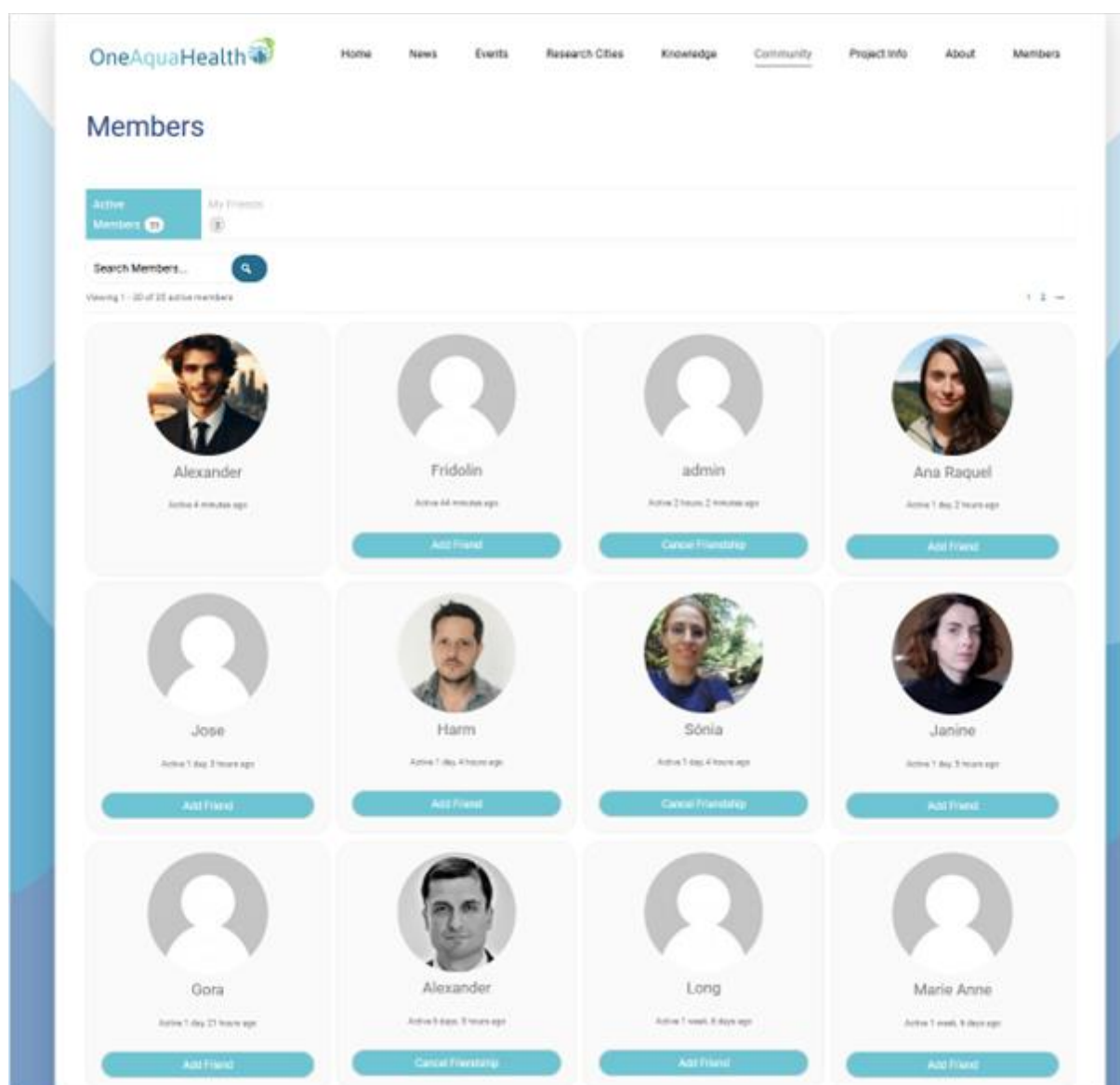


Figure 59 Snapshot of the "Members" section

The "Activity feed" (Figure 60) offers an overview of current activities, such as new comments or discussions, providing aggregated information to keep users updated on various topics. Freely accessible to the Community members, it showcases user activities, encourages participation, and supports the module's networking goals. As the Community grows, the Activity stream becomes increasingly valuable, making the OneAquaHealth Open Information Hub Community more resilient during and beyond the project.

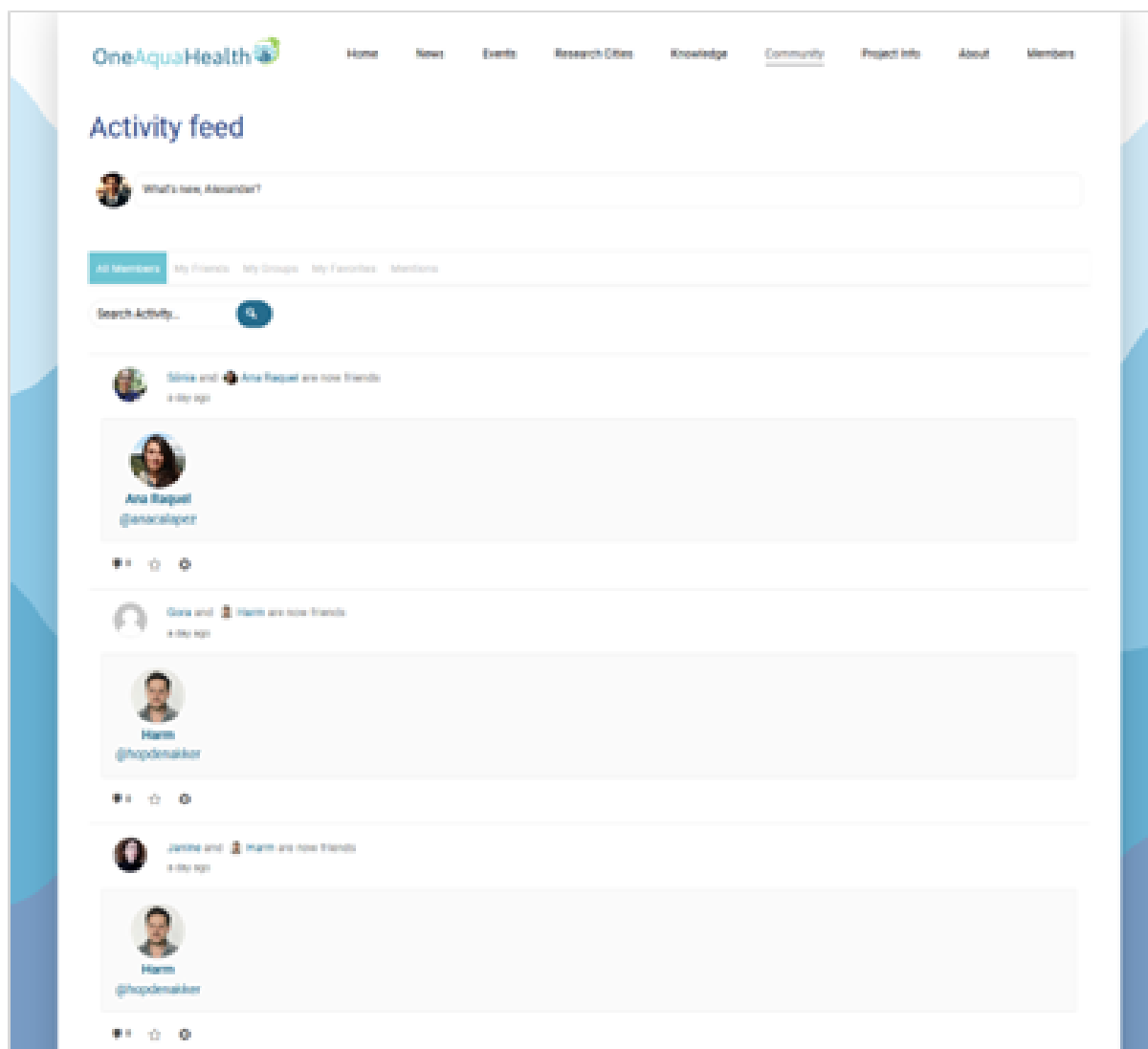


Figure 60 Snapshot of the "Activity feed" section

7.2.3 Future implementations

As the project's network expands and more knowledge and insights are gathered, the Community will grow in membership and output. A significant and planned development is the implementation of a dedicated page for visualising openly accessible City Dashboards (Figure 61). These dashboards are embedded via an iFrame directly into the Hub, allowing users to view real-time data and analytics without leaving the page. The iFrame is customised to fit the design and layout of the Open Information Hub, ensuring that the embedded City Dashboards align with the overall aesthetic and functionality of the page. Users can easily navigate and interact with the data, enhancing accessibility and engagement.

The OAH Citizen Science app, along with the dashboards, should be seamlessly accessible through the Hub. An installation instructions and a guide for the Citizen Science app will be provided. The content for the City Dashboards is currently under preparation, with the page set to launch in the next few weeks.

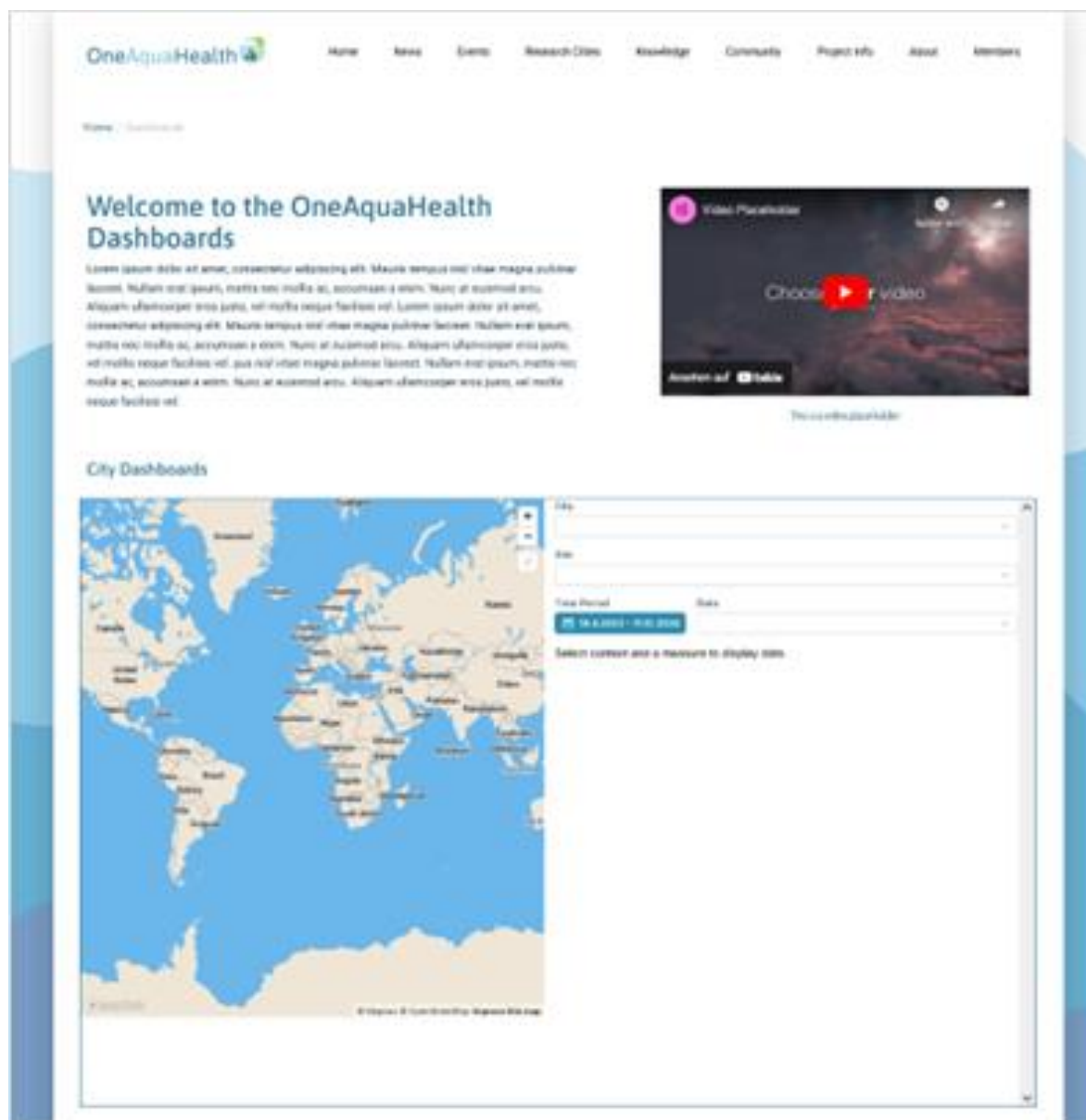


Figure 61 Snapshot of the preliminary design of the City Dashboards

7.3 Collaboration with OAH Decision-Support System Development Initiatives (Task 6.2)

Task 6.2 focuses on developing the OneAquaHealth Decision-Support System (DSS), a web-based Spatial Decision Support System (SDSS) to assist decision-makers and researchers. The system will generate spatial documents, maps, and reports and integrate multiple criteria decision analysis (MCDA). It will utilise early warning indicators and offer solutions for ecosystem rehabilitation to prevent health issues based on support from WP5. Harmonising the DSS with the OAH Geospatial and Satellite Information services will enable seamless data integration for better decision-making. EO services' geospatial data and spectral indices will support the DSS in assessing ecosystem health, simulating rehabilitation strategies, and providing actionable recommendations. This integration ensures the DSS is equipped to handle environmental challenges and support sustainable management. Furthermore, this will empower decision-makers to proactively address risks and

enhance the resilience of urban aquatic ecosystems through data-driven insights and predictive analytics.

7.4 Synchronising planned developments regarding OAH Predictive modelling and scenarios (Task 5.3)

Focuses on predictive modeling, Task 5.3 will assess the impact of environmental changes on ecosystem and human health. Using machine learning techniques and correlation analyses, the models will predict health risks based on environmental factors like water quality, vegetation, and hydrological features. These models will be tested across five research sites in Europe, helping to predict potential health risks and the effects of simulated environmental changes, such as water quality and climate variations. Supporting these models with OAH Geospatial and Satellite Information services ensures that real-time satellite data, such as spectral indices from EO services, can be integrated into predictions. This will enhance the models' ability to forecast environmental changes and their impact on human health, providing early warnings for risk management.

The models will also potentially incorporate citizen science app data, creating a dynamic system that continuously improves as new data is integrated. This ongoing process will make the models a key part of an early-warning system, supporting proactive health risk mitigation across Europe.

8 Ethical and Safety Considerations

8.1 Authentication mechanism

GEOSSIP employs a robust, multi-layered authentication and authorisation framework to safeguard data access and protect user privacy. It leverages OAuth 2.0 combined with JSON Web Tokens (JWT) for secure access delegation, enabling users to authenticate seamlessly. Authorisation is managed through Role-Based Access Control (RBAC), where user roles—such as admin, researcher, or citizen—define permissions and access levels. During the registration process, users create accounts by providing required details, ensuring their identity is established. Access controls are enforced across various layers, including API endpoints and data repositories, restricting interactions to functionalities and data aligned with each user's role. This comprehensive approach ensures both security and tailored access for all platform users.

8.2 Disclaimer and information of users for Citizen Science Application

The disclaimer of the Citizen Science App outlines the safety guidelines provided to the users of the Citizen Science App.

The developed disclaimer aims to safeguard that the use of the application's functionality aligns with the objectives of OAH and the safety and necessary precautions are regarded by the app's users..

By using the app, participants agree to participate voluntarily for scientific research, acknowledging that all activities, particularly those near water, carry inherent risks. The OneAquaHealth project is not liable for any injury or damage resulting from user participation. Health and safety guidelines are provided to ensure users' well-being during field activities, including precautions against potential exposure to harmful waterborne bacteria and plants. Users are informed that personal details are

register via Open Information Hub and their data are secured and stored offering them full control over which data they wish to share or only contribute for research purposes without sharing. Scientific research entails contributing to academic publications, environmental management tools, and educational resources. The draft disclaimer is provided in Appendix A of this document.

9 Scrum Summary

The development of GEOSSIP followed the SCRUM methodology, which was also applied to the Data Backoffice and the Citizen Science app for OneAquaHealth. This approach provided a flexible and efficient software development process within an agile framework, enabling iterative progress through time-boxed sprints. It allowed the team to effectively manage complex tasks, quickly adapt to changes, and incorporate feedback from stakeholders, researchers, and scientists.

To support SCRUM practices, JIRA was utilised for sprint planning, task tracking, and issue management, improving productivity and transparency by clearly highlighting progress and bottlenecks. BitBucket was also used for version control, facilitating smooth code collaboration and integration. Together, these tools helped maintain an agile development process aligned with project objectives. An example of this process is shown in Figure 62.

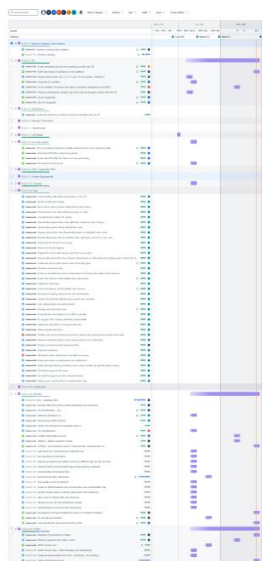


Figure 62 Jira tickets for the implementation

10 Next steps

Building on the results achieved so far, the following steps are to use the advanced visualisation tools of GEOSSIP to enhance the efficiency of the research methods and use EO data in the urban aquatic ecosystem research. The OAH model aims to establish a rigid method to assess the ecosystem health and the main priority for the coming period is to identify and validate the EO KPIs that could best support fluent ecosystem health assessment. The preliminary results obtained by the researchers are encouraging and show that certain phasmatic bands of the Sentinel-2 imagery seem promising in providing a preliminary assessment that is really close to the in-situ data analytics. The next version of

this deliverable will provide further details on this research. The refinement and improvement of the OAH Geospatial and Satellite Information Services API will continue in the coming period. The implementation of the local database to store calculations and establishing a local repository for satellite imagery of all sampling sites will significantly improve resource efficiency by reducing redundant computations and minimising repeated data fetching.

Another dimension that we are investigating is the use of AI-based preprocessing methods to enhance the accuracy and efficiency of data processing algorithms, improving the quality of the derived insights while reducing manual effort. Additionally, incorporating support for data streaming applications, such as Apache Kafka, will enable real-time data processing and analysis, ensuring that the system can handle high-velocity data and provide timely insights. Enhanced support and maintenance frameworks will be put in place to ensure the platform's reliability and scalability, including routine updates, performance monitoring, and user support to address any issues promptly. By implementing these steps, GEOSSIP will significantly enhance its capability to process and analyse geospatial data more efficiently, providing users with faster, more accurate, and actionable insights.

11 Conclusions and Future work

Future work for GEOSSIP should focus on expanding the platform's capabilities and enhancing its user experience. Efforts should be made to broaden the platform's scope to include more diverse data sources, create customisable dashboards for different user groups, and provide them to the OAH Open Information Hub. Collaboration with other research institutions and environmental organisations could be pursued to enhance the platform's relevance and impact. Finally, conducting user feedback sessions and usability studies would be important for iterative improvements and ensuring that GEOSSIP continues to meet the evolving needs of its users in urban planning and environmental monitoring.

The OAH Geospatial and Satellite Information services is an essential component of GEOSSIP solution aiming to support the processing, management, and visualisation EO data. Using modern technologies like Python for processing data, Docker Compose for creating consistent environments, and PostgreSQL for storing data, OAH Geospatial and Satellite Information services can efficiently handle large amounts of geospatial data. Key parts of the system, such as the Sentinel API for getting data, S3 Bucket for storing raw images, and services for converting images to GeoJSON format, work together to provide accurate and useful insights. The implementation is based on a modular design that uses microservices, supported by tools like Apache NiFi for managing data flow and Spring Boot for service deployment, ensuring it is scalable and easy to maintain. The data flow, from fetching and processing data to storing and displaying it, is optimised for efficiency and user experience.

Future improvements include improving the API, exploiting local databases and EO image repositories, integrating A.I. for better data processing, and supporting real-time data streaming with tools like Apache Kafka. Connecting OAH Geospatial and Satellite Information services with GEOSSIP and also with the Open Information Hub is expected to greatly promote collaborative research and the use of EO data.

References

- [1] OneAquaHealth - the potential of earth observations to monitor changes in urban aquatic ecosystems relevant for human health, Maria João Feio, Georgios Koutalieris, EuroGEO Workshop 2023, <https://egw2023.eurac.edu/>
- [2] Satellite observations as a proxy of urban streams ecological quality?, Maria João Feio, Janine Silva, Xenofon Karagiannis, Symeon Symeonidis, Sónia Serra, Ana Raquel Calapez, Georgios Koutalieris, XXII Congress of the Iberian Association Limnology, 2024.
- [3] Integrating multisource data for the assessment of urban aquatic ecosystems health, Maria João Feio, Georgios Koutalieris, Symeon Symeonidis, Xenofon Karagiannis, Sónia QR Serra, Ana Raquel Calapez, Janine P Silva, IGARSS 2024, <https://www.2024.ieeeigarss.org/>

Appendix

Appendix A

Draft Disclaimer – Citizen Science App

By using this citizen science app, you agree to the following terms and conditions:

Summary

You understand that your contributions are valuable to scientific research and agree to provide accurate and truthful information to the best of your ability. By using this app, you acknowledge that you are participating voluntarily in a citizen science project.

Health & Safety Advice

Activities near water always present a safety risk. Participating in the OneAquaHealth project by using the OneAquaHealth Citizen Science app is done at the user's own risk. The OneAquaHealth project and its partners are not liable for any damage or injury sustained by the app user or third parties.

We have designed the app activities and data collection to ensure that they are not dangerous or harmful in any way for the user. However, users are advised to exercise caution and common sense when participating in any outdoor or observational activities.

Please observe the following Health and Safety instructions.

Safety

Activities near water always present a safety risk. Please bear the following points in mind:

- Children should always be accompanied by an adult
- Keep a safe distance from the river or stream
- Rivers can rise quickly after heavy rain. Keep as far away as possible from a flood event.
- Banks may overhang the river and collapse
- River banks leading to water are often slippery so ensure you wear appropriate footwear
- Do not trespass on private property

- Do not climb or access barriers
- It is not recommended to enter the water
- Do not put yourself at risk in order to take a photograph

Health and hygiene

Areas of open water contain bacteria and other micro-organisms. The risk of contracting illness is small if sensible precautions are followed:

- Avoid swallowing river water
- Wash your hands before eating
- Cover all cuts and abrasions with waterproof plasters and avoid contact with the water

In addition, contact with some plants in the margins can cause burns or rashes. If you have such symptoms after touching a plant, seek medical advice.

Non-Harmful Intent

We affirm that this app will never intentionally request, encourage, or facilitate any actions that could cause harm to users, other individuals, animals, or the environment. If you encounter any content or requests that seem potentially harmful, please report it immediately through the contact on the OneAquaHealth website (office@oneaquahealth.eu).

User Responsibility for Submitted Images

We kindly ask you to ensure that the photos you submit include only the requested scenarios and adhere to the platform's guidelines, you are solely responsible for the images you submit. So, please, to protect privacy and uphold ethical standards, we encourage you to avoid uploading images of people under all circumstances. Failure to follow these rules may result in the removal of content or other appropriate actions.

Disclaimer for some AI-Generated Images

Some images showcased in the app were created using artificial intelligence tools, primarily Gemini AI, with a few generated using Journey AI Art, OpenArt Image AI, and Leonardo AI Phoenix. These images are subject to the terms of use of their respective platforms and are not original human-created works. Proper attribution has been provided to acknowledge the AI tools and platforms used in the creation process. If you have any questions about the image generation process or associated rights, feel free to get in touch.

Updates and Changes

We reserve the right to modify this disclaimer and the app's functionality at any time. Users will be notified of significant changes and continued use of the app after such modifications constitute acceptance of the updated terms.

By using OneAquaHealth citizen science app, you acknowledge that you have read, understood, and agreed to this disclaimer. If you do not agree with these terms, please refrain from using the app.